REDUCTION OF WIND INDUCED MICROPHONE NOISE USING SINGULAR SPECTRUM ANALYSIS TECHNIQUE

A SYSTEMATIC APPROACH TO IMPROVE SEPARABILITY IN NOISY ACOUSTIC SIGNALS FOR OUTDOOR ACOUSTIC SENSING

PhD Thesis

Author
Omar Hsiain Eldwaik

Supervisors
Dr. Francis F. Li
Prof. Trevor J. Cox

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

There are a few more people beyond the realms of the University whose friendship I would like to acknowledge; all my colleagues and friends back home. I must also pay a special thanks to some friends who I met here in the UK whose valuable advice and help were unforgettable.

On a more personal note, I would like to acknowledge with much appreciation the role of my whole family, brothers, wife, son and some other relatives for their endless support and patience. I am particularly grateful for their assistance and encouragement to get through this sustained and challenging period of time in most positive way. Their ability to always stay upbeat provided me with a balanced perspective that will always remain. Furthermore, this acknowledgement would not be completed without paying exceptional appreciation and heartfelt thanks to my older brother who passed away few years ago for all his support and assistance.

Lastly, dedicated to the memory of my parents, my beloved father and my dear mother who passed away last year and who always believed in my ability to be successful in the academic arena. You are both gone but your belief in me has made this journey possible. Without their care, moral support and untold number of sacrifices, I would not definitely be in the position I am today, and for that I am eternally grateful. Great eternal appreciation and enormous thanks are due to both of you.

DECLARATION OF AUTHORSHIP

I, Omar Eldwaik, hereby declare that this Thesis titled, ` Reduction of Wind Induced Microphone Noise Using Singular Spectrum Analysis Technique' and the work presented in it are entirely my own. Therefore, I confirm that:

- This work was performed mainly or wholly while in candidature for a PhD research degree at the University of Salford.
- In case where any part of this thesis has previously been submitted for a qualification or degree at the University of Salford or any other institution, this has been clearly indicated.
- In case where this thesis is based on work performed by myself jointly with others, this has been made clear and precisely stated to what extent the contribution of the others and what I have performed myself.
- Where any quotations have been taken from the work of others, the source is always provided. Apart from such quotations, this thesis is wholly my own work.
- Where the published work of others has been consulted, this is always clearly attributed.
- All main sources of assistance and help have been acknowledged.

Signed:

Date:

AUTHOR PUBLICATIONS

- Eldwaik, O. and Li, F.F., 2018. Mitigating wind induced noise in outdoor microphone signals using a singular spectral subspace method. *Technologies*, *6*(1), p.19. *MDPI journals.* Basel, Switzerland.

- Eldwaik, O. and Li, F.F., 2017. *Microphone wind noise reduction using singular spectrum analysis techniques*, in: 33nd Annual Conference and Exhibition. Reproduced Sound 2017: Sound Quality By Design, 21-23 November 2017, Nottingham, UK. In proceedings of the Institute of Acoustics, Vol. 39 Pt. 1.

- Eldwaik, O. and Li, F.F., 2017, August. Mitigating wind noise in outdoor microphone signals using a singular spectral subspace method. In *2017 Seventh International Conference on Innovative Computing Technology (INTECH)* (pp. 149-154). IEEE. This paper received a best paper award (see Appendix B for certificate).

LIST OF ABBREVIATIONS

AASP	Audio and Acoustics Signal Processing
BET	Behind-The-Ear
CNC	Computer Numerical Control
dBs	Decibels
DC	Direct Current
DSP	Digital Signal Processing
ECG	Electrocardiogram
EEG	Electroencephalography
EMD	Empirical Mode Decomposition
EOFs	Empirical Orthogonal Function
FFT	Fast Fourier Transform
HI	Hearing-Impaired
IMS	International Monitoring System
IoT	Internet-of-Things
IT	Information technology (related to computer-based information systems)
JSON	Java Script Object Notation
LPC	Linear Predictive Coefficients
MSSA	Multivariate Singular Spectrum Analysis
NH	Normal Hearing
NLOS	Non-Line-of-Sight
NR	Noise Residual
PAPR	Peak-to-Average Power Ratio
PCA	Principle Components Analysis
PBL	Planetary Boundary Layer
PCM	Pulse Code Modulation
PCs	Principle Components
PSD	Power Spectral Density
RCs	Reconstructed Components
RMS	Root Mean Square
SD	Standard deviation
SNR	Signal to Noise Ratio

SPL	Sound Pressure Level
SSA	Singular Spectrum Analysis
SVD	Singular Value Decomposition
WAV	Waveform Audio File Format
WSN	Wireless Sensor Networks
VAD	Voice Activity Detection
VAR	Vector Auto-Regression

SYMBOLS

As general remarks, mathematical formulations, symbols and notations are by the most part given in the time domain, unless otherwise specified. To clarify, italics and lowercase illustrations are used for signals in the time domain, whereas, italic uppercase is used to denote frequency domain notations. To indicate matrix and vector quantities, which are widely used in the text, bold uppercase is used for the former, whilst, bold lowercase is used for the latter. The standard typeface is used to show vector elements and individual matrix and the subscripts i and j are to indicate their position. For clarity, Roman and Greek capitals and lowercase letters are both used along with accents and other symbols.

The definitions of the symbols are tabulated below in which most the symbols implemented in the thesis are included except for the symbols used only in passing.

	Roman Capitals		*Lower-case Roman letters*
C	Covariance matrix	d	Number of elementary matrices
D	Dynamic range	ds	Statistical dimension
D	Eigenvalues diagonal matrix	f	Frequency
E	Matrix of right singular vectors	k	Delay shift (lag)
F_s	Sampling rate	k_1	Wave number in wind direction
I	Identity matrix	$\mathcal{L}$	Threshold specifies groups boundary
I	Specifies first group in grouping	m	Window length
L	Frame size	N	Trajectory matrix rows
Q	Crest factor	N_t	Time series length (data points)
S	Diagonal matrix of singular values	r	Limit of group selection
U	Matrix of left singular vectors	v	Kinematic (molecular) viscosity
V	Principle components matrix	z	Sensor height
Y	Trajectory matrix		*Accent and other symbols*
	Greek Capitals	$\bar{I}$	Specifies second group in grouping
Λ	Eigenvalues matrix	$s\,(n)$	Clean signal
P	Eigenvectors matrix	$\bar{u}$	Mean wind speed.
	Lower-case Greek letters	$w(n)$	Wind noise
λ	Eigenvalues	$X(t)$	Time series

ABSTRACT

Wind induced noise in microphone signals is one of the major concerns of outdoor acoustic signal acquisition. It affects many field measurement and audio recording scenarios. Filtering such noise is known to be difficult due to its broadband and time varying nature. This thesis is presented in the context of handling microphone signals acquired outdoor for acoustic sensing and environmental noise monitoring or soundscapes sampling. The thesis presents a new approach to wind noise problem. Instead of filtering, a separation technique is developed. Signals are separated into wanted sounds of specific interest and wind noise based on the statistical feature of wind noise. The new technique is based on the Singular Spectrum Analysis method which has recently seen many successful paradigms in the separation of biomedical signals, e.g., separating heart sound from lung noise. It has also been successfully implemented to de-noise signals in various applications.

The thesis set out with particular emphasis on investigating the factor that determines and improves the separability towards obtaining satisfactory results in terms of separating wind noise components out from noisy acoustic signals. A systematic approach has been established and developed within the framework of singular spectral separation of acoustic signals contaminated by wind noise. This approach, which utilises a conceptual framework, has, in its final form, three key objectives; grouping, reconstruction and separability. This approach is offered through introducing new mathematical models particularly for window length optimisation along with new descriptive figures. The research question has therefore been addressed considering developing algorithms according to updated requirements from method justification to verification and validation of the developed system. This thesis follows suitable testing criteria by conducting several experiments and a case-study design, with in-depth analysis of the results using visual tools of the method and related techniques.

For system verification, an empirical study using testing signals that introduces a large number of experiments has been conducted. Empirical study with real-world sounds has been introduced next in system validation phase after rigorously selecting and preparing the dataset which is drawn from two main sources: freefield1010 dataset, internet-based Freesound recordings. Results show that microphone wind noise is separable in the singular spectrum domain after validating and critically evaluating the developed system objectively. The findings indicate the effectiveness of the developed grouping and reconstruction techniques with significant improvement in the separability evidenced by w-correlation matrix. The developed method might be generalised to other outdoor sound acquisition applications.

1 CHAPTER ONE   INTRODUCTION

1

Introduction

1.1 Research Context

Recent and rapidly increasing innovation in acoustic sensing technology and Internet of Things (IoT) motivated the use of sound signatures to identify objects, sense the environmental variables and capture relevant events. Sound signal acquisition is one of the key stages in acoustics and audio engineering. In addition to common scenarios of audio recording; acoustic sensing assists daily activities, industrial operations and environmental management in many ways, including environmental noise and soundscapes monitoring as well as smart city applications. The ever-increasing research activities and demand of acoustic sensing in environmental sounds and soundscapes monitoring along with the hugely untapped potentials of acoustic sensing for added-value applications have driven the interest of this study. Outdoor acoustic sensing is particularly challenging as the transducers, typically microphones, are exposed to adverse weather conditions such as wind and rain; these might induce extraordinary noises in microphone signals.

Advanced acoustic sensing and signal processing is now being one of the key elements toward developing the field of audio and acoustics, giving birth to new technologies with tremendous exploitation potential as even innovations in high-tech are related to sound processing (Hancke, Silva and Hancke Jr, 2013). In outdoor data acquisition, there is a variety of environmental sound sources that produce different forms of sound as noise. One of the concerns is that much of acoustic data gets corrupted by different environmental sound sources such as wind and rain. Such noise-like sounds mask the useful content, hence capturing the target sound or the event of interest becomes extremely difficult. Improving de-noising techniques in order to reduce or remove unwanted sounds that affect target signals in outdoor acoustic monitoring still requires in-depth research and investigation.

Wind noise is a known problem that contaminates microphone signals in many field measurement and audio recording scenarios. Wind induced noise from microphone signals causes many problems in the subsequent use of acoustic information, however, this is one of the major concerns when applying acoustic sensing such as for environmental noise and soundscapes monitoring application. Wind noise problem is also an unsolved one in hearing aid applications and outdoor audio recordings such as field news broadcasting.

Microphone wind noise is a nuisance in audio recorded outdoor media applications. The problem can become more severe in acoustic sensing for environmental sounds and soundscapes monitoring and even smart environments applications: malfunctions can happen. For environmental sound monitoring, microphone wind noise contaminates and alters the soundscapes to be monitored, since the microphone noise is not what human listeners normally hear in the windy conditions, but the noise induced by the interaction between the wind and the microphones. For this reason, the terminology "microphone wind noise" and/or "wind noise in microphone signals" are used instead of wind noise in this thesis. This is to highlight that the noise of concern is induced by the presence of a microphone in windy conditions and such noise appears in the microphone signals.

In this thesis, a new method to mitigate wind induced noise in microphone signals has been developed. Instead of filtering techniques, wind induced noise is statistically separated in a singular spectral subspace. This thesis intends to identify whether the Singular Spectrum Analysis (SSA) method can mitigate wind noise artefacts to ensure separability for better handling microphone signals in outdoor data acquisition. By addressing this issue, this thesis explored the ways in which the SSA has been modified and developed to separate wind noise components in the eigen-subspace and reconstruct desired signals.

The SSA decomposes and projects time domain signals into the singular spectrum domain via Singular Value Decomposition (SVD), in which meaningful components such as trends and oscillatory components can be identified, isolated, re-grouped in a linear fashion, and finally time domain signals can be reconstructed. One of the advantages of this method is that it has the potential to retain wanted signals with less distortion when comparing with other known signal processing techniques. The SSA decomposes a time series into many component parts and reconstructs the series considering the meaningful components while leaving the noise component behind. Therefore, it is to track the changing spectrum of wind noise by considering the subspace corresponding to the higher-order eigenvalues in the singular spectra as a noise floor where the noise energy is concentrated. The SSA is a model free and non-parametric method, meanwhile, the only parameter that can be adjusted is the window length. However, along with the window length, grouping and reconstruction techniques are other key aspects to consider in the SSA.

This thesis aimed at identifying whether the SSA can effectively and usefully mitigate wind noise artefacts and enable better environmental sounds and soundscapes monitoring. The contribution in this study is targeting with systematic investigation and optimisation through

modifying and developing grouping and reconstruction techniques to improve the separability for enhancing wind noise separation and mitigation. Though, the separability may differ depending upon the dataset itself. In this thesis, an incremental methodology has been considered by gradually developing the method in terms of grouping and reconstruction techniques through performing a wide range of experiments in order to identify the separation approach and improve separability.

After the justification of the method when many experiments have been performed, an empirical study using testing signals has therefore been conducted for the verification of the developed system. Experiments started in the first phase with the development and implementation of the SSA for the separation of a mixture of deterministic signals such as sine wave, triangular wave, sweep tones, etc. In this phase, however, the method has been further developed to separate white noise and then wind noise from such testing signals.

Unlike the preceding procedure, in the second phase, the window length has been optimised and *w*-correlation matrix has been used to indicate the separability. Empirical study with real-world sounds has been introduced for system testing and validation. In this second experimental phase, the previously mentioned key elements in the SSA method have been considered in a different way. Therefore, SSA algorithm has been developed in many steps according to the technical requirements of developing grouping and reconstruction techniques to ensure the separability approach. In the system validation phase, it is to bring together the developed system for validation purposes and critical evaluation through adopting suitable dataset that links up to the application area of the thesis. The dataset consists of a variety of signals of interest such as car sirens, different alarm sounds, and birds' chirps as examples of the desired signals along with real measured wind noise samples.

1.2 Research Motivation

Overcoming some existing monitoring problems, specifically those related to excluding environmental noises such as wind from soundscapes signals, is an important research topic which still requires more investigations. The current complaints received from practitioners and researchers are mostly regarding the problem of not being able to obtain clean signals in outdoor acoustic soundscapes monitoring. Many existing noise reduction methods and algorithms have inherent limitations particularly for single-microphone wind noise reduction which will be explained in Chapter 3. However, such conventional algorithms provide

insufficient level of noise reduction (Jingdong Chen *et al.*, 2006; Dubbelboer and Houtgast, 2007; Hu and Loizou, 2007; Nelke *et al.*, 2014).

As previously mentioned, the implementation of acoustic sensing in outdoor acoustic data acquisition still has many limitations and challenges. Most of the recent applications of applying acoustic sensing such as soundscapes monitoring are commonly facing many problems mostly those caused by the presence of environmental noises. Therefore, the influence of the environmental noises on the sensed acoustic data is one of the main problems in the area, particularly when the noise sources are time-varying in nature and non-stationary such as wind noise. Furthermore, providing a quiet environment at the time of sensing is sometimes difficult if not impossible as such environmental noises occur naturally. Yet, this makes the effect of such environmental noises on the sensing process an unsolved problem; especially the existing methods did not come up with optimal and complete solution due to several limitations. This research is motivated by the enormously untapped potentials of acoustic sensing for added-value applications. It is also motivated by the recent dramatic increase of demand and activities of acoustic sensing in the context of soundscapes monitoring application.

The interest of this study has also been driven by investigating and developing noise reduction scheme for the SSA method under a fixed background noisy environment which is wind noise and different signals of interest. The research is also motivated to develop a new approach based on the SSA capabilities which is the separability instead of filtering. This will lead to a significant contribution to knowledge through the development of key elements in the method by establishing and developing a systematic approach of the method in the context of this thesis. In addition, the main points which increase the motivation of conducting this research can be outlined as follows:

- To get involved in developing methods for cleaning soundscapes signals by reducing wind noise as a serious issue of concern addressed by the research community. In this context, microphone wind induced noise reduction in particular, is an immensely interesting topic as wind noise intensely presents throughout a wide area compared to other most common noises and can last for long period during day and night times.

- As technology emerges and a variety of related supporting disciplines interact towards building smart environments with the contribution of acoustic and audio engineering

field, this makes the work in this area very exciting and could be collaboratively and competitively performed at the meantime.

- The enormous number of new expected applications, which aimed at improving the quality of life for inhabitants in many ways relying on the great advantages of acoustic sensing, makes conducting research in this area highly interesting.

1.3 Research Question, Aim, and Objectives

Research Question

Can Singular Spectrum Analysis effectively and usefully mitigate wind noise artefacts to ensure separability and enable better environmental soundscapes monitoring?

This research question has been formulated as the SSA is the proposed solution to the problem handled in this thesis. With the above in mind, the main aim of this study can be stated as follows:

Overall Aim

To identify whether a developed SSA can effectively and usefully mitigate wind noise artefacts to ensure separability and enable better environmental soundscapes monitoring.

Developing a more rigorous understanding of applying acoustic sensing in the context of environmental soundscapes monitoring and smart environments is among the desirable goals of this thesis since it is spawned from the lack of research in this area. Through developing the selected SSA methodology in terms of grouping and reconstruction techniques and the completion of the objectives listed below, this aim can be achieved.

Measurable Objectives

The central objective can be stated as follows:

Deploying an action plan by developing and modifying the SSA methodology to experimentally investigate its capabilities as a noise reduction method to be particularly developed for wind noise reduction in this thesis. Also, selecting suitable datasets for system verification and validation in a step by step manner considering developing key elements and approaches in the method.

1. Developing SSA algorithms for the separation of some testing signals from each other and from other types of noise in the system verification phase after justifying the method through particularly developing SSA algorithms using different examples.

- The method will be particularly developed in this thesis based on simplistic grouping technique and window length common calculation practice with no optimisation involved in the first phase.

- The method will be modified and extended via several state-of-the-art experimental methods using other types of noise as experiments; e.g., white noise before moving to wind noise so as to provide greater flexibility.

2. Developing experimental procedures to consider the important key aspects in the SSA in a different way.

    - The method will make use of window length optimisation approach as well as grouping and reconstruction techniques as key aspects in the SSA.

    - System verification, validation and evaluation will be performed by implementing the developed method using different wanted sounds from selected datasets through experimental and numerical studies.

3. Instead of filtering, developing and presenting the separation technique.

    - Introduce and develop the separation technique as a new approach based on the capabilities of the SSA and the statistical feature of wind noise.

    - Develop SSA algorithms to include this approach to separate noisy signals into wanted sounds of specific interest and wind noise.

    - Discuss the separability approach and include certain aspects such as the statistical dimension and nearly equal eigenvalues with a specified threshold for developing grouping and reconstruction techniques.

4. Combine the different methodologies, such as SVD, grouping and reconstruction techniques as a complete developed SSA system in a specific testing and validation platform particularly established for this thesis.

    - Provide an experimental demonstration and outline an experimental procedure by considering and adopting suitable datasets that link up to the application area contains real-world sounds (e.g., birds' chirps, car sirens, alarms, etc.).

    - Use of related digital signal processing techniques and objective measures.

    - Use of weighted correlation measure to demonstrate the results.

5. The verification, validation and critical evaluation of the developed system, (i.e. what expected to achieve and how to close the research loop).

- Ensure to what extent the developed system works for wind noise reduction through giving some comparative results and indicating its advantages and limitations.

- Make use of a specific wind noise dataset to systematically study and ensure the separation approach and the generalisability of the results.

- Show how the developed method can be generalised through linking up to the application area of the thesis using different samples from the dataset for critical evaluation.

- Report the results in such a way that they represent the experimental and numerical methods using multiple figures, detailed diagrams and tables that suitable to the case for systematic analysis and solid discussion.

1.4 Research Scope

In outdoor acoustic data acquisition, the environment is extremely uncontrolled, different types of noise will be picked up by acoustic sensors (microphones), and hence, the signal of interest gets contaminated by unwanted environmental sounds such as wind. In the presence of such environmental noise, the real picture of the sound can be immediately lost. Therefore, reducing the noise induced by the natural processes to enable long-term outdoor sensing of acoustic signals and improving urban soundscape design is a major problem.

Before establishing the systematic approach that has been developed for the SSA method and adopted for the empirical studies, a general layout of the framework of this thesis has been paned as presented in Figure 1.1. The focus of this thesis is mainly on wind noise reduction in the context of environmental soundscapes monitoring application particularly when developing the SSA method for this specific problem. The study sheds the light on some important concepts and key research components regarding developing and modifying the proposed method towards achieving the main aim of the thesis. This layout presented in the Figure 1.1 which summarises the problem of wind noise in microphone signals. Figure 1.1 also shows some desired signals taken as examples which will be used in the system validation phase as real-world sounds.
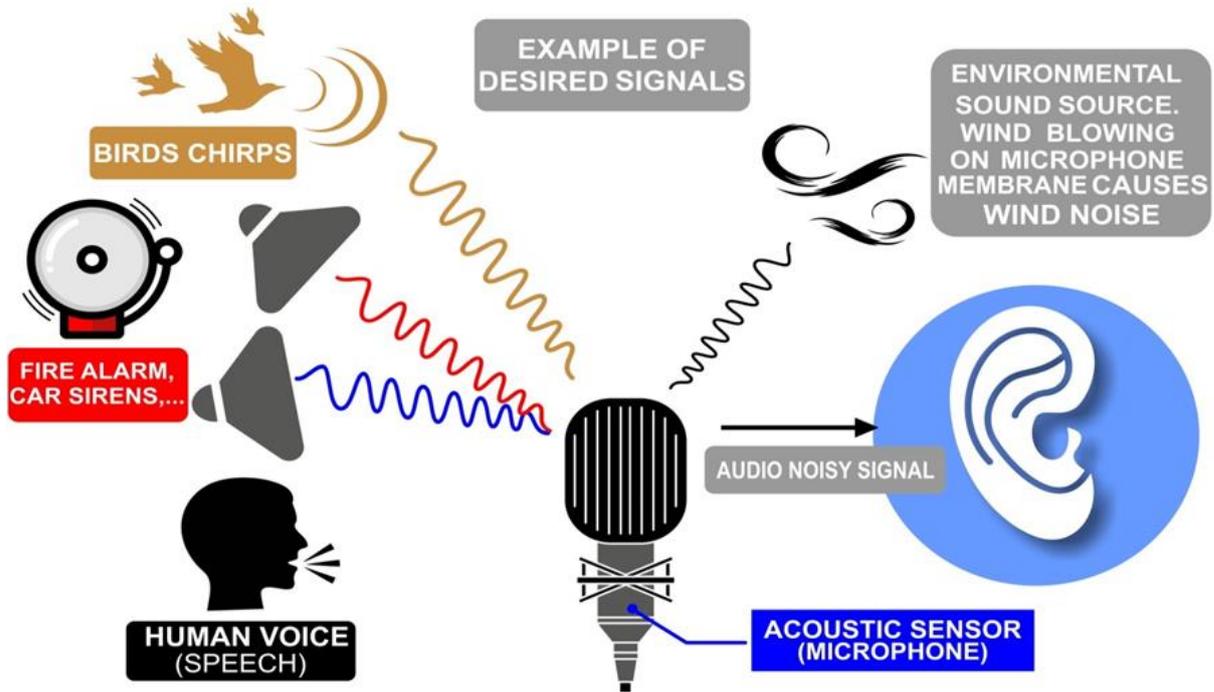
Figure 1.1. A general layout of the framework of the thesis indicating the problem of wind induced microphone noise

The long-term objective is to apply acoustics in smart city and outdoor soundscapes monitoring. Prior to that, it is essential to ensure that such acoustic sensing techniques are capable for sensing acoustic target data while reducing the effect of unwanted wind noise to enable further analysis. It is also important to decide which of the sensed acoustic data is considered as target acoustic events or signals of interest, however, this is mainly related to the application. Hence, within the context of this thesis, a large number of experiments has been carried out using examples of interesting signals from the environment such as birds' chirps and smart city signals such as alarm and sirens sound of ambulance cars or police cars.

It is worth mentioning that this thesis is not aimed at neither enhancing any of wind noise existing filtering methods nor including any comparison between these standards based on experimental investigation. Instead, the SSA has been developed and modified for the first time for wind noise reduction.

1.5 Contribution and Significance of the Research

This thesis attempts to respond to the recent call for research in the field of microphone wind induced noise reduction. In the meantime, it attempts to seek novelty through developing the SSA method and addressing such serious concern in the context of soundscapes monitoring

and smart city application. The focus is always on the stage where obtaining clean sensed acoustic data is required.

This thesis has contributed to the understanding of the singular spectral subspace method by developing the proposed method in terms of grouping and reconstruction techniques and introducing the separability as a new approach. This study makes a major contribution to research on grouping and reconstruction techniques by demonstrating the separability approach by which wind induced noise is statistically separated from wanted signals in a singular spectral subspace.

A systematic approach has been developed for the method with regards to all its key aspects and adopted in the framework of this thesis for wind noise separation as presented in Chapters 5 and 6. The uniqueness of this study exists in developing a novel and promising technique as well as giving much interest to a new interesting application area. This is the first study to undertake the SSA method to be developed and modified to particularly solve microphone wind noise problem. To demonstrate the potential of this approach and its suitability for the application, the *w*-correlation matrix has been used.

The long-term implications of this study might have an impact on the involvement of acoustics regarding smart city paradigm and monitoring applications towards a new aspect which is developing interactivity approach.

The significance of the research can be outlined as follows:

- This research has contributed to the understanding of the SSA method by expanding knowledge in this area and adding to a growing body of literature.

- This study has contributed to existing knowledge and will have a significant impact on the industry if the modified and developed version of the method can be implemented in real-world applications.

- This study has significantly improved acoustic sensing capabilities and helped in retaining wanted signals with no reconstruction errors which may lead to better enhancement of soundscape monitoring.

- It will be a sustainable impact if the method can be generalised to be applicable for other common outdoor data acquisition problems.

- It will be significant if the present study can make several noteworthy contributions to bioacoustics data analysis and automatic species recognisers as soundscapes signals contain various sorts of sounds to include sounds of birds and other kinds of animals.

- It will be substantial if the study can pave the way for pursuing further investigations to handle methods of manipulating acoustic data in the virtual new world objects and things that might be helpful for the improvement of urban environment soundscape design. This may lead to generalise a new approach for interactive urban environment which maybe significant as a long-term objective.

- As research is still carrying on for improving performance, efficiency and effectiveness and enhance functionalities regarding applying acoustic sensing and the improvement of sensing capabilities for detecting environmental sound sources and reducing environmental noises, this research will be beneficial to academics.

1.6 Research Methodology Implementation Procedure

Since the sensed acoustic data needs to be pre-processed to make it valuable for further analysis, the focus will be on the de-noising process by developing the SSA algorithm and improving certain key aspects in the method. Prior to that, a mixing model is required for generating mixed soundtracks for the simulation phases. Existing datasets such as published internet-based datasets that contain audio recordings which directly recorded from the field have been used. The verification process is extremely important to evaluate the specifications laid down during the development process.

The plan is to start with a simulation phase when adopting the developed systematic approach for the SSA in this thesis in a step by step manner. This phase is defined as a system verification when different testing signals generated and added together to have been used. Also, this phase has been gradually extended to involve white noise and finally wind noise. The next simulation phase was more sophisticated with real-world audio recordings and wind noise added for testing the developed system. This phase is defined as a system validation when realistic samples from another dataset have been used. A purposive sampling strategy has been adopted, for example, using realistic samples of field recordings from previous work at the University of Salford or internet-based published datasets. The research plan of this thesis is basically based on the following strategies.

- Use of existing datasets and realistic samples.
- Use of more than one dataset for multi-class problems and validation.
- Use of audio recording samples of same properties (e.g., length, sampling rate).
- Use of audio recording samples on interesting smart city sounds.
- Use of scholarly literature for theoretical and mathematical aspects.

- Use of software tools such as MATLAB and other useful tools to build a suitable working platform for system verification, testing and validation.
- Use of simulation and modulation techniques.

The research plan that has been outlined to implement the developed methodology of this research is mainly composed of four tasks; data collection and sample selection, the generation of mixed soundtracks to represent noisy data for the simulation phases by mixing desired signals with the defined wind noise, developing a systematic approach for the SSA method, developing and implementing SSA algorithms based on this approach for noise separation in different testing stages and experimental phases, and finally the validation and critical evaluation of the developed system and reporting the results as shown in Figure 1.2. However, further details will be given throughout the chapters of this thesis.

| **Dataset Preparation Testing Platform Requirements** | **Data Collection & Sample Selection (Audio Recordings)** |

| **Generation of Noisy Data (Before De-noising)** | **Developing and Applying Mixing Models, Averaging & Framing Methods** |

| **Developing SSA Algorithms based on Technical Requirements** |

| **Testing, Verification and Validation of the Developed SSA System (De-noising Process)** | **Justification of the Developed Method** |

| **System Verification** |

| **System Validation** |

| **Objective Measures and Analysis Methods (After De-noising)** | **Critical Evaluation, Reporting Results - Analysis & Comparison** |

Figure 1.2. Research methodology implementation

To experimentally investigate the developed method, the initial step is to focus on sample selection and dataset preparation, however, more details will be given in Chapter 7. The proposed data samples are mainly audio recordings samples (real-world sounds). The selected samples will be divided into two main categories, testing and validation datasets that contains multiple signals of interest is the first category while the second is wind noise samples. The next step is to develop the SSA for noise separation as the focus of the research. The research question of this study has been addressed considering an incremental methodology through developing the SSA algorithms according to updated requirements from one experimental phase to another. Also, the research question has been answered through performing many experiments using typical testing signals in the system verification phase followed by conducting several case studies using realistic samples for the final validation and critical evaluation. Figure 1.3 summarises the plan layout and overall strategy.

Singular Spectrum Analysis Method for Wind Noise Separation

Research Central Objective

- Deploying an action plan by developing and modifying the SSA methodology to experimentally investigate its capabilities as a noise reduction method to be particularly developed for wind noise reduction in this thesis. Also, selecting suitable datasets for system verification and validation in a step by step manner considering developing key elements and approaches in the method.

Plan Stages	Research Plan Sequence and Description
A) Investigation & In-depth Research	Reviewing literature, regularly updating and writing up the literature review and expanding knowledge in many related aspects
B) Training & Development Activities	Undertake interesting training courses and seminars for skills development along with group meetings activities
C) Research Participation	Research skills development by participating with the research community in the field, writing conference & journal papers
D) System Verification, Validation and Evaluation	SSA development and implementation, testing, verification, validation and system evaluation considering related techniques
E) Thesis Writing-up Phase	Writing-up the thesis in its final form by organising and updating the written materials performed throughout the program of study

Research Question

- Can Singular Spectrum Analysis effectively and usefully mitigate wind noise artefacts to ensure separability and enable better environmental noise and soundscape monitoring?

Research Aim

- This thesis intends to identify whether the singular subspace method can mitigate wind noise artefacts to ensure separability for better handling microphone signals in outdoor data acquisition for acoustic sensing applications.

Novelty and Main Research Contribution

The first study to undertake SSA to be developed and modified for wind noise separation through introducing separation instead of filtering as a new concept. A systematic approach includes new mathematical models and optimisation considering developing key elements which are grouping and reconstruction techniques in a new SSA version has been established and developed which improved separability and provided a satisfactory solution.

Figure 1.3. Research plan layout and overall strategy

1.7 Thesis Outline

Chapter 1
- This introductory chapter includes introduction to the thesis, research motivation and scope, research question, thesis aim and objectives, contribution and significance of the research, and research plan layout and methodology.

Chapter 2
- A literature survey chapter contains the current state-of-the-art and application of acoustic sensing concerning outdoor acoustic data acquisition challenges, acoustic signals, acoustic events and background noise, aspects regarding smart city real-world sounds and soundscapes monitoring. A comprehensive review of microphone wind noise, mechanisms of wind induced microphone noise and features, existing wind noise methodologies, limitations and main concerns.

Chapter 3
- This chapter is concerned with the methodology of this study. It includes concepts of Singular Spectrum Analysis, typical applications, discussion of the potentials of the SSA for microphone wind noise reduction and its significance as a relevant method to wind noise separation. Mathematical formulations and new models with working examples are also provided in this chapter .

Chapter 4
- This chapter is devoted to establish a framework of singular spectral separation of acoustic signals contaminated by wind noise .It also includes the theory and mathematical formulations that have been used to develop SSA algorithms along with some other details.

Chapter 5
- This chapter includes the development of central aspects (grouping, reconstruction, and separability) by adopting a systematic approach that has been established and developed within the framework of the SSA method of singular spectral separation of acoustic signals contaminated by wind noise.

Chapter 6
- Sample selection and dataset preparation with some other details related to the objective measures that have been used for evaluating the developed system.

Chapter 7
- System verification, an empirical study using testing signals and different types of noise, white noise and particularly wind noise for their separation from such testing signals. Results obtained from some selected experiments are presented in this chapter.

Chapter 8
- System validation, empirical study with real-world sounds and discussions are introduced in this chapter. This chapter includes results obtained from the validation phase using some selected experimental case studies along with results of window length optimisation followed by the critical evaluation using sound analysis, signal processing techniques, and w-correlation matrix.

Chapter 9
- This chapter is for overall conclusions and discussing suitable areas for future work.

Figure 1.4. A flowchart of the thesis outline

2 CHAPTER TWO LITERATURE SURVEY

2

Literature Survey

2.1 An Overview

An investigation into the state-of-the-art of acoustic sensing, acoustic signals, background information on acoustic events and background noise, urban environment aspects, and some other related concepts are offered in this chapter. The criteria used in analysing and comparing literature depends on several aspects such as ideas, specific purpose or objective, attitudes, structural design, techniques, theories, operating scenarios and performance, and conclusions of authors. Also, practical applications have been briefly considered to comprehensively assess the deployment of acoustics in real complementary target urban environments.

In spite of the considerable research work that has been conducted to automatically detect desired sounds and remove unwanted environmental sounds or reduce the harmful effect of environmental noises, microphone wind induced problem remains an issue of concern and requires more investigation. As previously mentioned, there is a recent call for research regarding monitoring problems in outdoor data acquisition for soundscapes monitoring and applications of urban environments. For this reason, important aspects regarding the application area have been included in this review chapter.

In this context, this chapter contains key elements regarding acoustic event and background noise as well as acoustic sensing challenges in urban environments. It also contains several state-of-the-art aspects and applications of acoustic sensing. The focus of the review in this chapter is basically on thesis topics and wind induced microphone noise. This chapter covers aspects regarding the mechanisms of wind induced microphone noise, wind theory; specifications and characteristics. The review also sheds the light on the most common principles in environmental noise reduction problem and existing wind noise reduction methodologies with their limitations.

2.2 Acoustic Event and Background Noise

Acoustics is known as an interdisciplinary field of science which deals with studying and analysing of sound waves and their propagation in different mediums, either in closed spaces, or in free spaces, or in channels and pipes. A wave that is propagating by means of vibrating

mediums such as solids, liquids, or gases is an acoustic wave. The impact of audio and acoustics can be seen in many applications as applied in almost all aspects of the society as it is the basis of many practical applications and fundamental phenomena (Kuttruff, 2006; Ballou, 2015).

For human beings, however, hearing is undoubtedly one of the main senses. Basically, acoustics relates to the sense of hearing. Therefore, there is no surprise to see the emergence of audio and acoustic science to become across several aspects of the modern society such as architecture, music, and industrial production, likewise, different species, such as birds for instance, use sound and hearing for marking territories or breeding rituals. Sound consists in mechanical vibrations within the audio frequency band of human hearing (Müller and Möser, 2012). An audio frequency is considered as periodic vibrations and the frequency of these vibrations is audible to human listeners. Environmental factors and the age can greatly influence the standard range of audible frequencies or in other words the range of frequencies that can be heard by humans. Frequencies below and above the audible frequency range are indicated by the terms infrasound and ultrasound, respectively (Kinsler *et al.*, 1999).

An acoustic event is timestamps in an audio stream and can be defined as a localised region or part of high intensity in a spectrogram. A sound event is also defined as a segment of audio that can be consistently labeled and distinguished by human listeners in an acoustic environment (Adavanne *et al.*, 2017). One spectrogram may contain several events where some of them are considered as calls of interest while the others are not. In environmental acoustic research, calls of interest are known as acoustic events. However, events that are not of interest are all named as background noise (Zhuang *et al.*, 2010). Background noise should be defined clearly because its definition is rather ambiguous. Sounds induced by rainfall, wind, distant traffic, and rustling of leaves are examples of continuous background noise. In outdoor data acquisition, much of the environmental noise are of different origin. For instance, wind noise, rainfall noise, rustling of leaves, etc., all are considered of physical origin; bird vocalisations and other animals are noises of biological origin; construction noise, airplane engine noise and highway traffic are considered as human generated sounds (Luther and Gentry, 2013). In this research, only continuous background noise is assumed through a time interval.

Noise is unwanted sound, hence, any acoustic event which is not of interest can simply describe noise. Measuring and quantifying the extent to which any given sound annoys any given individual is much more difficult (Ballou, 2015). Noise can be also defined as "a stochastic signal, emanating from an external source or phenomenon and disturbing the signal

of interest" (Sovijarvi *et al.*, 2000, p.604). These definitions lead to understand that the same acoustic source can be regarded as either noise or specific event. For instance, this implies if one is interested in birds' sound recognition, then wind and rain are considered as unwanted sounds. In fact, background environmental sounds, such as wind and rain, are not of interest in most of the applications and often rejected. In this study, wind noise in particular, represents the unwanted signal that affects the wanted acoustic event or the signal of interest.

In describing acoustic events, however, it has been found that acoustics deals with multiple steps, which start from the cause of the event to the final effects of sound while passing through the generation mechanism and control, transmission or propagation of mechanical vibrations, and reception of sound. The steps shown in Figure 2.1 can represent any acoustic event where the cause could be natural or volitional. There are various types of transducers that can transform energy into sonic energy and produce a sound wave. Sound waves carry energy through the propagating medium which later will be transduced to another form of energy in a natural or volitional way before the final effect. The presence of individual acoustic events characterises audio scenes. A multi-class description of an audio file can be managed by the detection of the categories of acoustic events that occur in the file. Acoustic event detection deals with processing acoustic signals and transforming them into symbolic descriptions that correspond to the perception of the listeners of different acoustic events presented in the signals and their sources (Temko *et al.*, 2006; Mesaros *et al.*, 2010).
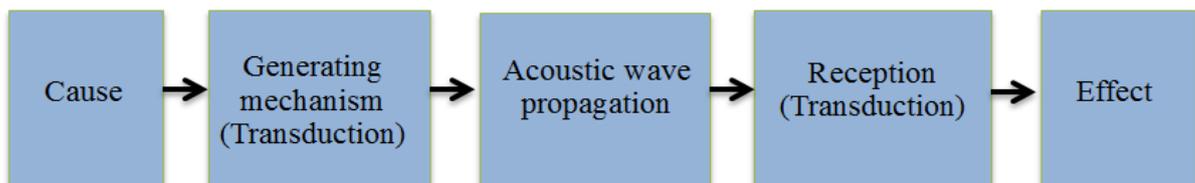
Figure 2.1. Acoustic event description

Soundscapes signals are composed of acoustic target of interest and background noise such as wind. So far, however, there has been little discussion about developing new contemporary methods for wind noise reduction in the scope of cleaning soundscapes signals and smart city and monitoring applications. In such applications, wind background noise undesirably masks valuable information and harmfully affects the acoustic event or the signal of interest, which is varying from one application to another, and makes the annotation of the target sound more difficult. If the background noise can be masked or separated, the detection of acoustic events of interest will be much easier. Hence, the perceived quality of the captured

acoustic data can be improved; subsequent processing and data analysis methods will effectively be deployed.

2.3 Acoustic Sensing in Urban Environments

In the last few years, research efforts in the field of Wireless Sensor Networks (WSN) have shown high potentials for surveillance applications and have paved the way to what is called ubiquitous or otherwise known as global sensing and smart cities paradigm that extends WSN to more generic Internet-of-Things (IoT) concepts. Applying acoustic sensing in urban environments to make cities such smart and probably interactive is considered as a great contribution of acoustics in attempting to improve the quality of life in many ways (Mao, Fidan and Anderson, 2007; Misra, Reisslein and Xue, 2008; Pham and Cousin, 2013). Edge computing and social IoT concepts have been introduced for the development of large-scale smart environments (Balakrichenan, Kin-Foo and Souissi, 2011; Cicirelli *et al.*, 2018).

Outdoor urban environments have been considered in previous literature and some researchers have worked on that. Developing and implementing acoustic sensing to collect data for outdoor applications, such as monitoring pollution, crowding areas, movement in restricted areas or some surveillance type of applications (e.g., route accidents and detection of unusual situations) are some of the pioneering applications that impact the daily lives of smart city inhabitants (Hancke, Silva and Hancke Jr, 2013; Luzzi *et al.*, 2013). In addition, there are some projects that have been recently launched to provide enhanced cyber-physical services when deploying acoustic sensing technology, such as traffic estimation and conjunction, detection of notable acoustic events in the city (e.g., sirens, accidents), localisation of events using multi-purpose acoustic devices, looking especially at surveillance type of applications. Also, there are some individual sound sources recognitions, environmental sound and noise recognition along with a considerable number of event sound and noise classifications. However, the existence of some monitoring problems due to the presence of environmental noises has made the implementation of acoustic sensing in urban environments for monitoring and smart city applications extremely challenging (Sanchez *et al.*, 2014).

In the context of this thesis, a graphic design that represents some city services, control and monitoring setup when acoustic sensing technology is involved is shown in Figure 2.2. This architectural design has been produced to be used in many ways as a model that incorporates some novel design features. Also, this design indicates microphone wind noise problem in terms of the presence of wind and its harmful effect on the sensed or recorded

acoustic data. The most interesting aspect in this design is that closer inspection of the main idea and research problem of this thesis can be revealed. Therefore, some real-world sounds that have been used in the system validation phase, such as police car siren, ambulance siren, fire alarm, and birds' chirps are presented. Figure 2.2 also shows a symbolic representation of speech signals in terms of press coverage and reports for example; however, speech is not considered in this thesis.

Figure 2.2. An architectural design of acoustic sensing topology in smart city and monitoring applications

Acoustic sensing technology can be applied in outdoor data acquisition for monitoring the soundscapes and background noise level as well as controlling the environment in many ways for some practical applications. However, controlling the whole urban environment is a broad topic and an interdisciplinary approach which go beyond the scope of this study as it requires implementation to be performed and consequently appropriate and extensive facilities will be needed. Acoustic sensing technology could derive benefits to the inhabitants of the new urban environments regarding socio-economic development. Hence, in this chapter, the study briefly sheds the light on how acoustic sensing is increasingly becoming important to deliver business efficiency and reliability and how it is promised to be widely implemented in many smart city applications and soundscapes monitoring.

2.3.1   Ubiquitous Computing and New Urban Environments

New urban environments, which can be considered as smart environments, are physical spaces enriched with various communication devices and sensors. Sensors are designed to detect different events and capture signals as raw data according to the type of the environment to manipulate and process this data by means of certain devices provided with specific software tools and signal processing techniques. This is to transform the captured raw data to meaningful information accumulated as knowledge to be exchanged among different communication devices to provide a variety of services. These services or applications are aiming at supporting the users and helping them in their activities. Smart urban environments can be considered as an application area of the paradigm of ubiquitous computing that could support collaboration, enhance productivity, or even boost creative activities (Komninos, 2009; Cicirelli, Fortino, *et al.*, 2017).

As stated in (Cook and Das, 2004), Ubiquitous Computing is one of the catching words along with Pervasive Computing and the new term Ambient Intelligence. The vision of the world has probably been paraphrased because of these terms that share the notion of smart, in which smart and intuitively operated devices surround the inhabitants of the smart city helping them in organising, structuring, and mastering their everyday life (Weiser, 1991). The introduction of these concepts causes major changes which could not leave residential environments unaffected. However, such environments border with many ubiquitous computing applications such as e-health or smart metering that have to be integrated even into the so-called smart home ecosystem (Fysarakis *et al.*, 2018).

A new paradigm regarding the interaction between a person and his everyday surroundings is characterised because of these new concepts. These surroundings are enabled by the smart environment to become aware of the goals and needs of humans who interact with it. A smart environment is formed by the continuously connected physical world and objects with computational elements (Ahmed *et al.*, 2016). In other words, inhabitants' activities could be proactively performed, their goals reached, and their tasks fulfilled in a user-centric manner with the assist of such environment where underlying technologies have transparently been integrated. The central themes of ubiquitous computing are user adopted behaviour and personalisation (Cook and Das, 2007; Fysarakis *et al.*, 2018).

Ubiquitous computing which is also known as global sensing including acoustic sensing is generally referring to technologies where many sensors, communication and mobile devices

and networks integrated into everyday artefacts in an environment. Such components interact seamlessly and spontaneously with each other and with users in a context-driven manner. Based on this vision, a variety of related paradigms have been developed. Each of these paradigms is emphasising on certain aspects of ubiquitous computing such as pervasive computing, ambient intelligence, and smart environments where acoustic sensing and other technologies are applied (Pham and Cousin, 2013; Fysarakis *et al.*, 2018).

Recently, acoustic sensing technology and sound processing are becoming leading research topic regarding developing new urban environments. There is a rising popularity of the topic and a growing desire for successful projects in the marketplace (Cook and Das, 2007). One trend in smart urban environments is the endeavour to enhance inhabitants' quality of life by investing in Information Technology (IT) solutions. A city that is smart needs to invest in, therefore modern communication infrastructures are required. To build a smart urban environment, involving a variety of applications, different techniques, architectures, algorithms, and protocols have been introduced in some previous studies and research as listed in (Cook and Das, 2004). However, sound processing and acoustic sensing technology have been involved in recent studies and projects (Sanchez *et al.*, 2014).

2.3.2 The Key to Success in New Urban Environments

Designing smart urban environments is a goal that appeals to researchers in a variety of disciplines and supporting fields, including pervasive and mobile computing, sensor networks, artificial intelligence, robotics, multimedia computing, middleware and agent-based software (Cook and Das, 2007; Cicirelli *et al.*, 2018). A tremendous increase in the number of smart city projects has been prompted due to the advances in these supporting fields. Smart city solutions need to be envisioned from both technical and sociotechnical perspectives to understand the impact of the technology on inhabitants' lives (Eckhoff and Wagner, 2018). According to (Schaffers *et al.*, 2011), smart cities are often considered successful when they are able to combine and balance factors such as economy, mobility, environment, people, living and governance.

Smart environment is defined as "one that is able to acquire and apply knowledge about the environment and also to adapt to its inhabitants in order to improve their experience in that environment" (Cook and Das, 2004, p.3). According to (Dey, Abowd and Salber, 2000, p.1) "one of the goals of a smart environment is that it supports and enhances the abilities of its occupants in executing tasks". A more user centric definition is stated in (Aarts and

Encarnação, 2006, p.322), "Smart environments are physical spaces that are able to react to the activities of users, in a way that assists the users in achieving their objectives in this environment". Large-scale smart environments are defined as pervasive and distributed dynamic systems that cover a wide geographical area when characterised by a large number of interacting devices of heterogeneous nature (Cicirelli, Guerrieri, *et al.*, 2017). In this context, agent-based approach has been introduced in (Cicirelli *et al.*, 2018) to leverage Edge Computing and social IoT paradigms for the development of large-scale smart environments.

Through the choices that citizens of new urban environments have and decisions they make, their participation along with businesses and other stakeholders is essential in shaping their future. The challenge therefore is to redefine the smart city as an environment of innovation and empowerment. Hence, the focus will be on the changes and transformations towards a smarter city in the sense of shaping a better and more participative, inclusive and empowering city, instead of imagining an ideal future (Schaffers *et al.*, 2011). The type of experience which individuals wish from their environment varies based on the individuals themselves and the type of the environment. Inhabitants may wish the environment to ensure safety for them or to reduce the cost of maintaining the environment, also, they may want tasks that they typically perform in the environment to be automated. Optimising the resource usage (e.g., utility/energy bills or communication bandwidth) could be also another wish (Das and Cook, 2005).

2.4 Outdoor Acoustic Data Acquisition Challenges

An acoustic wave from the physical standpoint is one that is transmitting and propagating as vibrations by means of vibrating a medium like air. Acoustic waveforms are complex by nature, however, for humans, these waves contain information that is significant to them (Ballou, 2015). A variety of valuable information can be extracted by analysing the data that is transmitted as physical waves from the large-scale data production of the real world. As audio and acoustics field continues to gather steam, the emergence of gadgets equipped with all kinds of sensors to help in improving our daily lives can be clearly seen.

Technologies for acoustic processing that handle general sounds including speech processing of human voices are promising media processing technologies that together with image/video processing are expected to contribute to the solution of various issues and could be very helpful in soundscapes monitoring application. In spite of the valuable research work and unlike other types of data, image or video for instance, there is still a huge untapped

potential for acoustic processing particularly in applying acoustic sensing towards creating new urban environments (Luzzi *et al.*, 2013).

Acoustic sensing technology can be applied outdoor to capture acoustic data for many applications, such as bioacoustics data analysis, monitoring pollution, and some other pioneering applications (e.g., surveillance type of applications) that impact the daily lives of smart city inhabitants (Luzzi *et al.*, 2013). For example, in bioacoustics data analysis, scientists who are interested in environment changes monitoring and birds' calls, which are of interest in outdoor sound sensing, will rely on deploying various types of acoustic sensors in the field. Distance outdoor microphones could help ecologists record whatever sound they want instead of conducting standard surveys. Although such method can bring some advantages over classical surveys, such as saving time and efforts, providing continuous recordings, and scaling over long period and huge area, however, the data collected will include much of background noises that make annotation of bird vocalisations more difficult. Hence, it is to consider microphone wind noise as a serious problem that affects the sensed acoustic data in outdoor sound acquisition. The effectiveness of bioacoustics data analysis will be increased if background noise can be masked or removed (Slabbekoorn, 2013; Grill and Schluter, 2017).

Improving and developing methods for cleaning soundscapes signals and reducing the harmful effect of environmental noises, such as wind noise, on the sensed or recorded acoustic data of interest by effectively separating such noises is a great challenge in outdoor data acquisition.

2.4.1   Motivations and Ambitions of Applying Acoustic Sensing

Using a technology that brings along with new enhanced services and developed applications towards the goal of creating smart cities, new solutions that could help in overcoming the existing socio-economic and financial problems and producing new products is now a great challenge for researchers in the field of acoustics. A technology that could even help in minimising potential hazards along with developing functionalities over the existing levels is another great challenge. This technology is represented in acoustic sensing technology, however, when combining with some other technologies in an interdisciplinary approach, it could help in achieving the central long-term aim which is building smart cities and monitoring the soundscapes provided that better solutions can be found for the existing problems regarding environmental noise reduction.

Integrating acoustic sensing with technologies that are emerging and getting developed day by day and making efforts to involve acoustic sensing when discarding environmental noises, such as wind, are all great ambitions that increase the motivation of conducting this research. The area of global sensing and smart cities witnessed a growing importance and a number of leading projects have been launched recently in which the objective is to develop and test acoustic sensing technologies in a city context (Sanchez *et al.*, 2014; Eckhoff and Wagner, 2018). Bringing acoustic sensing to large-scale outdoor wireless sensor networks where a large scale deployment of sensors in the city is provided is a big challenge in such projects (Pham and Cousin, 2013).

A smart city with its modern communication and information systems should be the perfect place where to exploit the existing technologies and develop awareness on environmental problems. The current vision about urban performance for most of cities depends mostly on the city's endowment of hard infrastructure, the ("physical capital"), which is decisive for urban competitiveness. Although, particular importance has been increasingly given to the availability and quality of knowledge exchange, information communication and social infrastructure which stated as ("intellectual and social capital") are facing many challenges (Moreno, Zamora and Skarmeta, 2014).

Recently, creating smart urban environment using acoustic sensing technology has drawn the attention of many researchers. There are several motivations behind involving acoustics in the development of smart city projects. However, one of these motivations is to overcome current limitations and introducing new methods to improve certain functionalities in this important sector. The availability of a huge untapped potential in sound processing and acoustic sensing, the advantages of acoustic sensing, the applications developed due to the development in certain related technologies such as wireless mobile communication, sensor networks, and future internet, all lead to consider the contribution of acoustics in developing smart city solution is by no means essential (Sanchez *et al.*, 2014).

The main idea of deploying acoustic sensing technology is to focus on effectively sensing sound signals produced by many sound sources that already exist in the city that considered of particular interest for several applications. The captured acoustic signals are composed of acoustic events of interest that occurred continually or even continuously in the urban environment and other signals considered as background noises. The automatic detection of environmental noises, such as wind noise, followed by applying methods to reduce these noises

and improve smart city and soundscape monitoring applications remain a great challenge in this field.

According to (Luzzi *et al.*, 2013), reducing noise or protecting a space from a noise source at receivers' level, who are generally inhabitants of a building or a specific area in the city, becomes a critical issue in introducing smart cities as an innovative solution for improving the quality of life in many ways. However, the involvement of acoustic sensing for more effective integration between energy, transport, and communication technologies sectors in smart cities becomes important. Along with developing other smart city dimensions, the main aim is to enable innovative as an attempt to place cities at the centre of innovation for acquiring and applying knowledge and soundscapes monitoring.

2.4.2   A New Prospective in Acoustic Sensing

One of the new concepts concerning the deployment of acoustic sensing in urban environment according to (Pham and Cousin, 2013) is based on the contribution to the so-called global sensing or situation-awareness applications as smart environments have emerged as an efficient infrastructure. Temperature and luminosity, for instance, were so far the traditional scalar physical measures proposed by most of the deployments for several environment-related applications. However, the new ideas of involving acoustics in smart city paradigm are promising to move a step further towards large-scale "real-life" experimentations of acoustic sensing as proposed for supporting high societal value applications.

New innovative range of services and applications mainly targeting to smart cities and even smart buildings are proposed to be delivered which all indicate such importance of applying acoustic sensing in urban environments (Pham and Cousin, 2013). For example, acoustic source localisation requires real-time simulation work using audio sensing technology for wireless sensor networks (Hollosi *et al.*, 2013). In (Kotus, Lopatka and Czyzewski, 2014), a method for automatic determination and localisation of sound events in the presence of sound reflections through employing acoustic vector sensors for smart surveillance applications is presented.

One of the important operating scenarios while applying acoustic sensing in smart city paradigm is the on-demand scenario which is typically intended for users requesting acoustic data on well-identified parts of the city. Acoustic data streaming feature can be used for many applications, such as surveillance systems, management of emergencies, event detection and some other applications as shown in Figure 2.2. Therefore, acoustic capture system has to be

26

developed by providing numerous of efficient and powerful sensors at different density that provide a large-scale coverage of the city. Also, the capability of these sensors within the environment could help the users when requesting acoustic data from a set of the capture system nodes to improve their understanding of the situation (Pham and Cousin, 2013). Such nodes are capable to respond in some ways to the enquires coming from the human operator through a mechanism known as streaming encoded acoustic data on lower sources devices. This means that streaming acoustic data can be realised in a multi-hop manner on such infrastructure (Mao, Fidan and Anderson, 2007; Luzzi *et al.*, 2013; Sanchez *et al.*, 2014).

In all cases and for whatever type of desired signals or application, removing unwanted sounds and reducing environmental background noise to obtain clear target acoustic data is a challenge that requires more investigations to improve the existing methods. Therefore, the Singular Spectrum Analysis has been proposed in this thesis to be developed for wind noise separation as a specific type of environmental noises. The de-noising process takes place in the preprocessing phase with regards to the whole system infrastructure and used for this specific purpose which is mainly the central focus of this thesis.

Setting up acoustic sensors which can detect target acoustic events and then trigger other specific sensors to track the source of the sound in some applications is the way to collect and send Metadata which then can be used to perform actions. In such applications, environmental background noises such as wind noise, which has a harmful effect on sensed acoustic data and a destructive influence on the target sound, needs to be discarded or reduced. With regards to the whole system when involving acoustic sensing technology in such applications, de-noising methods can be used in the low-level analysis in the preprocessing stage as shown in Figure 2.3 (Sanchez *et al.*, 2014).
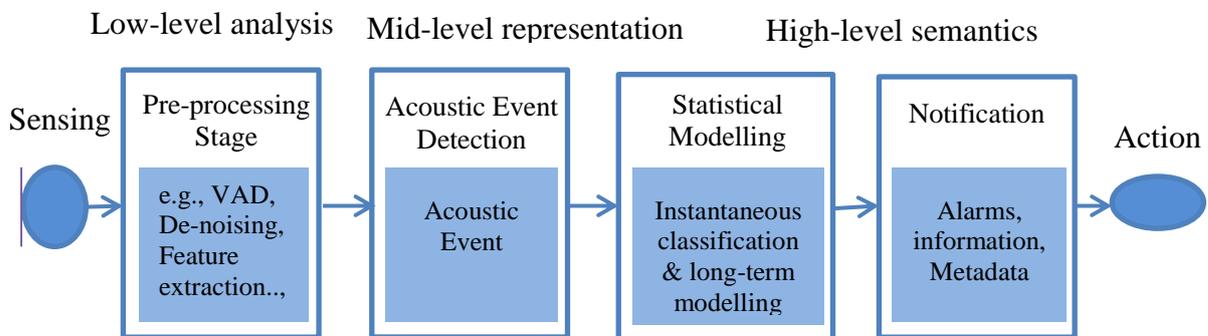
Figure 2.3. A systematic approach of implementing acoustic sensing technology in urban environments

The whole processes that start from sensing acoustic data using acoustic sensors to perform actions are presented in Figure 2.3. However, the focus is only on developing a specific method based on the SSA for developing the separation approach particularly for wind noise separation with no details provided regarding the other processes as they go beyond the scope of this thesis. The main purpose of presenting Figure 2.3 is to show where noise reduction methods can be involved with regards to the whole system of deploying acoustic sensing technology for smart city applications and soundscapes monitoring.

Although applying acoustic sensing techniques in new urban environments is very promising in terms of delivering a variety of innovative applications, however, some monitoring problems are still there. In spite of the considerable work that has been undertaken over the last few decades in the field of environmental noise reduction, still there exist significant gaps in the literature that impede the further development of wind noise reduction methods among other experimentally based broadband noise methodologies towards an optimal de-noising solution. Such gaps require more in-depth research and detailed investigation because of the limitations of the existing methods as will be discussed later in this chapter. For this reason, therefore, the proposed method for the separation of wind noise in this thesis is based on developing the novel SSA technique due to its successful paradigm as a de-noising method in some other applications.

2.5 The State-Of-The-Art and Applications of Acoustic Sensing

2.5.1   Acoustic Signals

In the literature, many state-of-the-art techniques that are based on acoustic signals have been applied in many applications. Recently, acoustic signals are used in diagnostic techniques of machines in the field of industry and engineering where many rotating machines, such as electric motors, are used. Diagnosis of such motors is considered as a normal maintenance process (Glowacz *et al.*, 2018). Nowadays, using acoustic signals is an up-to-date method for many applications of fault diagnosis and localisation in rotating machines. Therefore, to keep up the complex processes of industrial production, timely maintenance and fault identification of rotating machines, particularly electric motors, have been made possible by using acoustic signals. For example, in (Delgado-Arredondo *et al.*, 2017), a fault detection approach in induction motors in steady state operation by analysing acoustic sound and vibration signals has been introduced. The Complete Ensemble Empirical Mode Decomposition has been

proposed to decompose the detected acoustic signal into several intrinsic mode functions by which good fault detectability results have been obtained and more frequencies associated with the faults have been identified (Delgado-Arredondo *et al.*, 2017). Also, acoustic signals are used for early fault diagnosis such as for automatic bearing fault localisation (Jena and Panigrahi, 2015).

A pattern recognition system was proposed and developed based on acoustic signals for automatic damage detection on composite materials where Singular Value Decomposition (SVD) method was used to filter acoustic signals (O'Brien *et al.*, 2017). Analysis of using acoustic signal processing allows recognition of rotor damages in a DC motor (Głowacz and Głowacz, 2017). In practice, acoustic signals have been used in many cases for early diagnosis of electric motors such as rotor damages of three-phase induction motor (Glowacz, 2016). A description of monitoring acoustic emission-based condition is given in (Caesarendra *et al.*, 2016). Online monitoring of machines allows for intelligent maintenance with the optimised usage of maintenance resources (Glowacz *et al.*, 2018).

In work environment and industrial conditions where many machines are used, workers are exposed to noise levels that often exceed permissible values. At such workplaces, noise is defined as any undesirable sound that causes harmful or tiring to human health (Mika and Józwik, 2016). The measurement of noise becomes increasingly important to maintain noise levels under the permissible values. In (Mika and Józwik, 2016), a method for measuring noise in work environment based on the acoustic pressure of noise emitted by CNC machines has been developed. Using acoustic holography methods for identifying and monitoring of noise sources of CNC machine tools were presented in (Józwik, 2016).

In most of the applications of fault diagnosis of the motors when using acoustic signals, the idea is based on using condenser microphone or a group of microphones simultaneously for acoustic signal acquisition considering many factors that affect the fault frequencies such as motor speed along with the interferences of the environmental noises as the acoustical approach is sensitive to such noises. Importantly, calculating the frequency spectra of acoustic signals of the rotating machine, applying feature extraction methods such as correlation and wavelet transformation along with data classification methods are required for the complete procedure of fault diagnosis (Glowacz *et al.*, 2018).

Ultrasound characterises the region of acoustical phenomena that involves high frequencies which are not accessible to human perception in a band above 20 kHz and

continues up into the megahertz, and in industrial settings these acoustic frequencies are referred to as "ultrasonic" (Cheeke, 2016). Ultrasonic technology has been used in many applications. For instance, in (Duncan, Gaydecki and Burdekin, 1996), a track mounted ultrasonic scanning system was developed. Recently, this technology is applicable in many advanced industrial applications such as ultrasonic servo control drive (Shafik and Abdalla, 2013), and the development of a rotary standing wave ultrasonic motor based on piezoelectric technique along with many related applications (Shafik, Makombe and Mills, 2013).

2.5.2 Environmental Sounds

In this review, it is to particularly focus on the problem of environmental noises in sound signal acquisition using acoustic sensing in outdoor conditions. Environmental sounds are such a rich source of acoustic data, and at the same time they comprise much background noise (Ma, Milner and Smith, 2006). Environmental sounds are highly considered as non-stationary and underexploited source of data. Having difficulties in describing such sounds using common audio features as well as defining appropriate features for environmental sounds in automatic acoustic classification systems is an issue of concern. Several environmental sound sources such as wind and rain harmfully affect much of acoustic data of interest, however, these environmental sources produce unwanted sounds which considered as noise sources (Chu, Narayanan and Kuo, 2008).

An acoustic environment is a complex sound which made up of a combination of different events and every environment has its specific characteristic consistent along with a periodic background noise (Ma, Milner and Smith, 2006). In outdoor environmental recordings, much of the noise is of different origin. Generally, background noise can be divided into natural and artificial noises. Natural noises come from environmental sources such as wind and rain. Whereas artificial noises are human generated sounds that come from human activities, such as distant traffic, construction noise, airport noise, etc., (Kuttruff, 2006).

It is quite ambiguous to distinguish between the background noise and the signals of interest because background noise in some applications is considered as target signal in other applications (Luzzi, 2013). For example, if the aim of bird calls' classification in an audio recording is to find calls that belong to a specific species, therefore other species calls will be considered as background noise. Hence, background noise definition is dependent on the area of application. For better understanding, wind noise for instance represents background noise for ecologists when they are assessing species richness by sampling long outdoor acoustic

recordings. Avoiding durations that contain much background noise will help in improving the efficiency and quality of audio sampling (Ma, Milner and Smith, 2006).

The content of audio recordings can be automatically classified into different classes for the detection of environmental noises in outdoor data acquisition. Generally, the classification of the environmental sounds may include many sounds such as (rain, wind, human speech, animal and birds sounds, construction noise, and airport noise, etc.). Spectral and temporal entropies have been reported beneficial for detecting bioacoustics activities. Classification is based on combining pattern recognition methods with digital signal processing technology for audio classification and segmentation when using some machine learning techniques (Karbasi, Ahadi and Bahmanian, 2011). According to (Dillon, 2008), further classification of noise might be possible, wind noise in particular versus others. Numerous parameters can be used in such classification including spectral shape and rate of modulation across channels.

2.5.3 Potential and Challenges of Acoustic Sensing in Real-Life Settings

Acoustic sensing enables Non-Line-of-Sight (NLOS) sensing because sound can propagate around corners and obstacles/barriers. Acoustic sensing is also multipurpose by definition, which makes it such a powerful technology because of the amount of information that can be extrapolated from the environment. Acoustic sensors can capture all possible environmental information, being identified of events made on processing the captured information; their sensing capability is on the software level and on the algorithms that identify the events. They can be used virtually to sense anything, and they also can be complemented with other sensing approaches. Therefore, for added-value applications in new urban environments, it is important to be aware of the untapped power of sounds as a relevant data source (De Marziani *et al.*, 2009; Hancke, Silva and Hancke Jr, 2013).

To develop a sustainable and attractive city for its citizens, many challenges need to be handled. Environmental conditions have probably the same importance as housing, economy, culture, and social challenges that need to be handled for the cities (Riffat, Powell and Aydin, 2016). Using sound to picture the world in an innovative way and using acoustic sensing to collect acoustic data and explore the possibility with several new applications that impact on the daily lives of the inhabitants is very promising towards creating new urban environments. Despite all of these facts, such applications require great efforts regarding developing new noise reduction methods capable to overcome existing problems in outdoor acoustic data acquisition which presents an enormous technical challenge. However, in the presence of

environmental noises, the separation of such noises in the captured acoustic data in real-life soundscapes monitoring for outdoor recordings is very challenging.

The concept of smart cities has started to develop around the world in order to meet these challenges and create new smart, attractive and sustainable cities for its inhabitants (Kelly *et al.*, 2014). A smart environment is a small world where different kinds of smart interactive networked devices along with sensor-enabled components work continuously and collaboratively to improve the experience of its dwellers in order to make their lives more comfortable by supporting and enhancing the their abilities in automatically executing tasks (Ahmed *et al.*, 2016). The term ''smart'' or ''intelligent'' is defined as ''the ability to autonomously acquire and apply knowledge'' while an ''environment'' refers to the surroundings (Cook and Das, 2004, p.3). According to (Cicirelli *et al.*, 2018) in which large-scale smart environments have been introduced, the main aim of such environments is to provide enhanced-cyber services to the users.

The viability and potential of acoustic sensing in real-life settings can pave the way to the enhancement of functionalities and soundscapes design (Luzzi *et al.*, 2013). Using powerful acoustic sensors can lead to improved applications and services and a way to build and develop large-scale "real-life" experimentations of acoustic sensing for supporting high societal value applications and delivering new innovative range of services mainly targeting to smart cities (Pham and Cousin, 2013). In doing so, it is required to automate processes, prediction, and decision-making capabilities which are based on predefined strategies along with dynamic individual preferences of the citizens of smart environments. Also, enabling synchronisation of multiple sensors (microphones) to form wireless arrays and create a new topology in device communication techniques is needed (Petroulakis, Askoxylakis and Tryfonas, 2012; Petroulakis *et al.*, 2013).

2.6 Wind Induced Microphone Noise

For many decades, noise reduction algorithms have been developed and applied to suppress undesired components of the signal or at least to mitigate the impact of such components, which considered as noise, on the signal of interest. The objective of such methods is therefore to focus on the component of interest in the signal and make it much clearer for further manipulation or applications. Wind noise sensitivity of microphones for outdoor recording situations has been a major problem (Elko, 2007). Wind noise is known to have strong deleterious effects on the perception of the desired sounds due to its level which can

reach up to 85 dB SPL for wind speed of 3 m/s and as high as 100 dB SPL when the speed is doubled (6 m/s) (Zakis, 2011; Keshavarzi *et al.*, 2018).

Generally, all wind noise reduction technologies aimed at reducing all types of wind noise in both audible frequency and infrasound bands while taking into consideration preserving signal energy to the most possible level. Although there are several noise reduction methods which have been applied in many applications, wind continues to be a major source of annoying and dissatisfaction as many of these methods do not perform well for all users and applications. For example, in hearing aid application, with only around half of hearing aid users reported that their hearing aid devices cope well in windy environments (Kochkin, 2010). Besides the great deal of research conducted on developing noise reduction methods to improve the quality of life for listeners with hearing impairment, a new method that introduces emergency warning equipment based on a vibration measurement system has been developed (Abulifa *et al.*, 2013).

For environmental noise level measurement, especially long-term measurement or monitoring, microphone wind noise is added to the recorded noise level, giving inaccurate results. Wind shields are commonly used on microphones in outdoor sound acquisition to prevent the wind from exciting the membrane of the microphone, however, the effectiveness is limited, residual microphone wind noise remains problematic (Schmidt, Larsen and Hsiao, 2007). Moreover, with the fast development and use of small and smart high-technological consumer products like in mobile telephony contexts and hearing aids along with the introduction of smart city paradigm that based on acoustic sensing technology; every-day experience for people around the world has widely increased regardless the weather condition.

Additionally, in smart environment applications, acoustic sensing technology is supposed to be mainly deployed and then made interacted with the installed infrastructure. However, effective wind shielding is often difficult to apply to such acoustic sensors likewise with audio devices in applications such as hearing aids and mobile telephones (Schmidt, Larsen and Hsiao, 2007; Zakis and Tan, 2014). For example, the noise created by even a light breeze in the microphone of hearing aid is equivalent to the noise generated by a 100 dB SPL sound at the input of the hearing aid device (Dillon, 2008). As wind noise continuous to be a major concern in outdoor data acquisition applications, continued effort has to be targeted towards advancing wind noise reduction algorithms. Therefore, alternative solutions are sought to be addressed by the research community in this field to overcome the limitations of existing methods.

2.7 Wind Noise Problem in Outdoor Data Acquisition

In environmental noise and soundscapes monitoring, environmental sounds are such a rich source of acoustic data, and at the same time they comprise much background noise which hinder the extraction of useful information (Ma, Milner and Smith, 2006). Environmental sounds are highly considered as non-stationary and underexploited source of data. Having difficulties in describing such sounds using common audio features as well as defining appropriate features for environmental sounds in automatic acoustic classification systems is an issue of concern. There is a variety of environmental sound sources that produce different forms of sound. These environmental sources produce unwanted sounds which considered as noise sources (Chu, Narayanan and Kuo, 2008). The strong presence of environmental noises such as wind covers useful information and limits the usage of such data source efficiently to extract semantic information.

In the context of environmental noise and soundscapes monitoring, acoustic sensors are deployed outdoor, they are exposed to changeable and severe weather in uncontrolled conditions, and microphone wind noise becomes a major concern. Signals of interest are often corrupted by such environmental noises including wind as well as system-introduced noise like channel interference and quantisation. This, in turn, has an adverse effect on the perceived quality of the target sensed sound and posing difficulties to properly distinguish the sensed or recorded event of interest. Meanwhile, the performance of other processing algorithms such as speech or speaker recognition may also be affected (Nemer and Leblanc, 2009).

Wind noise is a problem in many applications such as smart city and soundscapes monitoring as well as for users who enjoy outdoor activities. Not only is wind noise annoying, but also it can create distortion by overloading the microphone and masking desired signals. Wind noise has been always seen as a nuisance and can be uncomfortable and loud. For example, in communication context, conversation becomes sometimes impossible even in only slightly windy listening conditions. In adverse acoustic environments, undesired signals like acoustic noise can make conversations virtually impossible without implementing sophisticated signal processing techniques along with suitably designed electroacoustic transducers. In fact, the presence of acoustic noise, such as wind noise, with all kinds of acoustic signal transmission in real-world environments is an ubiquitous problem (Teutsch and Elko, 2001; Benesty, Souden and Huang, 2012).

Wind noise can be attenuated through acoustic means which are mainly signal processing techniques without necessarily relying on methods of the prior art such as mechanical windscreens (Elko, 2007). For example, the reduction of the effects of wind noise in hearing aids application is based on signal processing techniques in addition to the acoustic design modifications. In order to reduce the harmful effect of undesired components of a signal, which considered as noise, on the wanted signal, a large number of noise reduction algorithms have been developed and implemented for many decades (Korhonen *et al.*, 2017; Keshavarzi *et al.*, 2018). Generally, the removal of noise is not an easy task as the problem is manifold, many unsolved issues owing to the different environmental sound sources that produce noise and particularly microphone wind noise. The reduction of wind noise is therefore a challenging task as wind noise consists of highly non-stationary components known as local short-time disturbances (Keith Wilson, Keith Wilson and White, 2010; Grimm and Freudenberger, 2018).Wind noise is highly broadband and non-stationary in time and even sometimes it resembles transient noise which makes it hard for an algorithm to estimate the noise from a noisy signal (Schmidt, Larsen and Hsiao, 2007).

A microphone diaphragm of a microphone used in an air flow can be deflected due to the pressure variations from low turbulence which exist in environmental wind. As a result of the deflection of microphone membrane, wind-induced noise appears in the microphone output signal (Morgan and Raspet, 1992; Zakis and Tan, 2014). Therefore, turbulent airflow over the microphone membrane causes wind noise which considered as a particular type of acoustic interference that creates an acoustic effect of a relatively high signal level (Nemer and Leblanc, 2009). Low frequency signal components of high amplitude are generated due to the airflow turbulence over the microphone membrane (Bradley *et al.*, 2003). The turbulence generated by the interaction of the wind and the microphone along with the fluctuations that occur naturally in the wind are the two components of wind noise (Nemer and Leblanc, 2009). Wind noise fluctuates rapidly and wind gusts might have very high energy (Schmidt, Larsen and Hsiao, 2007). The spectrum of the recorded wind noise has been described as a broadband but decreasing function of frequency, showing the bulk of the energy in the lower region of the spectrum (Schmidt, Larsen and Hsiao, 2007; Nemer and Leblanc, 2009).

Many attempts were made in the past using various filtering techniques to broadband noise suppression. These methods worked to some extent but have intrinsic limitations of distorting wanted signals and difficulties in retaining accurate signal energy levels to meet the measurement requirements. However, when the wind noise is removed, wanted signals are

distorted considerably (Nemer *et al.*, 2013; Nelke *et al.*, 2014; Zakis and Tan, 2014). Distortion describes the systematic undesirable change in a signal. However, noise and distortion limit the accuracy of the results in signal measurement systems and the capacity of data transmission. For event detection or decision making from acoustic signatures, distorted signals cause errors or misjudgements, mitigating the reliability, usability or even safety of such systems (Vaseghi, 2008; Zakis and Tan, 2014).

Methods that make use of two or more sensors known as multi-microphone and microphone array are relatively considered effective schemes for reducing wind noise as the difference in propagation delay between wind and acoustic waves is exploited. In other words, the propagation speed of wind is much slower compared to that of acoustic sound waves, and thus the multiple microphones receive correlating signals that help in detecting wind noise. In an attempt to accomplish a synchronisation, sufficient spatial aperture, simultaneous processing and calibration of all sensors in multi-microphone approach, the deployment of the microphones may be relatively distant, and this is expensive and technically difficult. However, the difficulty and high cost in deploying these complicated setups limit their prevalent usage (Brandstein and Ward, 2013; Nemer *et al.*, 2013). Single-microphone wind-noise reduction is still an open ended problem and a technical challenge for further extensive research (Nemer *et al.*, 2013; Nelke *et al.*, 2014).

The main concern discussed in this thesis is the destructive impact on the wanted sensed signals. Moreover, one of the greatest challenges for the future digital ecosystem is to better reduce such unwanted environmental noise in outdoor acoustic monitoring. Hence, improving de-noising techniques or developing and implementing new techniques that may lead to better solutions and be effective alternatives to the classic methods becomes increasingly important against the harmful effect of wind on the perceptual quality of the wanted signals. Linear separation in subspace seems to be potential solution to circumvent these problems. While the objective of the present study focuses on microphone wind noise reduction in the context of environmental noise management and soundscapes monitoring, the current research attempts to cover certain important aspects and develops a more rigorous understanding of the SSA technique as a proposed method.

2.8 The Need and Importance of De-noising Techniques

As society continues to promote applications related to smart city and soundscapes monitoring as critical to the overall quality of life, applications that rely on signal processing

algorithms for wind noise reduction such as hearing healthcare are also of similar importance. Therefore, researchers are increasingly interested in such applications because of their active lifestyle as to enjoy outdoor recreation to include different activities for all smart city inhabitants including hearing aid population. Along with the concern shown due to wind noise effects on the sensed acoustic data in smart city and soundscapes monitoring applications, wind noise has been reported as a series problem in hearing aids in which the utilisation of behind-the-ear (BTE) instruments means that wind noise is of major concern (Kochkin, 2010).

In signal processing, the separation of one signal from another is a key element and important aspect which embraces a variety of interesting applications. The desired signal is not observed directly in many applications as for several reasons; it might be distorted and noisy. For one reason, the signal gets noisy in digital communication systems for example due to the distortion occurs during its propagation form the source to the receiver. As another reason, when the resolution of the measuring equipment is limited, and the communication channels have non-ideal characteristics. Other reasons like multiple reflections and missing samples, echo, and signal fading reverberations could cause signal distortion (Vaseghi, 2008). In smart city monitoring, the most common reason is the presence of surrounding environmental noise or the interference of other signals in the environment where the signal of interest is observed (Hayes, 2009). In windy environments and for hearing aid users for example, wind noise is generated as the air passes the hearing device or the aid user's head as well as other obstacles (Dillon, 2008; Kochkin, 2010).

According to (Loizou, 2007), the ultimate goal of developing and implementing an algorithm for noise reduction problem for general applications is to help in significantly improving the perceptual quality of the captured acoustic data and intelligibility in case of speech enhancement for normal hearing (NH) as well as for hearing-impaired (HI) listeners. Research has been elusive for almost three decades regarding this critical issue. As stated by (Bentler *et al.*, 2008; Luts *et al.*, 2010), no significant intelligibility benefit has been revealed when implementing noise-reduction algorithms in single-microphone applications such as wearable hearing aids, although such algorithms are essentially found in hearing aid devices for HI listeners to ease of listening. Therefore, based on that, the actual need of developing and implementing new contemporary single processing methods for wind noise reduction that might come up with better solutions in outdoor data acquisition is now increased in importance.

Based on the application area, the desired signal could be an acoustic signal, speech, or even a radar signal, or an image in some other applications (Hayes, 2009). The desired signal in idealised and simple environments can be restored from the measured data by designing certain classical filters such as high-pass, low-pass, or band-pass. However, the best estimate of the signal cannot be produced in an optimum way using such filters in most of the cases. Therefore, producing the optimum estimate of a signal from a noisy observation or measurement was indicated as a problem that requires designing suitable de-noising methods or filters to overcome (Hayes, 2009).

Research has revealed that long-term average wind noise level in even a light breeze of 3 m/s (10.8km/h or 5.8 knots) is about 80 dB SPL. Wind noise level is limited to almost 115 dB SPL by saturation in the microphone for wind speed of 12 m/s (43.5 km/h or 23.3 knots). Wind noise spectrum is mainly dominated by lower frequencies which roughly less than 500 Hz. However, it has been shown that at 8 kHz, wind noise level can be greater than 60 dB SPL even though the saturation is presented at 12 m/s (Kochkin, 2005, 2010; Zakis, 2011). It can be clearly seen that at commonly encountered wind speeds, wind noise is at high levels and hence it needs to be suppressed. MarkeTrak data reinforced this fact for hearing aid users as indicated in (Kochkin, 2005, 2010), it shows that behind noisy situations, wind noise is the second-highest cause of dissatisfaction.

2.9 Factors Affect Noise Reduction Algorithms Performance

With general noise reduction algorithms, there are some factors that affect their performance and a quite number of reasons behind the lack of perceptual quality improvement for sensed acoustic data. According to (Dubbelboer and Houtgast, 2007), there are certainly some factors that contribute to the lack of perceptual quality with existing algorithms applied for single-microphone noise reduction. But in most cases, it is not entirely clear to what extent and how as specific parameter requires to be modified in noise reduction algorithm, so as to improve its performance.

The first reason is identified by (Loizou, 2007), it is related to the inability of accurately estimating the background noise spectrum, and such estimation is required for the implementation of most algorithms in single-microphone applications. For example, voice-activity detection or noise tracking algorithms are generally performing well in steady background noise (e.g., car environment), while in non-stationary noise environment (e.g., multi-talker babble) their performance is rather low. In some applications such as tracking time

delay between delayed signals contaminated with noise, it has been shown that the attenuation effect is also damaging and for better performance of such tracking algorithms, returned signals have to be amplified (ALdwaik and Eldwaik, 2012).

Introducing distortion by the majority of algorithms is the second reason as stated in (Hu and Loizou, 2007), which, in some cases, such algorithms might be more damaging compared to the background noise itself. For this reason, according to (Jingdong Chen *et al.*, 2006), many algorithms propose minimising target signal distortion (e.g., speech) by constraining the amount of noise distortion introduced which cause it to drop below a pre-set value or even below the auditory masking threshold. For example, as a special case, a family of algorithms related to speech enhancement applications have been described in (Loizou, 2007). However, improving speech quality and intelligibility by means of such algorithms can be achieved, but to some extent. Imposing some constraints such as across-time spectral constraints helped in improving speech quality. Yet, this approach has a main limitation which is not amenable to real-time implementation as reported by (Loizou, 2007).

Additionally, previous algorithms of wind noise detection do not always correctly distinguish between wind and non-wind causes of the differences in microphone signal as such algorithms generally assume these differences to indicate the presence or the absence of wind noise. Consequently, this procedure will lead to inappropriate engagement of wind noise reduction schemes. For example, in case of using two microphones for wind noise detection, existing algorithms assume that wind noise is indicated by large phase differences between the two microphones, whilst its absence is indicated by small phase differences. However, even without the presence of wind noise, these differences might exist due to acoustic reflections, unmatched microphones, or phase shift caused by microphone spacing. Accordingly, the performance of wind noise reduction processing might be affected (Petersen *et al.*, 2008; Zakis and Tan, 2014).

2.10 Wind Noise Theory

2.10.1 The Physics of Wind

To introduce wind noise resistant technologies and develop reliable and effective wind noise reduction methods, it is important to understand such wind-related noise by reviewing the physics of wind. From the physical stand point, global winds are initiated by the spatial differences in atmospheric pressure across the surface of the earth due to the uneven heating of

the earth by solar radiation and considered as a common element of the diurnal meteorological cycle that occurs frequently everywhere in the world (Panofsky and Dutton, 1984; Manwell, McGowan and Rogers, 2010).

Studying the physical characteristics of the "Planetary Boundary Layer" (PBL), which is about 1-2 km above the ground and also called "friction layer", is quite important to understand the nature of wind-related noise. The characteristics of this atmospheric boundary layer, which is the lowest part of the atmosphere, are influenced by the contact with the surface of the earth. Due to surface effects, many physical quantities such as heat, relative humidity, velocity, and vertical exchange of momentum can clearly define this layer. Here, these quantities can change rapidly in space and time. Furthermore, it contains multiple effects such as aerosols, dust and smoke which make it visible from the top. Using acoustic sounders in a range of 1–3 kHz can help in detecting the thickness of the PBL layer. At a 10 m height, this thickness can be predicted by a linear relationship with the speed of the wind (Koracin and Berkowicz, 1988; Manwell, McGowan and Rogers, 2010).

Atmospheric turbulence plays a vital role in describing the wind which is intimately related to its two types; convective and mechanical. As physically defined, convective turbulence in the troposphere is a predominant mechanism of mixing and determined by thermal instability. The other type, which is the mechanical turbulence, is produced by the interaction of the wind with ground-based objects and topography. Turbules are defined as "self-similar localised moving eddies" (De Wolf, 1983; McBride *et al.*, 1992; Goedecke and Auvermann, 1997). However, turbulence can be modeled with regards to a series of turbules with a distribution of sizes (Crocker, 2007).

Studying the so-called vertical profile of the PBL layer as a working definition in which temperature increases with height can lead to define this feature as "inversion" with regards to the changes in the temperature between day and night as a function of height. Figure 2.4 shows this relationship as the lower PBL or surface layer gets warmed by the solar heating during the day at a considerable height denoted by z_i. This height refers to the lowest inversion in which gradient changes in the sign of the temperature can be observed. Consequently, this is known as one mechanism of driving wind. Importantly, both mechanical and convective turbulences occur at multiple scales to produce wind since the warmer air near the surface is gravitationally unstable. Such scales are quite large in some cases since geographical differences affect surface heating which cause regional horizontal winds (Walker and Hedlin, 2010).
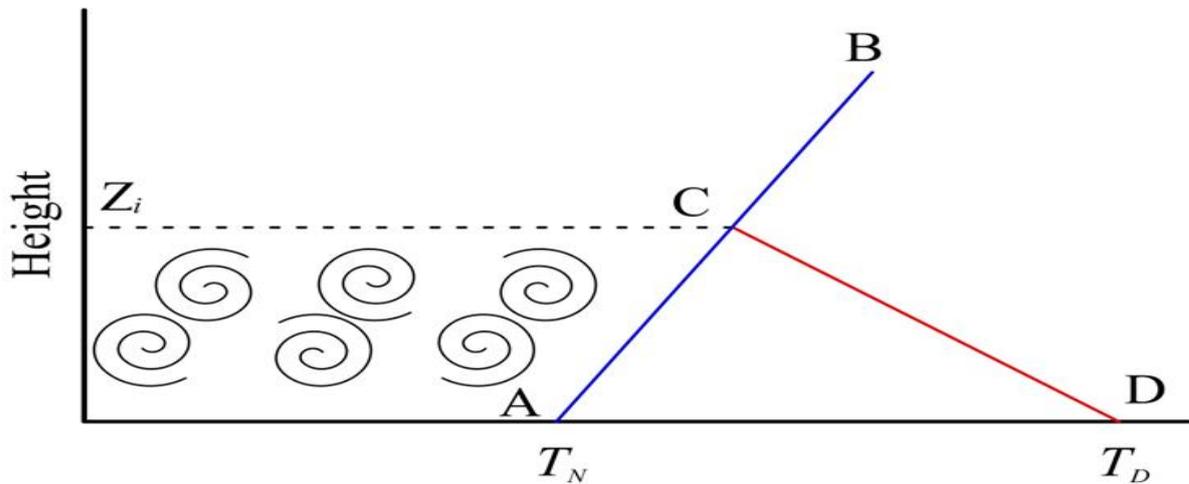
Figure 2.4. Temperature vertical profiles during the night (AB; inversion) when T_N is the night-time temperature. BCD indicates the vertical profile during the day when the inversion occurs between CB and T_D is the day-time temperature (Walker and Hedlin, 2010)

The mixing or surface layer is defined when smaller scale turbulence takes place. During the day, the thickness h of the surface layer is approximately equal to z_i. The air becomes gravitationally stable because an inversion often spreads down to the surface at night as illustrated in Figure 2.4. Accordingly, wind and turbulence are often less pronounced at night-time. Though, in the lower portions of the surface layer for a thickness around 100 m or less, slight mechanical turbulence still occurs with weak winds on clear nights (Manwell, McGowan and Rogers, 2010; Walker and Hedlin, 2010). The variation of horizontal wind speed with height above the ground is an important parameter in the characterisation of wind resource. This variation of wind speed with height is called the vertical profile of wind speed. Wind speed is very much dependent on ground cover variations and local topographical (Manwell, McGowan and Rogers, 2010).

Because of the surface friction, wind speed varies as a function of height. Winds in the surface layer increase logarithmically with height and the slowest winds are those at the ground level in case of pure mechanical turbulence (Thuillier *et al.*, 1964). In the case of convective turbulence, wind speed varies as a function of height as well, but in a much more complex scenario. This variation is often described using a power law over a specified height range. In both cases, the reduction in wind speed at ground level can be expressed as a function of what so-called surface roughness and the length that describes the size of mechanical turbulence. The surface roughness is known to be a measure used to indicate how efficiently momentum is transferred into the ground from the wind. For instance, craggy mountainous terrains or dense

forests have a high surface roughness that can lead to a significant reduction in surface wind speed. Wind speed increases significantly above relatively smooth surfaces, basically between 1 and 3 m height. For example, in light winds of speed (2–5 m/s), a noticeable difference form ~20 to 40% in the wind speed was recorded between these two indicated heights (Berman and Stearns, 1977).

It is worth noting that basic meteorological data can be used to calculate the potential for turbulence. A predictive measurement is required to superimpose the effects of both thermal and mechanical forcing as they have considerable influences on the turbulence. Two important scaling parameters in the surface layer which are the surface friction velocity u and the Monin-Obukhov length L were mathematically derived for many decades; details can be found in (Monin and Obukhov, 1954). Since z is the height above the ground, the ratio z/L has shown a key importance in superimposing the effects of mechanical and thermal forcing. This ratio has been used as a measure in characterising the atmospheric stability. However, smaller negative values indicate a dominance of mechanical turbulence, whereas, strongly negative values are associated with a dominance of convective turbulence. Pure mechanical turbulence is theoretically found at zero at the ground level. In the case of slightly positive values, the temperature stratification damps the mechanical turbulence in a direct proportion manner. Strong damping of turbulence is indicated by strongly positive values (Walker and Hedlin, 2010).

2.10.2 Influences of Geographic and Regional Variations on Wind

Studies of local wind patterns are highly beneficial and have a sustainable impact on carefully selecting sites as a strategy for the reduction of wind noise. As previously mentioned, the derivation of wind is based on the spatial differences in atmospheric pressure. These spatial differences are directly related to two factors that cause changes in temperature; solar heating and surface radiation. These differences do not occur at a same scale, but mainly at a variety of scales. Therefore, they interfere with each other making certain effects on the observed local conditions (Raspet, Webster and Dillion, 2006; Walker and Hedlin, 2010).

Most of the global circulation patterns can be seen in a general system at a global scale. This system drives such global circulation patterns to be eventually modulated by the local and regional influences. To be more precise, at a global scale, more solar heating affects the equatorial regions, whereas, the poles experience much less of such solar heating. In other words, the surface of the earth at the equatorial regions absorb greater amount of solar radiation

than at the poles. As a result, the air flows upward and towards the poles because of the warming through conduction and convection. In a simple flow model, air rises at the equatorial regions and sinks at the poles. At the poles, cooling causes the air to flow down to the surface and then back to the equatorial regions. The air flow is fixed in the forward flow and the so-called Coriolis Effect deflects the flow towards the left and the right in the southern and northern hemispheres, respectively. This scenario will lead to three circulation cells in each hemisphere, however, the global circulation patterns which eventually modulated by local and regional influences are driven by this system (Manwell, McGowan and Rogers, 2010; Walker and Hedlin, 2010).

In the continental interiors, the convection patterns get affected by the regional variations in solar heating during the day. Through the conduction property of heat energy, solar heating of the surface warms the air. This effect causes the air to ascend, and horizontal pressure gradients to form. In fact, air moves from high to low pressure due to the variations in the atmospheric pressure created by the spatial variation in heat transfer to the atmosphere of the earth. Moreover, the convection property helps in pulling in air from other regions that are cooler as the regions beneath cloud cover as shown in Figure 2.5 (Manwell, McGowan and Rogers, 2010; Walker and Hedlin, 2010).
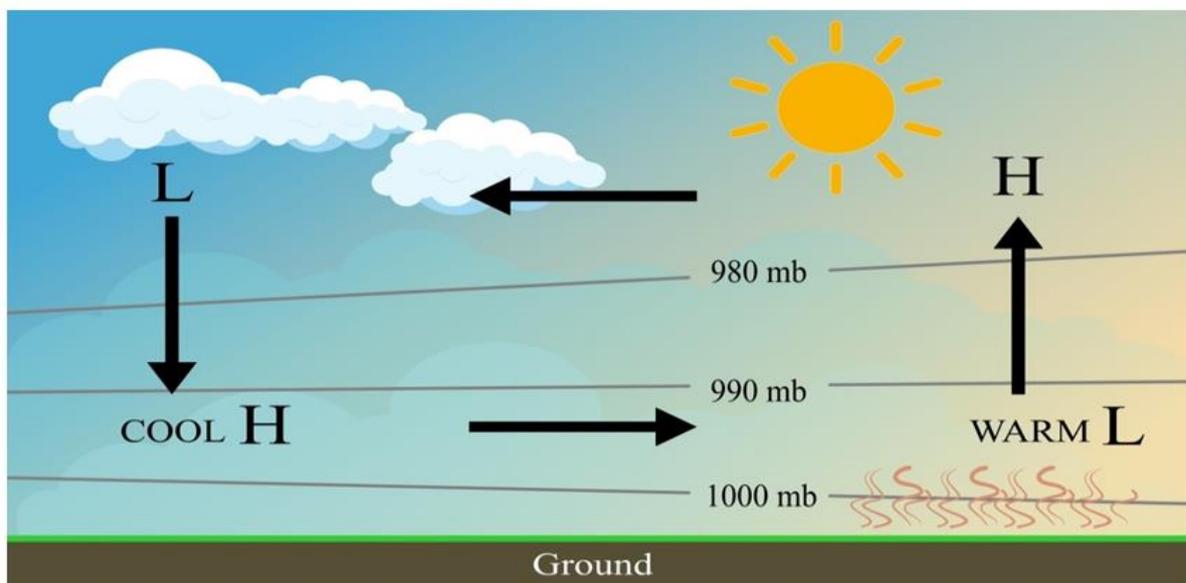
Figure 2.5. The influences of regional variations on wind in an intercontinental setting during daytimes; a closed convection system based on differences in solar heating at the surface that lead to horizontal air temperature and pressure gradients (Walker and Hedlin, 2010)

Due to the influences of global circulation patterns, horizontal winds do not usually travel in straight lines between high- and low-pressure regions. This explanation idealises a system that only works during daytime, however, variation in surface heating or cooling that could drive convection and surface winds are not remarkable at night (Panofsky and Dutton, 1984; Goedecke and Auvermann, 1997; Walker and Hedlin, 2010). A similar pattern of convection during daytime can be seen in coastal environments as illustrated in Figure 2.6. In such coastal areas, the surface of the water does not heat up to the level that of the land surface. The reasons behind that are the specific heat capacity of the water as well as the continuous mixing of the surface water with deeper layers. As a consequence, this leads to a differential air or surface heating, horizontal pressure gradient as well as onshore surface winds (Raspet, Webster and Dillion, 2006; Walker and Hedlin, 2010).

Unlike continental interiors, convection systems can be driven at night in coastal environments. To explain what happens in this situation, the third property of heat transfer, which is the radiation, clearly takes place after the sunset as the heat continues to be radiated from the heated land surface. Accordingly, the land surface becomes cooler than the surrounding air. On the other hand, the water surface continues to transfer heat to the air while it stays at a relatively fixed temperature. The pattern shown in Figure 2.6 is mainly reversed in this case, leading to derive offshore surface winds (Walker and Hedlin, 2010).
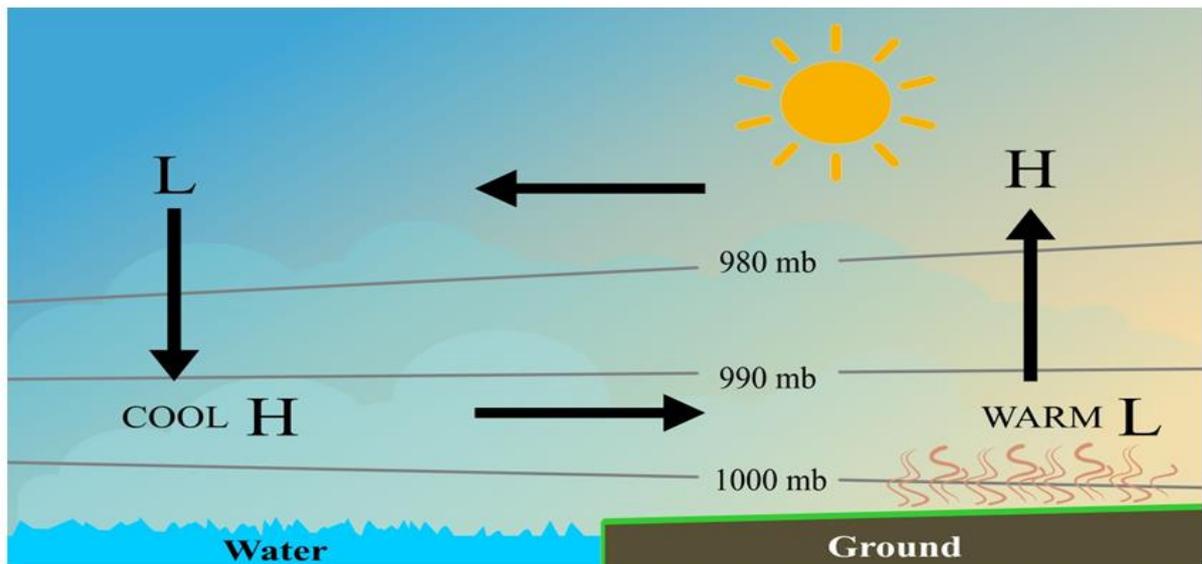
Figure 2.6. The influences of regional variations on wind in coastal setting during daytimes; a closed convection system based on the specific heat capacity of water that leads to differences in solar heating and cooling of the surface (Walker and Hedlin, 2010)

In addition to the above explained regional systems which are diurnal in nature, seasonal influences on wind over continental spatial scales are also exist. The so-called onshore "Monsoon" winds are derived during the summer season and can persist diurnally. Mainly, these humid winds are diurnally uniform in some sites and times of the year. The derivation of these onshore winds is scientifically explained by temperature difference between land surfaces and adjacent water surfaces. The average daily temperature of the adjacent water surfaces is less than that of the inner continental land surface. During the winter season, the situation is totally reversed and therefore dry offshore winds dominate the region (Panofsky and Dutton, 1984; Koracin and Berkowicz, 1988; Miles, Wyngaard and Otte, 2004).

The surface of the earth varies considerably with land masses and large oceans. However, these different surfaces affect the flow of air because of the absorption of solar radiation, variations in pressure fields, and the amount of moisture. The movement of air is affected by the circulation of oceans which act as a large sink of energy. Hence, these effects lead to differential pressures which have influences on global winds as well as many of persistent regional winds. Wind patterns are also affected by mountains which are natural obstacles to winds driven by global convection systems. Importantly, mountains also act as sources of heat throughout daytimes. The slopes of the mountains get heated by the solar heating making the air flows up and above the mountain up to the height of the surface layer where it is deflected. Instead of heat sources, at night, mountains become heat sinks, however, the convection pattern is reversed. In case of valleys, a circulation system is generated with vertical return flow above the axis of the valley when heating both its adjacent slopes (Manwell, McGowan and Rogers, 2010; Walker and Hedlin, 2010).

2.10.3 Turbulence Spreading

There have been numerous verifications of "Taylor's frozen turbulence hypothesis" which is of considerable significance to wind-noise theory. This hypothesis stated that, in a stationary anemometer or microphone, time-varying signals are supposed to result from a spatially varying field that is frozen at a velocity equals to the convection velocity in time moving across the anemometer. In other words, turbulences and their related observables are spatially fixed time-invariant anomalies. The turbulent structures move, under certain conditions, as frozen entities transported by the mean wind. Hence, the spatial pattern of turbulence is derived from its temporal description (Schlipf *et al.*, 2010; Walker and Hedlin, 2010; Higgins *et al.*, 2012).

Many studies have been conducted to verify and test this concept by comparing measurements on aircrafts with those taken from towers (Panofsky and Mazzola, 1971; Kaimal *et al.*, 1982). McDonald et al. (2010) conducted a study to test this concept, the obtained results were not totally valid Taylor's hypothesis in all turbulence cases. Their results thereby suggest that turbulences with short-wavelength structures move with the mean wind, whereas, turbulences with large-scale structures move at their own velocities. These results provided a striking contrast to Taylor's hypothesis. Even though such results showed that turbulences are not indefinitely time invariant, Taylor's hypothesis is generally considered to be valid at higher frequencies. Studies revealed that turbulences decline with the distance travelled by an amount that is proportional to their length scale. Smaller turbulences decay faster compared to larger ones over shorter distances travelled (Schlipf *et al.*, 2010; Walker and Hedlin, 2010; Higgins *et al.*, 2012).

In time/distance space, Douglas Shields (2005) described in his study wind noise measurements using a three-axis array of low-frequency sensors. The study aimed at investigating the correlation as a function of sensor separation based on recording signals from individual sensors. The procedure applied was based on locating one string of microphones on the ground along the wind direction while the other is perpendicular to it. It has been reported that expressing the sensor separation in wave numbers was extremely convenient to reduce the dependence upon wind speed. Therefore, wind velocity was decomposed to its two components; convection component, which is known as mean component, and fluctuating component. Interesting observations have been made after calculating the cross-correlation between the reference sensor, which was located at the intersection of the two strings, with the other sensors at greater distances in the downwind and crosswind directions. What is relevant to this review is that what observed by downwind sensors was observed by reference sensor some time later, however, with time/distance, the correlation is gradually reduced (Douglas Shields, 2005; Walker and Hedlin, 2010).

2.10.4 Turbulence Scales and Noise Spectra

For more than seven decades, turbulence velocity spectra have been divided into frequency ranges that are related to three spatial scales of turbulence. The large scales are associated with low frequencies and known as source region, the second is inertial subrange which is known as intermediate scales and associated with intermediate frequencies, and the small scales associated with high frequencies and identified as dissipation region. The large

scales comprise large eddies characterised by length scales from about tens of meters to some kilometres. This range is defined by spectral characteristics which are not isotropic. In addition to the wind, these spectral characteristics depend on many variables including the height of the surface layer and surface roughness (Raspet, Webster and Dillion, 2006; Walker and Hedlin, 2010).

Energy-containing eddies are produced from the source region that contains big eddies due to the mixing process within the PBL layer. These big eddies are fragmented into smaller eddies without energy dissipation. With no dissipation appears in the mixing process, the isotropic inertial subrange is defined with eddy length scales less than the height above the surface and larger than the so-called Kolmogorov microscale. This scale is known to be the smallest scale of turbulence as defined in Equation (2.1) (Shih and Lumley, 1993). Continued mixing brings about smaller eddies that are even smaller than Kolmogorov microscale by which the isotropic dissipation range is defined. In this dissipation range, molecular mixing dissipates energy in the surface layer over a length scale on the order of millimetres (Walker and Hedlin, 2010).

$$\eta = \left(\frac{v^3}{\varepsilon}\right)^{1/4} \tag{2.1}$$

where v is the kinematic (or molecular) viscosity and ε is the dissipation rate of turbulence kinetic energy into heat.

Wind generates acoustic energy; however, the interaction of wind with ground-based objects due to the ground topography can cause a radiation in the acoustic energy. For example, at large scales in the case of infrasound at lower frequency band 0.01–0.1 Hz, when wind interacts with mountain peaks; infrasound is radiated and can travel to large continental distances (Larson *et al.*, 2010). Also turbulent storm systems can radiate infrasound (Bowman and Bedard, 2010). Wind can also generate infrasound in an indirect manner. To explain that, when winds increase, the size of the ocean swells and this causes higher surf and in turn more energetic surf infrasound as well as interaction of intersecting swells patterns (Garcés *et al.*, 2004; Arrowsmith and Hedlin, 2005).

It is worth mentioning that wind noise in the infrasound band mainly relates to the source region and the inertial subrange. However, it is essential in all cases to identify these ranges in the spectra of recorded wind noise in order to fully understand what type of wind noise is being recorded which eventually helps in designing or selecting the optimum wind noise filtering or

separation method. Regarding the frequency that separates the source region on the low side from the inertial subrange on the high side in the case of stationary sensor and for wind speed fluctuations in the downwind direction, a simple and important mathematical formulation as defined in Equation (2.2) has been given in (Panofsky and Dutton, 1984).

$$\frac{fz}{\bar{u}} > 1, \tag{2.2}$$

where f is the frequency, z is the sensor height, and $\bar{u}$ is the mean wind speed.

Figure 2.7 shows this relationship which has been practically examined and demonstrated in (Panofsky and Dutton, 1984) using 10 logarithmically spaced wind speeds. As previously mentioned, the internal subrange is defined with intermediate frequencies, therefore it can be seen from Figure 2.7 that at a fixed wind speed and as the sensor height increases, the inertial subrange moves to lower frequencies. The sources region is defined with lower frequencies, however, for a fixed sensor height and as the wind speed increases, the sources region moves to higher frequencies. When the sensor is precisely located at the ground level, the inertial subrange becomes undefined, and the source region contains the entire infrasound spectrum. The International Monitoring System (IMS) arrays typically have sensor heights ranging from 5 to 40 cm and wind speeds that extend up to 5 m/s. Hence, for any recorded pressure or given spectra, it is expected to find this boundary above approximately 0.2 Hz. For instance, the boundary is above ~3 Hz for wind speeds of at least 1 m/s (Walker and Hedlin, 2010).
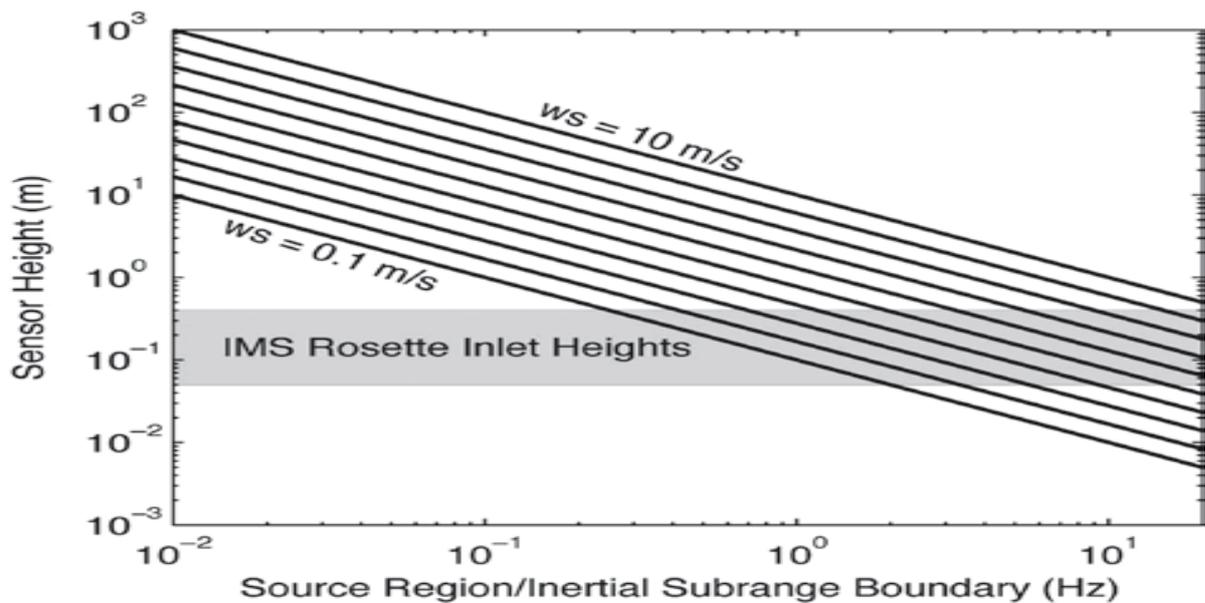
Figure 2.7. Frequency boundaries that separate the source region at the low side from the inertial subrange at the high side as a function of sensor height and wind speed for wind velocity spectra (Walker and Hedlin, 2010)

2.10.5 The Nature and Spectral Character of Wind Noise

In outdoor activities, microphones are subject to wind noise. For example, in hearing aid application when wind blows and hits an obstacle, such as the head of the user, the pinna, or the hearing aid device itself, turbulence is generated which inherently contains marked pressure fluctuations. Consequently, the microphone converts these fluctuations to sound which is an audible wind induced noise dominated by low- and mid-frequency components. Extremely high SPLs at the microphone input can be produced even at moderate wind speeds, however, this might be sufficient to overload the microphone (Dillon, 2008).

In audible frequency band, wind noise is a specific type of acoustic interference caused by the turbulent airflow over the microphone casing and diaphragm, however, such interaction generates an acoustic effect of a relatively high signal level (Nemer and Leblanc, 2009). In other words, wind turbulences around the microphone inlets make the air fluctuate and in turn this causes the vibration of microphone membranes that eventually create wind noise. For example, in systems using two or more inlet ports for sound, local air turbulence around the inlet openings generate wind noise. Due to this noise, uncorrelated sound signals will be received at the microphones. (Petersen *et al.*, 2008). Wind noise is bursty with gusts, broadband and non-stationary in nature with high amplitude that might go above the nominal amplitude of the wanted signal. Therefore, it is often irritating and leads to listener tiredness and user dissatisfaction (Schmidt, Larsen and Hsiao, 2007; Nemer *et al.*, 2013).

Air flow can be categorised into two categories as either laminar or turbulent. The first category describes the case when the air flows smoothly across a surface with no obstruction as a laminar flow field. When the air flows across any obstacle or resistance, it generates turbulence known as turbulent flow. For example, as previously mentioned in the case of hearing aid devices, turbulence can be created from the deflection of the air from the pinna or the end of microphone port which considered as an obstacle. These turbulences are referred to wind noise when they picked up by a microphone. Generally, wind noise represents a particular class of interference as it is created by turbulences in air stream on the microphone membrane or around the edges of a recording device leading to a non-stationary and fast changing noise signal (Dillon, 2008; Nelke *et al.*, 2014).

2.10.5.1 Wind noise components and spectral character

Numerous published work have been conducted to study the spectral character of wind noise such as in (Bradley *et al.*, 2003; Nemer and Leblanc, 2009) in which it has been shown that wind noise consists of two components.

- Fluctuations and flow turbulences occurring naturally in the wind.
- The interaction between the wind and the microphone generates turbulence that causes wind noise.

Both components have a significant impact on describing the effect of wind noise in many applications. Importantly, the second component has a dominant influence on wind-sensor interaction cases for handheld and ear-held devices such as in practical telephony contexts. In such applications, the effect of wind noise is very pronounced due to the additional significant turbulence which can be generated by the presence of users' hand and face (Nemer and Leblanc, 2009).

Generally, the spectrum of the recorded wind noise is shown as a decreasing function of frequency, however, in the lower spectrum, the bulk of the energy is concentrated. Research has also shown that the amplitude and bandwidth are function of wind speed and direction along with sensor or handset design. According to (Nemer and Leblanc, 2009), when wind noise is recorded at a 45º angle and relatively light breeze with different speeds ranging from 2mph (1m/s or 3.2km/h) to 8mph (3.6m/s or 13km/h), the power spectrum of wind noise can be seen as shown in Figure 2.8.
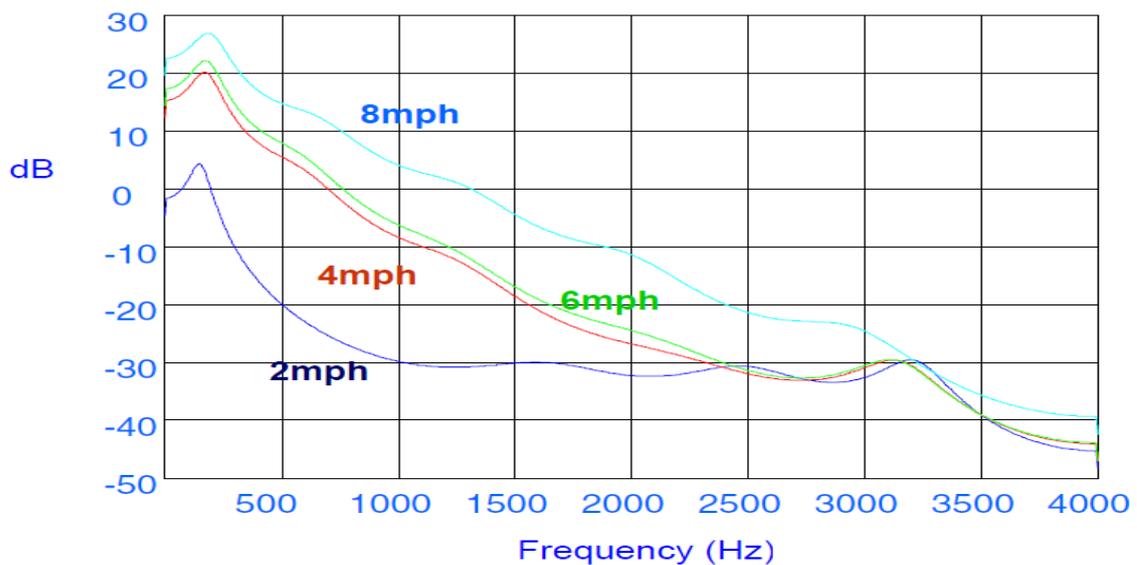
Figure 2.8. Wind noise spectrum at various speeds (Nemer and Leblanc, 2009)

Figure 2.8 shows an example of the spectral envelopes of wind noise generated by wind travelling at different speeds as indicated in the graph in which the x-axis represents the frequency and y-axis represents the logarithmic power magnitude. Since wind noise is non-stationary broad-band signal, most of the energy is in the low frequency range as shown in Figure 2.8 (Schmidt, Larsen and Hsiao, 2007; Nemer *et al.*, 2013).

The long-term average of wind noise spectrum varies mainly with the type of the acoustic sensor or the style of hearing aid device in case of hearing aid application. In addition to this factor, wind speed and azimuth or incident angle also affect wind noise spectrum (Chung, 2004; Dillon, 2008). However, wind noise severity depends greatly on these factors. Wind noise level is directly proportional with the velocity of air turbulence that flows across the microphone. Basically, in the low frequencies range, the microphone can be driven to the saturation due to air turbulences created by the strong winds and result in severe distortion. Wind noise is predominantly described in lower frequency; however, it diminishes at higher frequencies. The change of loudness growth contour can be indicated by minimal audible pressure curves to show that such loudness growth is more significant in high frequencies than in low frequencies. This concept may bear out the perception of the contribution of higher wind speed as it is more to higher frequencies (Raspet, Webster and Dillion, 2006).

2.10.5.2   Effects on wind noise level

As research has emerged in the past decades regarding wind noise problem in outdoor data acquisition, the research community has addressed certain important aspects that have influences on wind noise level. Such noise creates distortion and affects the desired signal as for example with respect to effects on speech understanding. Among these aspects is wind speed as discussed in (Grancharov *et al.*, 2008) as well as working memory, listening effort, and cognition (Pichora-Fuller, Schneider and Daneman, 1995; Arlinger *et al.*, 2009).

- **Wind speed**

The effect of wind speed on the noise level as previously shown in Figure 2.8, however, wind speed itself can be affected in two ways, either by an external breeze moving across the microphone or the movement at certain speed of the user carrying microphone (e.g., hearing aid users). The Beaufort Wind Scale, which ranges from 0-12, has interpreted various wind speeds. This scale is given in a table indicating the range or also called Beaufort number, wind speed, along with description and effects on land and sea. It is considered important as to

provide the kind of wind or description of the situation in defined terminologies with regards to different wind speeds and the effects that can be observed (Rudner and Lunner, 2014; Korhonen *et al.*, 2017).

- **Working memory and listening effort**

Working memory for humans, however, handles their memory and attention, and deciphers various processing schemes for listening understanding (Rönnberg *et al.*, 2013). Research has shown that listening in high annoyance conditions preoccupies and limits working memory capacity, and consequently reducing capacity for listening understanding and increasing communication difficulties (Meister *et al.*, 2013). It is said that Effortless Hearing is always critical amongst listeners when listening in noisy environments such as in windy condition which is seen as an adverse listening condition. Accordingly, listening understanding under such condition presumes higher working memory. For listeners with hearing impairment, the situation becomes even worst particularly those who experience reduced cognitive capability. In the field of cognitive neuroscience and for speech understanding as a particular case, effects of wind noise have been explored (Wagenmakers, Farrell and Ratcliff, 2004; Rudner and Lunner, 2014).

Listeners' working memory is involved to a large degree in speech processing. However, listening in any adverse condition such as in the presence of wind noise can quickly consume this memory as its capacity is limited. Therefore, advancements in noise reduction algorithms are supposed to eventually divert additional working memory towards better understanding while listening in windy conditions. Based on the advancements in digital platforms that allow innovative noise reduction algorithms to emerge, designers of such digital noise suppression algorithms have taken into account reducing listening effort as an ambitious goal to optimise the understanding of the wanted signal (e.g., speech understanding) (Chung, 2004; Korhonen *et al.*, 2017).

The working memory capacity has a key importance in language processing, however, when processing speech in noise, additional working memory is consumed as stated in (Rönnberg *et al.*, 2013). Hence, if noise reduction algorithms can help in easing listening in noisy environments, this would permit additional working memory capacity as well as improved language processing ability.

- **Annoyance**

Turbulent noise is commonly perceived as unacceptable and annoying when it goes beyond subjective tolerance levels as it leads to listener's dissatisfaction and harmfully impacts the proper understanding of desired signals. Annoyance and acceptable noise level have been investigated with respect to ease of listening (Chung, 2004; Korhonen *et al.*, 2017). As indicated in (Connolly *et al.*, 2013), a poor acoustic environment can negatively affect listeners' understanding and learning. Besides the annoyance even in case of intermittent noise, certain critical factors in the analysis for the ease of hearing such as the sensitivity to the noise and its consequences have been specified. difficult acoustical conditions can even lead to a significant effect on individuals' behaviour based on the analysis provided in (Connolly *et al.*, 2013).

There are two important linked concepts with regards to acceptable noise factor which are the understanding of the wanted signal (e.g., speech understanding) and listening comfort. Research has revealed that listening for understanding is highly associated with individuals' working memory and cognition (Rudner and Lunner, 2014). For this reason, listening for understanding requires a higher working memory capacity compared to comfort listening. Therefore, the combination between the two scenarios becomes important in order to bridge them without sacrificing either.

2.10.6 Types of Wind Noise in Inertial Subrange and Source Region

This section is to hold discussions on several types of interactions between turbulences among each other and in specific cases for pressure anomalies advected across the sensor along with wind-sensor interaction that all identify wind noise types, typically in the inertial subrange and source region. There are four types of wind-related noise discussed in the following sub-sections. More detailed discussions of this can be found in (Douglas Shields, 2005; Raspet, Webster and Dillion, 2006).

2.10.6.1 Wind-sensor interactions

Pressure energy is a result of the conversion of kinetic energy when an object deflects the wind. The so-called "stagnation pressure" is the pressure at the head of a body that is in front of the wind and is known to be the maximum pressure obtained as a result of the deflection of the wind. Fluctuating stagnation pressure on pressure-sensing surfaces is caused by fluctuating wind velocity. Turbulence–sensor interaction mainly characterises sensor

interference which leads to identify the first type of wind noise (Raspet, Webster and Dillion, 2006; Walker and Hedlin, 2010).

According to (Raspet, Webster and Dillion, 2006), the stagnation pressure depends on the bluffness and geometry of the sensor as these effects are seen to be easier to determine empirically than theoretically. Using recorded wind-velocity spectra in the inertial subrange to predict stagnation pressure as well as the upper limit of recorded infrasound wind noise has been suggested in (Raspet, Webster and Dillion, 2006). With regards to inertial subrange, the power density spectrum of wind velocity in the downwind direction has been shown in (Walker and Hedlin, 2010).

$$V_{11}(k_1) = a_2 \varepsilon^{2/3} k_1^{-5/3}, \qquad (2.3)$$

where k_1 is the wave number in wind direction assuming Taylor's hypothesis, ε is the dissipation rate of turbulence into heat, a_2 is a constant.

According to the recent observations made about power law indicated in Equation (2.3), it has been proved that it accurately describes the velocity spectra in the inertial subrange (Douglas Shields, 2005; Raspet, Webster and Dillion, 2006). As a function of the frequency f, and the mean wind speed $\bar{u}$, the wave number k_1 is given as in Equation (2.4).

$$k_1 = \frac{2\pi f}{\bar{u}}, \qquad (2.4)$$

The mathematical formulation regarding the derivation of some important equations for the stagnation pressure density spectrum in the inertial subrange can be found in (Raspet, Webster and Dillion, 2006).

Since predicting stagnation pressure eliminates the necessity for a reference pressure sensor, this procedure is potentially advantageous and helpful in testing of wind noise reduction methods. Importantly, when using the equations derived and explained in the above-mentioned study in the inertial subrange, it is required to use a research-grade anemometer that is capable of output sampling rates. The commonly used sampling rates for microphones are at least 20 Hz. Many of the existing IMS anemometers are not capable of such relatively high output rates. To derive new equations that are valid in the source range, the extension of the turbulence–sensor interaction theory into the source range has been considered in (Raspet, Webster and Dillion, 2006). In the inertial subrange, Raspet et al. (2006) verified that the predictions from the equation derived for the source range match with those from equations that previously introduced in the inertial subrange.

2.10.6.2 Advection of pressure anomalies across the sensor

1- Turbulence–Turbulence Interaction

Sensor interference discussed in the previous section is mainly characterised by turbulence–sensor interaction leading to the first type of wind noise. However, in the absence of such sensor interference and as predicted by Taylor's hypothesis, pressure differences which develop in turbulent flow may be advected with the mean wind speed across a sensor. This advection, which refers to the transport of certain properties of the atmosphere, will lead to another type of wind noise (Walker and Hedlin, 2010). As discussed in the literature, there are two possible sources of these advected pressure differences. Such sources of advected pressure anomalies are described based on a review of published identifications and measurements of different pressure fluctuations sources from a turbulent jet and in the case of no interferences from noise related with wind-sensor interaction (George, Beuther and Arndt, 1984).

The interaction between turbulences can generate pressure anomalies. In (Miles, Wyngaard and Otte, 2004), the authors describe that above some threshold sensor height, this "turbulence–turbulence interaction" is shown to be a dominant source of turbulence-induced pressure in the inertial subrange. As indicated in (Walker and Hedlin, 2010), some previous studies used dimensional analysis to derive mathematical formulation that shows the pressure power spectral density for turbulence–turbulence interaction, whereas (Hill and Wilczak, 1995) used analytical techniques to derive the same pressure power law. Though, for cases where the atmosphere is thermally stable and stratified (strongly positive ratio z/L) and when using "Large Eddy Simulation", Miles et al. (2004) found that this law needs further evaluation.

More recently, Raspet et al. (2006) derived two equations that permit the prediction of the turbulence–turbulence pressure spectrum in the inertial subrange from the velocity spectrum based on a previous study. The velocity and pressure power density have been calculated by Miles et al. (2004) from a large eddy simulation to fit their statistics and synthetic data from their simulations as well as for scaling laws based on some previous research work. The mathematical equations derived in the literature discussed above in this section are only valid in the inertial subrange. However, the theory has been extended to the source region to obtain an equation that fits the case as detailed in (Raspet, Yu and Webster, 2008).

2- Turbulence–Mean Shear Interaction

Near the ground region, where the vertical gradient of the average horizontal wind velocity acts as impedance, turbulences create another source of pressure fluctuation called

"turbulence–mean shear interaction". An empirical equation has also been developed by Raspet et al. (2006) to be valid in both source region and inertial subrange. It has been shown in (Raspet, Webster and Dillion, 2006) that the turbulence–mean shear interaction spectrum increases to the peak just before the transition to the inertial subrange where it significantly decays.

3- Correlation Distance of Turbulence

The type of wind noise discussed in here is identified in the low frequency range according to the study conducted by Douglas Shields (2005) who expanded his research upon previous results. The author examined pressure and wind data from multiple piezoelectric microphones on the ground at three field sites and located perpendicular and parallel to the dominant wind direction. To perform his experiments and develop a model for the narrow frequency band correlation of measured pressure as a function of sensor separation in downwind and crosswind directions, a frequency range (0.2–2.0 Hz) has been selected. The correlations are defined as the cross-correlations at zero lag time while the applied equations have no bearing on Taylor's hypothesis; however, they only describe the spatial structure of turbulence throughout a snapshot in time.

The results obtained in (Douglas Shields, 2005) showed that the spatial characteristics of the turbulence have a self-similar appearance in the relation between measured pressure correlation coefficients as a function of distance in wavelengths. In other words, the relationship between the spatial coherence length and the size of the turbulences is linearly proportional over a wide range of length scales. It is worth mentioning that the difference in the coherence length of wind noise between downwind and crosswind directions at specific wind speed and central frequency in the selected range is quite noticeable. Basically, it is a function of the central frequency and wind speed. The coherence length is always greater in the crosswind direction than in downwind direction with a periodicity to the correlation shown in the downwind direction (Raspet, Webster and Dillion, 2006; Walker and Hedlin, 2010).

The results obtained in (Douglas Shields, 2005) confirm the slightly frequency-dependent as there is an exponential decline in the correlation in all directions, nevertheless that there is an additional periodic factor oscillates about the zero axis in the downwind direction. As a prediction based on such results, which typically serves low-frequency sound applications, the spatial averaging of infrasound is along a line that is parallel to the wind direction and can help in considerably attenuating wind noise associated with the above-

mentioned wind noise types. Specially, for frequencies where coherence length is less than the sensor separation as explained above, notable reduction in wind noise for both crosswind and downwind directions can be obtained provided that the sensor separation is not so large as to exclude the negative portion in the correlation function. Another frequency range (0.5–5.0 Hz) was also examined in (Douglas Shields, 2005) by which notable wind noise reduction was obtained in the downwind direction for a sensor spacing in 4–8 m/s wind speed range when wind speed was recorded at 3 m height.

2.10.6.3 Differentiating between wind noise types

To design wind-noise reduction technologies that supposed to be adequately and effectively implemented with regards to the application, it is very helpful to understand the type of the recorded wind noise that one is attempting to reduce. Therefore, studying the different spectra regarding the previously mentioned wind noise types becomes increasingly important to distinguish between these types and to meet certain requirements in developing wind noise reduction methods. Going through detailed mathematical formulations that distinctly predict different spectra with unique slopes might be a principle objective towards achieving wind noise reduction aim (Walker and Hedlin, 2010).

Recent studies like (Douglas Shields, 2005; Raspet, Webster and Dillion, 2006) have been conducted to study the contribution to wind noise in the inertial subrange and the source range for turbulence–sensor, turbulence–turbulence, and turbulence–mean shear interaction wind noise types considering important key aspects such as the type of the acoustic sensor and its height above the ground. For example, a method for the attenuation of wind noise using spherical microphone wind screens with different size and at different height has been introduced in (Raspet, Webster and Dillion, 2006). Results show that in the inertial subrange and at a height of 1 m using 0.6 and 1.0 m spherical microphone wind screens size, wind noise was attenuated to the same level as that predicted by turbulence–turbulence interaction indicating that larger size would not provide further improvement.

Basically, Raspet et al. (2006) used a prediction methodology to compare between wind noises predicted by the indicated types with those measured by a variety of acoustic sensors. Therefore, the authors showed comparison results in a single diagram when using a diversity of sensors with different sizes and exterior shapes to record pressure spectra. As shown in Figure 2.9, the used sensors were; bare B&K 1/2-inch microphone (A), piezoelectric sensor (B), microphone in a 0.18 m windscreen (C), and microphone in a 0.90 m windscreen (D). The

aim was to compare power spectral densities of wind-noise pressure with six predictions of wind-noise pressure spectra based on the results of other studies. These predictions are for turbulence-sensor interaction (1, 2) (Raspet, Webster and Dillion, 2006), turbulence-turbulence interaction (Batchelor, 1951) (3) and (Miles, Wyngaard and Otte, 2004) (4), self-noise for the 0.18 m windscreen and self-noise for the 0.90 m windscreen (5, 6), respectively (Raspet, Webster and Dillion, 2006). The six predicted spectra are plotted as depicted in Figure 2.9 to be compared with the measured recordings using the four above indicated sensors (A-D).
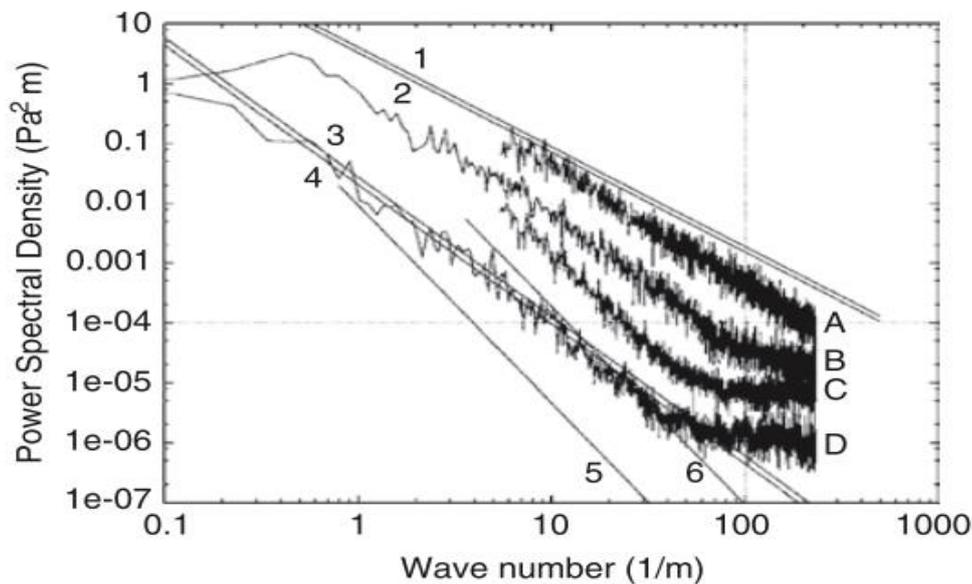
Figure 2.9. Recordings and predictions of power spectral densities of wind-noise pressure comparison. Recordings from four different sensors (A–D) are compared with predictions of wind-noise pressure spectra (1–6) (Walker and Hedlin, 2010)

Without going through deep details, such comparison indicates an important aspect which is that predicted spectrums can serve as an upper limit on the expected wind noise in some cases and this might help in identifying proper wind noise reduction method. For example, the spectrum of the bare microphone (A) is justly close in amplitude and slope to that predicted by the turbulence–sensor interaction (1, 2). There is a strong correlation between the measured recordings using the 0.9 m windshield sensor (D) with predicted wind noise by turbulence–turbulence interaction. Figure 2.9 also illustrated that the recorded pressure spectra using piezoelectric sensors (B) is about 20dB higher than that predicted by turbulence–turbulence interaction (3, 4) (Raspet, Webster and Dillion, 2006; Walker and Hedlin, 2010).

If such predictions are accurate, this indicates that the wind noise recorded by Douglas Shields (2005) is not caused by the advection of pressure anomalies across the microphone.

Douglas Shields (2005) put forward some arguments as the spectra obtained by both piezoelectric sensors and B&K were similar because both sensors were collected and inserted in a Quad Disk enclosure by which the pressure, which is independent of wind speed and direction, can be measured (Nishiyama and Bedard, 1991). These clarifications suggest that piezoelectric sensors used in (Douglas Shields, 2005) were not measuring turbulence–sensor noise. In other words, they are not greatly influenced by wind–sensor interaction. The author also provided details regarding the frequency band of the pressure data recorded in various winds.

It has been shown that Quad Disk should be insensitive to turbulence–sensor interaction through the developed technique that has been presented in (Wyngaard, Siegel and Wilczak, 1994) and used to predict the effect of velocity variations on pressure measurements. The turbulence–turbulence predictions showed a good fit with the recordings using windscreen microphones. Based on that, Raspet *et al.* (2006) interpreted the results to indicate that piezoelectric sensors are dominated by fluctuations of a smaller stagnation pressure and seen as quite aerodynamic sensors.

2.11   Wind Noise Reduction Methodologies

Microphones are quite often used in different outdoor data acquisition scenarios such as recording, sensing of acoustic data for monitoring purposes, and even for communication such as in mobile communication devices and this might happen in extreme acoustical environments. Wind noise which is picked up the microphone during such scenarios is an annoying factor. During the past, a considerable deal of research work has been conducted to develop methods that designed for wind noise reduction. A review of literature is provided in this section with regards to the methods applied for wind noise problem along with brief discussion of their relevance and limitations. However, no comparison will be made between these methods as such comparison requires experimental investigation and implementation which goes beyond the scope of this thesis.

To deal with microphone wind noises that interfere with the signals of interest, large numbers of standards or common approaches ranging from fixed, optimal to adaptive filtering have been applied to mitigate wind noise in microphone signals. Due to the broadband and time varying nature of microphone wind noises, such general noise suppression algorithms show some but limited effectiveness (Schmidt, Larsen and Hsiao, 2007). Microphone wind noise leads to listener fatigue as it is often annoying. It is impulsive and non-stationary in nature

with high amplitude that may exceed the nominal amplitude of the signal of interest. Due to that, conventional noise reduction schemes, such as spectral subtraction or statistical-based estimators, cannot effectively attenuate wind noise (Schmidt, Larsen and Hsiao, 2007; King and Atlas, 2008). Thus, special detection and processing might be required to better reduce the effect of wind noise when using such methods (Nemer and Leblanc, 2009).

Additionally, traditional remedy for wind noise compromises the quality of the sensed acoustic data. Wiener filter method for example, which is arguably one of the most well-established random noise optimal removal filters used for removing noise from a signal when the signal of interest and the noise have different frequency characteristics, shows some but limited performance (Hänsler and Schmidt, 2005). Wiener filter method has been widely used in speech enhancement applications. The method is mainly based on assuming that the second order statistics of the signal and noise processes are known. Hence, it works by attenuating frequencies where the noise is expected to be the most dominant. Wiener method assumes stationary signals, which might give inappropriate approximation for speech and wind noise when considering speech as the signal of interest (Dietrich and Utschick, 2005; Vaseghi, 2008).

Among the classic noise reduction methods that have been used for decades, spectral subtraction has been used for wind noise reduction (Boll, 1979). This method is essentially based on subtracting an estimate of the noise magnitude spectrum from the noisy signal magnitude spectrum. As a complete process and by using the phase of the original noisy signal, the method transforms the estimated noise magnitude spectrum back to the time domain. Using Voice Activity Detection (VAD), the noise estimate is mostly obtained during the silence period or speech pauses, similarly to Wiener method which also assumes that the noise is stationary. Therefore, the method is unable to obtain new noise estimates during speech or, generally, as long as the signal of interest is there (Boll, 1979). Wiener filtering and spectral subtraction methods are commonly used for noise reduction in single channel applications (Hänsler and Schmidt, 2005; Vary and Martin, 2006).

Adaptive versions have been developed for both methods which helped in relaxing the assumption of stationarity (Srinivasan, Samuelsson and and Kleijn, 2003), that is the noise can be estimated during the presence of the signal of interest. Besides the fact that they are robust and easy to implement, these methods have been comprehensively studied and generalised through many years which is considered one of their main advantages. An advanced method uses a hybrid model built around neural networks to quantify speech intelligibility was

developed by (Li and Cox, 2007). The authors indicated that the model offered more promise in terms of circumventing the use of artificial testing signals, reducing measurement channels and facilitating in situ assessment of speech intelligibility.

The approach of modelling the sources in the noisy signal independently has been involved in more recent methods in order to use these models to find the best estimate of the signal of interest and noise signal. However, individual signals can be separated by using binary masking as explained in (Cermak *et al.*, 2007). Based on this approach, many models have been introduced, such as Gaussian Mixture Models (Ding *et al.*, 2005), Hidden Markov Models (Roweis, 2001), Non-negative Sparse Coding (Schmidt, Larsen and Hsiao, 2007) and Vector Quantization (Ellis and Weiss, 2006). For speech signals cleaning, this approach often models an individual speaker rather than speaker independent.

Some methods have been used to attenuate the effect of wind noise when speech is the wanted signal such as comb filters. Such methods are based on reinforcing the harmonic nature of speech signals depending on accurate pitch estimation, which is difficult to achieve in noisy environments (King and Atlas, 2008). Always with regards to speech quality improvement in the presence of wind noise, post-filters have also been used in model-based speech coders (Juin-Hwey Chen and Gersho, 1995). In such filters, it is to emphasise the formant frequencies and deemphasise the spectral portion where noise contributes the most to the observed distortion (Nemer and Leblanc, 2009). Based on adapting the emphasis parameters, an extension of the conventional post-filter has been proposed in (Grancharov *et al.*, 2008). A time-domain adaptive post-filter was proposed in (Nemer and Leblanc, 2009) to reduce wind noise in corrupted speech. This filter is based on tracking the changing envelop spectrum of wind noise similarly to other post-filters except it deemphasises the wind 'resonance'. LPC (Linear Predictive Coefficients) analysis was used in the method explained in (Nemer and Leblanc, 2009) to distinguish between frames contain the most of speech energy and wind-only frames.

As quality and intelligibility of speech signals when superposed by wind noise can be greatly degraded, many methods have focused on wind noise reduction from speech signals in many applications such as with mobile devices (Nelke *et al.*, 2014). In such applications, which mostly equipped with a single microphone, wind screen cannot be offered; however, it is required to develop systems capable to reduce the effects of wind noise by using signal processing techniques. The accurate estimation of the noise Power Spectral Density (PSD) is

seen as a crucial part of such systems as indicated in (Nelke *et al.*, 2014). Based on this concept, many single microphone methods have been introduced for the estimation of noise PSD from noisy speech signals (Martin, 2001; Hendriks, Heusdens and Jensen, 2010; Gerkmann and Hendriks, 2011). Authors in (Nelke *et al.*, 2014) investigated the algorithms used in these methods and came up with a conclusion that the assumption of considering noise signal as slower varying over time than the speech signal is not true for wind noise signals. According to (Nelke *et al.*, 2014), due to inaccurate estimates of noise PSD, such conventional algorithms provide insufficient level of noise reduction.

Exploiting the spectral characteristics of the wanted signal (e.g., speech) and noise in order to estimate the wind noise PSD is the applied approach in (Nelke *et al.*, 2014). Nelke *et al.* (2014) considered the magnitude spectrum towards higher frequencies of wind noise and the harmonic structure of the wanted speech signal. Methods that provide an estimate of wind noise PSD are introduced in (Kuroiwa *et al.*, 2006; Hofmann *et al.*, 2012). Methods that dealing with wind noise reduction in single microphone signals can be found in (Kuroiwa *et al.*, 2006; King and Atlas, 2008; Nemer and Leblanc, 2009; Hofmann *et al.*, 2012) in which the concept relies on directly modifying noisy input signals.

In adverse listening conditions and for speech understanding in hearing aid applications for example, research has focused on the development of noise reduction algorithms. In such application, dual microphone technologies have been introduced using innovative designs which brought notable progress in noise reduction including wind noise and helped in enhancing overall speech perception in the presence of background noise. The coherence of dual microphone recording has also been exploited to develop methods that seen to be efficient as indicated in (Nelke *et al.*, 2014). In an exemplary hearing aid application, an algorithm that performs a statistical analysis of the microphone signals is presented in (Zakis and Tan, 2014). This algorithm is substantially more robust against non-wind causes differences, and therefore against false wind noise detection.

Although behind-the-ear (BTE) instruments for instance helped in thriving multi-microphone technology, however, its inherent microphone placement is susceptible to wind noise at the input of the microphone which is considered a downside of such instruments. Random turbulence at the end of the microphone port creates wind noise which leads to user dissatisfaction in spite of reporting such instruments as the preferred style for hearing care professionals and widely acceptable in recent years (Chung, 2004). This indicates that wind

noise still seen as a major problem in outdoor data acquisition in many applications. According to (Dillon, 2008), amplifying 100 dB SPL of unpleasant noise by a hearing aid is not worse than using two hearing aids to amplify this noise. For windy situations, it is not surprising that hearing impaired users do not rate bilateral hearing aids than unilateral as it is possible for the user to minimise wind noise at the microphone of the hearing aid device by orienting the head, however, achieving this is often impossible for two hearing aids at once. Therefore, this indicates that a substantial problem remains.

Microphone wind induced noise has been a problem in audio systems that comprise microphone arrays or directional microphones as it is generated at even low wind speeds. Typically, microphones are shielded using thick fuzzy materials or acoustically transparent foam. Placing mechanical windscreens in front of microphones sound inlet opening could not lead to optimal solution. In fact, this may result in reducing the overall performance of the microphone. The problem of wind noise is well known in hearing aid application with directional microphones and DSP systems that generate an output with directionality (Elko, 2007; Petersen *et al.*, 2008).

To filter out the noise in the case of multichannel with more microphones available, the correlation between the desired signals in the microphones can be used. Such approach has been used in methods that are computationally expensive like Independent Component Analysis (Knaak, Araki and Makino, 2007), directivity based applications like Beamforming (Dmochowski and Goubran, 2007), and also a combination between the two as in (Qiongfeng and Aboulnasr, 2007). Methods utilising two or more microphones are considered the most efficient and effective techniques for reducing wind noise as they are based on exploiting the difference in propagation delay between the acoustic signals and the wind. Such methods are computationally considered prohibitive as well as their complicated setups are difficult to deploy which limit their usage (Marshall, 1984; Brandstein and Ward, 2013). Multichannel approaches show further improvement with regards to speech quality. A frequency domain criterion has been developed for speech distortion based on multichannel Wiener filtering method (Doclo *et al.*, 2007). Also, for speech enhancement, distributed microphones approach has been used to develop the so-called blind-matched filtering (Stenzel and Freudenberger, 2012).

Regarding the low frequency band, most of the up-to-date techniques have focused on reducing the previously mentioned types (turbulence–sensor, turbulence–turbulence, and

turbulence–mean shear interaction). As discussed earlier, these types of wind noise depend on the length scales of the turbules. It has been seen that the coherence of turbules differs as a function of turbules' size and the distance traveled (Walker and Hedlin, 2010). Acoustic signals propagate at much faster velocities than wind with approximately 100 times, and might remain coherent at considerable separations of certain kilometers (Douglas Shields, 2005; Walker and Hedlin, 2010). Most of wind noise reduction techniques are essentially based on acoustic signals and the contrasting spatial coherence lengths of turbules. Also, such methods can be grouped into four categories; wind-sensor isolation strategies, acoustic integration filters, digital filtering with dense microphone arrays, and instantaneous integration sensors (Walker and Hedlin, 2010).

For such low frequency band, there are several filters of mechanical nature that have been developed using low-frequency microphones (micro-barometers). Among these filters is Daniels filter for wind noise reduction, details can be found in (Daniels, 1952, 1959; Walker and Hedlin, 2010). Certain developments to Daniels wind filter have been made with regards to study more isotropic pipe configurations such as circular pipes (Burridge, 2010; Grover, 2010). Daniels wind filter is the basis of designing and developing other filters such as rosette filter which is considered as a standard wind-noise filter used at IMS array sites. This filter was designed by Alcoverro in the late 1990s, details can be found in (Alcoverro and Le Pichon, 2005; Walker and Hedlin, 2010). Microporous hoses are also designed upon the basis of Daniels wind filter; however, their configurations may vary from linear to circular porous hoses. Many studies have been conducted to test this filter under several conditions considering various configurations of porous hoses in some empirical tests where it has fared well (Herrin *et al.*, 2001; Howard, Dillion and Shields, 2007). The optical fibre infrasound sensor is also used to reduce wind noise by directly measuring the integrated pressure change along a path in a different manner compared to the above mentioned filters although it gives similar results to the rosette filter; details can be found in (Zumberge *et al.*, 2003; Walker *et al.*, 2008).

A system known as "distributed sensor" or "adaptive processing with a dense array" has recently been developed. This system is based on extracting the signal of interest in post-processing by recording all the traces or using on-the-fly algorithms to reduce wind noise through weighted-averaging or filtering schemes that adapt to the changes in wind conditions, producing a single trace as an output. More details of this can be found in (Dillion, Howard and Shields, 2007; Walker and Hedlin, 2010). Instead of averaging over several sensing surfaces, a new approach that permits to isolate the sensor from the advected turbules has been

introduced. It is a completely different approach to wind-noise reduction which has been developed by introducing porous media filters and wind barriers that share similarities with each other, since they attempt to isolate the sensors from the wind. Details can be found in (Attenborough, 1983; Attenborough *et al.*, 1986; Sabatier *et al.*, 1986; Hedlin and Raspet, 2003; Shams, Zuckerwar and Sealey, 2005; Christie and Campus, 2010).

2.12   Research Problem Statement and Proposed Solution

The purpose of this review was to view gaps in the literature concerning wind noise sensitivity of microphones which is always seen as a major problem for outdoor applications. In this study, different criteria have been used to review the literature from the related studies within the past forty years and see why wind induced in microphone signals is an issue of concern for outdoor acoustic data acquisition applications as well as how methods to circumvent this problem have developed and why further developments are still required.

It is clear from the research reviewed that single-microphone wind noise is still a technical challenge and even dual microphone and multi-microphone or microphone array technologies are still facing problems concerning wind noise and having several limitations. Along with this, it is also clear that the field of applying acoustic sensing in outdoor applications, such as soundscapes monitoring and smart environments, just regarding the environmental noise removal is varied and continues to be studied and analysed in order to most benefit society at large. Wind noise sensitivity of microphones is still being debated, though, and continues to be problematic in outdoor applications as concluded from the different studies indicated in this literature review. This field is very important as at its centre is a concern with exploiting to the maximum of the untapped power of sound for added-value applications and helping users in many ways. Helping users become aware of the new smart environments and getting them to see the importance in developing new applications and overcoming many existing problems is also extremely important in the current society.

For environmental noise level measurement, especially long-term measurement or monitoring, microphone wind noise is added to the recorded noise level, giving inaccurate results, residual microphone wind noise remains problematic. However, many standards or common approaches have been applied to reduce wind noise in microphone signals as discussed in this chapter. What is of interest in this review is the signal processing standards that have been developed and applied to broadband wind noise suppression in audible frequency band. However, due to the broadband and time varying nature of microphone wind

noise these methods show some but limited effectiveness particularly for single-microphone which is the problem addressed in this thesis.

There have been much research and discussion conducted on wind noise problem including theories, procedures, advantages and limitations of the existing methods that have been proposed to reduce the harmful effect of wind noise based on different approaches through developing many signal processing techniques. However, none of these methods can be considered as an optimal solution as the problem is still there. These methods worked to some extent but have intrinsic limitations of distinguishing and separating wind noise components and difficulties in retaining accurate signal energy levels to meet the measurement requirements.

The main aim is to evaluate possible outcomes from this literature to reveal the limitations of the existing methods and indicate wind noise issue for outdoor applications and recording situations that eventually led to state the problem addressed in this thesis. Determining the expected outcomes of this literature was based on the results and reasons that indicate the existence of the problem as well as taking into consideration the recommendation for future work from the previous studies. More research and testing are therefore required to gain a better understanding of the nature of the problem after pointing out the major limitations of the existing methods as of not properly identifying wind noise components. Based on the key aspects discussed and the major limitations revealed in this review, the problem addressed in this thesis has been stated according to the following.

- Soundscapes signals are composed of acoustic target of interest and background noise (wind) which harmfully affects acoustic signals of interest.

- Although there are many existing methods, but single-microphone wind noise reduction is still a technical challenge and an open-ended problem for further extensive research.

- Wind induced microphone noise is also unsolved problem in hearing aid applications and outdoor audio recording such as the field of news broadcasting.

The literature has been comprehensively reviewed in this chapter regarding the problem of microphone wind induced noise including the application area and related concepts, wind noise theory and spectral character, and the methodologies and common approaches. This review included most of existing wind noise suppression methods which have been discussed earlier in this chapter. The review outcomes have been summarised and given in tables in the

Appendix A. In these tables, wind noise reduction methods are categorised into different groups that share similar characteristics, theory, approach and application, however, only examples were given. Also, these tables illustrate the main advantages and limitations of each category in an easy way to refer to when needed. Not all noise reductions methods have been mentioned in the summary, however, only those which have been implemented for wind noise even for specific purposes such as speech enhancement.

With the above in mind, microphone wind noise reduction problem in outdoor applications requires alternative solutions. Contemporary and powerful signal processing methods are sought to address the wind noise issues and can yield better results. The focus of this study is on the separation of microphone wind induced noise through acoustic means, however, methods of the prior art, such as mechanical windscreens, are not of interest. Linear separation in subspace seems to be potential solution to circumvent microphone wind noise problem. The proposed solution in this study is therefore based on developing a new separation signal processing technique based on Singular Spectrum Analysis.

The review outcome eventually led to suggest modifying and developing the existing SSA method which might be a solution to particularly wind noise problem and, of course, there are several reasons behind this selection. The SSA has been considered a non-parametric technique that works with arbitrary statistical processes, either stationary or non-stationary, linear or nonlinear, Gaussian or non-Gaussian (Hassani, 2010). The method is non-parametric which makes no prior assumptions about the data, however, this gives a great advantage for the SSA over traditional methods of time series analysis. Due to its potential capabilities, the SSA method has been developed, modified, and adopted to solve many problems as evidenced by its implementation in a wide range of applications which increases the motivation of its selection in this study. For example, the SSA has been recently seen many successful paradigms in the separation of biomedical signals (e.g., separating heart sound from lung noise). It has also been successfully implemented to de-noise signals in many various applications.

In recent years, traditional methods applied for time series analysis such as power spectra have been augmented by new methods. As an alternative to traditional digital filtering approaches, the SSA has been presented in many applications. The SSA has been introduced in a wide range of applications as a de-noising and raw signal smoothing method. It is mainly based on principles of multivariate statistics (Alonso, Castillo and Pintado, 2005). Basically, a

number of additive time series can be obtained by decomposing the original time series to identify which of the new produced additive time series is part of the modulated signal, and which is part of random noise (Hassani, 2010). However, the separation approach has been introduced, and this encourages the idea of further developing this approach instead of filtering and working on the separability for wind noise components in this research as one of the key concepts.

A great deal of research work has been conducted on the SSA to consider it as a de-noising method (Hassani, Dionisio and Ghodsi, 2010). The SSA has also been applied for extracting information from noisy dataset for biomedical engineering and other applications (Ghodsi *et al.*, 2009). The method has been employed and shown its capabilities for noise reduction for longitudinal measurements (Hassani *et al.*, 2009). The superiority of the SSA over other methods in biomechanical analysis was clearly demonstrated by several examples presented in (Alonso, Castillo and Pintado, 2005). With the above in mind, there is always a possibility to modify and develop the SSA in terms of its key aspects to the problem addressed in this thesis.

2.13 Summary

Improving signal processing methods to clean soundscapes singles and reduce the harmful effect of wind noise on the target acoustic events in outdoor data acquisition is an issue of concern that requires further systematic studies and research. The reduction or separation of such broadband noise from desired sounds is a vital issue to be addressed for the future design regarding soundscape monitoring and smart city applications. Therefore, high importance should be given to environmental noise reduction while developing functionalities in the future design of acoustic and audio monitoring solutions. In this chapter, many related key concepts and technical issues with regards to the application area of this study have been discussed.

Enabling a wide range of applications and enhanced services with high social and technological value require more intelligent usage of the audio modality (Hollosi *et al.*, 2013). To lead to success, system designers should be aware of all potential limitations in implementing acoustic sensing for audio monitoring in smart cities applications. Consequently, it is important that such solutions focus on outstanding events and environmental noise control and monitoring as well as providing attractive and fascinating applications to support high societal value applications and deliver new innovative range of services rather than storing audio data and voices of the inhabitants of smart environments.

Additionally, in this review, topics like wind noise problem in outdoor data acquisition, the importance of de-noising approach in signal processing, wind noise specifications and characteristics, and existing wind noise reduction methodologies have been reviewed. Also, wind noise theory and details including the physics and mechanism of the wind, a description of wind noise nature and spectrum in audible frequency band, and wind noise types induced by turbulence interaction in inertial subrange and source region specifying low frequency band are reviewed and presented in this chapter.

From this review, microphone wind noise has been mentioned as a serious problem in many applications for users who enjoy outdoor activities for normal hearing as well as for hearing-impaired listeners. Wind noise is annoying and can create distortion by overloading the microphone and masking desired signals. In this review, it has been shown that at commonly encountered wind speeds, wind noise is at high levels and hence it needs to be suppressed. It has been reported that for hearing aid users as indicated in (Kochkin, 2005, 2010), wind noise is the second-highest cause of dissatisfaction behind noisy situations. As discussed in this literature review, microphone wind-induced noise is still an issue of concern; however, as previously discussed, existing methods have many limitations. Single-microphone wind-noise removal is still a technical challenge for further extensive research and open-ended problem. A review outcome that led to propose modifying and developing the SSA method is also given in this chapter.

3 CHAPTER THREE   SINGULAR SPECTRUM ANALYSIS METHOD

3

Singular Spectrum Analysis

3.1 Introduction to Singular Spectrum Analysis Technique

In the recent years, Singular Spectrum Analysis (SSA), which is seen as a powerful technique, has been developed in the field of time series analysis. The SSA is a novel method that can be used to solve many problems in different fields such as in signal processing. This chapter introduces a thorough treatment of the SSA method in the context of wind noise separation with the author's own contribution towards establishing a systematic approach of the method in the framework of wind noise separation by developing a profound understanding of the separation concept and focusing on certain key elements throughout this chapter and the next two chapters.

As stated in (Elsner and Tsonis, 2013), the singular spectrum term came from the spectral (eigenvalue) decomposition of a given matrix $\mathbf{A}$ into its set (spectrum) of eigenvalues. These eigenvalues, denoted by λ, are specific numbers that make the matrix $\mathbf{A} - \lambda\mathbf{I}$ singular when the determinant of this matrix is equal to zero. Singular spectrum analysis is the analysis of time series using the singular spectrum. Spectral decomposition of matrices has many applications to problems in the natural and related sciences and is fundamental to much theory of linear algebra.

The widespread use of the SSA as a method for time series analysis is relatively recent. However, from applications of dynamical systems theory which is also called chaos theory, its usage has been emerged to a large extent. As indicated in (Elsner and Tsonis, 2013), the SSA was first used in biological oceanography by Colebrook in 1978 and was introduced into dynamical systems theory by Fraedrich in1986 and Broomhead and King in 1986 as well. The SSA is namely a linear approach that becomes attractive in the analysis and prediction of time series. The SSA method has a growing strength over classical spectral methods due to the data-adaptive nature of its basic functions which makes the approach even suitable for analysis of some nonlinear dynamics.

The SSA has a capability to provide useful insights into a range of systems and can be used to make predictions even when data amounts are modest. Throughout scientific research, a physical system can be described and characterised by means of measured time series. Consecutively, adequate descriptions can lead to useful forecasts of the behaviour in the future.

However, Information can be extracted from short and noisy time series based on the design of the SSA and therefore provide insight into the dynamics of the underlying system that generated the series. These dynamics could be unknown or only partially known (Ghil *et al.*, 2002).

The SSA has been applied in a wide range of possible application areas which diverse from mathematics, physics, and engineering to economics and financial mathematics. Also, the SSA is considered as a promising technique in some other application areas that may include meteorology, oceanology, social sciences, market research and medicine. Not only is the SSA limited to these application areas, but it also seems to be a promising method for noise reduction and even any seemingly complex time series could provide another example of a successful application of the SSA (Vautard, Yiou and Ghil, 1992; Hassani, 2010).

Decomposing the time series and then reconstructing the original series are the two complementary stages of the basis SSA method. The main concept in studying the properties of the SSA is how well different components can be separated from each other. It is often observed in series with complex structure the lack of approximate separability. For such complex series and series with special structure, there are many ways of modifying the SSA leading to different versions like SSA with single and double centring, Toeplitz SSA, and sequential SSA (Moore and Grinsted, 2006; Golyandina and Shlemov, 2015). By the presence of noise, the ability to extract valuable information from time-varying signals is limited. Since many decades, methods of noise reduction are a subject of widespread interest in several fields and applications such as physical systems, communication, experimental measurement and signal processing (Kostelich and Yorke, 1988).

To the best of the author's knowledge based on reviewing the literature, which increases the motivation of selecting this method, the SSA has not been used for wind noise reduction although it has generally been shown as a useful tool for noise reduction in some applications. This is the first study to undertake the SSA for further development in this context.

3.2 Applications of the SSA

The SSA is an innovative nonparametric technique and model-free method for time series analysis, namely classical time series analysis, multivariate statistics, dynamical systems and signal processing. It is basically a mixture of mathematical and statistical analyses (Elsner and Tsonis, 2013; Yang *et al.*, 2016). A large and growing body of literature has investigated the

application of the SSA in many disciplines. A comprehensive review on applying this technique, such as in digital signal processing, in oceanographic research and in environmental systems for air pollution studies has been given in many previous studies such as in (Fukuda, 2007; Hassani *et al.*, 2009; Chu, Lin and Wang, 2013; Golyandina and Shlemov, 2015).

The SSA is generally applicable for many practical problems such as the study of classical time series, dynamical systems and signal processing along with multivariate statistics and multivariate geometry. It is also an effective method for the extraction of seasonality components, extraction of periodicities with varying amplitudes, finding trends of different resolution, smoothing, simultaneous extraction of complex trends and periodicities, finding structure in short time series, etc., (Hassani, 2010; García Plaza and Núñez López, 2017; Xu, Zhao and Lin, 2017).

The basic capabilities of the SSA can lead to solve all these problems. Additionally, the method has many crucial extensions such as the multivariate version which permits the simultaneous expansion of several time series. Also, several forecasting procedures for time series can be examined by the SSA method (Launonen and Holmström, 2017; Rodrigues and Mahmoudvand, 2017). For change-point detection in time series, the same ideas are used as stated in (Moskvina and Zhigljavsky, 2003). The SSA has been employed for traffic load in large-scale WLAN infrastructure. However, it has been found that using a small number of leading principle components, the time records of traffic load at a given access point has a small intrinsic dimension and can be accurately modeled (Tzagkarakis, Papadopouli and Tsakalides, 2009).

According to (Kostelich and Yorke, 1988), generally the time series, in many cases, can be viewed as a dynamical system. The procedure described in (Kostelich and Yorke, 1988), was for reducing noise levels in some experimental time series related to dynamical systems. The authors described a noise-reduction method based on the principle of the SSA that works by taking numerous nearby points in phase space (corresponding to broadly varying times in the original signals) in order to find a local approximation of the dynamics.

As a method of prediction and forecasting specially for real-time records that usually have a complex structure, the SSA shows potential capabilities since it is non-parametric which makes no prior assumptions about the data. Ghil *et al.* (2002) reported the capability of the SSA as a useful method for extracting target components in climate prediction. In the SSA-based methods, some statistical and probabilistic concepts are employed. The SSA relies on

the bootstrapping to obtain the so-called confidence intervals for the predictions and forecasts (Patterson *et al.*, 2011; Hansen and Noguchi, 2017; Mahmoudvand, Konstantinides and Rodrigues, 2017).

Environmental systems are often considered as non-linear and ill-structured domains, and also involve multidisciplinary factors such as global and local ecological and socio-economic factors (Fukuda, 2007; Yang *et al.*, 2016). Therefore, Fukuda (2007) indicated that before investigating the dataset of interest and applying a process of knowledge discovery in databases (data mining), which found to be a useful tool in such systems, another tool should be used to reduce the noise because environmental dataset whether it is climate or pollution is commonly noisy and skewed. According to that, a primary step which is reducing the noise has to be taken, even though defining the noise component from such a noisy structure can be difficult. Hence, the SSA has been used for this purpose (Maddirala and Shaik, 2016; Traore *et al.*, 2017).

A great deal of research work has been conducted on the SSA to consider it as a de-noising method (Elsner and Tsonis, 2013; Jiang and Xie, 2016; Qiao *et al.*, 2017). The SSA has also been applied for extracting information from noisy dataset for biomedical engineering and other applications (Ghodsi *et al.*, 2009). It has been recently seen many successful paradigms in the separation of biomedical signals, e.g., separating heart sound from lung noise and many other applications (Ghodsi *et al.*, 2009). In (Ghaderi, Mohseni and Sanei, 2011), the authors set up the SSA in a commonly used diagnostic method using respiratory data to improve the pulmonary auscultation through separating and localising heart sounds. Using the same concepts, Sebastian and Rathnakara (2013) identify several advantages of the SSA in their study through introducing a method of selecting adaptive Eigen triples that correspond to heart sounds. Sanei *et al.* (2011) applied the SSA to detect a murmur from heart sound by working on the statistical properties of their data decomposed using the SSA. Always in the same area of application, the SSA was deployed to eliminate environmental sound in signals composed of heart sound signals by using eigenvalues spectra to select effective principal components based on the dominant eigenvalues produced by the SVD (Zeng, Ma and Dong, 2014).

The SSA has been implemented for extracting the rhythms of the brain of electroencephalography (EEG) (Tomé *et al.*, 2010; Hu *et al.*, 2017). In (Enshaeifar *et al.*, 2016), the SSA has been deployed for the decomposition of EEG data in sleeping analysis and a method for classifying sleeping into five levels based on this decomposition has been

introduced. Always in this field, the SSA has been implemented in (Mohammadi *et al.*, 2016) to optimise the efficiency of time and frequency domain analysis of sleep EEG. In the latter study, a new approach was proposed to automatically identify certain important factors related to brain waves from sleep EEG signals such as sleep spindles, brain waves and K-complexes. In another application, the method was used to discriminate the sound of walking into three groups which namely are walking upstairs, downstairs, and walking level (Jarchi and Yang, 2013). To examine this issue, Jarchi and Yang, (2013) used a triaxial sensor accelerometer placed on the ear to record the sound of walking. Furthermore, the SSA has been implemented for the detection of Parkinson's tremor from electromyograms (EMG) signals (Eftaxias *et al.*, 2015).

The SSA method has been employed and shown its capabilities for noise reduction for longitudinal measurements and surface roughness monitoring (Hassani *et al.*, 2009; García Plaza and Núñez López, 2017). It has also been implemented for structural damage detection (Lakshmi, Rao and Gopalakrishnan, 2017). The superiority of the SSA over other methods in biomechanical analysis was clearly demonstrated by several examples presented in (Alonso, Castillo and Pintado, 2005). Regarding potential classification accuracy and detecting weak position fluctuations, the SSA shows a great improvement and outperforms other popular methods such as Empirical Mode Decomposition (EMD) (Jiang and Xie, 2016; Qiao *et al.*, 2017; Xu, Zhao and Lin, 2017).

As highlighted by (Harmouche *et al.*, 2017), the SSA has been used as a non-stationary signal decomposition tool. It has been shown that the decomposed structures of the SSA can be applied to data mining techniques for image segmentation. Therefore, in (Traore *et al.*, 2017), the method has been used for de-noising and structure analysis as an application to acoustic emission signals produced from nuclear safety system. Moreover, the SSA was used in image processing for analysing the effect of the movement disability of a patient on their grasp as to determine the treatment type (Lee *et al.*, 2013). The SSA was also used to remove noise components in hyperspectral imaging for feature extraction (Zabalza *et al.*, 2014). The authors showed that the ability of the distinction of the features has been improved.

The SSA has been used for data pre-processing for further data mining application by removing noisy structures from climate and pollution dataset as in (Fukuda, 2007). In the latter study, the author reported that using the SSA provides two crucial benefits. Firstly, the SSA helped in identifying noise in the structures as the method decomposes the noisy time series

into a number of additive components for separating out high and low frequency signals from the original time series which later grouped and reconstructed to form a new time series. Secondly, the decomposed structures improve the results of the tree construction algorithm. Using the same concepts, the SSA was used to separate low frequency components from high frequency ones as in (Harris and Yuan, 2010; Mert and Milnikov, 2011).

Applying the SSA as a de-noising method is to identify which components can potentially be noise. However, according to the method introduced in (Fukuda, 2007), it has been suggested that such components should be examined on the classification accuracy. For instance, adding insignificant components which are generally high frequency components to the main structures which commonly low frequency can lower or have no influence on the classification accuracy. However, removing significant components, including some high frequencies, may lower the classification accuracy as well, which suggests that these components are unlikely to be noise. Improving separability has therefore been considered as a key element in the SSA using non-orthogonal decompositions of time series and independent component analysis as in (Hassani, Mahmoudvand and Zokaei, 2011; Golyandina and Shlemov, 2013; Golyandina and Lomtev, 2016).

The procedure of using the SSA as a noise reduction method is mainly depends on common steps regardless the type of dataset. As stated in (Fukuda, 2007), the SSA is specifically used in pre-processing phase for data mining approach which later required decision tree classifier to be applied on climate attributes to predict three carbon monoxide (CO) levels (high, medium, and low). The method relies on the classification accuracy to assess the improvement and effectiveness of applying the SSA in the pre-processing phase. In the study carried out by Fukuda (2007), the SSA was introduced as a noise reduction method for data pre-processing manner to later apply data mining technique and a decision classifier. Investigating noisy climate time series by using the SSA posited the way of forming a number of additive components to separate out noise from such time series. According to the specific purpose of the study in (Fukuda, 2007) which is the prediction of the different air pollution levels of CO, these components are used further to construct decision trees.

One of the most significant current discussions in the SSA application as stated in (Hassani, Dionisio and Ghodsi, 2010; Maddirala and Shaik, 2016) is that the signals obtained by the SSA decomposition differ from those obtained by filtering out frequency bands with the Fourier Transform for instance. However, the reason behind that is because such signals are

generated from eigenvectors which are not purely related to the frequency. Such characteristic gives the SSA a significant advantage which is the ability of adding or removing the additive components either low or high frequencies for certain applications such as decision tree construction.

Based on the literature, the SSA has been compared to many classical methods and algorithms used for time series analysis to show its potential capabilities and application areas (Hassani, Heravi and Zhigljavsky, 2009). As mentioned in (Hassani, 2010), in economic applications, the multivariate SSA has been put in application compared to an econometric model known as vector auto-regression (VAR) as well as for automatic methods of identification within the SSA framework and for recent work in the SSA new developments. Furthermore, the SSA was considered as a useful technique that could compete with more standard methods in even the area of nonlinear time series analysis (Vautard and Ghil, 1989; Jiang and Xie, 2016).

As mentioned previously, a considerable amount of literature has been published on developing and implementing the SSA in several fields and applications. In this review, however, it is important to discuss the methods of developing the SSA for noise reduction and data pre-processing in different studies rather than details of other specific application areas. A number of research works has been conducted on the SSA to consider it as a filtering method as in (Hassani, Dionisio and Ghodsi, 2010). The method has also been applied for extracting the noise information to therefore be used as a biomedical diagnostic test (Ghodsi *et al.*, 2009). According to (Hassani *et al.*, 2009), the SSA can be employed and shown its capabilities for noise reduction for longitudinal measurements.

The SSA has been introduced in a wide range of applications as a de-noising and raw signal smoothing method such as analysing and de-noising acoustic emission signals (Harmouche *et al.*, 2017; Traore *et al.*, 2017) and smoothing of raw kinematic signals (Alonso, Castillo and Pintado, 2005). Basically, a number of additive time series can be obtained by decomposing the original time series to identify which of the new produced additive time series can be part of the modulated signal and which be part of random noise (Alonso, Castillo and Pintado, 2005). It has been showed in (Fukuda, 2007), that using the SSA for data pre-processing is a helpful procedure that encourages improving the results of any time series for data mining. The SSA has been known as a two-step point-symmetric de-nosing method of time series. Noiseless signals can be obtained with minimum loss of data (Qiao *et al.*, 2017).

The SSA is considered as a suitable data pre-processing tool for data mining to make it more effective on noisy time series. According to the method introduced in (Fukuda, 2007) for a decision tree classifier for a noisy data, the SSA has also been explored as a noise reduction approach. The SSA can provide groups of additive components varying from low frequency to high frequency by decomposing the noisy time series. The method explained in (Fukuda, 2007) was to decompose climate data using the SSA and then to construct decision trees for the prediction of air pollution levels specifically the percentage of carbon monoxide when applying data mining approach. In the study conducted by (Alonso, Castillo and Pintado, 2005), biomechanical analysis was the application area where motion capture systems were used. Such systems introduce different systematic errors that appear in recorded displacement signals in the form of noise. Due to the unacceptable level of amplification of such noise when differentiating displacements in order to obtain velocities and accelerations, the SSA was used to smooth the displacement signal and reduce the noise introduced by the experimental system.

As another application area of implementing the SSA and still for noise reduction application is the field of seismic records. An interesting work has been carried out by (Oropeza and Sacchi, 2011) where a rank reduction algorithm that permits simultaneous reconstruction and random noise attenuation of seismic records was presented. Rank reduction has been proposed as a method to mitigate noise and recover missing traces. The proposed technique in (Oropeza and Sacchi, 2011) was mainly based on multivariate singular spectrum analysis (MSSA) and required organising spatial data into a block Hankel matrix at a given temporal frequency. In ideal conditions according to their proposed method, the matrix is of rank k, where k is the number of plane waves in the window of analysis. Notably, the rank of the block Hankel matrix of the data increases due to additive noise and missing samples. The proposed algorithm was an iterative one that resembles seismic data reconstruction with the method of projection onto convex sets along with the adaptation of randomised Singular Value Decomposition in order to accelerate the rank reduction stage of the algorithm.

The method presented in (Oropeza and Sacchi, 2011) was mainly based on applying MSSA reconstruction using synthetic examples and a field dataset where the purpose of synthetic examples was to assess of the method in two reconstruction scenarios: data contaminated with noise and a noise-free case. Using MSSA reconstruction method is to complete missing offsets as well as increase the signal-to-noise ratio of the seismic volume. Furthermore, Oropeza and Sacchi (2011) concluded that for both types of data they used, extremely low reconstructions errors were found that eventually led to an optimal recovery.

3.3 Rational and Justification of Selecting the SSA

In recent years, traditional methods applied for time series analysis such as power spectra have been augmented by new methods. As an alternative to traditional digital filtering methods, the use of singular spectrum analysis has been presented in many applications. The SSA has been introduced in a wide range of application as a de-noising and raw signal smoothing method. It is a novel non-parametric technique which is mainly based on principles of multivariate statistics (Alonso, Castillo and Pintado, 2005). Since the SSA is a nonparametric method, it works with arbitrary statistical processes such as linear or nonlinear, stationary or non-stationary, and Gaussian or non-Gaussian, however, these advantages increase the motivation of its use in different applications (Hassani, 2010).

The SSA decomposes a time series into its singular spectral domain components which are physically meaningful with regards to oscillatory components and trends and then reconstructs the series by leaving the random noise component behind. However, these fundamentals can be seen as great advantages to be exploited for developing the method for noise reduction applications (Hassani, 2010; Jiang and Xie, 2016; Wang, Liu and Dong, 2016). Furthermore and unlike many other methods, the SSA works well even for small sample sizes making it possible to quickly update the coordinator rotation to varying signals block by block in relatively small blocks (Hassani *et al.*, 2009; Hassani, Soofi and Zhigljavsky, 2010; Golyandina and Shlemov, 2015). Therefore, selecting the SSA to be investigated and developed for wind noise reduction in this study is basically motivated by its proven industrial applications and potential capabilities.

In principle, the SSA is a linear analysis and prediction method (Vautard, Yiou and Ghil, 1992). The SSA is mainly based on the data-adaptive character of the eigenelements which is the sense in which the method can use concepts from and can be useful even in nonlinear dynamics. This concept gives the SSA superiority over classical spectral methods. Therefore, the SSA can act as a data-adaptive filter, however, background noise can be removed and the leading statistically significant signals can be retained (Vautard, Yiou and Ghil, 1992; Gámiz-Fortis *et al.*, 2002). As stated in (Fukuda, 2007), using the SSA for data pre-processing is a helpful procedure that encourages improving the results of any time series for data mining, and many future environmental studies can probably adapt such studied approaches. Also, the superiority of the SSA over other methods in biomechanical analysis was clearly demonstrated by several examples presented in (Alonso, Castillo and Pintado, 2005).

The SSA has been successfully implemented to solve many problems based on its capabilities. When compared with other time series analysis methods, it shows certain superiority, potentials and broader application areas and competes with more standard methods in time series analysis (Hassani, Heravi and Zhigljavsky, 2009; Chu, Lin and Wang, 2013; Golyandina and Shlemov, 2015; García Plaza and Núñez López, 2017). For example, the method has superiority over classical techniques when it is used as a method of prediction and forecasting particularly for real time series that usually has a complex structure. As previously mentioned, the SSA is essentially a linear and prediction technique (Vautard, Yiou and Ghil, 1992). However, it also works with arbitrary statistical processes and non-parametric method which makes no prior assumptions about the data. Given that the dynamics of real time records typically go through structural changes during the time period under consideration and for prediction applications, one needs to ensure that the prediction method is not sensitive to the dynamical variations (Hassani, 2010).

As mentioned by (Kostelich and Yorke, 1988), the traditional noise reduction methods such as Wiener and Kalman filters are not applicable in case of reducing noise levels in certain experimental time series related to dynamical systems. Instead, developing the SSA and using local approximation of the dynamics collectively can lead to produce a new time series and make its dynamics more consistent with those on phase space attractor but slightly altered.

3.4 Time Series Analysis Using the SSA

The SSA is known as a method that is basically based on statistical approaches and elementary linear algebra. It is a general approach to time series analysis and can be used for a wide range of applications such as de-noising, forecasting, change-point detection, trend or quasi-periodic component detection and extraction. When applying the SSA to a time series analysis, it provides a representation in the so-called Eigen domain in terms of eigenvalues and eigenvectors of a matrix constructed out of the given time series (trajectory matrix) (Alexandrov, 2009). A set of orthonormal vectors can be produced from the decomposition of the covariance matrix when applying the SVD method. However, these eigenvectors are also known as the left singular vectors of the trajectory matrix and often called in SSA literature ''empirical orthogonal functions'' or simply EOFs and considered as axis of new coordinate system (Golyandina, Nekrutkin and Zhigljavsky, 2001; Sanei and Hassani, 2015). EOFs method is also known as Principle Component Analysis (PCA) in geophysical fields (Bjornsson and Venegas, 1997).

In principal, the idea of SSA is to embed a time series $X(t)$ into multi-dimensional Euclidean space and find a subspace corresponding to the sought-for component as a first stage. The second stage is to reconstruct a time series component corresponding to this subspace. To catch a clear glimpse for better understanding and characterising the underlying regular dynamical behaviour which is the signature of the dynamical system under study, the univariate case, which indicates a single time series, is considered. The SSA method can be applied to a single series or jointly to several records and hence in the former, it is referred to as univariate SSA and in the latter as multivariate SSA or MSSA (Launonen and Holmström, 2017; Rodrigues and Mahmoudvand, 2017). Since the concern in the context of this research is only with single channel signals, therefore only the univariate SSA has been considered.

Time records are defined as data vectors sampled over time in order and often at regular intervals and distinguished from randomly sampled data that form the basis of many other data analyses. They represent the time-evolution of a dynamic process and contain data points that are indexed, listed, or graphed. The linear ordering of time series gives them a distinctive place in data analysis with a specialised set of techniques (Elsner and Tsonis, 2013; Golyandina and Shlemov, 2015). To develop the SSA for the specific problem addressed in this study, it is important at first to handle with the method considering different key concepts through providing obvious explanation using several working examples. Therefore, the time series shown in Figure 3.1 has been selected as working example to be used throughout this chapter.
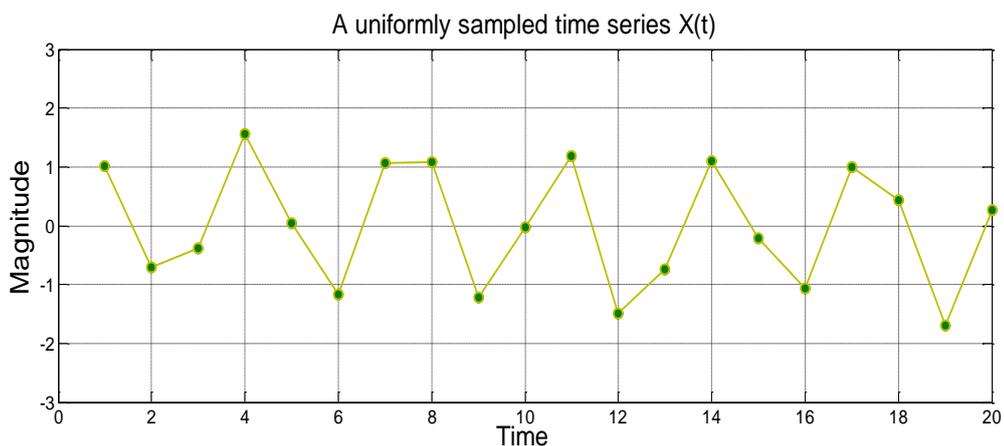
Figure 3.1. A time series sampled from a sinus function with Gaussian noise added

In this example, a uniformly sampled time series $\{X(t): t = 1,2,3 \dots \dots N_t\}$ in which the data points are sampled over time at given regular time intervals has been considered. Since the number of data points or the length of the time series is denoted by N_t , therefore from the

time series given in this example, it can be clearly see that there are 20 data points ($N_t = 20$) and hence they are represented by $x_1, x_2 \ldots\ldots\ldots\ldots x_{20}$ as graphed in Figure 3.1.

This example was taken for a variety of reasons. The most influential factor was that the main aspects of the SSA method can be better illustrated with examples drawn from practical and academic exercises which should have a real impact on the way the SSA works. Another reason was the relevance and importance of selecting a time series that is representative with regards to identifying its data points and short enough to suitably select a small window length. Consequently, this will have a significant impact on the size of the vector in which the data points are presented along with the size of the produced matrices to make them as small as possible. However, small matrices can be easily handled and manipulated regarding the numerical values and can also provide better explanations in all cases.

The selected time series also provided an opportunity to consider the main properties of the method as it has been indicated in the previous literature that the SSA works with arbitrary statistical processes such as linear or nonlinear, stationary or non-stationary, and Gaussian or non-Gaussian (Hassani, 2010). However, since the method has been particularly modified and developed to solve microphone wind induced noise as a non-stationary broadband noise, it has been developed in the system verification phase to be tested with a stationary noise (white noise). Also, this example can assess and reinforce at the mean time the understanding of these advantages since the time series represents a noisy series sampled from a sinus function with a defined period and a Gaussian noise added.

In addition to the multiple tests and case studies presented in the coming chapters with particular emphasis on the problem addressed in this thesis, this example can provide theoretical explanation using numerical values based on an experimental case like this. However, this case will be later used in Chapter 5 for the justification of the developed method before moving to system verification using different typical signal testing examples and system validation with real-world sounds for microphone wind noise problem.

3.5 Fundamental Concepts in the SSA Method

The idea of the SSA basically relies on the decomposition of a time series into oscillatory components and noise in the context of this thesis. The method is mainly based on statistical approaches and elementary linear algebra (Elsner and Tsonis, 2013; Harmouche *et al.*, 2017). The aim is to decompose the original time series into a small number of independent and

interpretable components such as a slowly varying trend, oscillatory components (harmonics), and a structure less noise (Ghil *et al.*, 2002; Hassani, 2010; Golyandina and Korobeynikov, 2014). The SSA method consists of two complementary stages which are:

1-Decomposition: this stage is related to the decomposition of the time series.

2-Reconstruction: in this stage, it is to reconstruct the original time series using estimated trend and harmonic components.

Each of these two stages consists of multiple steps. At the decomposition stage, the time series is decomposed into mutually orthogonal components after computing the covariance matrix from constructed trajectory matrix. The covariance matrix is needed to compute the eigenvalues and their associated eigenvectors. In the reconstruction stage or also known as estimation stage, the original series is reconstructed by selecting those components that reduce the noise in the series (Patterson *et al.*, 2011). Therefore, the decomposition of a time series and the reconstruction of desired additive components while separating out the undesired wind noise components are mainly the two main aspects to be considered in the SSA algorithms.

3.6 Theoretical and Mathematical Approaches

There are mainly several steps to go through to understand the SSA technique that can be applied in many applications including noise reduction problem. However, before going through that, it is important to be aware of how the data points that describe any given time series can be stored and how the so-called embedded time series or trajectory matrix can be created.

Time series can be stored in a vector denoted by **x** for example whose entries are the data points that describe the time series as a sequence of discrete-time data (Vautard and Ghil, 1989; Ghil *et al.*, 2002; Hassani, 2007; Elsner and Tsonis, 2013). Such a vector is an introductory element to the method because it includes the essential information about the given time series; in fact, it represents the original time series. The other processes and steps which lead to a complete SSA approach are mainly based on this vector.

3.6.1   Embedding Process

3.6.1.1   Vector representation

A systematic understanding of how to indicate a vector that represents the data points which describe any given time series is an initial step in the embedding process. Generally, a vector denoted by **x** for instance which includes N_t entries can define the time series as in

Equation (3.1). These entries are the data points that describe the time series under study at regular intervals. The given time series in this example has 20 data points and hence it can simply be represented as a column vector as shown in the vector equation (3.1).

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N_t} \end{bmatrix} \qquad if \ N_t = 20 \ then \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{20} \end{bmatrix}, \tag{3.1}$$

This column vector could also be seen as a matrix of dimension $N_t \times 1$ which shows the original time series at zero lag, that is when there is no delay $k = 0$. Essentially, when $k = 0$ this leads to $X(t + k) = X(t)$, where $X(t)$ is the original time series and k is the lag or "delay shift". At a given embedding dimension (window length m), the lag k can therefore be expressed in the range from 0 to $m - 1$ in a 1 lag shifted version as follows:

$$k = 0,1 \dots \dots, m - 1, \tag{3.2}$$

Hence, the window size m should be clearly specified to obtain the lag k which is needed to create a new matrix according to delay coordinates called in the SSA jargon "embedded time series" or trajectory matrix. It is worth mentioning that the window length is also called embedding dimension that represents the number of time-series elements in each snapshot as stated in (Elsner and Tsonis, 2013).

The whole procedure of the SSA method depends upon the selection of the window length which is the sole parameter in the embedding step. This parameter is very important because it plays a vital role for reconstructing noise free series from a noisy series of length N_t. The selection of the window length depends on the given data and the aim of the analysis. The improper selection of m may imply an inferior decomposition (Rukhin, 2002; Hassani, Mahmoudvand and Zokaei, 2011; Yang *et al.*, 2016).

3.6.1.2 Trajectory matrix

The so-called covariance matrix, which will be used later for computing eigenvalues and eigenvectors, can be computed by first creating a matrix denoted by **Y** and called "embedded time series" or trajectory matrix. This trajectory matrix contains the data points that represent the original time series in the first column, and a lag 1 shifted version of that time series in the second column, and so on. For more clarification, the column vector shown in Equation (3.1) is $X(t)$ when $k = 0$. As explained in (Claessen and Groth, 2002), according to delay coordinates, a total number of vectors equals m and each of size $N_t \times 1$ can be obtained when

the delay is considered in the last rows and supplemented by 0s. These vectors are similar in size to the first column vector but with a 1 lag shift at $k = 1, k = 2$, up to $k = m - 1$. This mathematically means that from all these vectors a matrix of dimension $N_t \times m$ can be created. However, in this matrix the vector **x** at $k = 0$, which represents the original time series $X(t)$, is the first column and $X(t + k)$ when $k = m - 1$ is the last column. This assumption is given when the last rows are supplemented by 0s based on the delay.

To simplify, when considering the time series example given in Figure 3.1 with $N_t = 20$ data points and assuming that the embedding dimension (window length) $m = 4$ for example, therefore, only lags of $k = 0, 1, 2$ and 3 will be considered according to Equation (3.2). As a result, a trajectory matrix **Y** of size $N_t \times m$ or 20×4 will be constructed as in Equation (3.4). For more clarification, the representation of the time series $X(t)$ when assuming the window length $m = 4$ given as an example in this section can be shown in Figure 3.2.
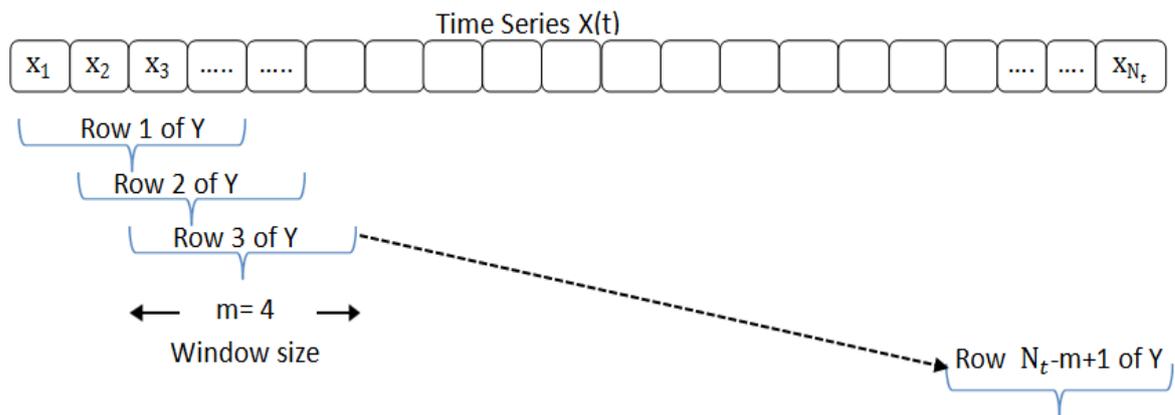
Figure 3.2. Representation of the time series of the example given in Figure 3.1 when considering window length (*m*=4) to construct the trajectory matrix

According to (Elsner and Tsonis, 2013), the coordinates of the phase space can be defined by using lagged copies of a single time series. The coordinates of the phase space will approximate the dynamics of the system from which the time record was sampled, and the number of lags is called the embedding dimension. For the purpose of the SSA, this procedure, which referred to a method of delay as it uses lagged (or delayed) copies of segments of a time series, takes a univariate time record and makes it a multivariate set of observations (Chu, Lin and Wang, 2013; Yang *et al.*, 2016).

The trajectory matrix **Y** corresponds to a sliding window of length *m* that moves along the time series $X(t)$. As illustrated in Figure 3.2, considering that the window length in this example $m = 4$ with an overlap equals $m - 1$ and values of $k = 0, 1, 2$ and 3 according to

Equation (3.2). The trajectory matrix can also be seen as a single time series of the delay coordinates that shape the structure of the matrix at all values of k in this example which are $(X(t), X(t+1), X(t+2), X(t+3))$.

Two important properties of the trajectory matrix are mentioned in (Golyandina and Korobeynikov, 2014). The first is that both the rows and columns of the trajectory matrix $\mathbf{Y}$ are subseries of the original series. The second is that $\mathbf{Y}$ has equal elements on anti-diagonals which make the trajectory matrix a Hankel matrix. In other words, all the elements along the diagonal $i + j = const$ are equal. A Hankel matrix is defined as a matrix in which each ascending skew-diagonal from left to right is constant.

Since the sliding window has an overlap equals $m - 1$ as shown in Figure 3.2, therefore the number of rows of the trajectory matrix $\mathbf{Y}$ which can be filled with the values of $X(t)$ denoted by N can be calculated according to the following equation.

$$N = N_t - (m - 1) = N_t - m + 1, \tag{3.3}$$

where N_t is the number of data points, m is the window size, and $m - 1$ is the overlap. Hence, in this case the trajectory matrix of dimension $N_t \times m$ can be seen as follows:

$$\mathbf{Y} = \begin{bmatrix} k=0 & k=1 & k=2 & k=3 \\ x_1 & x_2 & x_3 & x_4 \\ x_2 & x_3 & x_4 & x_5 \\ x_3 & x_4 & x_5 & x_6 \\ \vdots & \vdots & \vdots & \vdots \\ x_{17} & x_{18} & x_{19} & x_{20} \\ x_{18} & x_{19} & x_{20} & 0 \\ x_{19} & x_{20} & 0 & 0 \\ x_{20} & 0 & 0 & 0 \end{bmatrix} \begin{matrix} Row\ 1 \\ Row\ 2 \\ Row\ 3 \\ \vdots \\ Row\ N_t - m + 1 \\ Row\ 18 \\ Row\ 19 \\ Row\ 20 \end{matrix} \tag{3.4}$$

In Equation (3.4), only $(N_t - m + 1)$ rows, which equal to 17 rows in this example, can be filled with values of $X(t)$ and the last three rows are supplemented by 0s.

As explained in (Elsner and Tsonis, 2013), snapshots of the record for the same example explained above when considering only number of rows of the trajectory matrix $\mathbf{Y}$ which can be filled with the values of $X(t)$ according to Equation (3.3) are given as follows:

$\mathbf{v}_1^T = (x_1, x_2, x_3, x_4)$

$\mathbf{v}_2^T = (x_2, x_3, x_4, x_5)$

Up to $\mathbf{v}_{17}^T = (x_{17}, x_{18}, x_{19}, x_{20})$

$\mathbf{Y}$ can be obtained by arranging the snapshots in row vectors as in Equation (3.5).

$$\mathbf{Y} = \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_N^T \end{bmatrix}, \tag{3.5}$$

The constructed matrix $\mathbf{Y}$ is called the trajectory matrix and it includes the complete record of patterns that have occurred within a window of length m. For the example given in this section, $\mathbf{Y}$ can be shown as in Equation (3.6).

$$\mathbf{Y} = 1/\sqrt{N} \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 & x_3 & x_4 & x_5 \\ x_3 & x_4 & x_5 & x_6 \\ \vdots & \vdots & \vdots & \vdots \\ x_{17} & x_{18} & x_{19} & x_{20} \end{bmatrix}, \tag{3.6}$$

where $\sqrt{N}$ is the convenient normalisation.

To generalise, if $X(t)$ is a given time series where $t = 1,2,3 \dots N_t$, the augmented or trajectory matrix $\mathbf{Y}$ will be constructed as in Equation (3.7).

$$x_{ij} = x_{i+j-1}, \tag{3.7}$$

where $1 \leq i \leq N$ and $1 \leq j \leq m$. The arrangement of entries x_{ij} of the trajectory matrix depends on the lag, bearing in mind that the trajectory matrix has dimensions $N\ by\ m$. The trajectory matrix $\mathbf{Y}$ and its transpose $\mathbf{Y}^T$ are linear maps between the spaces R^m and R^N (Chu, Lin and Wang, 2013).

3.6.2 Covariance Matrix

3.6.2.1 Covariance definition and characteristics

Constructing what is called the variance-covariance matrix or can simply be named as covariance matrix as in SSA jargon is the first objective when the SSA is developed and used as a suitable technique for a specific purpose such as smoothing, extraction of seasonality components, and noise reduction applications (Ghil *et al.*, 2002).

The covariance matrix is essentially a matrix that shows the covariance between the values $X(t)$ and $X(t + k)$ which is mainly the covariance between lagged (or "delayed") values of the original time series $X(t)$. The covariance in probability theory and statistics is defined as a measure of the joint variability of two random variables change. The covariance could be positive or negative where the sign shows the tendency in the linear relationship between the variables. Hence, the idea is basically based on computing the covariance between the values

$X(t)$ and $X(t + k)$, where k is a lag (or "delay"). In other words, all the entries in the first column of the trajectory matrix indicate the values of the original time series at lag 0, and the other columns with a lag 1 shift. Therefore, when computing the covariance, it is mainly between the entries that appear in each row of the trajectory matrix (Vautard and Ghil, 1989; Vautard, Yiou and Ghil, 1992; Golyandina, Nekrutkin and Zhigljavsky, 2001).

From the definition above, if the covariance at lag k is positive, then the values $X(t)$ and $X(t + k)$ tend to vary together. For instance, from the original time series shown in Figure 3.1 and the first row in the trajectory matrix in Equation (3.4) that contains 4 entries in this example, it seems that the covariance is negative for $k = 1$ and $k = 2$, that is x_2 and x_3 values, and positive when lag $k = 3$, the value x_4. However, x_1 is representing the original data point at 0 delay and thus the covariance is computed with respect to the other entries in this first row, and similarly for x_2 with respect to the other entries in the second row and likewise for the other rows (Elsner and Tsonis, 2013).

3.6.2.2   Covariance matrix computation

According to (Vautard, Yiou and Ghil, 1992; Ghil *et al.*, 2002), there are two methods of computing the covariance matrix which denoted by **C**. The first method of computing the covariance, which referred to as correlation in engineering, is by computing a vector **x** of size m based on the data available in the trajectory matrix, where m is the window length. Given that $m = 4$ in the time series example used in this chapter where only lags of $k = 0,1,2$ and 3 are considered, then the produced vector **x** will contain only 4 entries according to the window size that can be generally expressed in the following form.

$$\mathbf{x} = [a \quad b \quad c \quad d], \tag{3.8}$$

The covariance between $X(t)$ and $X(t + k)$ with $k = 0, .., m - 1$ is represented in this vector which means that these values are at lags $k = 0,1,2$ and 3 consecutively. Since **x** is defined in such a way that it has variance 1, therefore the covariance at lag 0 equals 1. As a consequence of the choice of the signal, the covariance is either positive or negative as explained earlier. The values represented in this vector will be used to construct the covariance matrix by constructing a diagonal-constant matrix which is also known as Toeplitz matrix. The covariance matrix **C** is generally in the following form assuming that the vector **x** given in this example contains 4 entries.

$$\mathbf{C} = \begin{bmatrix} a & b & c & d \\ b & a & b & c \\ c & b & a & b \\ d & c & b & a \end{bmatrix}, \tag{3.9}$$

It is worth mentioning that there are some built-in MATLAB functions that can be used to facilitate calculating the vector **x** and the covariance matrix (e.g., "cov"). As stated in (Ghil *et al.*, 2002), the covariance matrix **C** of dimension $m \times m$, which can be estimated directly from the data, is considered as a Toeplitz matrix with constant diagonals. The entries c_{ij} of this matrix depend only on the lag $|i - j|$ according to the following equation.

$$c_{ij} = \frac{1}{N_t - |i - j|} \sum_{t=1}^{N_t - |i - j|} X(t)X(t + |i - j|), \tag{3.10}$$

where N_t represents the number of data points, the entries c_{ij} when $|i - j| = 0$ for $i = j$ are the entries across the main diagonal, that is, their values are typically close to 1.

The Second method to compute **C** is based on the trajectory matrix and its transpose. The covariance matrix can be computed form the trajectory matrix **Y** using a specific function in MATLAB. The variances of each column are represented in the main diagonal of the covariance matrix **C**. If N_t is large and much greater than m, these values are practically identical and very close to 1. It is worth mentioning that the remaining three values in the first row in the given example are the covariance between the values in the first column and those in the second, third, and fourth column, respectively. The variances that have been identified therefore assist in the understanding of computing the covariance matrix considering that the values in the other rows indicate the covariance between the columns as well (Elsner and Tsonis, 2013).

The symmetry of the covariance matrix is apparent from Equation (3.9) which shows a typical covariance matrix. However, the expected Toeplitz structure, in which all the values on the sub-diagonals should be identical and the main diagonal contains 1s, cannot practically obtained in every case because it mainly depends on some factors. The Toeplitz structure of matrix **C** can be determined by the length of the time series and the window length. The plausible explanation of not obtaining the expected Toeplitz structure lies in the ratio N_t/m. In other words, due to the shortness of the time series and high window size, the covariance matrix does not entirely have the expected Toeplitz structure (Claessen and Groth, 2002).

As stated in (Elsner and Tsonis, 2013), the lagged-covariance matrix $\mathbf{C}$ can be computed as the product of the trajectory matrix and its transpose $\mathbf{C} = \mathbf{Y}^T\mathbf{Y}$. Examining the trajectory matrix for repeating patterns that represent oscillations in the original time record or investigating the patterns that appear in the covariance matrix will lead to find that the elements of the covariance matrix are proportional to the linear correlations between all pairs of the patterns appearing in the m-window used to construct $\mathbf{Y}$. Hence the covariance matrix can be represented as in Equation (3.11) (Elsner and Tsonis, 2013; Yang *et al.*, 2016).

$$\mathbf{C} = \frac{1}{\sqrt{N}}\, \mathbf{Y}^T\, \frac{1}{\sqrt{N}}\mathbf{Y} \;=\; \frac{1}{N}\, \mathbf{Y}^T\mathbf{Y}\,, \tag{3.11}$$

Matrix $\mathbf{C}$ is the lagged-covariance matrix of the snapshots from the original time record and a square matrix of dimension m by m. However, its elements are all real numbers and $c_{ij} = c_{ji}$ for all i and $j,$ so it is symmetric, that is, $\mathbf{C} = \mathbf{C}^T$. As a general mathematical representation, the matrix $\mathbf{C}$ can be written as in Equation (3.12) (Elsner and Tsonis, 2013; Harmouche *et al.*, 2017).

$$\mathbf{C} = \frac{1}{N} \sum_{i=1}^{N} Y_i Y_i^T, \tag{3.12}$$

3.6.3 Eigenvalues and Eigenvectors

Eigenvalues and their associated eigenvectors are useful in many circumstances; however, it is quite important to simplify a matrix by preserving its eigenvalues. In this step applying the SVD method to the covariance matrix obtained from either of the two methods explained earlier is required to compute the eigenvalues λ's and eigenvectors $\mathbf{e}$'s. Specific MATLAB functions are offered to facilitate the computation process as will be seen later, but before that some mathematical descriptions are sequentially provided in this section.

3.6.3.1 Finding the eigenmodes

Finding the so-called eigenmodes (eigenvalues and eigenvectors) is based on the fundamental question of eigenvector decomposition. In general, this question is for what values is the matrix $\mathbf{A} - \lambda\mathbf{I}$ singular? In this equation $\mathbf{A}$ is a given matrix, λ represents the eigenvalues, and $\mathbf{I}$ is the identity matrix. Such question of singularity regarding matrices can be answered with determinants (Chu, Lin and Wang, 2013; Elsner and Tsonis, 2013). Using determinants, however, the fundamental question which just has been asked can be reduced to; for what values of λ is the determinant of the matrix $\mathbf{A} - \lambda\mathbf{I}$ equals to zero? or as in Equation (3.13).

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0, \tag{3.13}$$

This is called the characteristic equation for the matrix $\mathbf{A}$ where λ's that make the matrix $\mathbf{A} - \lambda\mathbf{I}$ singular are called eigenvalues (Elsner and Tsonis, 2013). The eigenvalues of this matrix can therefore be computed by solving the characteristic equation. For each of these special values there are a corresponding set of vectors associated with these values called the eigenvectors of $\mathbf{A}$ and should satisfy the equation below.

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{X} = \mathbf{0}, \tag{3.14}$$

Equation (3.14), in which $\mathbf{X}$ represents a set of eigenvectors $\mathbf{X}_n$ correspond to the eigenvalues $\lambda's$, can be written as:

$$\mathbf{A}\mathbf{X} = \lambda\mathbf{X}, \tag{3.15}$$

When representing the eigenvectors geometrically, they can be considered as the axis of a new coordinate system. Hence, any scalar multiple of these eigenvectors is also an eigenvector of matrix $\mathbf{A}$ (Golyandina and Shlemov, 2015). To simplify, all eigenvectors of the form $C\mathbf{X}_n$ (C is scalar) will form an eigenspace spanned by $\mathbf{X}_n$ which means that the eigenspace is one-dimensional and is spanned by $\mathbf{X}_n$. In this case, only the scale of the eigenvectors is changing while their direction remains unchanged.

The interpretation of eigenvalues and their associated eigenvectors mainly depends on the situation in which matrix $\mathbf{A}$ arises. The process of decomposition can be simplified as matrix $\mathbf{A}$ is usually symmetric with real coefficients. The eigenvalues and their associated eigenvectors can be seen as a way to express the variability of a set of data (Vautard, Yiou and Ghil, 1992). It is beneficial to mention some properties of eigenvalues and eigenvectors of an arbitrary matrix $\mathbf{A}$ before proceeding to the special case of the eigenvalues and eigenvectors of real, symmetric matrices. The trace of an n by n matrix $\mathbf{A}$ is given by the sum of the n diagonal entries which equals the sum of the n eigenvalues of $\mathbf{A}$ as shown in Equation (3.16).

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = a_{11} + a_{22} + \cdots + a_{nn}, \tag{3.16}$$

It is worth mentioning that performing SVD of the trajectory matrix produces a matrix $\mathbf{S}$ with entries s_i and each is equal to the square root of the eigenvalues of matrix $\mathbf{C}$ ($\sqrt{\lambda_i}$). The so-called singular spectrum is the stem-plot of these values against their index i (Clifford, 2005). The magnitudes of the eigenvalues are used in grouping step to divide the obtained elementary matrices into groups (Maddirala and Shaik, 2016).

3.6.3.2 Diagonal form of a matrix

Assume that $\mathbf{A}$ is an n by n matrix with n linearly independent eigenvectors ($\mathbf{e}_i$), $i = 1, n$, hence a matrix $\mathbf{E}$ whose columns are the eigenvectors of $\mathbf{A}$ can be constructed in a form that satisfies the following equation:

$$\mathbf{E}^{-1}\mathbf{A}\mathbf{E} = \mathbf{\Lambda}, \tag{3.17}$$

The product on the left side of Equation (3.17) is called the diagonal form of the matrix $\mathbf{A}$ and therefore $\mathbf{\Lambda}$ is a diagonal matrix whose nonnegative entries are the eigenvalues of $\mathbf{A}$. It is worth mentioning that the eigenvectors of matrix $\mathbf{A}$ should be linearly independent to make it diagonalisable in this way. Also matrix $\mathbf{E}$ is not unique because the eigenvectors can always be multiplied by a constant scalar preserving their nature as eigenvectors (Elsner and Tsonis, 2013). To conclude, $\mathbf{E}$ is a matrix whose columns are the eigenvectors of $\mathbf{A}$ and $\mathbf{\Lambda}$ is a diagonal matrix whose nonnegative entries are the eigenvalues of $\mathbf{A}$.

A working example using numerical values can lead to a better understanding. To verify Equation (3.17), the eigenvalues $\lambda's$ and their associated eigenvectors can be computed by assuming $\mathbf{A}$ as a square matrix as follows:

$$\mathbf{A} = \begin{bmatrix} 3 & 4 \\ -1 & -2 \end{bmatrix}$$

Using Equation (3.13) and when solving for λ, the eigenvalues $\lambda_1 = -1$ and $\lambda_2 = 2$ have been obtained. Also, their associated eigenvectors can be computed using Equation (3.14) when solving for $\mathbf{X}$, however, the following eigenvectors $\mathbf{X}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\mathbf{X}_2 = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$ have been obtained, and therefore $\mathbf{E}$ can be formed as $\mathbf{E} = \begin{bmatrix} 1 & -4 \\ -1 & 1 \end{bmatrix}$.

To verify, matrix product can be performed for the diagonal form of the matrix $\mathbf{A}$ according to Equation (3.17) using MATLAB to obtain $\mathbf{\Lambda} = \begin{bmatrix} -1.0000 & 0.0001 \\ 0 & 2.0001 \end{bmatrix}$, where the entries in the main diagonal represent the eigenvalues that exactly identical to those that have been computed using Equation (3.13).

3.6.3.3 Spectral decomposition

If our interest is of the eigenvalues and eigenvectors of real and symmetric matrices, then assuming $\mathbf{A}$ is a real symmetric matrix where $\mathbf{A} = \mathbf{A}^T$. Now, in this case, every eigenvalue of $\mathbf{A}$ is also real and if all eigenvalues are distinct, then their corresponding eigenvectors are

orthogonal. For more clarifications, if $\mathbf{X}_i's$ are the eigenvectors, then $\mathbf{X}_i^T \mathbf{X}_j = 0$ for all $i \neq j$. Normalising these eigenvectors can be performed as follows:

$$\mathbf{e}_i = \frac{\mathbf{X}_i}{\|\mathbf{X}_i\|},$$

where $\|\mathbf{X}_i\| = \sqrt{\mathbf{X}^2_{i1} + \mathbf{X}^2_{i2} + \cdots + \mathbf{X}^2_{in}}$

According to the above clarification, a set of orthonormal (orthogonal and normalised) eigenvectors can therefore be produced. Therefore, such vectors are orthogonal when $\mathbf{e}_i^T \mathbf{e}_j = 0$ for all $i \neq j$ and normalised when $\mathbf{e}_i^T \mathbf{e}_j = 1$ for $i = j$.

A real, symmetric matrix $\mathbf{A}$ can be diagonalised by an orthogonal matrix $\mathbf{E}$ whose columns are the orthonormal eigenvectors of $\mathbf{A}$. Now, the orthonormal eigenvectors of $\mathbf{A}$ can be placed as columns of an orthogonal matrix $\mathbf{E}$ in order to diagonalise $\mathbf{A}$ (Hassani, 2007). It is important now to state a principal theorem when there is a diagonalisable matrix $\mathbf{E}$ whose columns are orthonormal of a real and symmetric matrix $\mathbf{A}$ as follows:

$\mathbf{E}^T \mathbf{E} = \mathbf{I}$, and therefore $\mathbf{E}^T = \mathbf{E}^{-1}$

Hence, Equation (3.17) $\mathbf{E}^{-1} \mathbf{A} \mathbf{E} = \mathbf{\Lambda}$ can be written as follows:

$$\mathbf{A} = \mathbf{E} \mathbf{\Lambda} \mathbf{E}^T, \tag{3.18}$$

From Equation (3.18) the so-called spectral decomposition of $\mathbf{A}$ can be obtained

$$\mathbf{A} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^T + \lambda_2 \mathbf{e}_2 \mathbf{e}_2^T + \cdots + \lambda_n \mathbf{e}_n \mathbf{e}_n^T, \tag{3.19}$$

The spectral decomposition, which is fundamental to linear algebra, plays a vital role in the singular spectrum analysis method for time series analysis. For a give matrix $\mathbf{A}$ as in this demonstration, spectral decomposition expresses this matrix as a summation of the one-dimensional projections $\mathbf{e}_i \mathbf{e}_i^T$ (Elsner and Tsonis, 2013).

3.6.3.4 The singular spectrum

In this section, it is to move from general to specific by considering that our symmetric matrix of interest is the covariance matrix $\mathbf{C}$ instead of the arbitrary real symmetric matrix $\mathbf{A}$ given for explanation in the above sections. Recall from Equation (3.18), since $\mathbf{C}$ is real and symmetric, there is a diagonalisable matrix $\mathbf{E}$ whose columns are orthonormal and a diagonal matrix $\mathbf{\Lambda}$ such that

$$\mathbf{C} = \mathbf{E} \mathbf{\Lambda} \mathbf{E}^T, \tag{3.20}$$

Since $\mathbf{E}^T\mathbf{E} = \mathbf{I}$, and to reform Equation (3.20) in an easy way to handle, it is possible to multiply it from the right by $\mathbf{E}$ according to matrix multiplication rules, then Equation (3.20) can be rewritten as below in a form that called the spectral decomposition of $\mathbf{C}$.

$$\mathbf{CE} = \mathbf{E\Lambda}, \tag{3.21}$$

Recall from the previous explanation that the eigenvalues of $\mathbf{C}$ are the nonnegative entries of the diagonal matrix $\mathbf{\Lambda}$. Also, it has been mentioned that the lagged-covariance matrix $\mathbf{C}$ can be computed as the product of the trajectory matrix and its transpose $\mathbf{C} = \mathbf{Y}^T\mathbf{Y}$. However, replacing this form of $\mathbf{C}$ in the Equation (3.21), a new form can be found and further simplified as follows:

$$\mathbf{Y}^T\mathbf{YE} = \mathbf{E\Lambda}$$

Simplifying the above form by multiplying both sides from the left by $\mathbf{E}^T$

$$\mathbf{E}^T\mathbf{Y}^T\mathbf{YE} = \mathbf{E}^T\mathbf{E\Lambda}$$

Since $\mathbf{E}^T\mathbf{E} = \mathbf{I}$, then $\mathbf{E}^T\mathbf{Y}^T\mathbf{YE} = \mathbf{\Lambda}$. To simplify more

$$(\mathbf{YE})^T(\mathbf{YE}) = \mathbf{\Lambda}, \tag{3.22}$$

The matrix $\mathbf{YE}$ is the trajectory matrix projected onto the basis $\mathbf{E}$ and since $\mathbf{E}$ is composed of orthogonal vectors called the *singular* vectors of $\mathbf{Y}$. Therefore, the components of $\mathbf{Y}$ aligned along the basis $\mathbf{E}$ are uncorrelated. In fact, the eigenvectors of $\mathbf{C}$ are the singular vectors of $\mathbf{Y}$. The diagonal matrix $\mathbf{\Lambda}$ consists of ordered values arranged in a descending order of magnitude. The square roots of the eigenvalues of the matrix $\mathbf{C}$ are called the *singular* values of $\mathbf{Y}$. In short, these ordered singular values are referred to collectively as the *singular spectrum* (Hassani, Soofi and Zhigljavsky, 2010; Elsner and Tsonis, 2013; Golyandina and Shlemov, 2015). From the singular value decomposition, the trajectory matrix $\mathbf{Y}$ can be written as follows:

$$\mathbf{Y} = \mathbf{USE}^T, \tag{3.23}$$

where $\mathbf{U}$ and $\mathbf{E}$ are left and right singular vectors of $\mathbf{Y}$, and $\mathbf{S}$ is a diagonal matrix of singular values.

From the definition of the covariance matrix $\mathbf{C} = \mathbf{Y}^T\mathbf{Y}$ and by substituting Equation (3.23) in this form of the covariance matrix, the decomposition of $\mathbf{C}$ can be shown as follows:

$$\mathbf{C} = \mathbf{Y}^T\mathbf{Y} = (\mathbf{USE}^T)^T(\mathbf{USE}^T) = \mathbf{ESU}^T\mathbf{USE}^T$$

Since $\mathbf{U}^T\mathbf{U} = \mathbf{I}$, then $\mathbf{C} = \mathbf{ES}^2\mathbf{E}^T$

For the decomposition being unique, it follows that $S^2 = \Lambda$. In short, the right singular vectors of $\mathbf{Y}$ are the eigenvectors of $\mathbf{C}$ and the left singular vectors of $\mathbf{Y}$ are the eigenvectors of the matrix $\mathbf{YY}^T$ (Chu, Lin and Wang, 2013; Golyandina and Shlemov, 2015). As a demonstration, the following example using numerical values is provided.

Assume the time series example given in the beginning of this chapter as previously shown in Figure 3.1 and suppose to take only the first 6 values as indicated in the table below. To simplify, the window length is assumed to be $m = 3$.

t	1	2	3	4	5	6
x_t	1.014	- 0.711	- 0.39	1.57	0.044	- 1.166

From the above given data, $N_t = 6$ and then $N = N_t - m + 1 = 4$, and by solving with Matlab, the following matrices have been found

$$\mathbf{Y} = \begin{bmatrix} 1.0140 & -0.7110 & -0.3900 \\ -0.7110 & -0.3900 & 1.5700 \\ -0.3900 & 1.5700 & 0.0440 \\ 1.5700 & 0.0440 & -1.1660 \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{\sqrt{N}}\mathbf{Y}^T \frac{1}{\sqrt{N}}\mathbf{Y} = \frac{1}{N}\mathbf{Y}^T\mathbf{Y}$$

$$\mathbf{C} = \begin{bmatrix} 1.0377 & -0.2467 & -0.8399 \\ -0.2467 & 0.7811 & -0.0793 \\ -0.8399 & -0.0793 & 0.9946 \end{bmatrix}$$

A built-in MATLAB function [P, D]=eig (C) can be used to compute the eigenvectors presented in the matrix $\mathbf{P}$ (RHO) and the eigenvalues λ_n (Lambda) presented in matrix $\mathbf{D}$.

$$\mathbf{P} = \begin{bmatrix} 0.7221 & 0.1439 & 0.6767 \\ 0.1139 & -0.9401 & 0.3214 \\ 0.6823 & 0.3092 & 0.6624 \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} 1.8702 & 0 & 0 \\ 0 & 0.8450 & 0 \\ 0 & 0 & 0.0982 \end{bmatrix}$$

The spectral decomposition of $\mathbf{C}$ yields the eigenvalues as in matrix $\mathbf{D}$ and the eigenvectors in matrix $\mathbf{P}$ considering that the ith column of $\mathbf{P}$ is the eigenvector corresponding to the eigenvalue in the ith column of $\mathbf{D}$. Importantly, the number of eigenvalues is equal to the window size, in turn the number of the associated eigenvectors that matrix $\mathbf{P}$ contains. Consequently, this makes matrix $\mathbf{P}$ a square matrix of size $m \; by \; m$.

The amplitude of each eigenvector will be comparable to the amplitudes of all other eigenvectors with the normalisation used in the calculation of the trajectory matrix. The first eigenvectors represent a high-frequency oscillation if successive elements of the eigenvectors over time are considered, while the others capture the lower-frequency components of the time series (Elsner and Tsonis, 2013).

Since $\mathbf{C} = \mathbf{E \Lambda E}^T$, therefore this equation can be verified using matrix multiplication as by definition $\mathbf{E} = \mathbf{P}$ and $\mathbf{\Lambda} = \mathbf{D}$. The singular spectrum of the trajectory matrix $\mathbf{Y}$ consists of the square roots of the eigenvalues of matrix $\mathbf{C}$ that are given in matrix $\mathbf{\Lambda}$ or $\mathbf{D}$. These square roots are called the *singular* values of $\mathbf{Y}$ with the *singular* vectors being identical to the eigenvectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ in the given working example which are aligned in matrix $\mathbf{E}$ or $\mathbf{P}$. It is worth mentioning that the desired and undesired subspaces that represent the signal of interest and unwanted components, respectively, can be identified from the pattern of the eigenvalues (Sanei and Hassani, 2015).

3.6.4 Principle Components

The eigenvectors of matrix $\mathbf{C}$, which are measured at different lags, indicate the temporal covariance of the time series. Also, they can be used to compute the principal components $v^{k'}s$ by projecting the original time series onto the new coordinate system represented by these eigenvectors. In other words, it is simply a matrix multiplication process where the principal components matrix $\mathbf{V}$ results from multiplying the trajectory matrix $\mathbf{Y}$ by matrix $\mathbf{P}$ that contains the eigenvectors of $\mathbf{C}$. The kth PC can be defined as the orthogonal projection coefficient of the original time record onto the kth EOF (Vautard, Yiou and Ghil, 1992). This process can be given as follow:

$$v_i^k = \sum_{j=1}^{m} x_{i+j-1} e_j^k, \tag{3.24}$$

where v_i^k represents the principle components arranged in columns k and rows i for $[i = 1,2 \dots N]$ as each element of $\mathbf{V}$ results from multiplying a row of $\mathbf{Y}$ by a column of $\mathbf{P}$. x_{i+j-1} represents the incremental elements in the rows of $\mathbf{Y}$, e_j^k represents the jth component of the kth eigenvector in matrix $\mathbf{P}$ (Hassani, 2010; Golyandina and Shlemov, 2015).

For each value of i the summation will stop when $j = m$, and this repetitive procedure ends when $i = N$. The new produced matrix $\mathbf{V}$ will have a number of rows i equals to N and a

number of columns j equals to m or simply of size N by m. Considering the example provided in the previous section where $N = 4, m = 3$, Equation (3.24) can be demonstrated as follows:

$$
\mathbf{V} = \mathbf{Y} \times \mathbf{P} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \end{bmatrix} \times \begin{bmatrix} e_1^1 & e_1^2 & e_1^3 \\ e_2^1 & e_2^2 & e_2^3 \\ e_3^1 & e_3^2 & e_3^3 \end{bmatrix} = \begin{bmatrix} v_1^1 & v_1^2 & v_1^3 \\ v_2^1 & v_2^2 & v_2^3 \\ v_3^1 & v_3^2 & v_3^3 \\ v_4^1 & v_4^2 & v_4^3 \end{bmatrix}, \tag{3.25}
$$

It is worth mentioning that the principal components are again time series and of the same length as the embedded time series introduced in the trajectory matrix and ordered in the same way as the eigenvectors. However, this indicates that the first column is the first principle component *PC1*, the second column is the second principle component *PC2* and so on. The principle components are different from the embedded time series $\mathbf{Y}$ as a different coordinate system is chosen to plot each point. With MATLAB the principle components *PCs* are computed by a matrix product: $\mathbf{V} = \mathbf{Y} \times \mathbf{P}$ as mentioned above.

The columns of $\mathbf{V}$ do not correspond to different time lags as in the trajectory matrix $\mathbf{Y}$. Rather, in the SSA, the principle components matrix is introduced as the projection of the embedded time series onto the eigenvectors. The original values of $\mathbf{Y}$ have been projected in a new coordinate system for gathering the variance in the *PCs* (Elsner and Tsonis, 2013). As previously mentioned, the eigenvectors, which are needed to compute the principle components as they represent the axis of projection, can be simply obtained using the covariance matrix $\mathbf{C} = \frac{1}{N} \mathbf{Y}^T \mathbf{Y}$. The eigenvectors are presented in an $m \times m$ matrix that corresponds to a number of eigenvalues of $\mathbf{C}$ presented along the main diagonal of the eigenvalue's matrix. Determining the eigenvalues in this way is by performing the singular value decomposition method. To give a full picture, the example provided in the previous section with numerical values is used for such clarification, however, with MATLAB, matrix $\mathbf{V}$ has been computed as follows:

$$
\mathbf{V} = \mathbf{Y} \times \mathbf{P} = \begin{bmatrix} -1.0793 & 0.6937 & 0.1993 \\ 1.5403 & 0.7497 & 0.4336 \\ 0.4905 & -1.5184 & 0.2699 \\ -1.9243 & -0.1760 & 0.3041 \end{bmatrix}
$$

The covariance matrix of the principle components matrix $= \frac{1}{N} \mathbf{V}^T \mathbf{V}$. Since $N = 4$ in the given example, the matrix of eigenvalues has been computed results in the following matrix.

$$
\begin{bmatrix} 1.8702 & -0.0000 & -0.0000 \\ -0.0000 & 0.8450 & -0.0000 \\ -0.0000 & -0.0000 & 0.0982 \end{bmatrix}
$$

In this matrix, the diagonal contains the eigenvalues, and the off-diagonal elements are all zeros. By computing the variance-covariance matrix of the matrix **V**, two important aspects can be demonstrated. The first is that the variance of each *PC* equals to the eigenvalue of the corresponding eigenvector. The second is that, at lag zero, no covariance can be obtained between the *PCs*. When the normalisation is used, the variance of each principal component is λ_1 seeing that the amplitudes increase with the increase in k. It is possible to isolate and probe each principal component independently from the remainder of the time series. An approximate geometric visualisation of a possible underlying attractor can be provided by the principal components if the embedding dimension (window length) is high enough (Ghil *et al.*, 2002). To get back to the original time series, it is required to convolve the principal components with their associated eigenvectors.

After decomposing the original time record and generating the so-called elementary matrices of the trajectory matrix when considering $\mathbf{C} = \mathbf{Y}\mathbf{Y}^T$, the singular vectors can be used to compute the principal components of the time series. However, for more clarification, the principle components matrix will be defined in terms of the transpose of the trajectory matrix, matrix **S** that represents the square roots of the eigenvalues presented in matrix **Λ**, and the eigenvectors matrix, producing the collection $(\sqrt{\lambda_i}e_i v_i)$ which is called the ith eigentriple of the SVD. More details will be given in the next chapters when explaining the frame work of singular spectral separation established for acoustic signals contaminated by wind noise.

3.6.5 Reconstruction of the Time Series

The principal components represent a projection in a different coordinate system as their interpretation is different from that of $X(t)$. Therefore, they cannot be compared to the time series $X(t)$. Now, however, it is possible to return from the Eigen domain to the time domain by projecting the *PCs* back onto the eigenvectors. Using such projection therefore can help in obtaining time series in the original coordinates referred to the "Reconstructed Components" *RCs*, as in the SSA literature. Each of the obtained time series in the reconstruction step corresponds to one of the *PCs*. The *RCs* that resemble the desired signals are the ones with higher variance which are mainly the reconstructed components correspond to the low-order eigenvalues in the desired subspace. In other words, the *RCs* that correspond to the most dominant principle components associated with dominant eigenvalues are the ones that can be used to reconstruct the signal of interest (Elsner and Tsonis, 2013).

The grouping of SVD components used for reconstruction are quite important in this stage as the selected eigenvectors associated with the most dominant eigenvalues have absolutely a great influence on the output time series. The magnitudes of eigenvalues can be used for estimating the signal subspace (Maddirala and Shaik, 2016). Signals of interest are located in the lower subspace, while noise components arise in the higher subspace (Mohammadi *et al.*, 2016). It is significant, however, to choose the group *I* of indices of the eigenvectors based on the standard SSA recommendations to avoid missing part of the signal or adding part of the noise (Hassani, 2010; Golyandina and Shlemov, 2015).

Among the reconstruction techniques is to reconstruct a new matrix denoted by **Z** in a similar way of constructing the embedded time series or the trajectory matrix **Y**, but with two main differences. The first is that $X(t)$, of course, will no longer be used in this stage to construct **Z**; instead, the principal components will be used. The second is that time delay runs in the opposite direction. Matrix **Z** can be constructed for each principal component. Returning to the example given at the beginning of this chapter and used throughout the different sections to build up a detailed and clear picture of the method developed for wind noise separation and provide further demonstration. Matrix **Z** can be constructed as follows:

$$\mathbf{Z} = \begin{bmatrix} v_1^1 & 0 & 0 & 0 \\ v_2^1 & v_1^1 & 0 & 0 \\ v_3^1 & v_2^1 & v_1^1 & 0 \\ v_4^1 & v_3^1 & v_2^1 & v_1^1 \\ \vdots & \vdots & \vdots & \vdots \\ v_{20}^1 & v_{19}^1 & v_{18}^1 & v_{17}^1 \end{bmatrix}, \tag{3.26}$$

The opposite direction of the time delay compared to the trajectory matrix **Y** is depicted in Equation (4.26). In this illustration, only one principle component is used. Also, it can be seen that the first column of **Z** is simply *PC1* (the first column in matrix **V**). The other columns of **Z** are *PC1* at different time delay ranging from t_1 to t_3 in opposite direction compared to matrix **Y** and always preserving the same window length for the coordination of matrix dimensions. In a similar way to matrix **Y**, zeros have been put in where data is not available (Claessen and Groth, 2002).

It is worth mentioning that the dimensions of the used matrices should match to enable matrix multiplication considering that the selected window length plays a vital role in deciding the size of the different matrices as well as the way of computing the covariance matrix.

According to (Ghil *et al.*, 2002), the part of a time series that is associated with a single or several EOFs can be reconstructed by combining the associated *PCs* as follows:

$$RC_{\mathcal{K}}(t) = \frac{1}{m} \sum_{k \in \mathcal{K}} \sum_{j=L_t}^{U_{\mathcal{K}}} v(t - j + 1)e_k(j), \tag{3.27}$$

where $\mathcal{K}$ is a set of eigenvectors on which the reconstruction is based, m is the normalisation factor which is the window size, $U_{\mathcal{K}}$ and L_t are the upper and lower bound of summation which may differ from the central part of the time series and its end point as stated in (Vautard and Ghil, 1991), $v(t - j + 1)$ are rows entries of $\mathbf{Z}$, $e_k(j)$ are columns entries of $\mathbf{P}$.

With MATLAB, the *RCs* can be computed in an operation that analogous to Equation (3.27). Generally, the first reconstructed components are of interest because they represent the oscillatory components that can be obtained by recovering the time series when returning from Eigen domain to time domain. Therefore, the first reconstructed component *RC1* for example can be obtained in a similar way as above but by multiplying matrix $\mathbf{Z}$ by the first eigenvector only of matrix $\mathbf{P}$ and considering the normalisation factor m which can be specified in MATLAB operations (Hassani, Mahmoudvand and Zokaei, 2011; Chu, Lin and Wang, 2013; Golyandina and Korobeynikov, 2014).

The reconstructed components *RCs* can be computed by inverting the projection using the principle components $\mathbf{PC} = \mathbf{Y} \times \mathbf{RHO}$ onto the eigenvectors transpose matrix as in the operation; $\mathbf{RC} = \mathbf{Y} \times \mathbf{RHO} \times \mathbf{RHO}^T$. All the *RCs* can be put together into a single matrix denoted by $\mathbf{RC}$ in which the columns are the reconstructed components. This can be accomplished by constructing matrix $\mathbf{Z}$ for the other *PCs* in a similar way as above and projecting the new matrix on the eigenvector's matrix. However, this procedure depends on the selection of the eigenvectors associated to other eigenvalues that might be seen as dominant ones. In this way, a set of *RCs* will be produced to form the matrix $\mathbf{RC}$. Such *RCs* are related to specific and selected principle components that corresponds to the eigenvector associated to dominant eigenvalues as explained in the given example. To complete the transformation back to one-dimensional time domain vector that describes a specific signal component, the so-called diagonal averaging approach, which will be explained in the next chapter, has been adopted.

In addition, in this thesis, it is to develop the SSA method based on its two complementary stages which are decomposition of a given noisy time series and the reconstruction of the signal of interest to further improve the method to separate wind noise components and reconstruct the desired signal free of noise. In other words, this can be achieved by developing grouping and reconstruction techniques to improve the separability approach and separating wind noise components out from the signal. However, details about grouping and reconstruction techniques to improve separability will be explained in the next chapters.

3.7 Summary

Contemporary and powerful signal processing methods are sought to address the wind noise issues and might yield better results, e.g., (Schmidt, Larsen and Hsiao, 2007). Therefore, the SSA method has been selected for further development particularly for microphone wind noise problem and proposed to be the main method in this study as justified in this chapter. Also, so far, no attention has been devoted concerning specifying wind noise reduction in the context of environmental noise and soundscapes monitoring and design using novel and powerful techniques such as the SSA. This leads to consider this study the first to undertake developing the SSA for wind noise problem in an interesting application area of outdoor data acquisition.

The SSA has been developed and implemented to overcome many practical problems including noise reduction in which the method shows potential capabilities. Many aspects regarding time series analysis have been covered in this chapter with working examples. Also, several important theoretical and mathematical approaches have been adopted to handle with the method for further development as a noise separation method, particularly for microphone wind noise problem. This includes explanation of the different steps involved in its two complementary stages.

The SSA utilises a representation of data in a statistical domain which called Eigen domain rather than a time or frequency domain. However, instead of using a set of axes that represent discrete frequencies such as in Fourier Transform, data is projected onto a new set of axes that fulfil some statistical criterion. This statistical technique depends on the structure of the data being analysed, whereas, in Fourier-based techniques, Fourier components onto which a data segment is projected are fixed. Therefore, in the SSA, data is projected onto axis that might change according to the change of the structure of the data over time (Clifford, 2005).

One of the most important discussions in the SSA is that the signals obtained by the SSA decomposition are generated from eigenvectors which are not purely related to the frequency when compared to those obtained by filtering out frequency bands with the Fourier Transform for example (Fukuda, 2007). This property gives the SSA an obvious and key advantage which is the ability of adding or removing the produced additive components either low or high frequencies based on the applied grouping criterion and reconstruction techniques to therefore improve the separability approach. To demonstrate what have previously mentioned in this chapter with useful clarifications, a framework has been established regarding singular spectral separation of acoustic signals contaminated by wind noise in the next chapters.

4 CHAPTER FOUR SINGULAR SPECTRAL SEPARATION OF ACOUSTIC SIGNALS CONTAMINATED BY WIND NOISE

4

Singular Spectral Separation of Acoustic Signals Contaminated by Wind Noise

4.1 Overview

The SSA is introduced in a wide range of applications as a de-noising and raw signal smoothing method such as analysing and de-noising acoustic emission signals (Harmouche *et al.*, 2017; Traore *et al.*, 2017). Basically, a number of additive components can be obtained by decomposing the original time series to identify which of the new produced additive subseries be part of the modulated signal and which be part of random noise (Alonso, Castillo and Pintado, 2005). It is showed in (Fukuda, 2007; Saito *et al.*, 2011), that using the SSA for data pre-processing is a helpful procedure that encourages improving the results of any time series for data mining and signal discrimination.

The idea of the SSA is based on the decomposition of a time series into several subseries, so that each subseries can be identified into different groups as; oscillatory components, a trend or noise. The SSA is a model free and non-parametric method consists of two complementary stages (decomposition and reconstruction) as shown in Figure 4.1. The method is mainly based on statistical approaches and many operations of the SSA algorithm are elementary linear algebra (Claessen and Groth, 2002; Chu, Lin and Wang, 2013; Yang *et al.*, 2016; Launonen and Holmström, 2017).
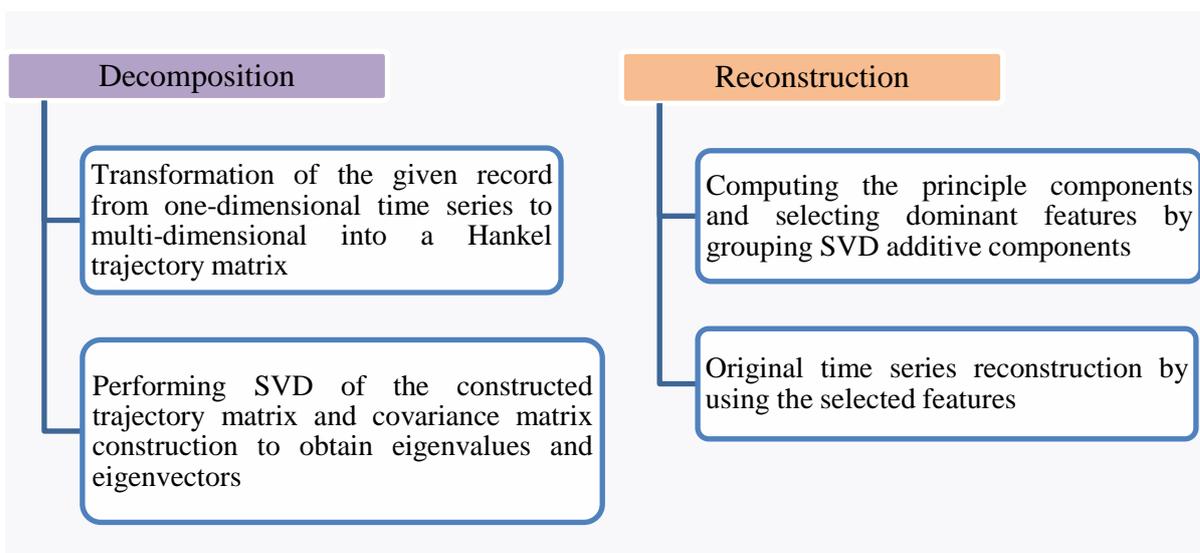
Figure 4.1. The two complementary stages of the SSA method

The SSA has been successfully implemented to solve many problems based on its capabilities. When compared with other time series analysis methods, it shows certain superiority, potentials, and broader application areas and competes with more standard methods in the area of time series analysis (Hassani, Heravi and Zhigljavsky, 2009; Chu, Lin and Wang, 2013; García Plaza and Núñez López, 2017). The SSA has been known as a two-step point-symmetric de-nosing method of time series. Noiseless signals can be obtained with minimum loss of data (Jiang and Xie, 2016; Yang *et al.*, 2016; Qiao *et al.*, 2017).

In this study, a framework of singular spectrum separation of acoustic signals contaminated by wind noise has been established. Within this framework, as systematic approach has been adopted to effectively develop a wind noise separation method based on the singular spectrum analysis through its two complementary stages which are composed of multiple steps. In the developed SSA algorithms, it is to focus on wind noise components to be separated out from the signal and reconstruct the desired signals as described in the subsequent sections. It is to consider developing the method towards achieving the aim of the study through developing grouping and reconstruction techniques as key elements to improve the separability which will be explained in the next chapter.

4.2 Key Concepts and Useful Insights in the SSA Method

The key element in the de-nosing process is to remove the noise without losing a significant portion of the signal and this can be accomplished with the SSA. The SSA can provide an important concept from the time series analysis known as statistical dimension. The statistical dimension of the process from which the time series was taken is defined as the number of eigenvalues before the noise floor. This concept develops the use of the SSA as a de-noising technique (Elsner and Tsonis, 2013; Yang *et al.*, 2016). In this thesis, the main aspect in studying the properties of the SSA is the separability which describes how well different components can be separated from each other. If the dataset is separable, it is then to focus on how to apply a proper grouping criteria after identifying and optimising suitable window size (Hassani, 2010; Golyandina and Shlemov, 2015; Golyandina and Lomtev, 2016).

In the SSA method there are some useful insights to be observed. A harmonic component commonly produces two eigentriples with close singular values; however, a pure noise series typically produces a gradually decreasing sequence of singular values. Furthermore, checking breaks in the eigenvalue spectra is also another useful insight. This can be accomplished by

using visual SSA tools in the eigenvalue spectra. In other words, checking breaks is to ensure the availability of eigentriples with nearly equal and close singular values when specifying a threshold of such equality. Large breaks among the eigenvalues pairs in the eigenvalue spectra cannot lead to produce a harmonic component (Hassani, 2007; Chu, Lin and Wang, 2013; Yang *et al.*, 2016).

In addition to the perception of the typical shape of the eigenvalue's spectra, the analysis using *w*-correlation matrix is used to distinguish between frames containing mostly the energy of the wanted signal and wind-only frames (in the case of this study) presented in the higher subspace of the singular spectra.

4.3 The Established Procedure of the SSA Method

From the SSA theory, the spectral (eigenvalue) decomposition of a given matrix **A,** for example, into a set (spectrum) of eigenvalues and eigenvectors provided a clear theoretical and scientific understanding of the term "singular spectrum" which came from such spectral decomposition. Also, these eigenvalues are specific numbers that make matrix $\mathbf{A} - \lambda \mathbf{I}$ singular when the determinant of this matrix is equal to zero. In this mathematical illustration, matrix **I** is the identity matrix and $\lambda's$ are the eigenvalues. Singular spectrum analysis, per se, is, the analysis of time series using the singular spectrum. Therefore, as previously mentioned, the time series under investigation needs to be embedded in a so-called trajectory matrix as a first step (Elsner and Tsonis, 2013; Golyandina and Korobeynikov, 2014).

The representation of the data in the SSA method is in a statistical domain which is called Eigen domain rather than a time or frequency domain. However, in the SSA, data is projected onto a new set of axes that fulfil some statistical criterion and might change according to the change of the structure of the data over time (Clifford, 2005; Golyandina and Shlemov, 2015).

In principal, the idea of the SSA on which the developed SSA algorithms are based, is to embed a time series $X(\text{t})$ into multi-dimensional Euclidean space and find a subspace corresponding to the sought-for component as a first stage. At the first decomposition stage, the time series is decomposed into mutually orthogonal components after computing the covariance matrix from the embedded time series or the constructed trajectory matrix. The trajectory matrix is obtained from the real observations of the time series and decomposed into elementary matrices or also called additive components using SVD method (Chu, Lin and Wang, 2013; Golyandina and Shlemov, 2015; Harmouche *et al.*, 2017). The SVD method is

also used to determine the principal components of a multi-dimensional signal (Clifford, 2005). The aim is to decompose the original time series or the time signal captured by sensors into a small number of independent and interpretable components such as a slowly varying trend, oscillatory components (harmonics), and a structure less noise (Ghil *et al.*, 2002; García Plaza and Núñez López, 2017). Figure 4.2 illustrates a descriptive procedure of the SSA method.
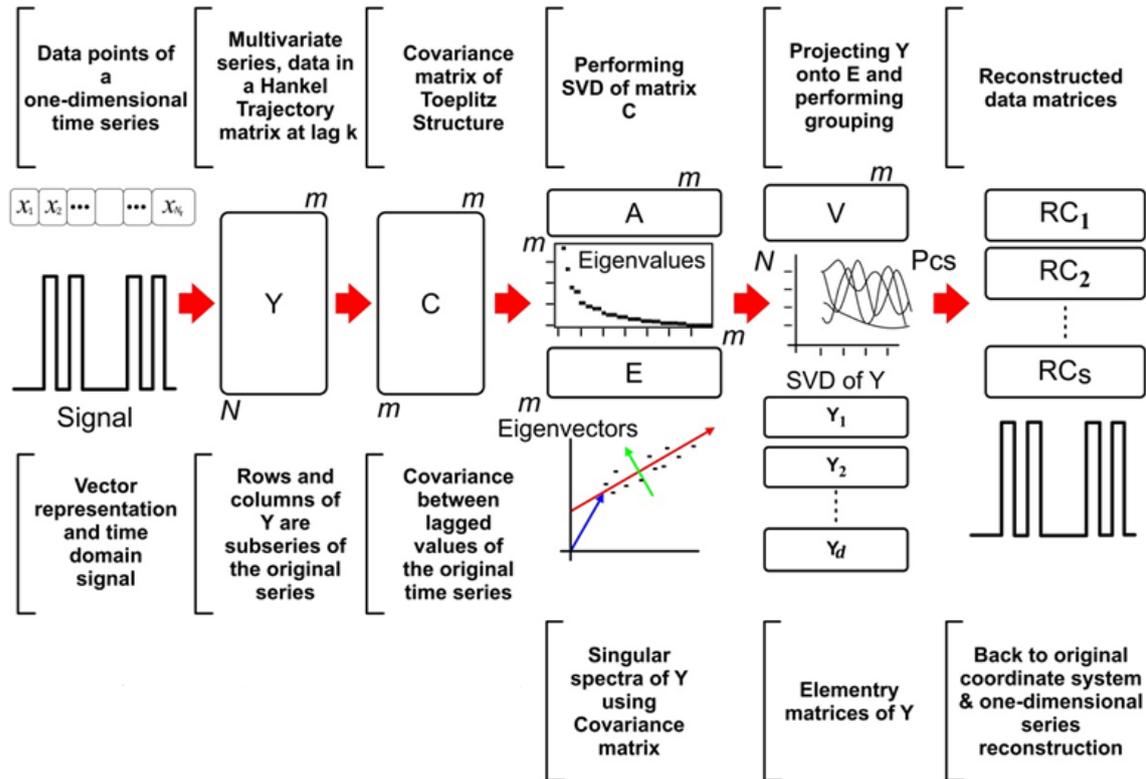
Figure 4.2. A descriptive procedure of the SSA method

Once the SSA decomposes noisy signals in the eigen-subspaces, it selects and groups the principle components according to their contributions to noise and wanted signals in the singular spectrum domain. However, this requires a suitable grouping technique to separate noise components that located in the undesired subspace from the signal of interest in the desired subspace (Sanei and Hassani, 2015). Eventually, the second stage is to reconstruct a time series component corresponding to this subspace (Chu, Lin and Wang, 2013; Golyandina and Shlemov, 2015). The second stage entails the reconstruction of the signal into an additive set of independent time record (García Plaza and Núñez López, 2017). The SSA reconstructs the wanted components back to the time domain resulting in the separation of noise and wanted signals. The reconstruction of the original time series is accomplished by using estimated trend and harmonic components (Ghil *et al.*, 2002). The time series is reconstructed by selecting those components that reduce the noise in the series (Patterson *et al.*, 2011).

4.4 Mathematical Formulation and Algorithm Description

Time series can be stored in a vector denoted by **x** for example whose entries are the data points that describe the time series as a sequence of discrete-time data (Claessen and Groth, 2002). Such vector is an introductory element to many further steps as it includes the required information about the time series. The SSA method is basically consists of main four aspects (Jiang and Xie, 2016; Maddirala and Shaik, 2016; Xu, Zhao and Lin, 2017). Figure 4.3 shows these aspects.
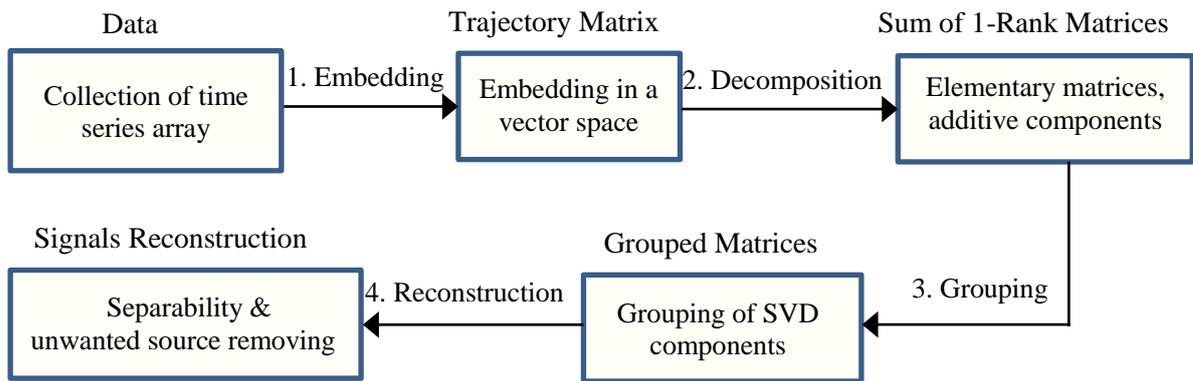
Figure 4.3. Main aspects in the SSA method

The SSA algorithm involves the decomposition of a time series and reconstruction of a desired additive component as the two main stages with multiple steps as illustrated in the Figure 4.3. These steps mainly describe the SSA algorithm and will be explained in more details while considering many important mathematical aspects in the next subsections.

4.4.1   Signal Decomposition Stage

4.4.1.1   Step one: Embedding in a vector space

In practice, the SSA is nonparametric spectral method based on embedding a given time series $\{X(t): t = 1, \dots, N_t\}$ in a vector space. The vector **x**, whose entries N_t are the data points of a time series, can clearly define and describe this time series at regular intervals (Chu, Lin and Wang, 2013; Golyandina and Shlemov, 2015; García Plaza and Núñez López, 2017). When considering a real-valued time series $X(t) = (x_1, x_2, \cdots, x_{N_t})$ of length N_t and $x_1, x_2, \dots, x_{N_t}$ data points, therefore, the given time series can simply be presented as a column vector as shown in Equation (4.1).

$$\mathbf{x}^T = (x_1, x_2 \dots \dots x_{N_t}), \qquad (4.1)$$

This column vector shows the original time series before considering a 1 lag shifted version at a given window length m as the time record can be presented in the vector space. Hence, this vector represents the time series at zero lag (i.e. when there is no delay, $k = 0$). According to that, the window length m should be suitably identified to obtain the lag k which is needed to construct a new matrix according to delay coordinates. In the SSA jargon, this matrix is called "embedded time series" or trajectory matrix and denoted by **Y** as previously mentioned. The window length is also called embedding dimension and it represents the number of time-series' elements in each snapshot (Chu, Lin and Wang, 2013; Elsner and Tsonis, 2013; Golyandina and Shlemov, 2015; García Plaza and Núñez López, 2017).

The whole procedure of the SSA method depends upon the best selection of this parameter as well as the grouping criterion. These two key aspects are very important to develop the concept of reconstructing noise free series from noisy records. Different rank-one matrices obtained from the SVD can be selected and grouped to be processed separately. If the groups are properly partitioned, they will reflect different components of the original time record (Chu, Lin and Wang, 2013; Golyandina and Shlemov, 2015).

4.4.1.2   Step two: Trajectory matrix production

The algorithm generates a trajectory matrix from the original time series $X(t)$ by sliding a window of length m. With the SSA, every time series can be decomposed into a series of elementary matrices after mapping it into a trajectory matrix. Each of these matrices shows glimpses of particular signature of oscillation patterns. The trajectory matrix is then approximated using SVD. A one-dimensional time series $X(t) = x_1, x_2, \ldots \ldots, x_{N_t}$ can be transferred into a multi-dimensional series in the embedding step which can be viewed as a mapping process. The multi-dimensional series contains vectors $\mathbf{x}_k$ which are called m-lagged vectors (or, simply, lagged vectors) as in Equation (4.2).

$$\mathbf{x}_k^T = (x_{k+1}, x_{k+2}, \ldots \ldots, x_{k+(N_t-m+1)}), \tag{4.2}$$

The trajectory matrix **Y** is then considered as multivariate data with m characteristics and N observations. In this embedding step, the single parameter is the window length which is an integer such that $2 \leq m \leq N_t$. The result of the embedding process is a Hankel trajectory matrix **Y** with entries $(x_{ij})_{i,j=1}^{N,m}$. The columns $\mathbf{x}_j$ of **Y** are vectors which lie in m-dimensional space R^m (Hassani, Heravi and Zhigljavsky, 2009). Then the embedded time series can be written as follows:

$$\mathbf{Y} = [\mathbf{x}_{k=0}, \quad \mathbf{x}_{k=1}, \quad \dots \dots \quad , \mathbf{x}_{k=m-1}], \tag{4.3}$$

The trajectory matrix contains the original time series in the first column and a lag 1 shifted version of that time series for each of the next columns. It can be obviously seen from Equation (4.3) that the first column is $X(t)$ when $k = 0$ and the last is $X(t)$ when $k = m - 1$. However, as explained in (Claessen and Groth, 2002), according to delay coordinates, a total number of column vectors equals m will be obtained. Importantly, these vectors are similar in size to the first column vector but with a 1 lag shift. This is seen as a first method when last rows of the produced matrix are supplemented by 0s based on the delay. In the second method, arranging the snapshots of any given time series as row vectors can lead to construct the trajectory matrix when the last rows are not supplemented by 0s (Chu, Lin and Wang, 2013; Golyandina and Shlemov, 2015). However, in this case, a trajectory matrix $\mathbf{Y}$ of size $N \times m$ will be constructed. For more clarification, a representation of any given time series $X(t)$ can be shown as in Figure 4.4 since there is a sliding window with an overlap $m - 1$ used to construct the trajectory matrix in the embedding process. An example with numerical values is also provided in Figure 4.4 to show the construction of $\mathbf{Y}$ considering the delay shift from lag 0 to lag $N - 1$.
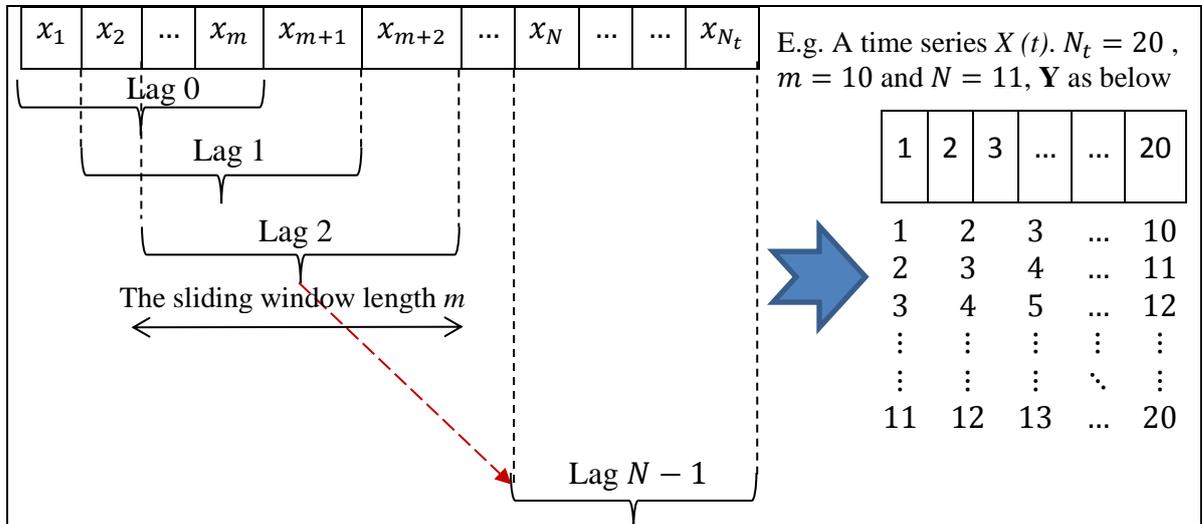
Figure 4.4. Embedding process to construct the trajectory matrix

The coordinates of the phase space can be defined by using lagged copies of a single time series (Elsner and Tsonis, 2013). The trajectory matrix corresponds to a sliding window of size m that moves along the time series $X(t)$ (Chu, Lin and Wang, 2013; Yang *et al.*, 2016). The sliding window has an overlap equals $m - 1$ as shown in Figure 4.4 and the values of the lag k are given by $k = 0,1, \dots \dots, m - 1$. In case of considering the arrangement of column vectors

as in Equation (4.3) with respect to the original time series presented in a column vector, however, the number of lags will be limited by $N - 1$. This is also correct as the snapshots of the given time record can be arranged in row vectors considering the 1 lag shifted version as depicted in Figure 4.4. In all cases, it is to choose the way that gives an adequate understanding of how to manipulate the matrix through its rows or columns while keeping the same principle. The embedding dimension and the lagged time are the main principles applied in this step in which the delays represent the transformation of the time series presented in the vector space into a multivariate set of time observations (Elsner and Tsonis, 2013). The number of rows N of $\mathbf{Y}$, which can be filled with the values of $X(t)$, indicates the number of embedding vectors and is given by $N = N_t - m + 1$ as derived in the previous chapter, recall from Equation (3.3).

Using the trajectory matrix transformation method is a way to make a multivariate statistical analysis possible from a univariate time series. The coordinates of the system can be defined when implementing lagged time of a single time record. However, the window length is seen as the time spanned by each embedding vector when considering the overlap. As explained in (Elsner and Tsonis, 2013), the snapshots of a given record when considering only the number of rows of $\mathbf{Y}$ that can be filled with the values of $X(t)$ are called embedding vectors. These vectors can be seen as $\mathbf{v}_1^T = (x_1, x_2, \dots, x_m)$, $\mathbf{v}_2^T = (x_2, x_3, \dots, x_{m+1})$, and so forth for the rest of the embedding vectors up to $\mathbf{v}_N^T = (x_N, x_{N+1}, \dots, x_{N_t})$. Hence, the trajectory matrix can be constructed by arranging the snapshots as row vectors and only $N_t - m + 1$ rows can be filled with values of $X(t)$ as in Equation (4.4).

$$\mathbf{Y} = 1/\sqrt{N}\begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \mathbf{v}_3^T \\ \vdots \\ \mathbf{v}_N^T \end{bmatrix} = 1/\sqrt{N}\begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_m \\ x_2 & x_3 & x_4 & \cdots & x_{m+1} \\ x_3 & x_4 & x_5 & \cdots & x_{m+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_N & x_{N+1} & x_{N+2} & \cdots & x_{N_t} \end{bmatrix}, \tag{4.4}$$

where $\sqrt{N}$ is the convenient normalisation, N the number of embedding vectors, N_t denotes the length of the time record that is being processed with the SSA at a selected window length.

The constructed trajectory matrix includes the complete record of patterns that have occurred within a window of length m. To generalise, assuming that $X(t)$ is a given time series for $t = 1,2,3, \dots, N_t$, the augmented or trajectory matrix is constructed as in Equation (4.5).

$$\mathbf{Y} = [\mathbf{x}_1: \quad \dots \quad :\mathbf{x}_m] = (x_{ij})_{i,j=1}^{N,m}, \tag{4.5}$$

where $x_{ij} = x_{i+j-1}$ are the entries or time domain samples of the column vectors for $1 \leq i \leq N$, $N = N_t - m + 1$, and $1 \leq j \leq m$, and the column vectors $\mathbf{x}_i$ of this matrix can be presented as $\mathbf{x}_i^T = (x_i, x_{i+1}, \dots, x_{i+N-1})$, for the values of k from 0 to $N - 1$ in a 1 lag shifted increment and $i = 1, \dots, m$ is the index of the columns.

The arrangement of entries x_{ij} of the trajectory matrix depends on the lag. For example, when considering that the trajectory matrix has dimensions $N \times m$, at 0 lag, the trajectory matrix $\mathbf{Y}$ represents the first N elements from time record vector $\mathbf{x}_1^T = (x_1, x_{1+1}, \dots, x_{1+N-1})$, while at lag 1, $\mathbf{x}_2^T = (x_2, x_{2+1}, \dots, x_{2+N-1})$ and so on up to $i = m$ as illustrated in Equation (4.5) and Figure 4.4. As explained in (Golyandina, Nekrutkin and Zhigljavsky, 2001), the successive vectors $\mathbf{x}_i$ are supposed to be long enough to characterise the dynamic of the discrete time record. However, based on the selection of the length of the sliding window, which plays a vital role in calculating the number of the embedding vectors N, and if $N > m$, the number of rows of matrix $\mathbf{Y}$ will be greater than the number of its columns. The dimensionality of the trajectory matrix is highly dependent on two main factors which are the window length and the length of the time record under investigation. These two factors also affect the processing time of the SSA. For longer time series and bigger window size, the dimensionality of $\mathbf{Y}$ will be larger and consequently the overall computation and processing time increases.

It is worth mentioning that $\mathbf{Y}$ can be constructed with different dimensions when using the transpose vectors of the snapshots and arranging these snapshots as column vectors, however, all the above-mentioned equations are always valid in this case as well. Consequently, the size of the covariance matrix will be changed, but without affecting its nature as a square matrix. The trajectory matrix and its transpose $\mathbf{Y}^T$ are linear maps between the spaces R^m and R^N (Chu, Lin and Wang, 2013). Two important properties of the trajectory matrix are stated in (Golyandina and Korobeynikov, 2014), the first is that both the rows and columns of $\mathbf{Y}$ are subseries of the original series. The second is that $\mathbf{Y}$ has equal elements on anti-diagonals which makes it a Hankel matrix (i.e. all the elements along the diagonal $i + j = const$ are equal) as shown in Equation (4.4).

4.4.1.3 Step three: Computing the lagged-covariance matrix

This step is the preparation for applying the SVD in the decomposition stage. The covariance matrix $\mathbf{C}$, which is also known as diagonal matrix, is basically a matrix that shows

the covariance between the values $X(t)$ and $X(t + k)$ which is mainly the covariance between lagged (or "delayed") values of the original time series $X(t)$ (Chu, Lin and Wang, 2013; Golyandina and Shlemov, 2015). According to (Vautard and Ghil, 1989), there are two methods of computing the covariance matrix. Estimating the lagged-covariance matrix directly from the data is one method. Measuring the relation between the entries of the trajectory matrix can produce this lagged-covariance matrix. However, the entries in the (i, j) position of the covariance matrix imply the covariance between the entries in the position (i^{th}, j^{th}) and the elements in the opposite position (j^{th}, i^{th}) of the trajectory matrix **Y**. The repeating patterns reflect the oscillation in the time record as explained in (Elsner and Tsonis, 2013).

The covariance matrix is a square matrix of dimension $m \times m$ and can be generally seen as a Toeplitz matrix with constant diagonals as explained in the previous chapter. Recall from Chapter 3, Equation (3.10), the entries c_{ij} of this matrix depend only on the lag $|i-j|$ (Ghil *et al.*, 2002). As a second method, the lagged-covariance matrix can be computed form the trajectory matrix **Y** and its transpose (Elsner and Tsonis, 2013). Therefore, the product of the trajectory matrix and its transpose can lead to compute the covariance matrix. The variances of each column of **Y** are presented in the main diagonal of the covariance matrix. As previously mentioned, repeating patterns of the trajectory matrix represent oscillations in the original time series. However, by examining these repeating patterns and patterns that appear in **C**, it can be seen that the entries of the covariance matrix are proportional to the linear correlations between all pairs of the patterns appearing in the m-window used to construct the trajectory matrix.

Recall from Equation (3.12), it shows a general mathematical representation of computing the covariance matrix of the snapshots from the original time record using the trajectory matrix and its transpose. It is worth noting that the elements of **C** are all real numbers and $c_{ij} = c_{ji}$ for all i and j, therefore, the covariance matrix **C** is a symmetric matrix, and hence $\mathbf{C} = \mathbf{C}^T$. As the SSA algorithms have been developed using MATLAB working environment and used in several experiments performed in this study, however, several built-in functions are very beneficial in either way of computing the covariance matrix. For example, a specified MATLAB function that gives a vector of size m can be used to construct **C** based on this vector as explained in the previous chapter.

4.4.1.4   Step four: Performing the SVD

- **Spectral decomposition and eigenmodes**

It is a step of computing the diagonal values and their corresponding vectors where each presented in a separate matrix; diagonal and orthogonal matrices, respectively. This step is prior to obtaining the additive components or elementary matrices by performing the SVD of the trajectory matrix and computing the principle components. To obtain spectral information on the time series, the SSA proceeds by diagonalising the covariance matrix. Spectral decomposition is a factorisation of a diagonalisable matrix into a canonical form whereby the representation of the matrix is in terms of its eigenvalues and eigenvectors (Chu, Lin and Wang, 2013; Elsner and Tsonis, 2013). Since the lagged-covariance matrix $\mathbf{C}$ is square symmetric matrix of order $m \times m$ with m linearly independent eigenvectors $(e_i), i = 1, \dots, m$; therefore, a matrix $\mathbf{E}$ can be obtained. As any matrix can be factored into many pieces (Elsner and Tsonis, 2013), hence $\mathbf{E}^{-1}\mathbf{C}\mathbf{E} = \mathbf{\Lambda}$. The columns of $\mathbf{E}$ are the eigenvectors of $\mathbf{C}$.

The product $\mathbf{E}^{-1}\mathbf{C}\mathbf{E}$ is called the diagonal form of $\mathbf{C}$ and therefore $\mathbf{\Lambda}$ is a diagonal matrix whose nonnegative entries are the eigenvalues of $\mathbf{C}$. The eigenvectors of $\mathbf{C}$ should be linearly independent, so as to make $\mathbf{C}$ diagonalisable in this way. Importantly, matrix $\mathbf{E}$ is not unique because the eigenvectors can always be multiplied by a constant scalar preserving their nature as eigenvectors (Elsner and Tsonis, 2013).

As the covariance matrix is assumed to be a real symmetric matrix where $\mathbf{C} = \mathbf{C}^T$, then every eigenvalue obtained by the spectral decomposition of $\mathbf{C}$ is also real. Since all eigenvalues are distinct, however, their corresponding eigenvectors are orthogonal. The real, symmetric matrix $\mathbf{C}$ can therefore be diagonalised by an orthogonal matrix $\mathbf{E}$ in which the columns are the orthonormal eigenvectors of $\mathbf{C}$ (Hassani, 2007). Matrix $\mathbf{E}$ is a diagonalisable matrix whose columns are orthonormal of the real and symmetric covariance matrix. The spectral decomposition of the covariance matrix produces two matrices $\mathbf{\Lambda}$ and $\mathbf{E}$.

Since $\mathbf{E}^{-1}\mathbf{C}\mathbf{E} = \mathbf{\Lambda}, \mathbf{E}^T\mathbf{E} = \mathbf{I}$, and $\mathbf{E}^T = \mathbf{E}^{-1}$,then, $\mathbf{C} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^T$.

Matrices $\mathbf{\Lambda}$ and $\mathbf{E}$ can be written as follows:

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m \end{bmatrix}, \mathbf{E} = \begin{bmatrix} e_1^1 & e_1^2 & \dots & e_1^m \\ e_2^1 & e_2^2 & \cdots & e_2^m \\ \vdots & \vdots & \ddots & \vdots \\ e_m^1 & e_m^2 & \cdots & e_m^m \end{bmatrix}, \tag{4.6}$$

Matrix $\mathbf{\Lambda}$ is symmetric with entries λ_i along the leading diagonal for $i = 1, \ldots\ldots, m$, however, $\mathbf{e}^k$ is the corresponding normalised column eigenvectors aligned in the matrix $\mathbf{E}$ as $\mathbf{E} = [\mathbf{e}^1, \ldots, \mathbf{e}^k]$ for $k = 1, \ldots\ldots, m$ as well. The eigenvectors matrix consists of a set of column vectors with entries e_j^k that represent the jth component of the kth eigenvector. Once these matrices are conserved as square matrices then $j = k = 1, \ldots\ldots, m$, and each single eigenvector is of length m. The diagonal matrix $\mathbf{\Lambda}$ consists of ordered values $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$ (Schoellhamer, 2001; Ghaderi, Mohseni and Sanei, 2011). The $m \times m$ matrix $\mathbf{E}$ that presented the eigenvectors of $\mathbf{C}$ reflects the temporal covariance of the time record at different lags. In fact, the extracted eigenvectors are seen as axes of a new coordinate system considering that any scalar multiple is also an eigenvector (Golyandina and Shlemov, 2015; Harmouche *et al.*, 2017).

- **The singular spectrum**

The so-called singular values of the trajectory matrix are basically the square roots of the eigenvalues of the covariance matrix and these ordered singular values are referred to collectively as the singular spectrum (Hassani, Soofi and Zhigljavsky, 2010; Elsner and Tsonis, 2013; Golyandina and Shlemov, 2015).

From the SVD, the trajectory matrix can be factored in the form $\mathbf{Y} = \mathbf{USE}^T$, where $\mathbf{U}$ is the matrix of the left singular vectors of $\mathbf{Y}$, $\mathbf{E}$ is a square matrix contains the right singular vectors of $\mathbf{Y}$ for $\mathbf{C} = \mathbf{YY}^T$, and $\mathbf{S}$ is a diagonal matrix of singular values. The singular spectrum of $\mathbf{Y}$ consists of the square roots of the eigenvalues of $\mathbf{C}$ which are called the singular values of $\mathbf{Y}$ with the singular vectors being identical to the eigenvectors that given in matrix $\mathbf{E}$ (Elsner and Tsonis, 2013). Therefore, the factorisation of $\mathbf{Y}$ using SVD produces three components which are left singular vectors, diagonal matrix of singular values, and right singular vectors presented in the transpose matrix $\mathbf{E}^T$ whose entries also the eigenvectors of $\mathbf{C}$. These components can be physically interpreted as rotation, scaling, and reverse rotation. The spectral decomposition of matrix $\mathbf{C}$ can also be performed by substituting the form of the trajectory matrix resulted from the factorisation using SVD which is $\mathbf{Y} = \mathbf{USE}^T$ in the equation $\mathbf{C} = \mathbf{YY}^T$, results

$\mathbf{C} = (\mathbf{USE}^T)(\mathbf{USE}^T)^T = \mathbf{ESU}^T\mathbf{USE}^T$,

Since $\mathbf{U}^T\mathbf{U} = \mathbf{I}$, then

$$\mathbf{C} = \mathbf{ES}^2\mathbf{E}^{\mathrm{T}}, \tag{4.7}$$

It has been found that $\mathbf{C} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^T$ and for the decomposition being unique it follows that $\mathbf{S}^2 = \mathbf{\Lambda}$. From the SVD of the trajectory matrix, the right singular vectors of $\mathbf{Y}$ presented in matrix $\mathbf{E}$ are the eigenvectors of $\mathbf{C}$ and the left singular vectors of $\mathbf{Y}$ presented in $\mathbf{U}$ are the eigenvectors of the matrix $\mathbf{Y}\mathbf{Y}^T$ (Chu, Lin and Wang, 2013; Golyandina and Shlemov, 2015). Importantly, the number of eigenvalues is equal to the window length and in turn the number of the associated eigenvectors that matrix $\mathbf{E}$ contains (Elsner and Tsonis, 2013).

4.4.1.5   Step five: Contribution of *PCs* and grouping

A matrix that contains the principle components *PCs* is introduced as a projection of the embedded time series onto the eigenvectors in the eigen-subspace. Therefore, the eigenvectors of $\mathbf{C}$ can be used to compute the principal components vectors with entries $v^{k\prime}s$ by performing such projection. The selected groups of the principle components are presented in vectors and can be aligned in a single matrix. Recall from Chapter 3, this process was given in Equation (3.24). The resultant matrix $\mathbf{V}$ will be of dimension $N \times m$ in which each individual vector of the principle components is presented in each column as in Equation (4.8).

$$\mathbf{V} = \mathbf{Y}\mathbf{E} = \begin{bmatrix} v_1^1 & v_1^2 & \cdots & v_1^m \\ v_2^1 & v_2^2 & \cdots & v_2^m \\ \vdots & \vdots & \ddots & \vdots \\ v_N^1 & v_N^2 & \cdots & v_N^m \end{bmatrix},$$
(4.8)

The principle components are ordered in matrix $\mathbf{V}$ in a similar way to the eigenvectors. The length of the *PCs* is the same of the embedded time series and they are seen as time series, however, unlike the embedded time series introduced in the trajectory matrix, a different coordinate system is used to plot each point. The columns of $\mathbf{V}$ do not correspond to different time lags as in the trajectory matrix. Rather, in the SSA method, the principle components matrix is introduced as a projection of the embedded time series onto the eigenvectors. The original values of $\mathbf{Y}$ have been projected in a new coordinate system for gathering the variance in the principle components (Elsner and Tsonis, 2013; Golyandina and Shlemov, 2015).

The variance between any consecutive principle components (e.g., $\mathbf{v}_1$ and $\mathbf{v}_2$) is identical to that of the eigenvalues of the corresponding eigenvectors. Importantly, due to the orthogonal characteristics of the principle components, it is possible to isolate each individual one from the others to be independently investigated as each principle component contains a part of the oscillation information (Elsner and Tsonis, 2013; Golyandina and Shlemov, 2015). Based on the separability approach introduced with the SSA method and which has been developed in

the case of the wind noise in this study, it is possible to eliminate unwanted components of the signal by selecting their corresponding *PCs*. Usually the selection of the unwanted principle components can be made with the eigenvalue's spectra. However, the dominant *PCs* are the ones that correspond to the significant oscillation or the wanted signal (Hassani, Mahmoudvand and Zokaei, 2011; Golyandina and Shlemov, 2013; Golyandina and Lomtev, 2016).

The principle components can be computed after decomposing the original time series and producing the so-called additive components or elementary matrices of matrix **Y** when considering $\mathbf{C} = \mathbf{Y}\mathbf{Y}^T$. Basically, it is again a projection of the embedded time series onto the eigenvectors. The resultant matrix of the principle components is given in Equation (4.9) in terms of the transpose of matrix **Y**, the eigenvectors matrix **E**, and the singular values matrix **S** which is the square root of the diagonal matrix **Λ** (Golyandina, Nekrutkin and Zhigljavsky, 2001; Saito *et al.*, 2011).

$$\mathbf{V} = \frac{Y^T \mathbf{E}}{\sqrt{\mathbf{\Lambda}}}, \tag{4.9}$$

The collection $(\sqrt{\lambda_i}, e_i, v_i)$ is called the ith eigentriple of the SVD (Golyandina, Nekrutkin and Zhigljavsky, 2001). Any individual *PC* vector can be generally expressed as:

$$\mathbf{v}_i = \begin{bmatrix} v_{ij} \\ \vdots \\ \vdots \\ v_{im} \end{bmatrix} \tag{4.10}$$

where v_{ij} for $(i = 1 \dots N; j = 1 \dots m)$ are the elements of the principle component vectors that form matrix **V**, therefore the multiple *PCs* vectors can be seen as:

$$\mathbf{v}_1 = \begin{bmatrix} v_{11} \\ v_{12} \\ \vdots \\ v_{1m} \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} v_{21} \\ v_{22} \\ \vdots \\ v_{2m} \end{bmatrix}, \text{and so forth until } \mathbf{v}_N = \begin{bmatrix} v_{N1} \\ v_{N2} \\ \vdots \\ v_{Nm} \end{bmatrix}$$

Hence, when the complete set of vectors are aligned together, matrix **V** will be as

$$\mathbf{V} = \begin{bmatrix} v_{11} & v_{21} & \cdots & v_{N1} \\ v_{12} & v_{22} & \cdots & v_{N2} \\ v_{13} & v_{23} & \cdots & v_{N3} \\ \vdots & \vdots & \ddots & \vdots \\ v_{1m} & v_{2m} & \cdots & v_{Nm} \end{bmatrix}, \tag{4.11}$$

Grouping corresponds to splitting the elementary matrices $\mathbf{Y}_i$ into several groups and summing the matrices within each group (Sanei and Hassani, 2015). The SVD of $\mathbf{Y}$ can be written as a set of elementary matrices and performed as in Equation (4.12).

$$\mathbf{Y} = \mathbf{Y}_1 + \mathbf{Y}_2 + \cdots + \mathbf{Y}_d, \tag{4.12}$$

where $\mathbf{Y}_i = \sqrt{\lambda_i} e_i v_i^T$ are the additive components or elementary matrices of $\mathbf{Y}$ for $i = 1, \dots, d$ and d is the number of non-zero eigenvalues of $\mathbf{Y}\mathbf{Y}^T$ which equals to the window size.

The definition of $\mathbf{Y}_i$ is equivalent to the elementary matrix. If only two groups are required when applying the grouping criterion, the number of elementary matrices in the first group will be denoted by r while the rest $m - r$ will represent the matrices in the second group. The set $(\sqrt{\lambda_i} e_i v_i)$ is the ith eigentriple of matrix $\mathbf{Y}$ of the SVD when considering $\mathbf{e}_1, \dots, \mathbf{e}_d$ which are the orthonormal system of the associated eigenvectors and the principle components $\mathbf{v}_1, \dots, \mathbf{v}_d$ (Golyandina, Nekrutkin and Zhigljavsky, 2001; Sanei and Hassani, 2015; Xu, Zhao and Lin, 2017). Using the eigentriple as in Equation (4.9), the following elements are obtained.

$$\mathbf{Y}^T = \sqrt{\Lambda}\,\mathbf{E}\mathbf{V}^T, \tag{4.13}$$

$$= \sqrt{\lambda_1}\,e_1 v_1^T + \sqrt{\lambda_2}\,e_2 v_2^T + \dots \dots \sqrt{\lambda_m}\,e_m v_m^T$$

$$= \mathbf{Y}_1 + \mathbf{Y}_2 + \cdots \mathbf{Y}_m$$

The decomposed matrix $\mathbf{Y}_i$ is given as follows:

$$\mathbf{Y}_i = \sqrt{\lambda_i}\,e_i v_i^T = \begin{bmatrix} y_{11}^i & y_{21}^i & \dots & y_{m1}^i \\ y_{12}^i & y_{22}^i & \dots & y_{m2}^i \\ y_{13}^i & y_{23}^i & \dots & y_{m3}^i \\ \vdots & \vdots & \ddots & \vdots \\ y_{1N}^i & y_{2N}^i & \dots & y_{mN}^i \end{bmatrix}, \tag{4.14}$$

where y_{ij}^i are the entries of $\mathbf{Y}_i$ for $(i = 1 \dots m; j = 1 \dots N)$.

The SSA is a well-known method for time series analysis and generally seen as an adaptive noise reduction and signal discrimination technique in many applications based on its capabilities for decomposing and filtering noisy signals (Sivapragasam, Liong and Pasha, 2001). The decomposition stage in the SSA method can deliver significant results if the produced additive components of the embedded time series are separable from each other (Golyandina, Nekrutkin and Zhigljavsky, 2001). In this study the separation approach has been

introduced to develop the SSA for separating noise components from given signals contaminated with wind noise. In the developed SSA algorithms, it is to perform spectrum analysis on the given input data, eliminate irrelevant features, and invert the remaining desired components to yield a noise-free signal.

4.4.2   Diagonal Averaging and 1-Dimensional Series Reconstruction

Recall from Chapter 3, Equation (3.27), the reconstructed components *RCs* can be computed by projecting the principle components presented in matrix **Z** onto the eigenvector's matrix **E**. The *RCs* can also be calculated by inverting the projection of the principle components onto the eigenvectors transpose matrix as in Equation (4.15). The aim is to generate reconstructed components whose length are the same as the original record. Each *RC* can be generated by the convolution of one principle component with the corresponding singular vector (Vautard, Yiou and Ghil, 1992; Wu, Chau and Li, 2009).

$$\mathbf{RC} = \mathbf{YEE}^T, \tag{4.15}$$

Each *RC* re-translates its corresponding *PC* into the original units of the time series $X(t)$. Now, the comparison between the reconstructions based on the selection of the *PCs* and $X(t)$ is possible. As explained in (Elsner and Tsonis, 2013), the singular vectors can be used to compute the *PCs* of the time series. Therefore, to reflect oscillatory modes of interest, the original time series can be filtered through a convolution when selecting a small number of *PCs* and their associated eigenvectors. Importantly, from these reconstructions, the time series has been reduced to *oscillatory components* that correspond to the most dominant eigenvalues with high variance and *noise components* that correspond to the rest of eigenvalues.

The produced *RCs* can be used to filter the time series by using less of the total number of *RCs* (Elsner and Tsonis, 2013). The part of the time series which supposed to be noise and mainly represented in the higher subspace of the eigenvalue's spectra can be separated out when considering a proper grouping criterion and wind noise spectrum explained in Chapter 2. In other words, the first *RCs*, however, are generally the most dominant that defining an oscillatory signal due to the phase quadrature of the corresponding *PCs* and their eigenvalues.

A group of *r* eigenvectors with $1 \leq r \leq d$ defines an *r*-dimensional hyperplane in the *m*-dimensional space R^m of vectors in $\mathbf{Y}_i$. The projection of **Y** into this hyperplane will approximate the original matrix **Y**. Once the boundaries of each group have been determined, in this second complementary stage, the next step is then to sum-up the matrices in each cluster

together. The selected group I of the additive matrices that represent the oscillatory components, of which the time series has been reduced to, corresponds to the most dominant eigenvalues. All the entries along the diagonal $i + j = const$ in the additive matrices presented in groups are not equal as such matrices are not Hankel. Therefore, to reconstruct the signal, it is required to perform the diagonal averaging approach over the diagonals $i + j = const$ (Hassani, Soofi and Zhigljavsky, 2010; Golyandina and Korobeynikov, 2014).

Averaging over the diagonals of matrices gathered in a specific group I will enable completing such transformation back to one-dimensional time domain vector which tends to characterise a specific signal component. If the noisy signal consists of two components, the other group will represent the residual noise. The ith reconstruction components x_j^i of the reconstructed subseries are given by the diagonal averaging (Saito *et al.*, 2011; Elsner and Tsonis, 2013; Mohammadi *et al.*, 2016). Therefore, from Equation (4.15) and after summing up the elementary matrices through computing the contribution of each to the wanted signal, matrix $\tilde{\mathbf{Y}}$ can be found as follows:

$$\tilde{\mathbf{Y}} = \begin{bmatrix} y_{11}^i & y_{21}^i & y_{31}^i & y_{41}^i & \cdots & \cdots & y_{m1}^i \\ y_{12}^i & y_{22}^i & y_{32}^i & \cdots & \cdots & \cdots & y_{m2}^i \\ y_{13}^i & y_{23}^i & y_{33}^i & \cdots & \cdots & \cdots & y_{m3}^i \\ y_{14}^i & y_{24}^i & y_{34}^i & \cdots & \cdots & \cdots & y_{m4}^i \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & y_{m,N-1}^i \\ y_{1N}^i & y_{2N}^i & y_{3N}^i & \cdots & \cdots & y_{m-1,N}^i & y_{mN}^i \end{bmatrix} \tag{4.16}$$

$$x_1^i = y_{11}^i$$

$$x_2^i = (y_{12}^i + y_{21}^i)/2$$

$$\vdots$$

$$x_{N_t-1}^i = (y_{m,N-1}^i + y_{m-1,N}^i)/2$$

$$x_{N_t}^i = y_{mN}^i$$

The averaging ends when $j = N_t$ for x_j^i entries. The transformation back to the univariate time record is accomplished and the desired series is now reconstructed into N_t components.

$$\breve{X}(t) = (x_1^i, x_2^i, \dots, x_{N_t}^i), \tag{4.17}$$

where N_t represents the length of one-dimensional reconstructed time series and in the given vector, x_j^i represents the entries of the subseries i for j from 1 to N_t.

4.5 The SSA Analysis Interpretation

The simplest explanation of the idea of the SSA lies in its two complementary stages which are the decomposition of a time series into oscillatory components and noise and then the reconstruction of the desired components. As mentioned earlier, the aim is to make a decomposition of the original time series into a small number of independent and interpretable components such as a slowly varying trend, oscillatory components (harmonics), and a structure less noise. The idea of the embedded time series is based on the chaos theory which indicates that the parameters specifying a system at a 1 lag shifted version of time t+1 can be seen as given functions of those parameters at time 1 (Canavier, Clark and Byrne, 1990). In fact, using lag shifted versions of time of a single time record can help in defining the coordinates of the dynamic system. The trajectory matrix results from the method of delays in which the dynamic of the system can be approximated by the coordinates of the phase space when using lagged copies of the time series. With a careful choice of the window length, the trajectory matrix can reflect the evolution of the time record (Wu, Chau and Li, 2009).The embedding dimension or the window length is the time spanned in the vector space by each embedding vector.

When using the SSA for noise reduction, the reconstruction can be viewed as a separation method when such reconstruction is partially made. In other words, not all information is used, only a selected portion which is the essential portion of the time series that represents the signal of interest in the desired subspace. For this reason, only eigenvectors that indicate a large portion of the variance measured by their eigenvalues are used. The eigenvectors matrix reflects the temporal covariance of the time record at different lags and a new coordinate system in the Eigen domain is used to produce the required principle components.

The SSA reconstruction can be interpreted in such a way that for a specific time t, observed values of the time series $X(t)$ that fall within the window of size *m* around time t are considered. However, in this window and as a particular kind of weighted average of the values of $X(t)$, the reconstructed components at time t are computed. Hence, the eigenvectors define the weights for this averaging. The reconstruction, however, relates to a kind of moving average of a given time series. The moving average is identical to the original time series when all of eigenvectors are considered.

4.6 Summary

In this chapter, a framework of the SSA method has been established and developed based on its capabilities. The chapter covered the multiple steps of the SSA algorithm while defining and clarifying all related mathematical formulations. The SSA procedure consists of two main complementary stages which are decomposition and reconstruction. The first step in the decomposition stage is called the embedding process in which the given signal presented in a vector is transformed into an embedded time series known as a trajectory matrix. This step is a transformation from one-dimensional time series to multidimensional. Then the trajectory matrix and its transpose are used to compute the lagged-covariance matrix. The second step is considered the most significant in which two important matrices are computed using the Singular Value Decomposition method. These two matrices are the diagonal square matrix with non-negative values called the eigenvalues aligned along the main diagonal and the matrix that contains their corresponding eigenvectors. The latter matrix is also a square matrix represents the right eigenvectors; however, each eigenvector is presented in a column of this matrix.

The third step is to compute the principle components using these two matrices along with the embedded time series presented in the trajectory matrix. In fact, the principle components can be computed after decomposing the original time series using the SVD method as well to produce the elementary matrices or also called the additive components. In this case, when considering the eigentriple of SVD of the trajectory matrix, a mathematical formula has been devised to compute the principle components includes matrices $\mathbf{E}$ and $\mathbf{S}$, where $\mathbf{S}$ is the square root of the diagonal matrix $\mathbf{\Lambda}$ produced from the singular spectrum of matrix $\mathbf{C}$. Computing the principle components in this way is seen as a projection of the embedded time series onto a new coordinate system, which is the eigenvectors, in the eigen-subspace.

Grouping which corresponds to splitting the elementary matrices into several groups and summing the matrices within each group is therefore performed. The SSA visual tool which is based on the singular spectrum that shows the eigenmodes and indicates the eigenvalues in descending order versus their index can be used for the selection of the principle components that correspond to the most dominant eigenvalues. These principle components are added together, however, the undesired ones have to be eliminated. In the reconstruction stage and after computing the reconstructed components, it is to map the wanted target signal back to the time domain. The reconstructed series is therefore a time domain signal and can be presented in a vector similar to the original series.

Significant results can be obtained if the resulting additive components of the time series from the decomposition stage are approximately separable as stated in (Golyandina, Nekrutkin and Zhigljavsky, 2001). Importantly, the performance of the SSA algorithm is highly dependent upon the selection of the length of the sliding window along with the applied grouping criterion. Furthermore, such aspects have a key importance for the whole SSA procedure towards improving the separability approach. These key aspects will be discussed in the next chapter.

5 CHAPTER FIVE   GROUPING, SEPARABILITY, AND RECONSTRUCTION TECHNIQUES

5

Grouping, Separability and Reconstruction Techniques

5.1 Overview

The SSA method works based on the principle of how well the components of a given signal can be taken a part or separated from each other (Mohammadi *et al.*, 2016). As mentioned previously, two stages that complement each other are entailed in the SSA approach; decomposition and reconstruction. In turn, each of these two stages includes some distinct steps. To decompose the signal, embedding process accompanied by SVD method are the two main steps involved in the first stage. The reconstruction stage involves applying the grouping criterion and the diagonal averaging for reconstructing the one-dimensional series that might be exploited for further analysis.

It remains difficult to solve microphone wind induced problem due to several reasons, namely the broadband and time varying nature of wind noise. Developing the adaptive selection approach of eigentriples in the singular spectrum analysis might be a key to success to improve the separability and a better performance regarding microphone wind induced removal. Decomposing noisy signals into additive components or elementary matrices after producing embedded time series will lead to decompose the given time series into numbers of oscillations that correspond to the signal of interest and unwanted wind noise components. This can be performed by using SVD method and exploiting wind noise features based on wind noise spectrum to generate groups of oscillatory and wind noise components. Therefore, in this chapter, it is to introduce and develop the separation approach while a high significance is given to develop grouping and reconstruction techniques to improve the separability.

Since the SSA represents a non-parametric statistical method which enables it to be used with arbitrary signals nevertheless of their processes or distribution such as stationary or non-stationary signals, however, these advantages and capabilities motivate its development for microphone wind induced separation. Moreover, the SSA is a well-known method for time series decomposition and analysis as well as the categorisation of the oscillation signatures of the time record over time. All these specific characteristics might lead to better separation of the wanted signal and noise with less distortion imposed on the signal of interest.

With the SSA, every time series can be decomposed into many additive components or a series of elementary matrices which can be achieved by mapping the time record under test into a trajectory matrix and then processing it using the SVD method. These elementary matrices can be grouped in smaller groups by applying a proper grouping criterion based on the aim of the decomposition. After arranging such matrices in different groups, however, it is now to sum-up matrices that show glimpses of a specific signature of oscillation patterns while leaving the random noise behind. This can be accomplished by using many tools such as the eigenvalues spectra and the selection of most dominant principle components. It is therefore the reconstruction stage in which computing the reconstruction components is performed to map the desired signal back to the time domain and as a result a time domain reconstruction series that is similar to the original one can be produced.

Basically, as explained in (Sanei, Ghodsi and Hassani, 2011), in which a clear demonstration of the decomposition of a time structure into a filtered signal and noise is given, the decomposition approach is seen as a common practice towards noise removal aim. In fact, the SSA works by decomposing signals into oscillatory components, trends and noise by deploying the SVD method. However, the SSA has been successfully applied for separating desired signals and noise with an obvious advantage of slight distortion being imposed on the desired signals. The SSA method has been applied in a variety of signal processing applications such as EEG and ECG for the purpose of the separation and localisation of a combination of signals produced from amplitudes/frequencies that are generally different (Sanei and Hosseini-Yazdi, 2011; Bonizzi *et al.*, 2015; Wang, Liu and Dong, 2016).

5.2 Parameters of the SSA Algorithm

The SSA technique depends upon two important parameters which are the window length m, which is the sole parameter in the decomposition stage, and the number of elementary matrices r. The selection of the optimal value of the window length is crucial for an increased accuracy of the SSA method (Tzagkarakis, Papadopouli and Tsakalides, 2009). According to (Yang *et al.*, 2016), the window length is highly related to spectral information or frequency width that corresponds to each principle component. Moreover, the performance of the SSA algorithm very much depends on the window length. In spite of the diversity in selecting the best values of the window length in relation to the length of the given time record, however, it is still important to follow the standard SSA recommendations. Also, such recommendations might be helpful in choosing the group of indices of the eigenvectors in the grouping stage.

The length of the sliding window represents the number of columns of the trajectory matrix. In practice, the size of the embedding dimension should make a compromise between the quality of the information and the computational complexity. Also, the embedding dimension should be adequately large to capture one period or more of the expected periodic signal. In addition to the methods of selecting the window length as a percentage of the length of the time series, it can also be selected according to the lowest frequency of interest (Mohammadi *et al.*, 2016).

There is no general rule for the choice of the window length since the selection depends on the initial information on the time record and the problem of interest (Traore *et al.*, 2017). As stated in (Alexandrov, 2009), the choice of the window length parameter used for the decomposition and grouping of SVD components, which eventually will be used for reconstruction, can totally affect the output time series. It is important, however, to select values of m and groups of the eigenvectors to ensure better separability (Launonen and Holmström, 2017). The performance of the SSA algorithms is highly dependent on the selection of the window length (Harmouche *et al.*, 2017; Rodrigues and Mahmoudvand, 2017).

In spite of the greater computation burden due to bigger sized matrices, it has been recommended that m should be large enough but not greater than $N_t/2$ to significantly represent separated components and obtain satisfied results (Rukhin, 2002; Yang *et al.*, 2016). Whereas, other studies reported that it should be larger than $N_t/2$ when $N > m$ (i.e. the trajectory matrix has many rows greater than columns) (Harris and Yuan, 2010). It has been stated in (Hassani and Mahmoudvand, 2013; Golyandina and Shlemov, 2015) that the window length should be sufficiently large to provide the information about the data variation. It can be always assumed that $m \leq N_t/2$ as this value has been regarded as the most interesting case in practice. If m is too large, this will leave too few observations from which to estimate the covariance matrix of the m variables. Generally, large values of m induce longer period oscillations to be resolved.

Previous research has established that that the window length m should be relatively bigger because when it is considerably smaller than N_t, the results are not greatly sensitive to m (Penland, Ghil and Weickmann, 1991). It has been reported that the influence of the variations of the window length about a sufficiently large m can only be on stretching or compressing the spectrum of eigenvalues while leaving the relative magnitudes of the individual eigenvalues unchanged (Elsner and Tsonis, 2013).

In spite of the considerable attempts and various methods that have been considered for choosing the optimal value of m, there is inadequate theoretical justification for such selection (Patterson *et al.*, 2011; Chu, Lin and Wang, 2013; Yang *et al.*, 2016). Whereas, according to (Elsner and Tsonis, 2013), the window length used to construct the trajectory matrix can be computed as $(N_t/4)$ and considered as a common practice. Smaller values can also be considered if the purpose is to extract the trend even when the time series is short (small N_t) (Alexandrov, Golyandina and Spirov, 2008). Therefore, in the latter study a small window length was used to extract trend for a short time series, however, the separability was insufficient. A method of selecting the window length is described in (Hassani, Mahmoudvand and Zokaei, 2011; Yang *et al.*, 2016) and a detailed description of selecting this parameter is given in (Golyandina and Korobeynikov, 2014).

Achieving appropriate separability of the components is a cardinal rule in selecting the window length. Hence, the decomposition stage of the SSA delivers significant results if the resulting additive components of the time series are approximately separable from each other (Golyandina and Lomtev, 2016; Harmouche *et al.*, 2017). The improper selection of m would imply an inferior decomposition, and in turn inaccurate results will be produced. The improper selection of r will also have its influence on the results although some visual inspection of the SSA decomposition of the whole series or some large parts of the series is recommended. Alternatively, if a preliminary study of the time series is not possible, then it is advised to use all visual SSA tools in the first part of the series to choose r (Hassani, Mahmoudvand and Zokaei, 2011; Golyandina and Shlemov, 2015).

The group $I = 1, \dots r$ which is based on the number of eigenvalues is such that the first r values describe well the signal while the lower components correspond to noise. A part of the signal will be missed if r is smaller than the true number of eigenvalues (under fitting). As consequence, the reconstructed series becomes less accurate due to the underestimated components. Otherwise, if r is too large (over fitting), then a part of noise together with the signal will be approximated in the reconstructed series and hence finding a change in the signal becomes more difficult (Hassani and Mahmoudvand, 2013).

The choice of m depends on some criteria such as complexity of the data, the aim of the analysis as well as the forecasting horizon in case of using the SSA for forecasting studies. Still there is lack of enough rationalisation regarding the selection of the optimal values of m and r (Patterson *et al.*, 2011; Yang *et al.*, 2016).

Since the aim of this study is to separate out signals of interest and wind noise components not yet examined with the SSA through exploring and developing the grouping and reconstruction techniques as key important aspects in the method, it is important to determine the optimum window length. However, more experimental investigation has been carried out for the selection of the optimal length of the sliding window following a heuristic method which will be shown in the next chapters.

5.3 Singular Value Decomposition in the SSA Algorithm

Similar to the Principal Component Analysis technique that uses Singular Value Decomposition SVD, the SSA utilises a representation of the data in a statistical domain which called Eigen domain rather than time or frequency domain. To be more precise, however, instead of using a set of axes that represent discrete frequencies such as with the Fourier Transform, the data is projected onto a new set of axes that fulfil some statistical criterion. This statistical technique depends on the structure of the data being analysed, whereas, in Fourier-based techniques, Fourier components onto which a data segment is projected are fixed. Therefore, in this statistical method, data is projected onto axes that might change according to the change of the structure of the data over time (Clifford, 2005; Golyandina *et al.*, 2013).

With the SSA, de-noising any given signal contaminated with noise can effectively be performed by discarding the projections that correspond to the unwanted sources such as the noise and inverting the transformation. Therefore, a wind noise separation method based on singular spectrum analysis has been developed in this research. In the developed method, it is to use the SVD method for the decomposition of the given noisy time record embedded in the trajectory matrix and then the reconstruction of a noise-free series through developing the adaptive selection approach of eigentriples in the singular spectrum analysis.

Data contained in the components that correspond to the eigenvectors associated to the higher-order eigenvalues in the singular spectra is assumed to be mostly represent noise (Hassani, 2007). Unwanted source can be removed from the original signal as long as data can be transformed back to the original observation space using matrix manipulation in SSA-based techniques. SVD can be used to determine the principal components of a multi-dimensional signal (Clifford, 2005). Following the previous stage, which is obtaining the trajectory matrix from the real observations of the time series, it is now to perform the SVD to decompose matrix $\mathbf{Y}$ into its eigen subspaces, however, $\mathbf{Y}$ is decomposed in the form $\mathbf{Y} = \mathbf{USE}^T$.

In this mathematical formula, $\mathbf{S}$ is a matrix with nonzero entries along the leading diagonal and zeros elsewhere and its entries s_i are arranged in descending order of magnitude in a similar way of the eigenvalue's matrix. In fact, each entry of $\mathbf{S}$ is equal to the square root of the eigenvalues $\sqrt{\lambda_i}$ of the covariance matrix $\mathbf{C}$. These square roots of the eigenvalues of the matrix $\mathbf{C}$ are called the singular values of the trajectory matrix $\mathbf{Y}$ and they are referred to as the *singular spectrum*. Hence, $\mathbf{S}$ is called a diagonal matrix of singular values. The eigenvectors are presented in the columns of matrix $\mathbf{E}$ which represents the right singular vectors of $\mathbf{Y}$ while matrix $\mathbf{U}$ represents the left singular vectors of $\mathbf{Y}$.

As previously mentioned, the principle components matrix is basically a matrix of projection. In other words, it is the matrix that computed by projecting the trajectory matrix $\mathbf{Y}$ onto the eigenvectors of $\mathbf{C}$. Hence, the matrix $\mathbf{YE}$ is the trajectory matrix projected onto the basis $\mathbf{E}$ since $\mathbf{E}$ is composed of orthogonal vectors called the singular vectors of $\mathbf{Y}$. Importantly, the eigenvectors of $\mathbf{C}$ are the singular vectors of $\mathbf{Y}$ (Golyandina *et al.*, 2013).

From the singular value decomposition, the matrix that results from the projection of the trajectory matrix $\mathbf{Y}$ onto the eigenvectors of $\mathbf{C}$ can be seen in the form of the matrix $\mathbf{U}$ that represents the left singular vectors of $\mathbf{Y}$ by applying the principle of the singular spectrum. Therefore, from the SVD of $\mathbf{Y}$, which is indicated above, matrix $\mathbf{U}$ can be derived using the factorisation formula of $\mathbf{Y}$ as follows:

$$\mathbf{Y} = \mathbf{USE}^T, \tag{5.1}$$

Since $\mathbf{E}^T\mathbf{E} = \mathbf{I}$, then multiplying both sides of Equation (5.1) by $\mathbf{E}$ from the right according to matrix multiplication rules $\mathbf{YE} = \mathbf{USE}^T\mathbf{E}$.

The simplified form becomes as $\mathbf{US} = \mathbf{YE}$.

Since $\mathbf{S}$ is a symmetric matrix, hence $\mathbf{S}^T = \mathbf{S}^{-1}$, therefore by multiplying both sides of the above form by $\mathbf{S}^{-1}$, matrix $\mathbf{U}$ can be represented as follows:

$$\mathbf{U} = \mathbf{S}^{-1}\mathbf{YE}, \tag{5.2}$$

Finding the first r column vectors of $\mathbf{U}$ of elements u_i , where $u_i = s_i^{-1}y_ie_i$, $(i = 1:r)$, s_i^{-1} are the elements of $\mathbf{S}^{-1}$ and e_i are the elements of $\mathbf{E}$. The rest of $m - r$ vectors are not of interest as they represent noise. Recall from Chapter 4, Equation (4.12), the SVD of the trajectory matrix can be written as a set of elementary matrices $\mathbf{Y} = \mathbf{Y}_1 + \mathbf{Y}_2 + \cdots + \mathbf{Y}_d$. The grouping, which corresponds to splitting the elementary matrices $\mathbf{Y}_i$ into several groups and

summing the matrices within each group, can then be performed. If a truncated SVD of **Y** is performed, which means only the significant *p* eigenvectors are retained, then the resultant matrix **Y`** in which the columns are the noise-reduced signal is as follows:

$$\mathbf{Y}^{\grave{}} = \mathbf{U}\mathbf{S}_p\, \mathbf{E}^T,$$
(5.3)

It is worth mentioning that for the smaller eigenvalues, the energy represented along the corresponding eigenvectors is low. Consequently, the smallest eigenvalues are located in the higher subspace where commonly less powerful signals occur, and noise components generally arise. Whereas, more powerful signals correspond to lower-order eigenvalues are located in the lower subspace (Mohammadi *et al.*, 2016).

5.4 Grouping

5.4.1 Elementary Matrices Selection in Grouping Criterion

Generating the elementary matrices is a convolution of the principle components with their corresponding eigenvectors and weighted by the eigenvalues; however, each elementary matrix represents a specific oscillation of the signal including what can be considered as noise. As explained in the above section, performing the SVD of the trajectory matrix produces a set of elementary matrices. The adaptive selection approach of eigentriples in the singular spectrum analysis has been developed for the best reconstruction of the desired signal. However, the projections that correspond to the unwanted sources have to be discarded before inverting the transformation (Golyandina *et al.*, 2013).

The eigenvalues spectra are often used for detecting the boundary of the subspace that belongs to the signal of concern as the spectra of the eigenvalues are computed based on the variance. The noise is autocorrelated, which means that more variance project onto the lower-frequency oscillations. However, with the increase in the lag k, higher associated frequencies fall down monotonically (Elsner and Tsonis, 2013). In other words, this indicates that higher variance in the frequencies project onto the lower subspace of the eigenvalue spectra.

The effectiveness of the SSA method depends largely on the selection of the elementary matrices in suitable groups through a criterion that is convenient to reconstruct the desired components without losing any portion of the signal. Notably, each matrix in the selected groups is supposed to share similar harmonic characteristics (Hassani, 2010). However, different grouping techniques have been reported in some previous studies for detecting the

boundaries of the groups. For example, in their study that is related to recognise walking patterns and level walking using the SSA, Jarchi and Yang (2013) used only the second and third elementary matrices for the reconstruction. Their selection was based on considering that these matrices were related to the dominant oscillations since their corresponding eigenvalues were remarkably similar and the signal has a periodic pattern.

Another important grouping technique is to consider pairs of eigenvalues that have nearly equal values for each pair. However, this means that each pair of similarly equal values corresponds to a significant oscillation pattern (Vautard and Ghil, 1989). This grouping criterion has been adopted by (Mohammadi *et al.*, 2016) and has been used to extract sleep spindles, brain waves, and the so-called K-complexes from a sleeping EEG signal. It is worth mentioning that finding an appropriate group of eigentriples to reconstruct the desired components of a given signal is one of the major concerns in the SSA method (Enshaeifar, Sanei and Took, 2014; Mohammadi *et al.*, 2016).

5.4.2 Separation Boundaries and Threshold Setting

Initially, the SSA can be used to split the elementary matrices into only two groups, which is the case of the selected signals in this study, through computing the contribution of each elementary matrix $\mathbf{Y}_i$ (Ma *et al.*, 2012). For developing a noise removal and separation method based on the singular spectrum analysis, which is the case of this research, higher and lower subspaces have to be identified and carefully distinguished. High-order eigenvalues can be isolated as they represent low oscillations. However, detecting the isolated boundaries can be completed based on the analysis of the eigenvalue's spectra.

Generally, low-order eigenvalues, which are related to the wanted components, can be determined by standard SSA recommendations. When only two groups are considered, applying such recommendations by grouping the first r elementary matrices in one group and leaving the rest $m - r$ that represent wind noise in another group is the primary step in the grouping stage towards developing and establishing a proper procedure. This procedure can lead to separate wind noise components out from the desired signals as in Equation (5.4).

$$\frac{\sum_{i=1}^{r} \lambda_i}{\sum_{i=1}^{m} \lambda_i} \geq \mathcal{L} \tag{5.4}$$

where $\mathcal{L}$ is a threshold that specifies the boundary between groups I and $\bar{I}$, r represents the first selection of the elementary matrices, m is the window length which exactly equals to d elementary matrices produced from SVD, d has been indicated in Equation (4.12).

The threshold $\mathcal{L}$ has been defined by many authors as to equal or more than 0.85 which mostly defines the "elbow" point in the eigenvalue's spectra as a boundary. However, immediately after this point the singular spectrum turns to the lower region that mostly represents the noise floor (Mamou and Feleppa, 2007; Ghaderi, Mohseni and Sanei, 2011). In fact, this changeover in the eigenvalue's spectra is from one subspace to anther or form lower- to higher-order eigenvalues. In some other studies, this threshold has been specified by 90% of the total variance of the signal as in (Mohammadi *et al.*, 2016). The identification of this ratio is to discard the components that most likely correspond to the noise floor in the subspace.

The first group I composes the desired elementary matrices after aggregating them together to implement each group in a single matrix of the same dimension as of the trajectory matrix. The unwanted components which reflect the wind noise component are now in the other group $\overline{I}$. These groups can be mathematically expressed as in Equation (5.5).

$$I = \sum_{i=1}^{r} \mathbf{Y}_i , \quad \overline{I} = \sum_{i=r+1}^{m} \mathbf{Y}_i \tag{5.5}$$

The grouping criterion that has been implemented in this study is based on the selection of the eigenvalues that seem to be pairs, which means those eigenvalues with nearly equal values. However, it has been considered to firstly compute the variances between the successive eigenvalues and then selecting the eigenvalues that have the smallest difference.

Splitting the elementary matrices into several groups I_t is seen as an initial step in the reconstruction stage. The elementary matrices in each group are then summed up in one matrix (Golyandina, Nekrutkin and Zhigljavsky, 2001). In fact, each group is displayed by the associated matrix $\widetilde{\mathbf{Y}}_I \subset \mathbb{R}^{N \times m}$ and given in a general mathematical expression as follows:

$$\widetilde{\mathbf{Y}} = \sum_{I=1}^{I_t} \widetilde{\mathbf{Y}}_I \tag{5.6}$$

where the sum of the elementary matrices within the group I is represented by $\widetilde{\mathbf{Y}}_I$, I_t indicates the total number of groups, and I is an index that refers to the Ith subgroup of eigentriples.

A specific $\widetilde{\mathbf{Y}}_I$ is selected after completing the splitting stage; however, the Hankelization procedure is then required to reconstruct the subseries. As mentioned previously in section 4.4.2 in Chapter 4, this procedure is the averaging along entries $i + j = const$. Recall from Chapter 4, Equation (4.16), the elements y_{ij}^{i} of the matrix $\widetilde{\mathbf{Y}}_I$ are under the procedure of anti-

diagonal averaging. The nth term x_n^i of the new reconstructed series $\breve{X}(t)$ for $n = 1, 2, \ldots, N_t$ is computed according to this procedure by calculating the average of all along all i, j in a manner that $i + j = n + 1$. Therefore, in the reconstructed series for $n=1$, the first entry can be obtained as $x_1^i = y_{11}^i$, for $n=2$ the second entry can be computed as $x_2^i = \frac{y_{12}^i + y_{21}^i}{2}$ and so on until the procedure ends when $n = N_t$ which gives $x_{N_t}^i = y_{mN}^i$ to produce $\breve{X}(t)$ that represents the reconstructed time series with length N_t (Golyandina, Nekrutkin and Zhigljavsky, 2001; Golyandina and Shlemov, 2015; Mohammadi *et al.*, 2016).

5.4.3   The Applied Constrains in the Grouping Criterion

To separate wind noise components out from the desired signal and extract the signal of interest using the SSA method, there are several constrains have to be considered in the development of the grouping criterion.

5.4.3.1   Eigenvalues rejection in the subspace

In the decomposition stage, the variance of the signal can be represented by each eigenvalue in the direction of the corresponding principle components. Based on the SSA visual tools, the eigenvalues spectra show that the lower-order eigenvalues are related to the more powerful components of the signal. Whereas, the eigenvalues located in the higher subspaces, where the noise components typically arise, are the higher-order eigenvalues that represent the undesired components which are the wind noise. Therefore, in this study the desirable subspace is the lower one. However, this procedure is valid for the selected signals in the dataset for all the conducted case studies and experiments, even though the separability might slightly differ depends on the dataset.

This rule has been set out among the other rules and constrains in the grouping stage which demonstrated by the conducted case studies and the experiments in the next chapters. Therefore, the focus is on separating the wind noise part which is found in the higher subspace from the lower subspace that indicates the lower-order eigenvalues. To remove the noise part, a proper procedure has been established for the grouping criterion as indicated in Equation (5.4), however, the specified threshold can be written using this Equation as follows:

$$\mathcal{L} = min\left\{r : \frac{\sum_{i=1}^{r} \lambda_i}{\sum_{i=1}^{m} \lambda_i} > 0.85\right\} \tag{5.7}$$

where the number of the eigenvalues whose overall energy is 85% of the total energy is defined by *r*.

The principle components associated with the eigenvalues above the elbow point in the spectra which correspond to 85% of the total variance of the signal are omitted. In other words, this percentage indicates that the sum of the eigenvalues equals to the total variance of the original time record. From Equation (5.7), the eigenvalue λ_i are rejected if $i > \mathcal{L}$.

5.4.3.2   The extraction of periodic component

Since the aim is to extract the oscillatory components and leave the wind noise components behind or otherwise separate the noise components from the signal of interest in the noisy signal, the periodicity nature of the oscillatory components has been used to select the best subgroup of principle components for reconstruction a noiseless signal. It is known by now that a pseudo-periodic time record can be factorised into some eigenvalues pairs using the SSA (Vautard, Yiou and Ghil, 1992; Elsner and Tsonis, 2013; Hu *et al.*, 2017). Therefore, using the lower subspace as previously mentioned, only the eigenvalues that appear as pairs will be selected.

Importantly, in case of not obtaining eigenvalues pairs, this simply means that the component does not exist in the given time domain signal as stated in (Mamou and Feleppa, 2007). However, when selecting some eigenvalues, the highest peak of the associated eigenvectors is relevant to the frequency of the periodic components and the power spectrum density can be used to estimate the related frequency (Tao, Lam and Tang, 2001). There are several important points have to be considered when selecting the eigenvalues pairs (Mamou and Feleppa, 2007; Ghaderi, Mohseni and Sanei, 2011; Mohammadi *et al.*, 2016).

- Because of the effect of noise, the possibility of having two equal eigenvalues is quite low. Therefore, the focus is on the nearly equal eigenvalues in each selected pair.
- Selecting pairs of eigenvalues that belong to noise components might happen, however, it is important to be very selective to avoid missing any portion of the signal or adding any portion of noise. To acquire the actual pairs that represent the signal of interest, eigenvalue pairs λ_i and λ_j are selected as a pair only if all the conditions listed below are satisfied:

1. Discarding all the eigenvalues that assumed to be associated with wind noise, that is when i and j are less than $\mathcal{L}$, where $\mathcal{L}$ is the threshold as defined in Equation (5.7).
2. For λ_i , λ_{i+1}, then $|\lambda_{i+1} - \lambda_i| = min \; |\lambda_j - \lambda_{j+1}| \quad \forall \; 1 < i < m$
3. $\left| 1 - \frac{\lambda_i}{\lambda_j} \right| < \mathcal{K}$

According to the amplitude of the waveforms and the oscillatory components of the signal, the value of $\mathcal{K}$ might be changed, however, for each principle component, a specific value of $\mathcal{K}$ can be set (Mohammadi *et al.*, 2016). In this study, it is to focus on wind noise components that have to be separated out from the signal while considering a particular threshold value $\mathcal{K}$ based on the analysis of the eigenvalues spectra for the noisy signal including the desired signal using the SSA visual tools.

5.5 Separability

The representation of an observed series as a sum of interpretable components, which mainly are trends, periodicals or harmonics with different frequencies, and noise along with the separation of such components is always considered as an issue of concern in time series analysis. With a developed SSA, this problem can be solved, trends can be extracted, and harmonic signals can be separated out from noise. Concretely though, and when the SSA is used for noise reduction, it shares the goals of de-nosing and filtering, with two major advantages. First, the SSA allows easier intuitions in the selection of the additive components to be integrated in the reconstruction of the desired signal as these components can be rated according to their eigenvalues. Second, the decomposition of the given noisy signal and reconstruction of the desired signal can be successfully performed even in situations where the sampling rate is relatively low (Del Pozo and Standaert, 2015).

The idea is based on developing suitable grouping technique to the SVD components matrix in order to transform back to time series expansion from the expansion of grouped matrix. The separability of the components of the time series can therefore be defined as the ability of allocating these components from an observed sum when appropriate grouping criterion is applied (Golyandina and Shlemov, 2015; Golyandina and Lomtev, 2016; Hansen and Noguchi, 2017). The SSA decomposition relies on the approximate separability of the different components of the time record (Harmouche *et al.*, 2017; Traore *et al.*, 2017).

For splitting the indices $1 \leq r \leq d$ of the r group of eigenvectors into I groups that is adequate to achieve the separability, r must be clearly specified. In this study, only two groups have been considered; one associated with the signal and the other associated with wind noise. In this case, group $I = \{1, \dots, r\}$ with the related elementary matrices, which will be $\mathbf{Y}_I = \mathbf{Y}_1 + \mathbf{Y}_2 + \cdots + \mathbf{Y}_r$, are associated with the first group that represents the signal. The second group $\overline{I} = \{r + 1, \dots, d\}$ and the related elementary matrices $\mathbf{Y}_{\overline{I}} = \mathbf{Y}_{r+1} + \mathbf{Y}_{r+2} + \cdots + \mathbf{Y}_d$, represent

the noise. In other words, the rest of vectors $(d - r)$ is not of interest as they represent noise (Hassani, 2007; Chu, Lin and Wang, 2013; Golyandina and Lomtev, 2016; Mahmoudvand, Konstantinides and Rodrigues, 2017).

A quantity known as weighted correlation or simply (*w*-correlation) and defined as a natural measure of the dependence between the reconstructed components can be used to achieve the separability. Well-separated reconstructed components are the ones that have zero *w*-correlation. Whereas, reconstructed components with large values of *w*-correlation should be considered as one group as this corresponds to the same component in the SSA decomposition (Harmouche *et al.*, 2017; Rodrigues and Mahmoudvand, 2017; Xu, Zhao and Lin, 2017).

The plot of the singular spectra, which shows the eigenvalues $\lambda_1, \dots, \lambda_d$, can give an overall observation of the eigenvalues to decide where to truncate the summation of the additive components in Equation (4.12) for building a good approximation of the original matrix. Notably, similar values of the eigenvalues $\lambda's$ can give an identification of the eigentriples which correspond to the same harmonic component of the time series. Furthermore, using periodogram analysis of the original time series can also lead to select the groups. As mentioned previously, the higher subspaces of the singular spectra typically show a slowly decreasing sequence of eigenvalues and mainly related to the noise component (Hassani, 2007).

It is worth mentioning that in the higher subspaces where higher-order eigenvalues are located, the energy represented along the corresponding eigenvectors is low. Consequently higher-order eigenvalues are commonly considered to be noise. However, it is possible to remove the unwanted source from the original signal if the data can be transformed back into the original observation space using matrix manipulation in the SSA-based techniques. If a truncated SVD of the trajectory matrix $\mathbf{Y}$ is performed (i.e. when only the significant eigenvectors are retained), then the columns of the resultant matrix $\widetilde{\mathbf{Y}}_I$ are the noise-reduced signal (Chu, Lin and Wang, 2013; Golyandina and Shlemov, 2015).

5.6 Example for the justification of the developed SSA Method

I. Description and Procedure

Before moving to the experimental phases which started with a simulation phase using different signals and noise including wind noise, a working example has been selected for the

justification of the developed method. This example has been selected to frame the main concepts in developing SSA algorithms. The development process of SSA algorithms went gradually based on the requirements of the testing, verification and validation phases. In the selected example, a simplistic grouping and reconstruction techniques have been applied using a short time record to simplify the way of handling the produced matrices as of smaller- sized. However, it is an example with the dual purpose of justifying the method and showing the nature of the produced matrices using numerical values.

Many different testing experiments have been carried out in this study regarding the verification of the developed system using different signals in the first experimental phase. However, some selected experiments from the first phase will be reported in Chapter 7, whereas the second experimental phase, which is more sophisticated when realistic samples of real-world sounds for the system validation have been used, will be described in Chapter 8. Therefore, in this section and to justify the developed method, many points regarding the main aspects of the SSA technique will be illustrated with this example.

In principle, it is to work on the theory that the SSA is a statistical and mathematical approach based on elementary linear algebra following the developed SSA algorithm steps as explained in the previous chapters. Developing an algorithm that illustrates the principle of time series reconstruction using the SSA and includes many important aspects in the method with a given time series is the main objective towards achieving the following aims of giving this example.

- To frame the important concepts of the SSA technique including matrix construction and manipulation.
- To introduce the SSA approach in a worked example from the experiments with a given time series explained it in a step by step manner.
- To demonstrate the SSA capabilities in recovering the original time series.

Consider a uniformly sampled time series $X(t)$, $t = 1,2,3 \dots N_t$ in which the data points are sampled over time at given regular time intervals. Since the number of data points or the length of the time series is denoted by N_t, therefore from the time series given in this example, it can be clearly seen that there are 20 data points ($N_t = 20$) and $X(t) = x_1, x_2 \dots \dots \dots x_{20}$ as graphed in Figure 5.1.

The algorithm has been implemented using MATLAB platform, however, the purpose is to recover the time series by selecting the dominant *PCs*. Such selection, which will be

explained later, corresponds to the dominant eigenvalues based on its associated eigenvectors as a procedure of grouping in order to construct matrix **Z**. The main purpose of this example is to apply the principle components approach based on their selection to reconstruct the original time series shown in Figure 5.1. The time series in this example is sampled from a sinus function with a defined period and with Gaussian noise added. This time series example has been used for different explanations in Chapter 3 in which a justification of its selection has been provided. The sampled signal is the original time series with a specified length indicated through the data points shown in Figure 5.1. Also, this experiment is aimed at studying the effect of the selection of the principle components on the reconstruction process.
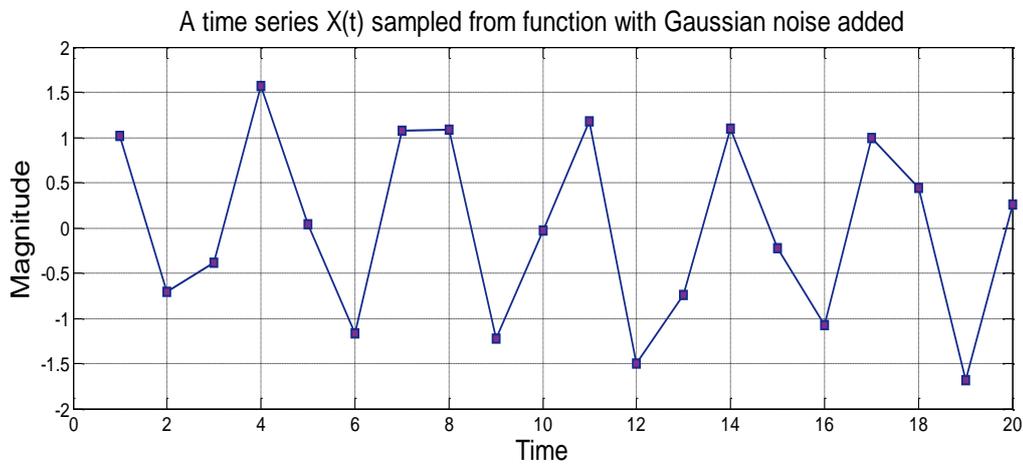
Figure 5.1. An example of a time series sampled with 20 data points

In this example, the grouping and reconstruction techniques are based on the proper selection of the principle components after decomposing the one-dimensional time series into multi-dimensional time series. However, the concepts explained earlier are quite important in the SSA method for reconstructing the time series and returning to the original coordinate system that representing the reconstructed one-dimensional time series in the time domain.

In this demonstration and while comparing the reconstructed components with the original time series $X(t)$, it is assumed that the length of the time series presented in the Figure 5.1 is $N_t = 20$ and the window length $m = 4$. To provide a considerable justification, the main steps of the developed SSA algorithm in this experiment can be outlined in the following points with obvious explanations.

1. **Trajectory Matrix Construction**

Since $N_t = 20$, $m = 4$, the trajectory matrix **Y** can be constructed considering a lag (or "delay shift") of 1. This delay shift is provided as $k = 0,1, ..., m - 1$. The first column of **Y**

is a column vector denoted by $\mathbf{y}_0$ and represents the original time series $X(t)$ at lag $k = 0$ as shown in Figure 5.1. Using numerical values from the experiment, this vector can be illustrated as follows:

$\mathbf{y}_o^T = $ [1.0135518 -0.7113242 -0.3906069 1.565203 0.0439317 -1.1656093 1.0701692 1.0825379 -1.2239744 -0.0321446 1.1815997 -1.4969448 -0.7455299 1.0973884 -0.2188716 -1.0719573 0.9922009 0.4374216 -1.6880219 0.2609807]. The constructed trajectory matrix can be presented in the following form.

$$\mathbf{Y} = \begin{bmatrix} k=0 & k=1 & k=2 & k=3 \\ x_1 & x_2 & x_3 & x_4 \\ x_2 & x_3 & x_4 & x_5 \\ x_3 & x_4 & x_5 & x_6 \\ \vdots & \vdots & \vdots & \vdots \\ x_{17} & x_{18} & x_{19} & x_{20} \\ x_{18} & x_{19} & x_{20} & 0 \\ x_{19} & x_{20} & 0 & 0 \\ x_{20} & 0 & 0 & 0 \end{bmatrix}$$

2. Covariance Matrix Computation

Since the number of embedding vectors $N = N_t - m + 1$, then $N = 20 - 4 + 1 = 17$. The covariance matrix $\mathbf{C}$ can be computed as $\mathbf{C} = \frac{1}{N} \mathbf{Y}^T \mathbf{Y}$. This diagonal constant matrix is a square matrix and symmetric about its leading diagonal (top left to bottom right) which is known as a line of symmetry or a mirror line. Therefore, matrix $\mathbf{C}$ is of Toeplitz structure in which each descending diagonal from left to right is constant as shown in Figure 5.2.
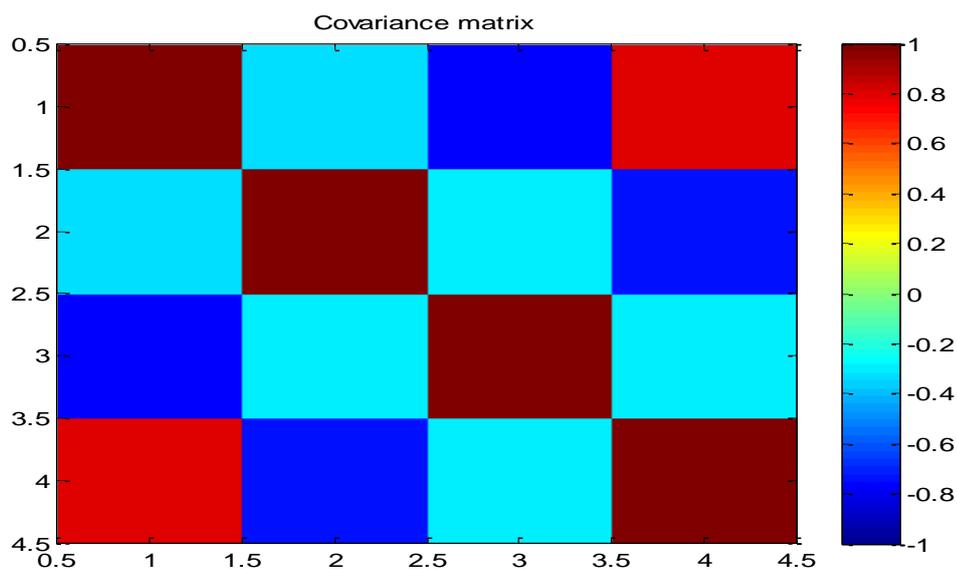
Figure 5.2. The spectrogram of matrix $\mathbf{C}$ shows the line of symmetry (leading diagonal)

The covariance matrix is a square matrix of size $m \times m$ or 4×4 (in this example) as $\mathbf{Y}^T$ is of dimension 4×20. From the experiment, matrix $\mathbf{C}$ has been obtained as:

$$\mathbf{C} = \begin{bmatrix} 1.1765 & -0.3373 & -0.7701 & 0.8086 \\ -0.3373 & 1.1160 & -0.2949 & -0.7468 \\ -0.7701 & -0.2949 & 1.0863 & -0.3113 \\ 0.8086 & -0.7468 & -0.3113 & 1.0773 \end{bmatrix}$$

3. SVD for Eigenvalues and Eigenvectors Computation

the eigenvalues and their associated eigenvectors of the covariance matrix can be computed using the MATLAB function [P, D] = eig (C), or [RHO, Lambda] = eig (C). Matrix $\mathbf{\Lambda}$ is the eigenvalues matrix in which the eigenvalues are along the main diagonal and zero elsewhere and it is a square matrix of size $m \times m$ or 4×4 (in this example). For each eigenvalue there is a corresponding eigenvector represented in a column of matrix $\mathbf{P}$, therefore $\mathbf{P}$ is of the same dimension as $\mathbf{\Lambda}$. These two matrices are shown as follows:

$$\mathbf{P} = \begin{bmatrix} 0.6218 & -0.2325 & 0.6618 & 0.3483 \\ -0.3654 & -0.6759 & -0.2119 & 0.6039 \\ -0.3675 & 0.6578 & 0.2583 & 0.6045 \\ 0.5871 & 0.2373 & -0.6711 & 0.3854 \end{bmatrix}$$

$$\mathbf{\Lambda} = \begin{bmatrix} 2.5933 & 0 & 0 & 0 \\ 0 & 1.5492 & 0 & 0 \\ 0 & 0 & 0.1639 & 0 \\ 0 & 0 & 0 & 0.1496 \end{bmatrix}$$

The eigenvalues spectra depicted in Figure 5.3 illustrates only 4 values in this example because the window length has been selected as 4. However, Figure 5.3 typically shows the eigenvalues spectra with higher and lower subspaces.
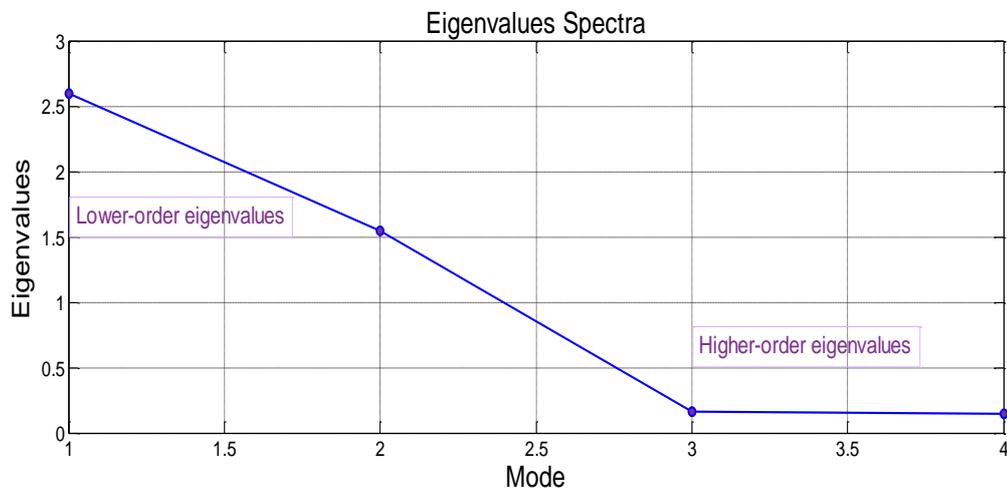
Figure 5.3. Eigen-mode indicate the subspaces in the eigenvalue's spectra

The variance of the signal can be represented by each eigenvalue in the direction of the corresponding principle components. The associated eigenvectors of these eigenvalues are shown in Figure 5.4. The eigenvectors represent the axes of projection and can describe trend and phase. As explained earlier in this chapter and in the previous two chapters that eigenvalues spectra are typically divided into higher and lower subspace. The lower-order eigenvalues are mostly the dominant ones and represent the significant oscillations, whereas the eigenvalues located in the higher subspace are generally represent noise components. Although pairs of eigenvalues cannot be clearly seen in the spectra due to the short time series and small window length selected in this example, however, the lower-order eigenvalues can be used for reconstructing the time series.
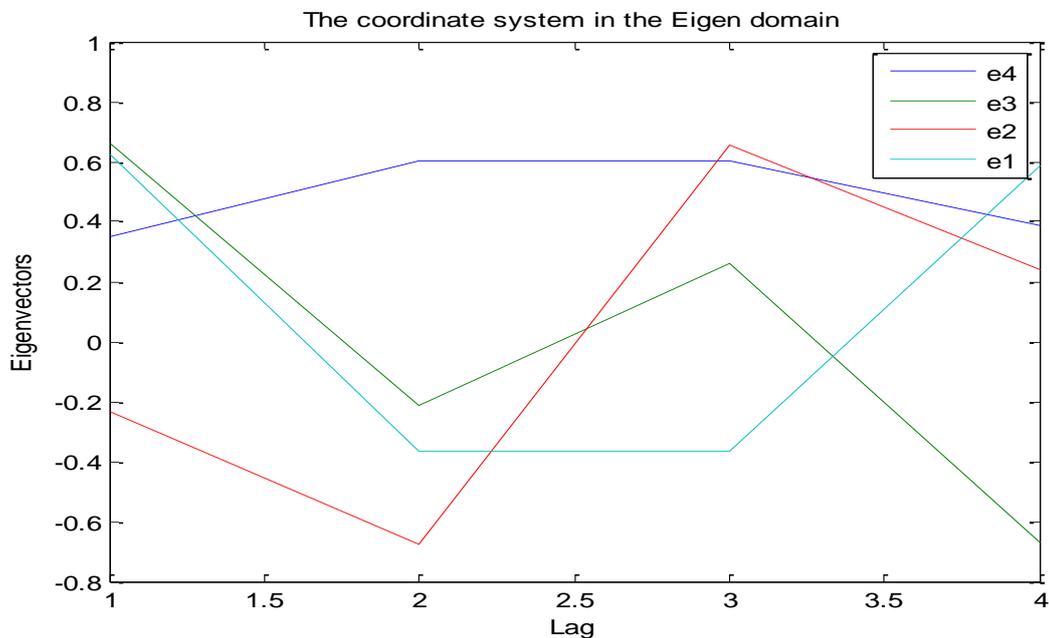
Figure 5.4. Eigenvectors graphical representation

4. Principle Components Computation

The *PCs* can be computed with MATLAB by projecting the trajectory matrix onto the eigenvectors using the operation $\mathbf{V} = \mathbf{Y} \times \mathbf{P}$. This operation yields the principle components matrix $\mathbf{V}$ of the same dimension as $\mathbf{Y}$. The principle components ordered in the same way as the eigenvectors and resemble the embedded time series introduced in the trajectory matrix; however, they are presented in a different coordinate system. Therefore, the columns of the principle components matrix do not correspond to different time lags when compared to the trajectory matrix. In other words, the original values of the trajectory matrix have been projected in a new coordinate system for gathering the variance in the *PCs*.

As mentioned previously, the principle components are aligned in a proper sequence in their matrix in which the first column is the *PC1*; the second is *PC2*, and so on. The graphical representation of the principle components is illustrated in Figure 5.5. Notably, this is not a time domain representation, however, the principle components can indicate dominance and can also give a clear observation about the reconstruction.
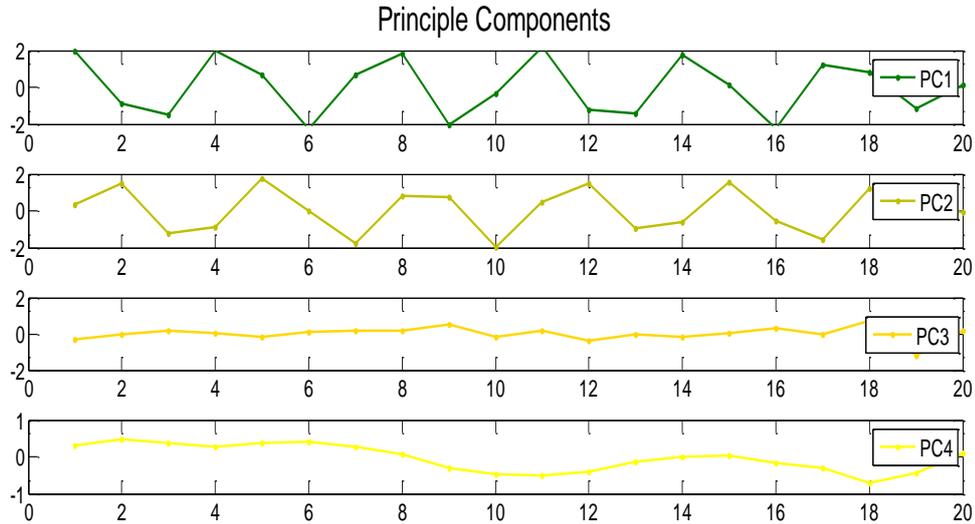
Figure 5.5. Principle components graphical representation

5. Reconstructed Components Computation

This step can be performed by firstly constructing a matrix $\mathbf{Z}$ in the grouping step. This matrix is constructed based on the selection of the principle components aligned in matrix $\mathbf{V}$ when considering the principle components that correspond to the eigenvectors associated to the lower-order eigenvalues as explained above. Matrix $\mathbf{Z}$ is similar to $\mathbf{Y}$ and has the same size which is 20×4 in this example as the number of the entries in each column is equal to N_t. The reconstructed components *RCs* can be computed by projecting the *PCs* presented in matrix $\mathbf{Z}$ onto the eigenvector's matrix $\mathbf{P}$ when the normalisation is considered $\mathbf{RC} = \frac{1}{m}\mathbf{Z} \times \mathbf{P}$. In doing this, the comparison between the reconstructed series and the original one is possible.

Figure 5.6 shows the reconstructed components where *RC1* corresponds to the eigenvector associated with the lower-order eigenvalues, whereas *RC4* corresponds to the eigenvector associated with the higher-order eigenvalue applied for a group related to the first principle component. The reconstructed components indicate how to build the reconstruction and the dominance in terms of retaining accurate signal energy in the reconstruction. The *RCs* that correspond to the first eigentriples are commonly defining oscillatory signal due to phase

quadrature. Figure 5.6 also illustrates that the first two *RCs* are in phase and practically contain all variance of the time series which can already be known based on the eigenvalues. *RC4* seems to describe a trend in the data.
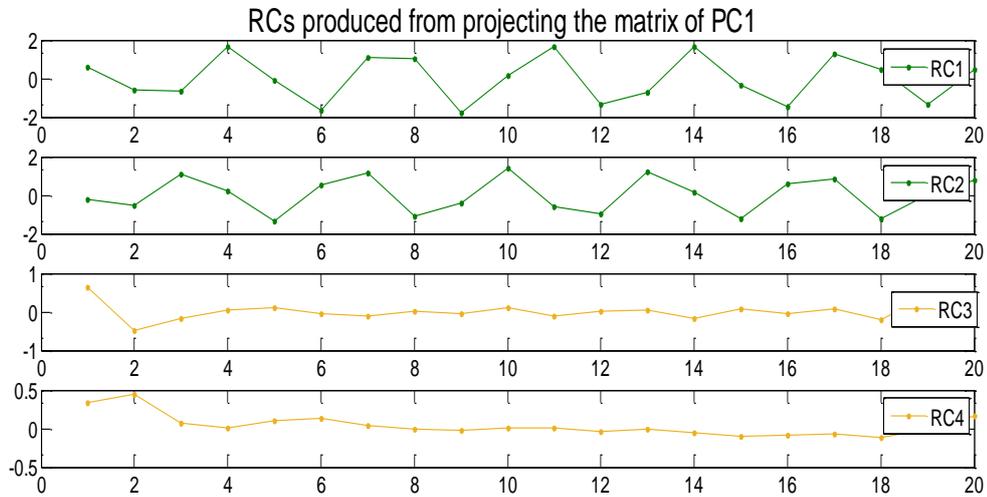
Figure 5.6. A graphical representation of the reconstructed components for *PC1*

II. Discussion and Conclusion

When considering the group related to the last principle component *PC4* for example, the reconstruction will be completely different from the original time series as illustrated in Figure 5.7, and hence no reconstruction of the time series can be obtained. This eventually means that when considering groups related to selecting the principle components, not all of these principle components are dominant and consequently only few of the reconstructed components are required.
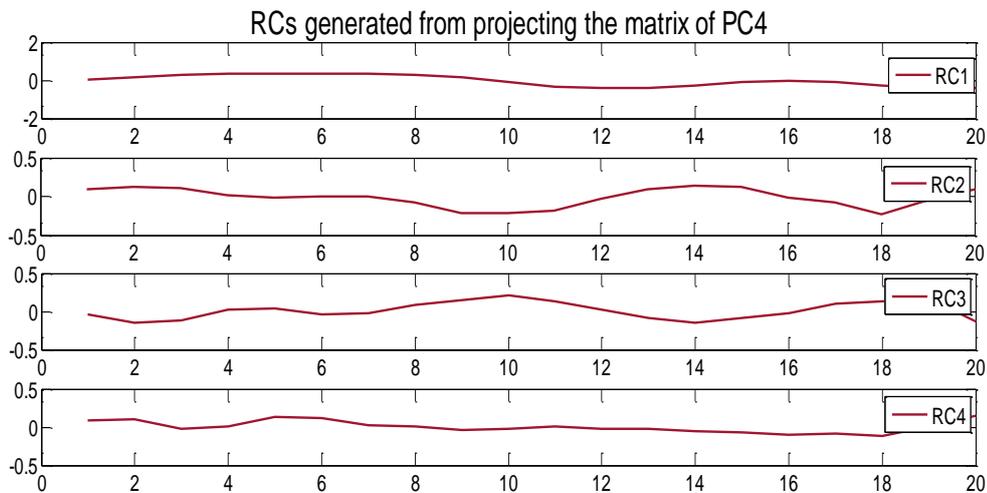
Figure 5.7. A graphical representation of the reconstructed components for *PC4*

This result becomes very clear when looking at Figure 5.8 which illustrates a comparison between the reconstructed components and the original time series including the last two reconstructed components. Even though the selected *RCs* have all been produced from projecting the matrix of the selected principle component *PC1* only onto the eigenvector's matrix, it is important to consider the axis of projection. Therefore, the adaptive selection approach of eigentriples has been developed for the best reconstruction of the desired signal.
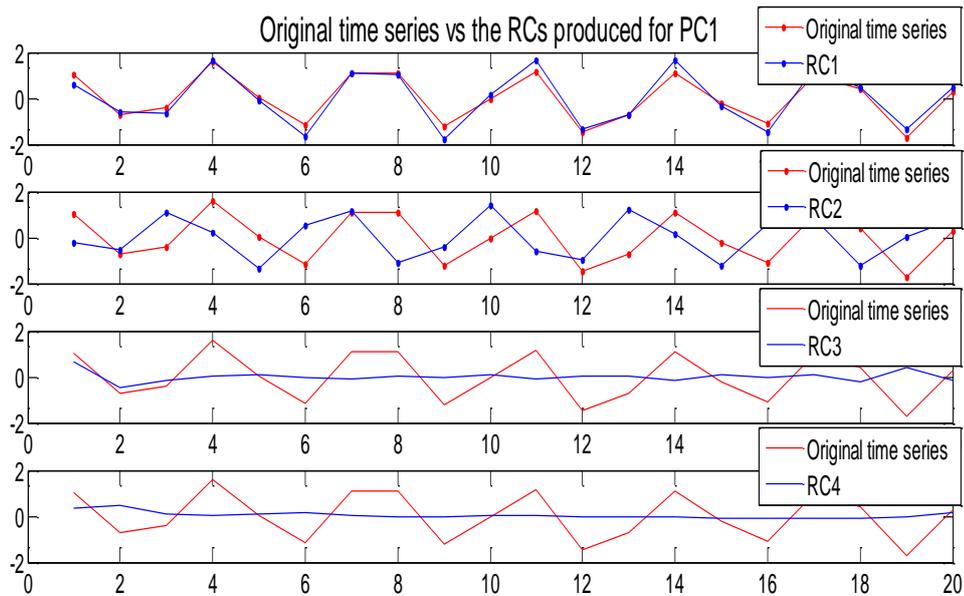
Figure 5.8. Comparison between the *RCs* for *PC1* and the original time series

The comparison between the reconstructed time series based on the first principle component and the original time series is illustrated in Figure 5.9.
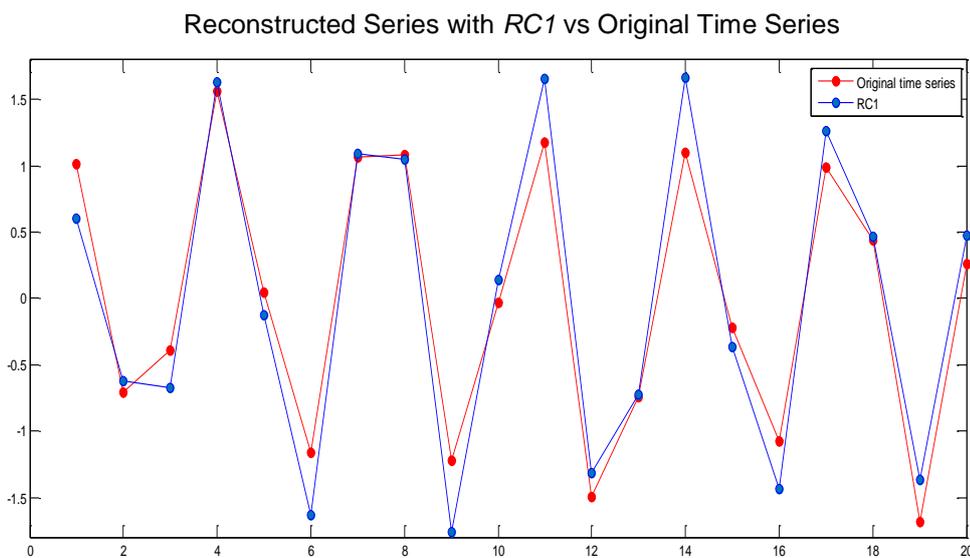
Figure 5.9. Comparison between the *RC1* for *PC1* and the original time series

145

Analysing the results demonstrated in Figure 6.9 can approve the effectiveness of developed SSA particularly with regards to the selection of the principle components and the *RCs* generated based on this selection. In fact, it is a satisfactory outcome and very encouraging result that can give a considerable justification as the reconstruction of the original time record can be performed with the selection of only the dominant *PCs* and few numbers of the reconstructed components. Some important points can be outlined from the previous explanation and the results reported from this experiment as follows:

- Short records can make differences to the results, affect the ordering of the eigenvalues in the spectra, and affect the estimated matrix.

- The principal components represent a projection in a different coordinate system as their interpretation is different from that of $X(t)$. Therefore, they cannot be compared to the time series $X(t)$.

- Each *RC* re-translated its corresponding principal component into the original units of the time series $X(t)$. Now, the comparison between the reconstructions (based on the *PCs*) and $X(t)$ is possible.

- To reflect oscillatory modes of interest, the original time series can be filtered through a convolution when selecting a small number of *PCs* and their associated eigenvectors.

- The produced *RCs* can be used to filter the time series by using less of the total number of *RCs*. The part of the time series which supposed to be noise is mainly represented in the higher subspace of the eigenvalues spectra and can be separated out. In other words, the part with higher-order eigenvalues is the one that supposed to be noise. The first two *RCs*, however, are generally defining an oscillatory signal due to the phase quadrature of the corresponding *PCs* and their eigenvalues.

A predefined time series has been used, however, from this example, it has been found that using *RC1*, the defined time series can be reconstructed as in Figure 5.9 and this satisfies the SSA concepts for both its two complementary stages; decomposition and reconstruction. In this example, it can be obviously seen that the first two eigenvalues are meaningfully higher than the others and take more of the energy of the given signal, while the others in the higher subspace represent the noise floor. Consequently, the corresponding *PCs* will be significant for reconstructing the series as they provide glimpses of the periodic components, however, the others represent random components. Importantly, the first two *PCs* contain all the variance of the given time series which is consistent with the first two eigenvalues.

As explained above, the pre-calculated *PCs* can be used to reconstruct each elementary matrix by projecting the principle components matrix back onto the eigenvector coordinates. The grouping criterion applied in this example depends on the singular spectra which have been divided into two subspaces along with principle components approach. It is worth noting that the grouping and reconstruction are key elements in the SSA-based techniques. Therefore, they have systematically developed and discussed in this chapter where the aim is to improve the separability. However, the development of the SSA algorithms in terms of these two concepts has passed through different steps during the time slot of conducting this research based on the aims and requirements of the testing experiments and the planned experimental phases.

This example provides a clear illustration of the usefulness of the SSA method for reconstructing the decomposed components based on the eigenvalue's spectra and the principle components and omitting the undesired ones. It also gives a practical demonstration of the capability of the SSA method towards further improving the separability for omitting the undesired components. From this example, it can be concluded that the anticipated subspace belongs to the desired components and the subspace corresponds to the undesired component, which is the additive noise, can be defined. The main purpose of this example was to practically introduce the SSA main concepts and perform a systematic manipulation of the produced matrices in a working practical example rather than giving strong focus on noise separation approach which will be experimentally illustrated in the next chapters.

6   CHAPTER SIX   DATASET AND PREPARATION

6

Dataset and Preparation

6.1 Overview

In the previous chapters, a systematic approach has been developed to use the SSA for wind noise removal and separation that might guide to success and drive to creating academic knowledge and claiming certain novelty and uniqueness in this field. This systematic approach has been developed in many ways including the mathematical equations that have been written in new and specific forms. However, that was mainly based on reviewing the literature in which it has been found that the SSA has never been implemented in the past for such specific purpose. This might lead to consider the developed and modified SSA version in this study as an alternative method to many existing wind noise methodologies.

After justifying the developed method through giving some working examples using results from the performed experiments in this thesis, it is now to adopt this systematic approach for more testing. Therefore, the SSA has been explored based on this approach using typical testing signals. Several case studies including wind noise have been conducted as a first experimental phase which is related to the verification of the developed system. In the second phase, it is to adopt this developed system of the SSA in a validation phase with a complete selected dataset of real-world sounds that includes several interesting environmental and smart cities sounds as signals of interest along with wind noise samples from real measurements.

The long-term objective is to pave the way for applying acoustic sensing in outdoor data acquisition applications for monitoring smart city soundscapes by improving further development to best tailor these methods for practical uses. In addition, exploring smart cities as environments of open and user-driven and studying the influence of applying acoustic sensing in such urban environments can also be seen as a long-term aim which might be achieved in further research work based on the work conduced in this study.

This chapter will describe and outline the basics of the selected dataset and review the benchmark and Freesound recordings and freefield1010 dataset samples that have been selected for this study. Furthermore, this chapter also covers some aspects of the experimental procedure for this research and points out certain important concepts related to the criteria required for testing, comparing, and analysing results. Also, the adopted design, which is needed to meet the research requirements to achieve the main aim of this study, will be

described in this chapter as well. A model of generating mixed soundtracks from these datasets using different mixing ratios is also presented. Indeed, this chapter is an effective presentation to provide an obvious explanation of the audio samples and mixed soundtracks that are used for the verification, validation, and evaluation of the developed SSA system in the subsequent chapters.

6.2 Data Collection and Dataset Preparation

Since cleaning up the microphone signal from wind noise as a particular type of environmental noises to enable further analysis is the main aim, hence, for experimental investigation purposes, the whole procedure is divided into several stages. The first stage is to generate noisy signals that are required for exploring the method, testing the algorithms in the system verification phase. The second is further developing the SSA algorithms in many ways by performing a large number of experiments and conducting several case studies with real-world sounds.

Due to the difficulties of doing urban soundscape synthesis to get field recordings, existing datasets will be used. However, realistic samples which have been collected for different published datasets and urban soundscape synthesis for some internet-based dataset sources such as Freesound recordings and freefield1010 dataset will be valuable for testing purposes and performing different experiments as well as for validation and evaluation purposes.

6.2.1   Concepts for Dataset Preparation

During the data collection and sample's selection stage, it is always important to be very selective, particularly with regards to the sample type and size that should be representative. Also, it is important to consider some factors that may affect how well the selected samples can precisely reflect the specific aspects of concern and consequently how to draw valid and reliable conclusions. Furthermore, a selection that can lead to a conclusion that generalising the aspect of concern relating to the objective of the study.

There are some key concepts that have been considered in the data collection stage such as the required level of precision which in known as the margin of error. According to (Lenth, 2008), the sample size has a clear effect on how precise the estimate is. Basically, the size of the sample dictates the amount of information that can be made available. Hence, in part, it determines the level of confidence and precision that is required in the selected samples.

As stated in (Sarah, 2014), other concepts such as the power and the effect size are considered important in data collection stage. The power is a concept used to show evidences of any noticed differences between the groups (when there is a variety in the dataset given that there are different categories where the data falls in) and this requires larger sample size. The effect size can be expressed in terms of the deviation which is generally required being as smaller and closer to the mean as possible. As a conclusion, the narrower margin of error, the greater power, and the smaller effect size require a larger sample size and representative data. It is essential, however, to consider how the sample is selected to make sure that it is unbiased and representative of the target items. In this research, the above-mentioned concepts have been considered in dataset selection stage regarding wind noise dataset and the multiple signals and audio recording samples used in the different experimental phases.

6.2.2   The Structure of the Dataset

The selection of the samples of audio recordings included in the dataset is based on considering examples of important and interesting smart city sounds that have been indicated significant in many applications. The samples in the dataset are realistic and required for the testing, validation and evaluation of the developed system after exploring the developed method by performing many experiments in the verification phase using typical testing signals. The results of some selected experiments in the verification phase will be shown in the next chapter, whereas the results of system validation will be reported in Chapter 8.

A freefield1010, which is a published dataset of standardised 10 seconds excerpts from Freesound field recording, has been selected. The freefield1010 database collection is adequate since it consists of a large number of tracks including many wanted real-world sounds required for this study such as birdsongs, city, nature, train, etc., (Stowell and Plumbley, 2013; Grill and Schluter, 2017). A benchmark database consists of the required signals of interest such as birds' chirps, alarms, and car sirens along with wind noise sample, has been set up for this study. Accordingly, it is particularly important to carefully preparing the working platform and the required tools to perform the practical work in a methodical and systematic manner.

It took considerable effort to make the samples suitable to the experiments by using automatic methods where possible to ensure that all the samples included in the dataset are pure sounds (e.g., birds' chirps or alarms). In addition, all silent gaps have been removed from the samples. The dataset also includes the samples of wind noise as the samples from this

dataset have been mixed to generate mixed soundtracks. As mentioned previously, the processing time of the SSA depends on the length of the given time series and the selected window length. To reduce the processing time of the SSA, the frame by frame processing method has been applied to process the soundtrack samples instead of processing the whole audio file directly. The average method has been used, however, an average 100 ms frame size has been selected to be a representative sample from five thousand frames in the dataset of audio recordings of the desired signals.

The data is divided into two main categories depending on the requirements of the experiments which are wanted signals sample and wind noise sample. Moreover, multiple recordings that contain single pure sounds (e.g., alarms) have also been considered in the dataset which is the case of some possible city recordings as such audio recordings will be subjectively easier to identify with regards to the enhancement level of noise reduction. It seems apparently important to have such arrangement as a step toward specific applications for adopting the developed systematic approach and direct it for improving further development. Also, to best tailor this method for practical uses in environmental smart city application and soundscapes monitoring. Different typical testing signals have been used to perform many experiments in the system verification phase for exploring the capabilities of the developed method and testing new functionalities. The distribution of the data sources is based on dividing the samples into signals of interest and wind noise as follows:

I.   Realistic sample of audio recordings

For testing and validation purposes, different audio recordings of real-world sounds such as real birds' chirps, alarms, and car sirens have been used as signals of interest.

II.   Wind noise sample

As explained, real measured wind noise samples have been used to be mixed with the real audio samples of desired signals to produce mixed soundtracks.

The main purpose of organising the dataset in this way is to make the dataset easy to handle when manipulated and to help in reporting, analysing and discussing the results. It is important, at first sight, to bear in mind that the main aim of the research is to develop noise reduction algorithms based on the singular spectrum analysis to improve separability when using real-world sounds of audio recordings from the city in the presence of wind noise. At this stage, the main concern is to decide to what extent the developed system is a valid one for wind noise separation regardless for what application the cleaned signals will be used for.

6.2.3   Descriptions and Specifications of the Dataset

Clearly specifying the path of the research could always help in selecting the required tools and specifying the dataset as well as indicating the long-term objective of the study. Therefore, the first priority was to think of something manageable and doable and at the same time should be justified for the academic. Wanted signals and wind noise samples were required to establish a benchmark database for the testing and validation of the developed SSA particularly for wind noise separation and removal. Using such published and standard datasets is highly beneficial as to accurately report and compare the results even with results from other related work. The sample selection from such datasets might be seen as a careful and right choice because they are popular and standard databases used, validated and cited in many research works. Freefield1010 dataset is an open dataset for research on audio field recording archives (Stowell and Plumbley, 2013; Grill and Schluter, 2017).

In the early stages of conducting this research and in many experiments, the Audio and Acoustics Signal Processing challenge (AASP) dataset was used. This dataset, which is published by IEEE and was collected in many different places in the London area, provide one set for soundscapes classification and includes many event types such as alert (beep sound), speech, etc. For the later stages of this study, however, after discussing the results obtained in the first phase, expanding the dataset or using different signals of interest was essential to prove or disprove the developed SSA system in the final phase which is the validation.

The size of the samples taken for conducting the multiple experimental investigations phases is big enough to satisfy the testing and validation requirements. Audio recording files with their metadata including descriptions and contents are saved in JSON format (Java Script Object Notation) in the original dataset. Table 6.1 below shows the specifications of the samples as standardised file format selected for this research.

Table 6.1.   The specifications of the samples

Audio file format	Duration (s)	Sampling Rate samples/second (Hz)	Total Samples	Bit depth /resolution) Bits/Sample	Channels
WAV	10	44100.0	441000	16-bit PCM	Single

All audio files in the dataset need to be passed through the mixing stage to mix their contents with wind noise to produce noisy signals. However, it is to work out with the average for reporting the results after producing noise reduced audio files which represent the output of the de-noising process of the SSA algorithms.

6.2.4 Mixing Recorded Sounds with Wind Noise

The wanted signals/wind noise mixing strategy requires an audio mixer which is an algorithm developed in MATLAB to mix the pre-recorded samples. Importantly, to avoid misinterpretation, it is to consider adding the two signals in a correct proportion. Therefore, the normalisation has been addressed in the algorithm as a default method to ensure the reliability that the mixed signals are to the same perceived level. The main goal of the mixing stage is to generate soundtrack benchmark database.

A variety of environmental sounds have been used in the experimental investigation and testing phases. The scenario of mixing multiple signals recorded from the environment that contain desired sounds as signals of interest with the under testing unwanted wind noise signals was based on selecting a desired SNR ratio in a reasonable range (from -20 dB to 20 dB). The SNR was mainly the variable that has been changed to produce a number of soundtracks ready for testing purposes in the proposed dataset. It is also considered as an objective measure to evaluate noise reduction methods and compare results before and after de-noising.

Selecting multiple predetermined mixing ratios is for showing the behaviour and assessing the performance of the developed SSA noise reduction algorithms under different conditions. Meanwhile, this procedure helps in examining the features of the signals as the content of the noise changes in a proportional manner with the content of the signal to optimise the noise reduction algorithms of the developed system.

6.3 Algorithm, Testing and Implementation

Figure 6.1 is a flowchart for the implementation of the developed method in the experimental investigation's phases. Figure 6.1 shows the sequence of performing the experiments starting from defining the input to the analysis and evaluation. The input differs from one experimental phase to another. In the system verification phase, typical testing signals have been used before using different recordings containing multiple real-world sounds recorded from the city in the system validation phase as shown in the explanation of the dataset. The second step is the mixing stage using a mixing model to produce noisy signals by mixing signals together or the desired signals with wind noise. Passing the signals through the mixing stage is for the simulation purposes which can be achieved by specific MATLAB codes for mixing signals at desired SNR. The output of this stage is noisy signals that required to be tested by the noise separation and reconstruction technique developed for this research.
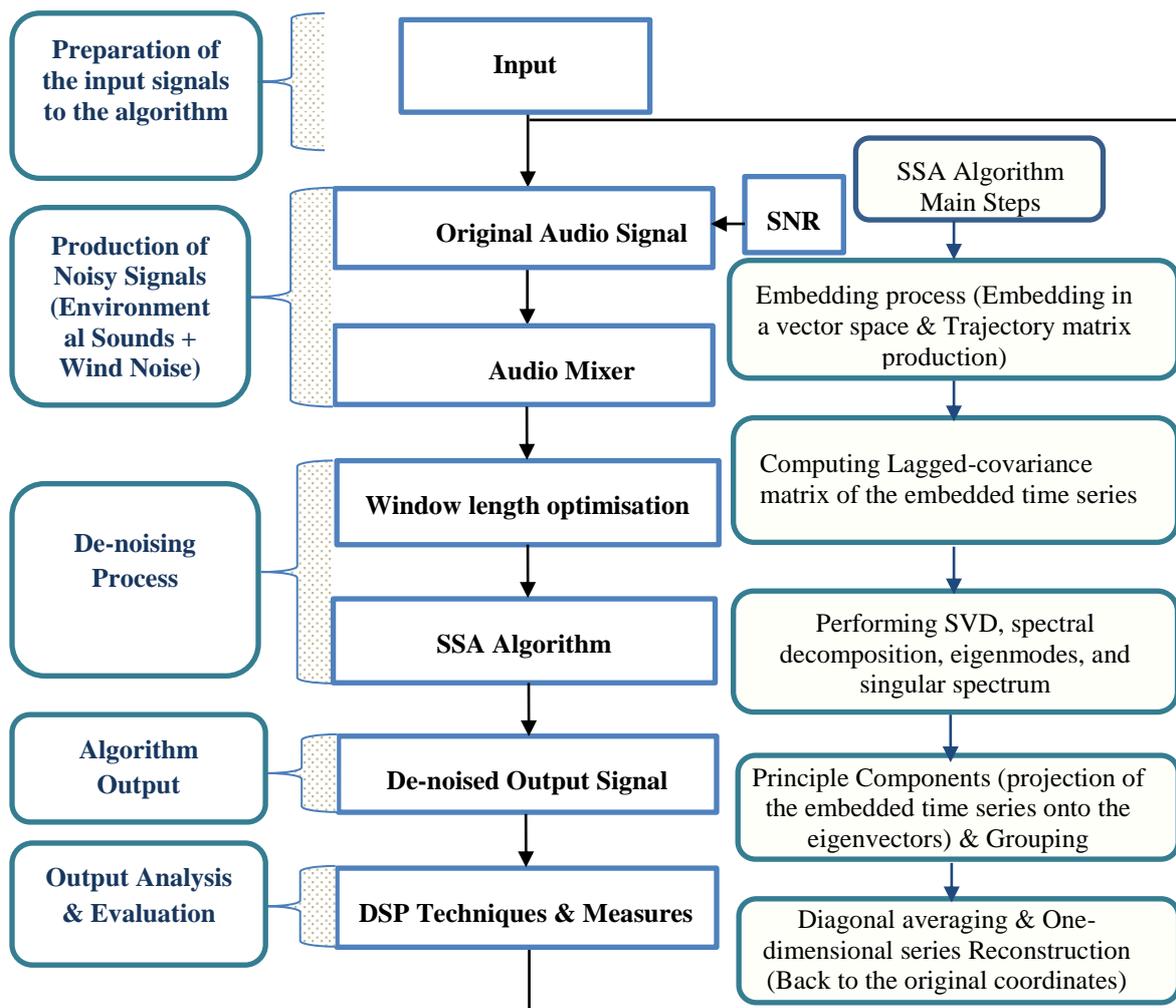
Figure 6.1. A flowchart of the experimental investigation phases

After obtaining such noisy signals, however, the next step is to implement the developed SSA algorithms for wind noise reduction and separation. It is important to consider the optimisation of the sole parameter of the embedding process in the SSA algorithms because of its vital role in the decomposition stage. Also, developing and applying an appropriate grouping criterion is a key element for testing and validation purposes to produce the output which is supposed to be noise reduced signals. Furthermore, based on the developed grouping and reconstruction techniques, desired signals can be reconstructed, and further improvement of the separability can be achieved.

The de-noised signals are the resultant signals and the final output of the SSA wind noise reduction algorithms of this experimental procedure. To avoid complexity, the SSA noise reduction algorithms do not include a complete set of objective measures or any analytical procedure. Therefore, the output of the SSA algorithms needs to be manipulated by other tools

155

and codes to report and compare the obtained results after indicating suitable evaluation measures. All SSA algorithms development and implementation, data analysis, and producing figures have been performed in MATLAB.

6.4 Analysis, Validation, and Evaluation Methods

There are different ways to evaluate the output of a noise reduction algorithm such as objective evaluations, subjective evaluations and an objective perceptual evaluation. The latter is an automated methodology used to capture the perceptual quality when applying subjective evaluations. Using relevant measures is quite important to evaluate how a method performs. Signal-to-Noise ratio is one of the objective evaluation measures that often considered as an appropriate measure of how the applied methods perform. However, rarely has been the perceptual comprehension of the output of an applied technique that needs to be evaluated subjected to an objective evaluation.

Numerous of different subjective tests are available and have been used in some research work. Using subjective test evaluation method is also considered as a good way to grade noise reduction algorithms and evaluate the sound output, which is planned to be used by real people, however, it is very time consuming. Therefore, for a noise reduction algorithm optimisation, subjective test evaluation is not highly recommended. Alternatively, some objective evaluation measures, which are known as objective perceptual evaluation, that give predictions closed to those of a subjective listening test are indicated as good measures (Rohdenburg, Hohmann and Kollmeier, 2005; Ding, Lee and Soon, 2012). It is not necessary to use all the evaluation methods at a time to evaluate a noise reduction algorithm; however, it could be sufficient to use some of them. The subjective test evaluations are time consuming, instead, the most commonly used evaluation measures are objective measures. In this research, however, only objective evaluation methods have been used.

After experimentally verifying and testing the developed SSA algorithms and reporting related results, it is significant to validate the developed system and applying comparison and evaluation approaches to ensure how efficient it is. Verification is an activity used to address whether such algorithms appropriately reflect the specified requirements. Whereas, validation is the process of evaluating a product, method or an algorithm to determine whether it satisfies specified requirements. It is also an activity used to demonstrate whether a method or an algorithm can fulfil its intended use. However, similar approaches such as testing, analysis, inspection, demonstration, and simulation are used for validation activities. Validation

activities can sometimes be concurrently run with the verification and portions of the same environment (Maropoulos and Ceglarek, 2010; Sarode and Deshmukh, 2011).

Validation activities can be applied to different aspects of a method being developed in any of its intended environments, such as training and pre-integration of integration phases. Generally, to accomplish validation, the methods employed for that can be applied with regards to requirements, designs, and prototypes which should be selected based on how well the method or algorithm will satisfy the specifications and the needs. Therefore, the selected environment for validation activities should represent the intended environment suitable with the deployment of the method (Maropoulos and Ceglarek, 2010; Sarode and Deshmukh, 2011). Figure 6.2 illustrates a flowchart of software verification and validation processes applied to the developed method in this research.
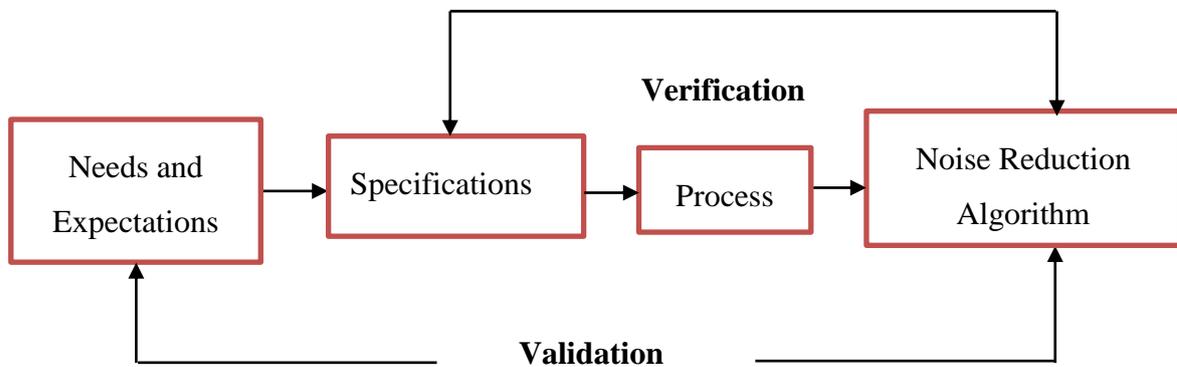
Figure 6.2. Algorithms verification and validation activities

The verification activity is often to evaluate the specifications during the development process. The validation activity is always prior to the final deployment and performed with realistic data. The validation of the results of this study will be performed using the samples of the second dataset as proposed.

6.5 Objective Evaluation Method

Among all the objective measures available for use, few can particularly give a clear indication on the amount of the reduced noise especially if the desired signal of interest is speech (Ding, Lee and Soon, 2012). The amount of the reduced noise along with the signal distortion are two key metrics to evaluate the enhanced perceptual quality of the de-noised signal. It is important therefore to have such balance while applying objective measurement tools to evaluate the capability of the noise reduction algorithm with regards to reducing the noise and at the mean time recovering desired signals with as much less distortion as possible.

For evaluating noise reduction algorithms, the most commonly used objective measure is the Signal-to-Noise measure including its different variations. To evaluate the performance of the developed SSA wind noise reduction algorithms in this research, it is important to select suitable criteria and consider multiple objective measures for analysing and evaluating the results. Each criterion gives useful indications and measures a specific means of the de-noising method by comparing the output to the input. The input represents the signals before passing through the noise reduction algorithm which are the noisy signals, while the output represents the de-noised signals. The original clean signals are also used for comparison purposes. It is worth mentioning that simulated noisy signals have been generated by adding the target environmental noise to wind noise-free audio signals.

6.5.1 Signal-to-Noise Ratio

To assess the effect of noise on a signal, the Signal-to-noise ratio measure is generally used. SNR is defined as the ratio of the power of a signal which is the meaningful information and the power of background noise which is the unwanted signal. The ratio between the signal powers of $s(n)$ to the noise $w(n)$, where w stands for wind noise in this context, defines the SNR which is given in a logarithmic scale in decibels (dB). That is to consider the logarithmic perception of loudness in humans and to capture the wide range of potential SNR values (Villanueva-Luna *et al.*, 2011). The input Signal-to-Noise ratio is defined as:

$$\text{SNR}_i = 10 \log \frac{\sum_n s(n)^2}{\sum_n w(n)^2}, \tag{6.1}$$

where $s(n)$ is the original clean signal and $w(n)$ is the wind noise.

The signal estimate which is given the notation $\hat{s}(n)$ is the output of the noise reduction algorithm. Hence, the output error is given as:

$$s_o(n) = s(n) - \hat{s}(n), \tag{6.2}$$

The output error $s_o(n)$ contains the unfiltered wind noise and the part of the signal that has been removed by estimation errors in the algorithm. The output SNR is given as follows:

$$\text{SNR}_o = 10 \log \frac{\sum_n s(n)^2}{\sum_n s_o(n)^2}, \tag{6.3}$$

The difference between SNR_i and SNR_o is therefore the noise as the desired signal is the same in both cases.

6.5.2   Noise Residual

Applying the same concept and using a similar measure which is known as Noise Residual (NR) can show how the output resembles the wind noise. Noise Residual NR can be computed as in Equation (6.5). Given that:

$$w_o(n) = w(n) - \hat{s}(n), \tag{6.4}$$

where $w(n)$ is the true wind noise and $\hat{s}(n)$ is the output

$$\text{NR}_o = 10 \log \frac{\sum_n w(n)^2}{\sum_n w_o(n)^2}, \tag{6.5}$$

Using SNR along with NR measure can evaluate the noise reduction method. However, a good method is the one with high SNR_o and low NR_o.

6.6 The Measurement of Audio Signals

Along with SNR measure, there are certain measurements that have been suggested to be used such as root mean square values, crest factor, dynamic range, and sound pressure level. Sound analysis methods are sometimes applied and seen as useful data analysis tools. The Root Mean Square RMS is a mathematical way of expressing the amount of energy in a signal and its measurement is time dependent. It is particularly useful as it allows comparing signals in equal terms. RMS calculation for complex waveforms such as audio signals which are made of multiple frequencies is far more complex although finding the peak level is relatively easy. If two signals have the same peak level, it would not necessarily that these signals have the same RMS. In other words, even if two signals have the same peak level, one might sound louder than the other. Depending upon the type of the signal, the RMS level can be estimated by using the crest factor if the peak level is known (Ballou, 2015).

- **Crest factor**

Crest factor is defined as a measure of a waveform and it indicates the ratio between the peak value and the effective value. Since it represents the ratio of peak value to the root mean square value, therefore a crest factor 1 such as for direct current indicates no picks; however, higher crest factors indicate peaks such as sound waves. Generally, sound waves tend to have high crest factors in a considerable range between 10dB to 20dB for most typical sounds. Accordingly, crest factor indicates how extreme the peaks are in a waveform. It can be useful in some applications to judge overall perceived loudness especially in loudspeaker testing

standards it is fairly an important issue of the test signals as well as in modulation techniques (Crocker, 2007; Ballou, 2015). It can mathematically be given as the peak amplitude divided by the RMS value as a dimensionless quantity. Therefore, crest factor is defined as a positive real number and sometimes stated as a ratio of two whole numbers.

$$\text{CF} = \frac{|x|_{peak}}{x_{rms}}, \tag{6.6}$$

Crest factor can be also expressed in decibels as in Equation (6.7). However, in this case it is equivalent to peak-to-average power ratio (PAPR) which is mostly used in signal processing applications due to the way decibels for amplitude ratios and power ratios are calculated (Crocker, 2007).

$$\text{CF}_{dB} = 20 \log_{10} \frac{|x|_{peak}}{x_{rms}}, \tag{6.7}$$

The minimum possible crest factor is 1, 1:1 or 0dB, that's for a square wave. Theoretically and even in the real world, other waveforms have a non-zero crest factor since there is always some difference between their peak and RMS values. For instance, a sine wave has a crest factor of 3dB when all peak magnitudes have been normalised to 1. That means the RMS level is 3dB below the peak level. For real life sound the crest factor of human voice is greater than 14dB according to statistical values, which means RMS level is about 14dB below the peak level (Crocker, 2007; Ballou, 2015).

- **Dynamic range**

    Dynamic range and signal-to-noise ratio are closely related concepts. Basically, SNR measures the ratio between an arbitrary signal level and the noise considering that it is not necessarily the most powerful signal possible. Dynamic range measures the ratio between the maximum strongest un-distorted signal and the minimum discernible signal which for most purposes is the noise level. According to this definition, the dynamic range can be calculated as in Equation (6.8) (Crocker, 2007; Ballou, 2015).

$$\text{DR}_{dB} = 20 \log_{10} \frac{x_{max}}{|x|_{min}}, \tag{6.8}$$

The dynamic range often describes the ratio of the amplitude of the loudest possible to the RMS noise amplitude. According to (Huber and Runstein, 2013), the level of a signal can vary broadly from one moment to the next which indicates the dynamic range of the signal. For instance, the variance in an audio signal from an impassioned scream following a soft

whispery passage is a good indication of a jump from the optimum recording level into severe distortion. According to (Kendrick *et al.*, 2015), exerting some form of control over a signal's dynamic range becomes obviously required by using various techniques and dynamic controlling devices. However, this is another issue which is not directly fit in this context. In short, as stated in (Huber and Runstein, 2013), the dynamics of an audio program's signal resides somewhere in a continuously varying realm between three categories. These categories are saturation which occurs when an input signal is so large, average signal level which represents where the overall signal level resides, and the third one is the ambient noise.

- **Sound pressure level**

For environmental noise measurement, the so-called A-weighting curve has been broadly adopted and considered as standard in many sound level meters. This system is used in any measurement of environmental noise such as rail noise, roadway noise and aircraft noise. A-weighting is the most commonly used among the curves defined in the International standard IEC 61672:2003 relating to the measurement of sound pressure level. Low frequency signals are given very low weight while high weight is given to higher frequencies. Wind noise, which is the noise of concern in the context of this research, has several of low-frequency content. Also, A-weighted SPL measurements of noise level are increasingly found for domestic appliances (Kuttruff, 2006).

The A-weighting measures could be useful for comparing the output of a noise reduction method when putting more emphasis on the frequency range where most of the content of the signal of interest exists. For this purpose, sound pressure level SPL measure has been used in some experiments in this study. SPL, which is also known as acoustic pressure level, is a logarithmic measure of the effective pressure of a sound relative to a reference value. It is denoted by L_p and measured in dB and can be defined as in Equation (6.9) (Rienstra and Hirschberg, 2003; Kuttruff, 2006; Crocker, 2007; Müller and Möser, 2012).

$$L_p = 20 \log_{10} \frac{P}{P_o}, \tag{6.9}$$

where P is the root mean square sound pressure and P_o is the reference sound pressure.

The commonly used reference sound pressure in air is $P = 20 \mu Pa$ which is often considered as the threshold of human hearing. Most sound level measurements are made relative to this reference (i.e.1 Pa equals an SPL of 94dB).

6.7 DSP Analysis Techniques and Visual SSA Tools

This section discusses some Digital Signal Processing DSP analysis techniques that are commonly used in sound analysis. It is just a brief introduction to these methods rather than detailed mathematical description. Appling the concepts of time-domain analysis and spectral analysis are quite important to visually present audio signals. Such common DSP techniques are considered useful to demonstrate valuable interpretation while analysing any given signals regarding viewing important aspects such as frequency content, noise presented, silence periods, etc. Not only will they help in visualising sound signals and their related concepts but also observing the changes occurred to the signals during different processes for comparison purposes. For instance, showing the spectrogram can help in interpreting the changes occurred to the signals after the de-noising process when compared to the original signals and noisy signals before applying the noise reduction algorithms as the case for this research. Therefore, MATLAB plots will be introduced and used to show the results such as time and frequency domain analysis as well as spectrogram analysis and visual SSA tools such as the eigenvalues spectra in the Eigen domain.

To do such sound analysis for numerous of signals used in the multiple experiments performed in this study according to the requirements of the experimental investigation phases along with the proper presentation of the results, specific MATLAB algorithms have been developed. The "audio read" function in MATLAB has been used to load the data files from the dataset. Complete codes have been developed for computing the measurements explained in section (6.6) along with plotting certain important figures regarding the analysis techniques introduced in this section. SNR calculation and SPL presentation require separate codes.

Time domain analysis is a mode used for analysing data over time. It gives the behaviour of the signal over time and allows predictions and regression models for the signal. A time domain graph shows how a signal changes over time which therefore indicates time versus amplitude. Frequency domain is a method refers to analysing a signal or a mathematical function with respect to the frequency rather than time. All signals have a frequency domain representation and this method is mostly used to signals or functions that are periodic over time. Transformation is the most important concept in the frequency domain analysis which is used to convert a frequency domain function to a time domain function and vice versa. A frequency-domain graph shows how much of the signal lies within each given frequency band over a range of frequencies indicating frequency versus amplitude (Sovijarvi *et al.*, 2000).

Spectrogram is a visual representation of sound or any signal. It shows the amplitude of the frequency components of the signal over time. Compared to another visual representation of sound known as the waveform, spectrogram is a record or graphic representation that indicates the variation with time of the resonance of the sound or series of sounds. The spectrogram reflects the change in the frequency in the signal and is useful with complex signals especially those that contain more than one frequency component. Frequency components of the signal cannot be discerned from a visual examination of a waveform, but they are obvious when using spectrogram. It is produced by a sound spectrograph and based on the mathematical algorithm Fast Fourier Transform FFT which takes the signal and decomposes it into frequency components. In spectral analysis, the spectrogram is always seen as a useful tool and it presents interpolated colours of magnitude versus time. Therefore, the spectrogram plot not only allows seeing which frequencies are present in the signal, but also with reference to time and amplitude (Sovijarvi *et al.*, 2000; Johnson, 2011).

As stated in (Johnson, 2011), the spectrogram represents the amplitude by colour rather than on a separate axis in the graph. The spectrum can be drawn by replacing the vertical axis with a colour scheme as a single line that changes colour to show the approximate amplitudes of the components. Therefore, instead of drawing a peak, a dot is drawn on the frequency axis in (yellow) to represent that the amplitude at that frequency is high. The spectrogram is a method that adopts a colour scheme such that any frequency component with amplitude in the top tenth of the amplitude scale is indicated in (yellow). Consequently, frequency component with amplitude in the next tenth of the amplitude scale is indicated by another colour (red), and so on (Johnson, 2011).

The spectrogram shows time in the horizontal x-axis, and frequency along the vertical y-axis as in waveform visual representation, and intensity indicated by varying shades of darkness of the pattern. In this way a spectrogram is 3-D plot. In many cases, the components that make up a complex signal do not share the same amplitude value. Therefore, differences in amplitude are shown in the spectrogram by shading, that is the frequency components with the highest amplitude values are shown in darkest region of the colours, components of lower values are shown in ladder shad of light colours up to the one that signifying very low amplitude or silence (Johnson, 2011). Among the SSA visual tool that uses such DSP analysis techniques, the spectrogram is also used to show the covariance matrix for example.

7 CHAPTER SEVEN SYSTEM VERIFICATION: EMPIRICAL STUDY WITH TESTING SIGNALS

7

System Verification: Empirical Study with Testing Signals

7.1 Overview

Wind noise has been a problem for many decades and numerous methods have been applied, yet not many effective and realistic solutions have been obtained. This was the main reason behind developing a wind noise separation method based on singular spectrum analysis in this study. A noise reduction technique is the process of removing noise in the signal, and generally, such techniques are conceptually quite similar regardless the signal that is being processed (Sarode and Deshmukh, 2011). Based on this definition many experiments have been performed and several case studies have been conducted. After the justification stage that intended to develop and test the method to investigate its potential capabilities with regards to the decomposition and reconstruction of signals, the method has been further developed in the system verification phase for signals and noise separation using typical testing signals.

The contribution is targeting with systematic investigation and optimisation along with more advanced grouping techniques to improve the separability for wind noise separation and mitigation. Though, the separability may vary depending upon the dataset itself. However, in this study, the methodology has been developed incrementally to identify the potential and capability for wind noise removal with testing signals before using real-world sounds in the system validation phase. The experimental procedure includes testing new functionalities within the framework established for this study when adopting the developed systematic approach for the method. This strategy was successful in drawing the algorithm flowchart and develop the method in a step by step manner in the empirical study phases.

The method was systematically developed regarding grouping and reconstruction techniques as key elements and in system verification and validation phases. In the system verification phase, different experiments have been performed for the separation of a mixture of deterministic signals such as sine wave and triangular wave, etc., as a first stage. In the second stage, the method has been developed to separate white noise and wind noise from such deterministic signals. Results of some selected experiments which have been carried out during the experimental investigation of system verification are presented in this chapter.

7.2 Principle Framework Description

In the system verification phase, it is important to develop and investigate the SSA for signals and noise separation by first applying the principles of the method through performing many experiments using different testing signals and noise before developing a complete SSA algorithm when using real audio recordings. As previously mentioned, developing and testing new functionalities and developing the method in many ways based on the systematic approach explained in the previous chapters is the central objective. The main focus while experimentally investigating the developed SSA method is on how to improve the separability. Among the long-term aims is to perform further development and best tailor the selected method for practical uses in environmental smart city application and soundscapes monitoring. The developed SSA algorithms have been implemented using MATLAB platform.

In the system verification phase, the principle of the SSA as a noise reduction method has been applied including all the steps of the SSA algorithm when two signals are generated and mixed together. The principle of recovering and reconstructing one of the signals and excluding the other at the end has been developed and applied as a kind of separating signal from noise based on the capabilities of the SSA for reconstructing signals. The development of the SSA algorithms related to wind noise separation have been performed by examining such algorithms in a simulation phase when wind noise is added to some typical testing signals. The main purpose is extracting information by specifying the oscillation that represents the signal of interest while excluding the noise. Such simulation phase can be seen as a pre-integration phase for testing the developed functionalities and ensuring the capabilities of the modified SSA method for signals and noise separation.

In addition, such simulation process can help in obtaining a full descriptive and clear idea about the nature of the obtained findings when running SSA algorithms and carrying on further experimental investigations with realistic samples in the system validation phase. Further development of the key aspects in the method and optimisation of the window length have been considered in the validation phase which includes producing mixed soundtracks or noisy audio recordings based on establishing a dataset of audio files containing different environmental sounds as explained in Chapter 6.

The key words in this experimental work are; the SSA method as a statistical technique based on linear algebra, typical testing signals, wind noise, digital signal processing, sound analysis including audio and acoustic approaches, and MATLAB coding.

7.3 Experimental Investigation Procedure

The simulation procedure is firstly based on generating different signals such as sine wave and triangular waveform for instance. Secondly, mixing two signals together considering recovering one signal as a signal of interest when applying the multiple sequential steps of the SSA algorithm as explained in the previous chapters. In other words, this system verification phase is an activity used to address whether the developed algorithms appropriately reflect the specified requirements, and also to evaluate the specifications during the development process. Such verification activity is required as an introductory step towards the system validation phase and implementing the developed method with real audio recordings. Figure 7.1 shows a flowchart of the system verification phase presenting the simulation procedure of the experiments using typical testing signals. Figure 7.1 also illustrates the sequential steps of the developed SSA algorithm in case of two different generated signals that are mixed together.
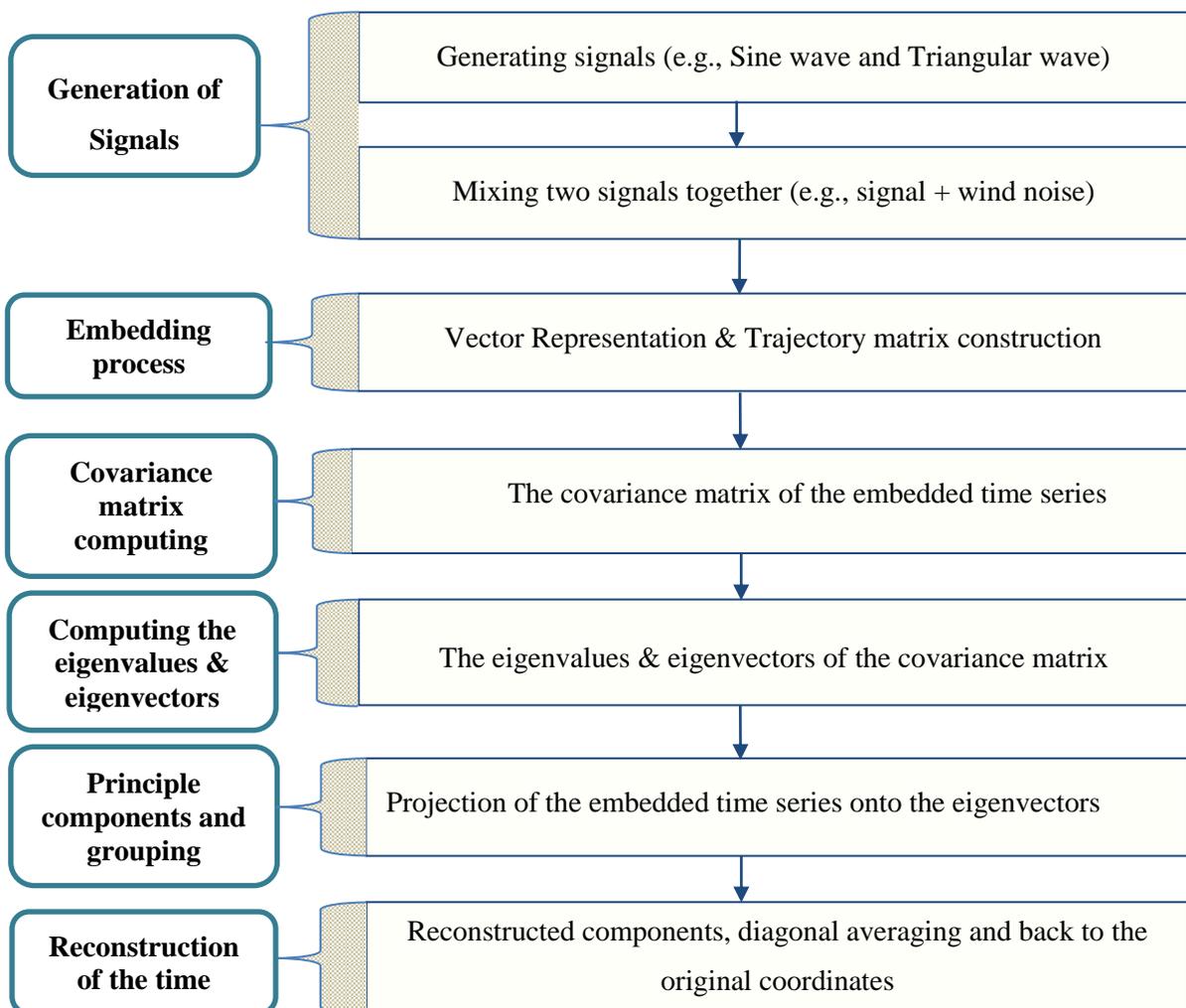

Figure 7.1. A flowchart of the system verification phase presenting the simulation procedure for typical testing signals and wind noise

To enable this simulation process, the SSA algorithm has been written and developed using MATLAB platform in such a way to perform the whole steps and plot figures related to each step. Numerical values regarding the different matrices used in the method are not provided in the results. However, only findings from the experiments that provide practical assistance for the purpose of this verification phase are indicated. Also, it is to avoid giving many matrices with enormous numbers which by themselves will have no much significant meaning in interpreting and demonstrating the SSA method. Instead, demonstrating the performance of method in this phase is by presenting multiple figures. The working examples given in the previous chapters which provided with numerical values and performed with MATLAB are sufficiently enough to show the nature of the different matrices used in the SSA as well as to give a clear idea of the construction of the matrices, their arrangement, manipulation, and the values they contain. The working environment of such simulation and testing platform is with MATLAB. The different steps of this procedure will be discussed with more details in this chapter.

7.4 Proposed Algorithm Framework

The SSA consists of two complementary stages as explained in the previous chapters which are; decomposition and reconstruction. Decomposing the original series into a sum of small number of independent and interpretable components is the first step that aimed to obtain such components like oscillatory components, and a structure-less noise. The reconstruction stage is to reconstruct the time series based on the selection of the principle components that correspond the eigenvectors associated with the dominant eigenvalues.

The SSA has been used for noise reduction, however, when developing grouping and reconstruction techniques it removes noise in broadband and it can be seen as a source separation. Wind noise is a broadband noise that covers a wide range of frequencies. Generally, broadband noise is defined as sound that does not have any tonal character and is not dominated by any particular frequency; however, the energy of such noise is distributed over a wide sector of the audible range. Since the decomposed additive components can be identified as to distinguish between signal components and wind noise components in the Eigen domain, therefore these wind noise components can be separated out from the others.

The experiments in the experimental investigation phase described in this chapter are gradually performed in a step by step manner starting from implementing typical testing signals that are mixed together to the next stage when such signals mixed with white noise and wind

noise. As previously mentioned, the developed SSA system that performs all the required functions has been validated and evaluated using mixed soundtracks of real audio recordings towards achieving the central objective of this study.

The SSA algorithm has been written and developed in the development process to meet the specifications laid down according to the needs and expectations of each experimental phase as explained in the previous chapter. Therefore, the algorithm has been written in different versions to satisfy such conditions considering the contribution of the study which is developing the method specifically with regards to the key concepts in a methodical and systematic manner. The systematic approach developed for the SSA within the framework of this study has been explained in Chapters 4 and 5.

7.5 Practical Work Phases

As previously mentioned, the SSA decomposes a time series into many component parts and reconstructs the series considering the meaningful components while leaving the noise component behind. However, this fundamental principle has been applied for developing the SSA algorithms during the different experimental phases. The method has been incrementally developed and implemented in a logical sequence according to the requirements and aims set for each experimental phase. The experimental investigation procedure of this research is divided into three main phases which are the justification, verification, and validation of the developed SSA system.

The first phase, however, was an introductory phase to investigate the SSA by studying and applying its principles and concepts through performing a large number of experiments to handle with the method. In addition, the first phase was to justify the developed method to implement it as a wind noise separation technique in the subsequent phases. The focus was on generating different signals and applying the principles of the method for decomposition and reconstruction of signals. Also, the idea was to provide, as outcomes, with working examples to show the two main stages of the method, how to manipulate with the multiple types of matrices, and the way of recovering and reconstructing signals which all given in the previous chapters. Meanwhile, this phase also included testing new functionalities which are gradually developed in the adopted systematic approach of the method.

In the second phase which is presented in this chapter, experiments started with the implementation of the developed SSA for the separation of a mixture of deterministic signals

and noise as previously mentioned. This phase is divided into two stages, the first is to implement the developed method with such deterministic signals as typical testing signals when mixed together (e.g., sine wave and triangular wave, etc.). The second stage is a crucial stage in which the method developed and deployed to separate white noise followed by wind noise from such deterministic signals.

As known, the window length is the sole and important parameter in the embedding stage of the SSA algorithm. However, unlike the proceeding phases, its optimisation is left to the third experimental phase. Also, *w*-correlation matrix has been used to indicate the separability. In the third phase, it is to bring together the different aspects of the developed SSA system for validation and critical evaluation considering and adopting suitable dataset that links up to the application area of the study.

7.6 System Verification Phase Using Typical Testing Signals and Noise

Many different testing experiments have been performed in this phase regarding the verification of the developed SSA algorithms using typical testing signals. Only three experiments have been selected as examples to illustrate the main aspects of the developed SSA technique in this study as explained above.

The selected experiments presented in this chapter cover all the previously mentioned aspects when working on the theory that the SSA is a time series decomposition method used for signals separation and noise reduction. To describe these experiments, a specific style has been established and followed. This style includes setting the aims for each experiment, introducing a relevant part of the general SSA theory that fulfill the requirements of the experiment, description and procedure, and eventually results and discussion.

7.6.1   Experiment 1: The SSA Algorithm for Separating Two Signals

The Implementation of the developed SSA algorithm with two signals mixed together (e.g., a sine wave and a triangular wave) is presented in this experiment.

Aims:
- To practice the SSA aspects on two signals generated and mixed together.
- To apply the essential concepts of the SSA regarding its two complementary stages.
- To demonstrate the SSA capabilities for separating signals out from each other.
- To understand the basis of the complete reconstruction and the dominant reconstructed components.

Theory:

The developed SSA system is based on the two complementary stages of the method as the SSA is a two stand-point method with multiple steps involved in the developed algorithm. The fundamentals of generating signals and mixing them are required for this experiment. Window length selection in the embedding stage along with the eigenvalues spectra that have to be considered for selecting the significant eigenvalues and their associated eigenvectors to identify the dominant principle components are parts of the general SSA theory.

Procedure and description:

Generating a time series and defining parameters that are required to adequately perform the experiment is the first step in writing the algorithm of this experiment. A time series $X(t)$ is created by first generating a time series composed of a sine wave of length N and added to another generated triangular wave. The resultant time series is a mixed signal composed of these two signals. Figure 7.2 illustrates the original signals (sine wave and triangle wave) generated and mixed together to be used in the experiment.
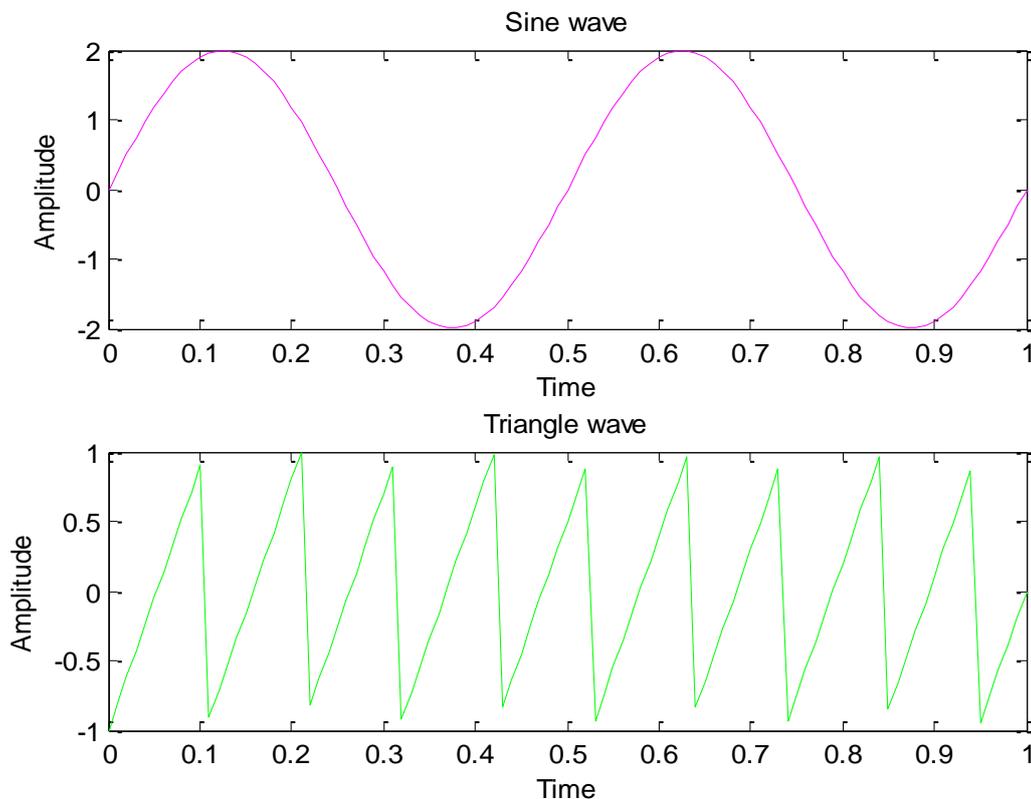

Figure 7.2. Original sine function and triangle wave used in the experiment

The embedding dimension (window length), the length of the generated time series, the period length of the sine wave, and the triangular wave should be set and clearly defined. Recall from Chapter 3, window length is generally in the range $\frac{N_t}{4}$ to $\frac{N_t}{2}$. For this experiment, it is assumed that $N_t = 100$, $m = 30$ where the selection of the length of the time series and the parameter m is based on an assumption for appropriate simulation. A sample of the mixed wave is shown in Figure 7.3.
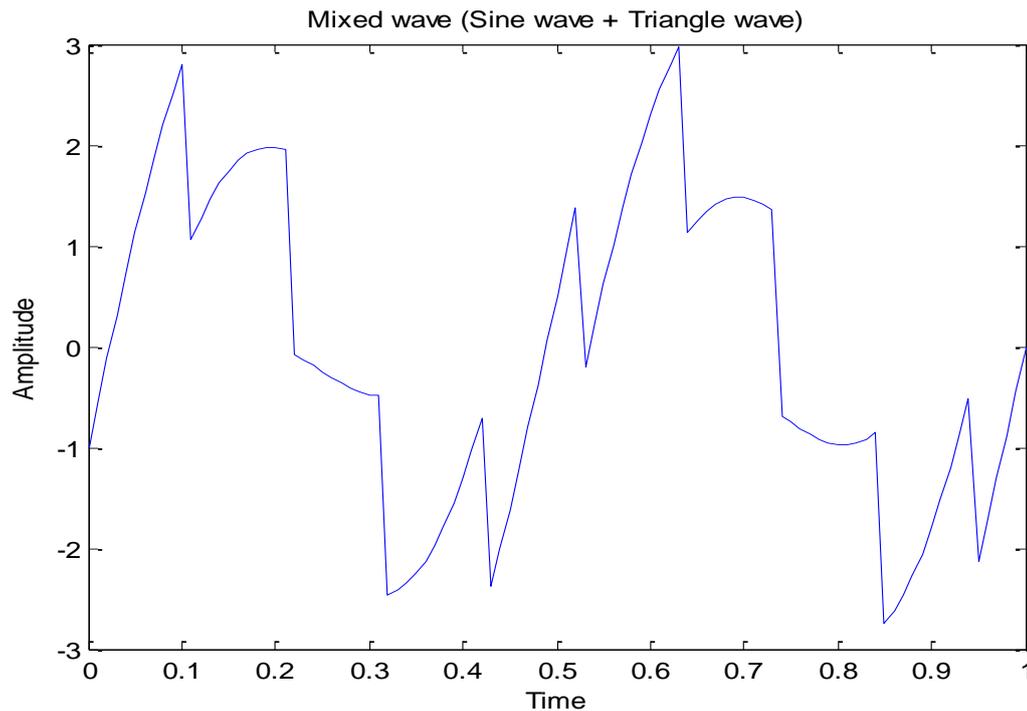
Figure 7.3. A time series sampled from sine wave mixed with triangle wave

To provide a detailed description of this experiment, the steps of the algorithm can be explained and outlined as follows:

I. Constructing matrices **Y** and **C**

After constructing the embedded time series or the trajectory matrix **Y**, the covariance matrix **C** can be calculated based on the (Toeplitz approach) or it can be computed based on the trajectory approach directly by multiplying the trajectory matrix **Y** by its transpose. There are some built-in MATLAB functions (e.g., "cov") to compute the covariance matrix and a vector **x** whose entries can be used to construct the covariance matrix which is a diagonal-constant matrix which is also known as Toeplitz matrix. Figure 7.4 shows the spectrogram of the covariance matrix.

Figure 7.4. The spectrogram of the covariance matrix

II. Computing matrices LAMBDA and RHO

Figure 7.5 illustrates the eigenvalues spectra and first four eigenvectors for comparison where the most significant are the first two eigenvectors which are in phase.
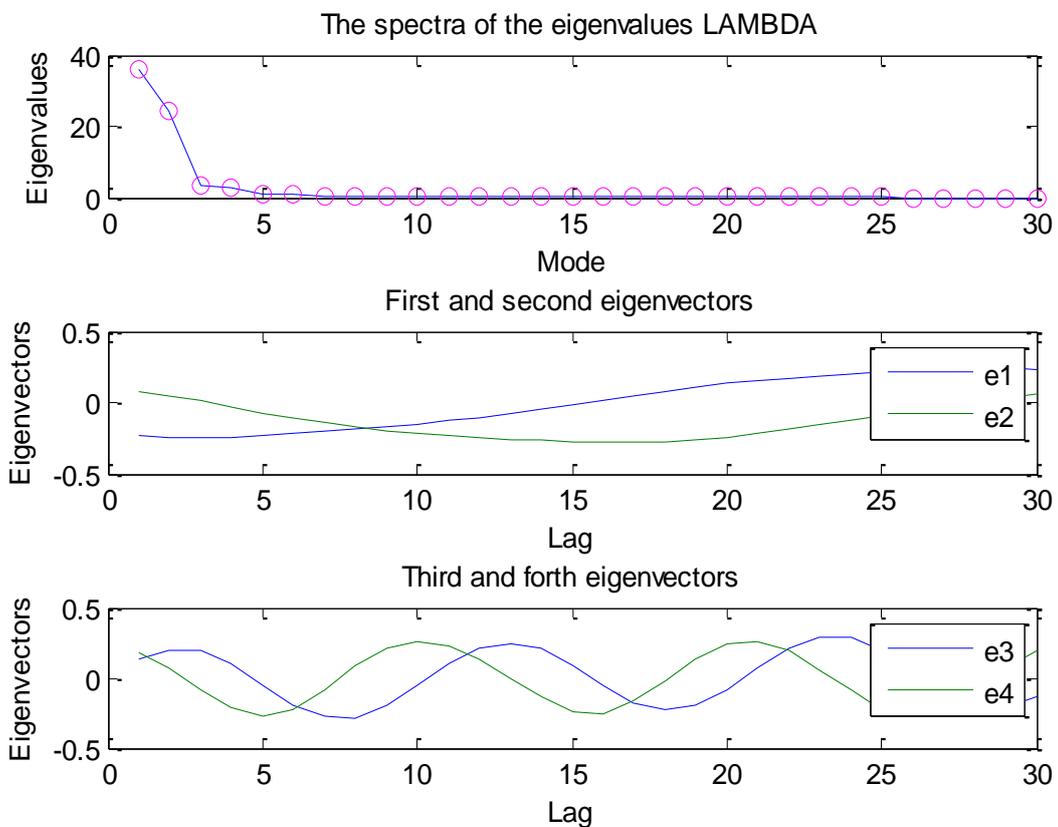
Figure 7.5. The eigenvalues and first four associated eigenvectors

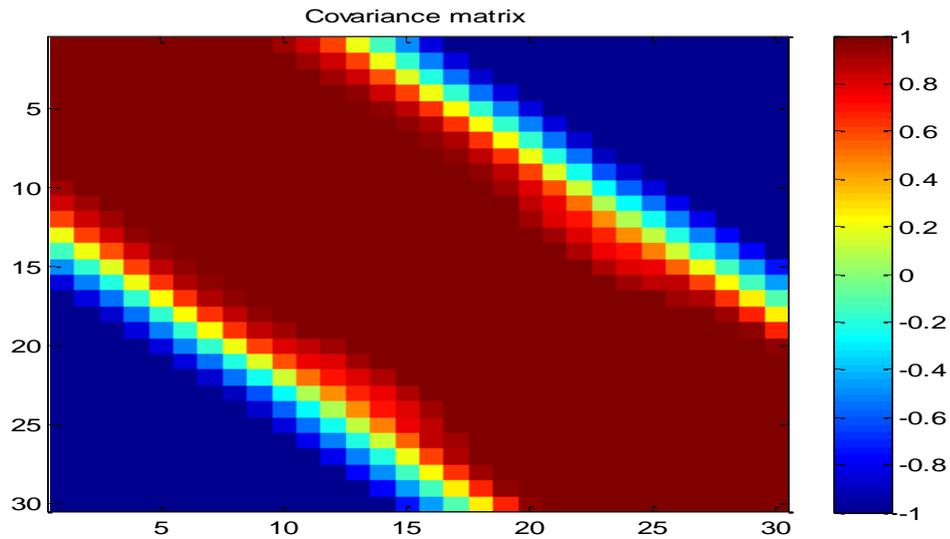The eigenvalues and their associated eigenvectors of the matrix **C** can be calculated using a MATLAB function (EIG). This function returns two matrices, the matrix LAMBDA with eigenvalues along the leading diagonal and the matrix RHO with eigenvectors arranged in columns. In this step, it is to extract the diagonal elements of matrix LAMBDA and sort their associated eigenvectors from matrix RHO. It is worth mentioning that the eigenvectors represent the axis of projection in the Eigen domain and they can describe trend and phase.

III. Principal components

The principal components *PCs* can be computed by projecting matrix **Y** onto and the eigenvectors matrix. The *PCs* represent the variance of the signal and give a clear observation about the reconstruction. The first four principle components are shown in Figure 7.6.

Figure 7.6. The first four principle components

IV. Reconstructed Components

The reconstructed components can be computed by inverting the projection of principle components in Equation (7.1) onto the eigenvectors transpose matrix as Equation (7.2).

$$\mathbf{PC} = \mathbf{Y} \times \mathbf{RHO}, \tag{7.1}$$

$$\mathbf{RCs} = \mathbf{Y} \times \mathbf{RHO} \times \mathbf{RHO}^{T}, \tag{7.2}$$

Averaging along anti-diagonals provides the *RCs* for the original input $X(t)$. It is possible to completely reconstruct the original time series $X(t)$ by the sum of all recon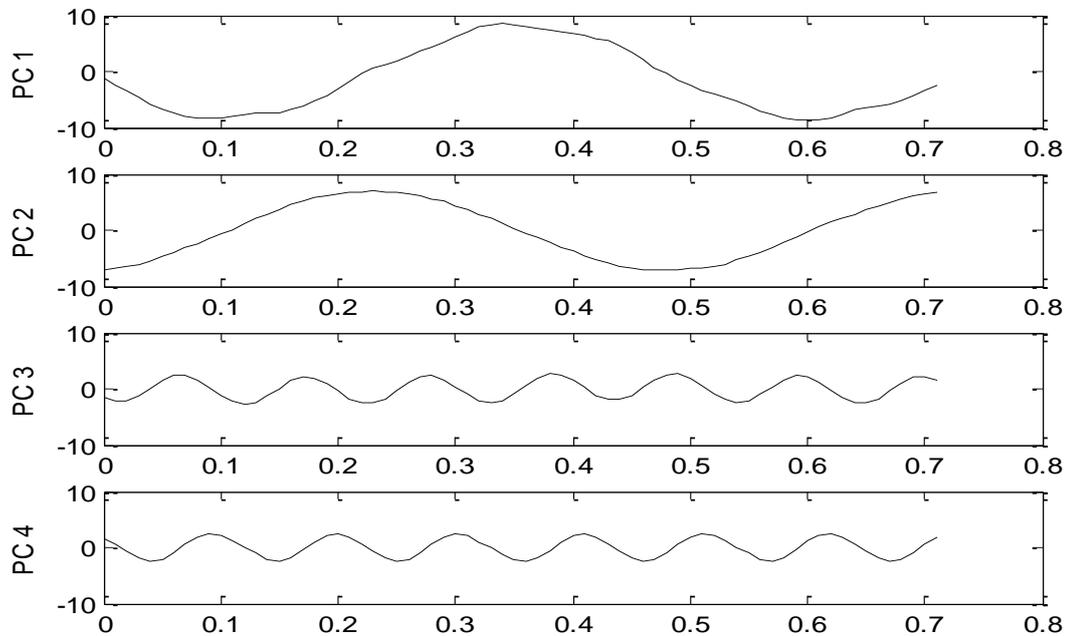structed components *RCs* or the signal of interest by the first two *RCs*. Figure 7.7 presents the first four *RCs* for comparison purposes.
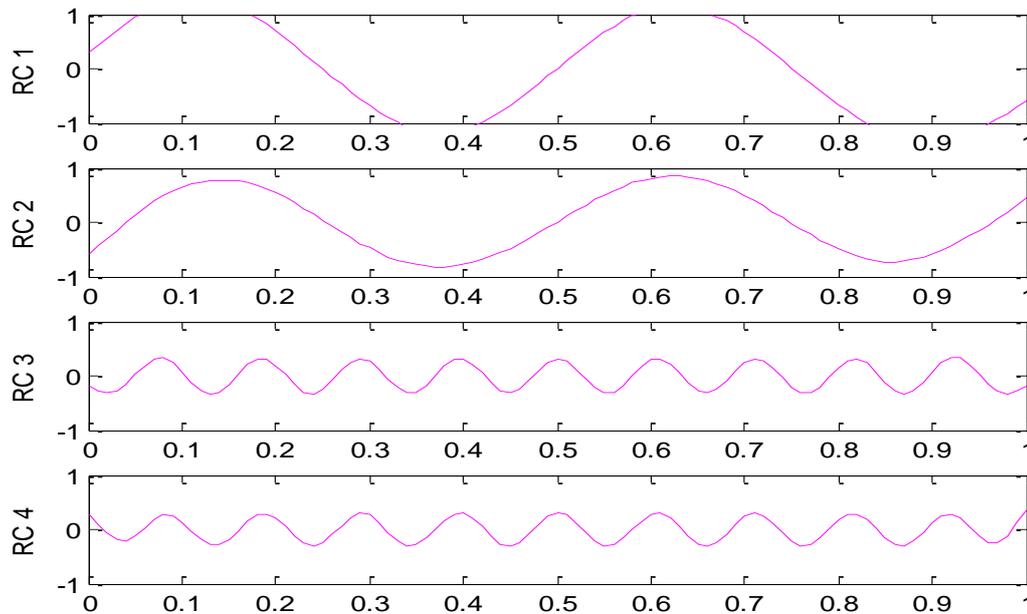
Figure 7.7. The first four reconstructed components

Results, discussion, and conclusion:

Since the main aim set in this experiment is to implement the developed SSA algorithm for separating two generated signals (e.g., sine wave and triangular wave) and to work on the theory of the SSA with regards to its two complementary stages, however, from the experiment, some important points can be outlined as follows:

- The time series has been decomposed to oscillatory components that correspond to the first two eigenvalues located in the lower subspace of the eigenvalues spectra and other components that correspond to the higher-order eigenvalues.

- From the reconstructions shown in Figure 7.8, the time series $X(t)$ can be completely reconstructed using the whole set of the reconstructed components while the signal of interest, which is assumed the sine wave, can be reconstructed using the first two *RCs*.

When comparing the original time series $X(t)$ to the reconstruction with *RC1-2*, which is the sum of the first two *RCs*, it can be found that the signal of interest can be reconstructed while excluding the others.

175
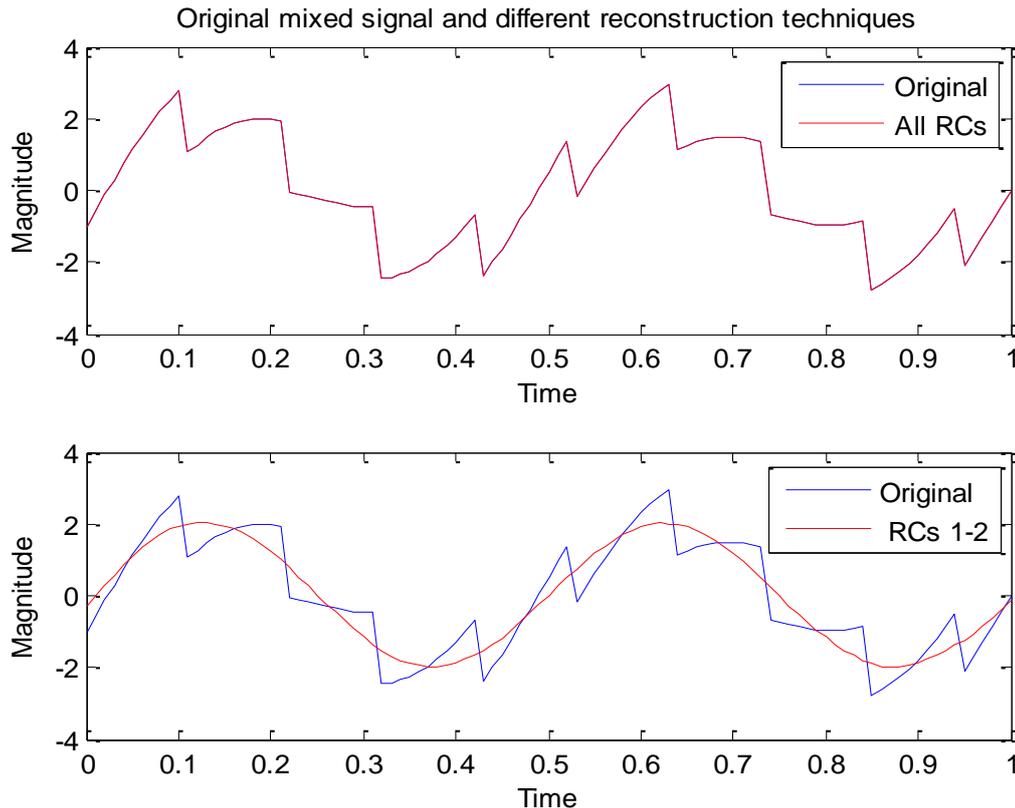
Figure 7.8. Reconstruction with *RCs* 1-2 vs the complete reconstruction

It can be concluded from this experiment that the time series can be decomposed and reconstructed either completely for the whole set of RCs (with no selection or specifying boundaries in the subspace) or partially regarding the signal of interest for the first two ones that correspond to the most dominant principle components. This principle will be applied in the next experiment when considering noise added to signals instead of two periodic signals by using less of the total number of *RCs*. A part of the given noisy time series (the mixed signal) that supposed to be noise can be separated out.

7.6.2   Experiment 2: SSA Algorithm for Separating Sine Wave and White Noise

This experiment is an implementation to demonstrate the developed method in step by step manner by analysing a generated time series composed of a sine function and white noise.

Aims:
- To illustrate the two main complementary stages in the SSA method.
- To analysis a generated time series composed of a sine function and white noise.
- To apply a developed SSA algorithm for noise reduction by separating the noise out and reconstruct the signal of interest.

Theory:

Decomposing the original time series into a small number of independent and interpretable components, such as a trend, oscillatory components (harmonics), and a structure less noise, is the first stage in the SSA method. The original time series can be completely reconstructed by the sum of all reconstructed components. The signal of interest can generally be reconstructed with the first *RCs* that correspond to the most dominant principle components. In the decomposition stage, the variance of the signal can be represented by each eigenvalue in the direction of the corresponding principle components. Based on SSA visual tools, the eigenvalues spectra show the eigenvalues found in the lower subspaces (the lower-order eigenvalues) are related to the more powerful components of the signal. Whereas, eigenvalues located in the higher subspaces, where the noise components typically arise, are the higher-order eigenvalues that represent the undesired components which are noise components.

Procedure and description:

As a first step in writing the algorithm is to set the needs and expectations and lay down the specifications, however, it is also required to define all the parameters to carry out the particular functions in the algorithm. A time series $X(t)$ composed of a sine function of length N with observational white noise added is created in this experiment. It can be generated when removing the mean (at zero mean) and normalised to standard deviation equals one. White noise has an equal energy distribution and its frequency spectrum is completely flat.

As known, the sole parameter in the embedding stage of the SSA method is the embedding dimension (window length) which has to be selected along with the other variables required in the algorithm. These variables are the length of the generated time series, the period length of the sine function, and the signal to noise ratio (SNR) for the white noise added which all should be set and clearly defined in the algorithm. Recall from Chapter 3, window length should be generally in the range from $\frac{N_t}{4}$ to $\frac{N_t}{2}$.

For this experiment, it is assumed that $N_t = 200$, $m = 30$, period length of sine function (T) equals 22 and the SNR is set at 1dB for instance. The selection of these variables is based on an assumption for appropriate simulation. The time series is therefore sampled from a sinus function with white noise added, and has mean=0, SD=1, and a length of 200 data points. A time series sampled from sine function with white noise is shown in Figure 7.9.
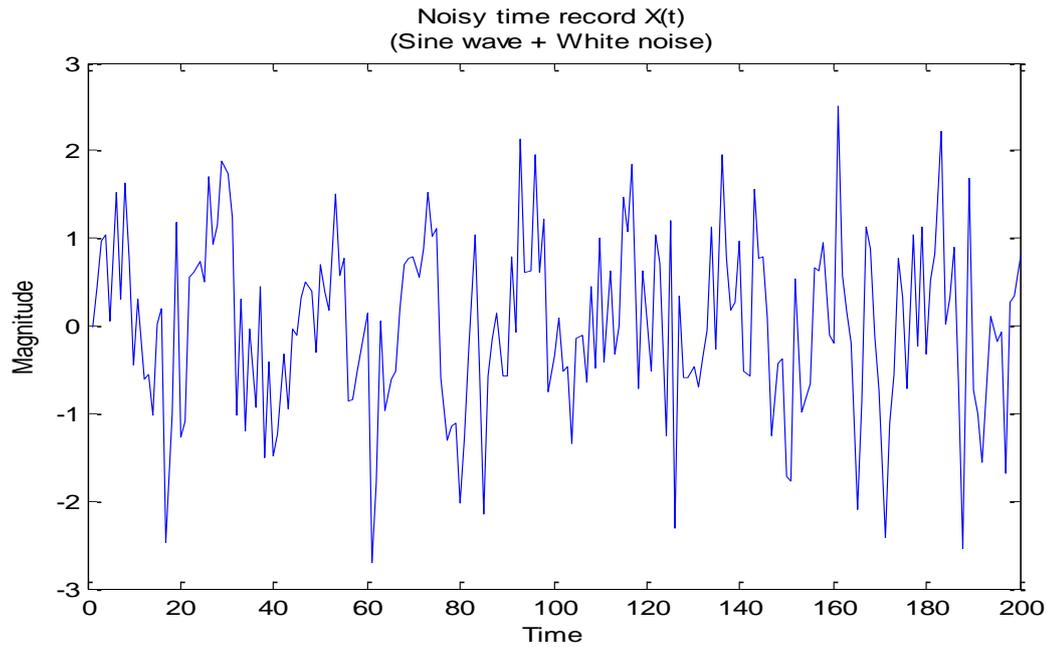
Figure 7.9. A time series sampled from sine function with white noise added

To provide a full description of this experiment, the steps of the algorithm can be explained and outlined as follows:

I. Constructing matrices **Y** and **C**

In the embedding process, it is to construct the embedded time series or trajectory matrix **Y** as a first step and then the covariance matrix **C** can be calculated based on the (Toeplitz approach). Figure 7.10 shows a spectrogram of the covariance matrix in this case.
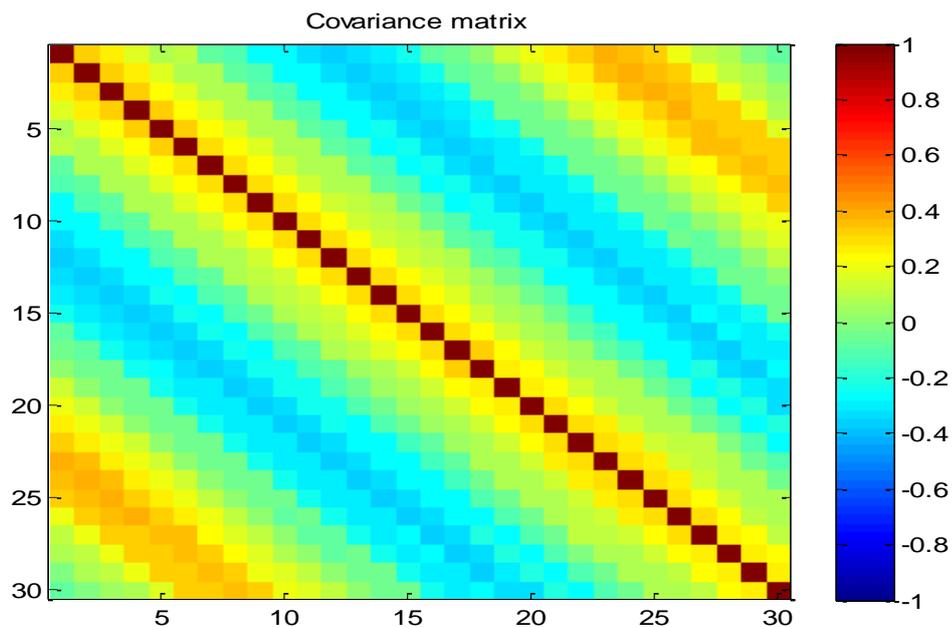
Figure 7.10. The spectrogram of the covariance matrix

There are several numerical approaches to estimate **C** such as calculating the covariance function in MATLAB and computing **C** with the function TOEPLITZ. Alternatively, the covariance matrix can be computed based on the trajectory approach directly from the scalar product the time-delayed embedding of $X(t)$ which is the trajectory matrix **Y**.

II. Computing matrices LAMBDA and RHO

Here, the eigenvalues LAMBDA and eigenvectors RHO of the matrix **C** can be computed and as an essential step in the SSA algorithm. In doing so, a MATLAB function (EIG) is used. This function returns two matrices, the matrix LAMBDA with eigenvalues along the main diagonal and matrix RHO with eigenvectors arranged in columns. In this step, it is to extract the diagonal elements of matrix LAMBDA and sort their associated eigenvectors from matrix RHO. The eigenvalues spectra and the first four eigenvectors which have been selected for comparison purposes are all illustrated in Figure 7.11. It is worth mentioning that the eigenvectors represent the axes of projection in the Eigen domain and can describe trend and phase. Therefore, it can be clearly seen that the eigenvectors in the bottom part of Figure 7.11 are not in phase and thus they cannot be considered as axes of projection compared to the first two eigenvectors.
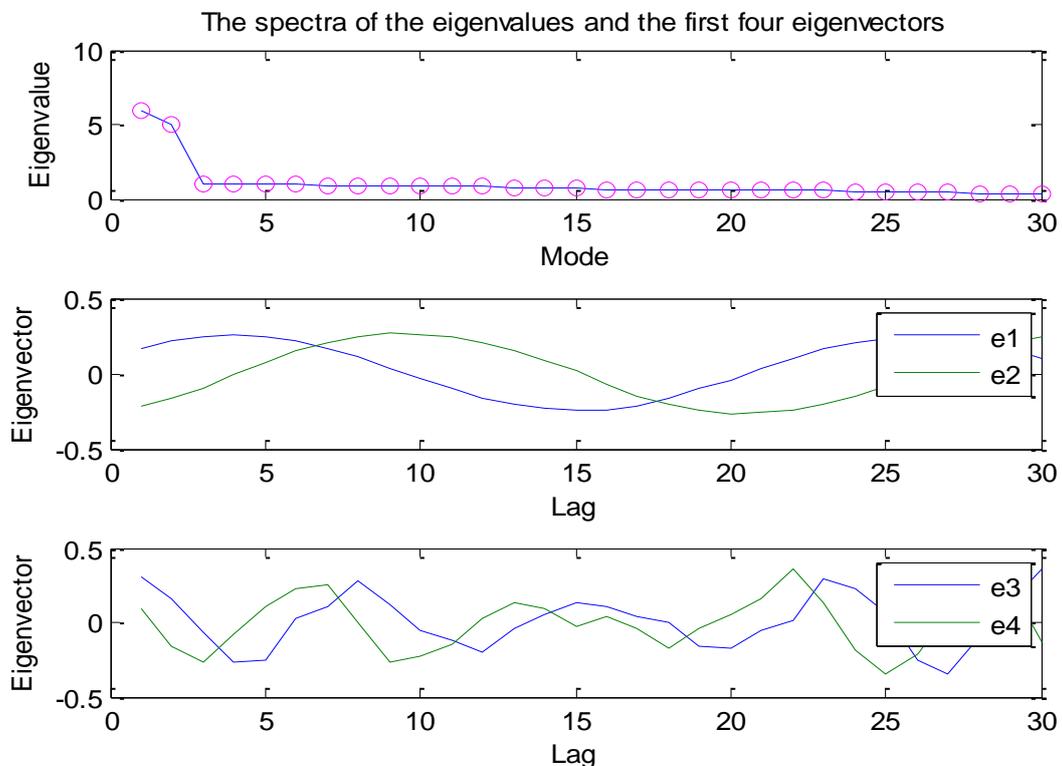
Figure 7.11. The eigenvalues and the first four associated eigenvectors

III. Principle components

The principal components are given as a projection of the time-delayed embedding of the time series $X(t)$ which is given in matrix **Y** onto and the eigenvectors matrix RHO as axis of projection presented in a new coordinate system. In other words, it is the scalar product between **Y** and eigenvectors matrix RHO. The graphical representation of the *PCs* is not a time domain representation, but an Eigen domain representation that can indicate dominance. Figure 7.12 illustrates the first four principle components for comparison and selection of the dominant ones. The dominant principle components can give a clear observation about the reconstruction as for *PC1* and *PC2*, whereas the others cannot.
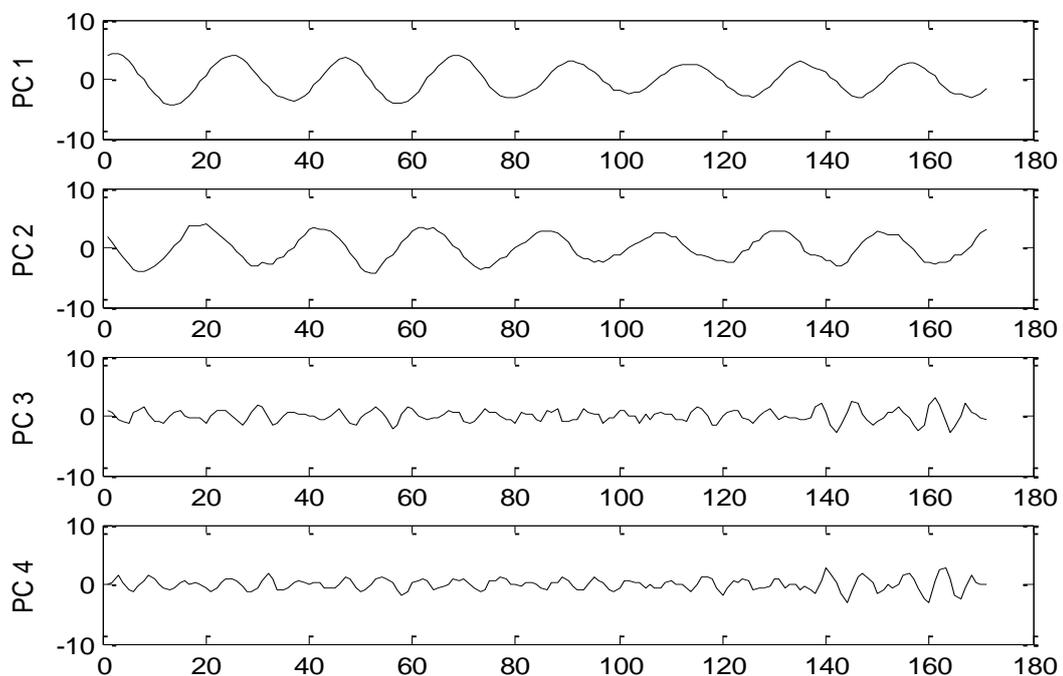
Figure 7.12. The first four principle components

IV. Reconstructed components

Computing the reconstructed components *RCs* can be performed by inverting the projection of the principle components onto the eigenvectors transpose matrix as explained in the previous experiment. Averaging along anti-diagonals provides the *RCs* for the original input $X(t)$. The signal of interest which is the sine function in this example can be reconstructed with the first reconstructed components. The original time series $X(t)$ can be completely reconstructed by the sum of all reconstructed components. The first four reconstructed components are shown in Figure 7.13.

The reconstructed components indicate how the reconstruction can be built. They also indicate the dominance with regards to retaining accurate signal energy in the reconstruction stage. The first reconstructed components are commonly defining an oscillatory signal due to phase quadrature as shown in Figure 7.13. The first two *RCs* indicate higher variance compared to the others.
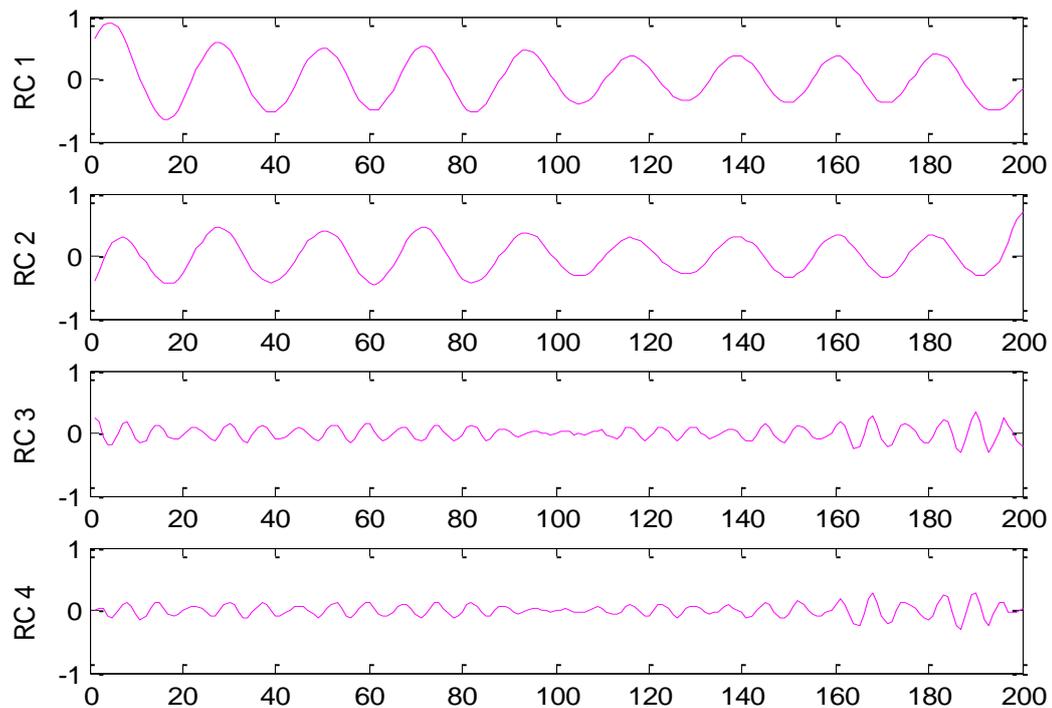
Figure 7.13. The first four reconstructed components

Results, discussion, and conclusion:

The reconstructed components that resemble the original time series are the ones with higher variance. They are mainly the ones related to the principle components that correspond the eigenvectors associated with the eigenvalues located in the lower subspace. In other words, the powerful signal corresponds to the first eigenvalues which are pairs of nearly equal values. From the developed reconstruction technique, the time series has been reduced to oscillatory components that correspond to the first two eigenvalues with high variance and noise components that correspond to the rest of eigenvalues.

Figure 7.14 shows the possible reconstruction when using complete reconstruction procedure using all the RCs and the best selection for comparison. The top part of Figure 7.14 shows the complete reconstruction with the sum of all *RCs*. The signal of interest can be

reconstructed with the *RCs* that correspond to the first pair of nearly equal eigenvalues as in the bottom part of Figure 7.14.
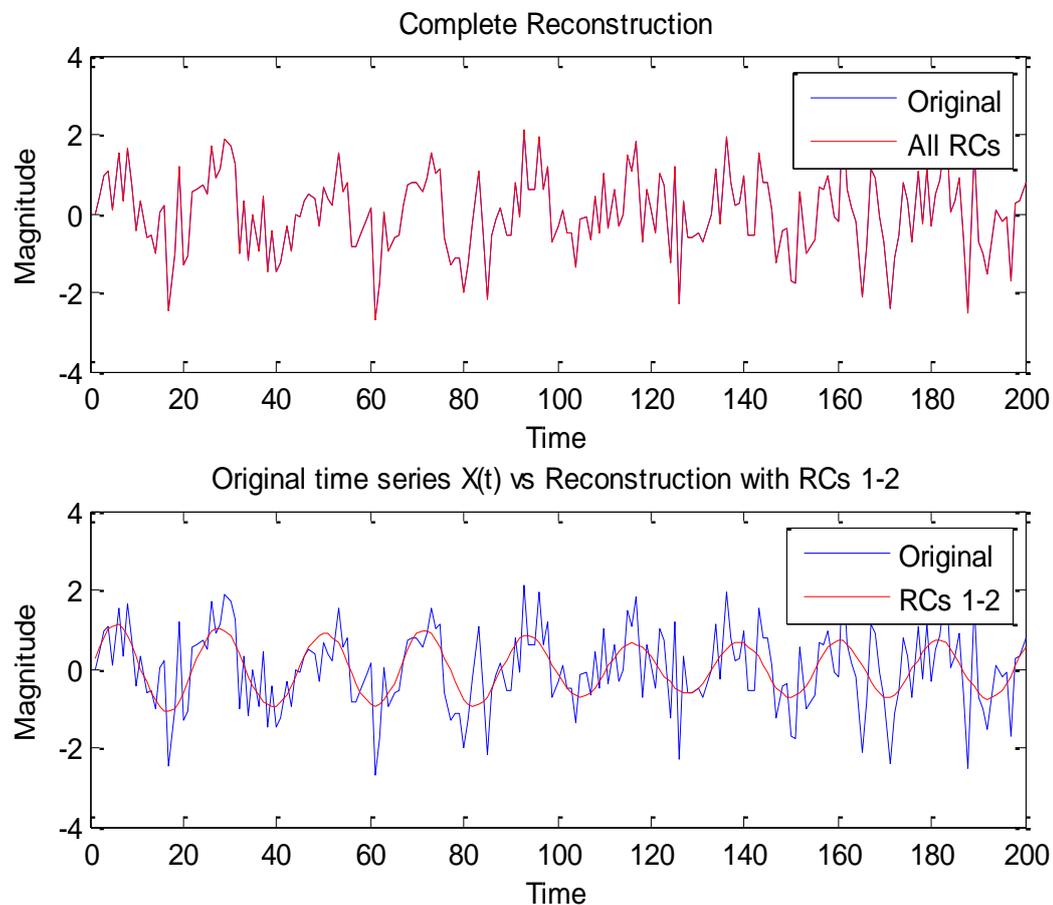
Figure 7.14. Reconstruction with *RCs* 1-2 vs the original series

When comparing the original time series $X(t)$ to the reconstruction *RC1-2* which are the sum of the first two *RCs*, it can be found that the reconstruction of the signal of interest can be performed in this way while excluding the other components that represent the noise. According to the SSA recommendations and what have been demonstrated by using less of the total number of *RCs* in this experiment, a part of the time series that supposed to be noise can be separated out.

This verification activity of the developed SSA as noise reduction method indicates the separability approach of the method in many ways. In this experiment, the time series has been reduced to oscillatory components and noise components. Oscillatory components correspond to the first two eigenvalues as the only pair appears in the lower subspace in this case. However, noise components correspond to the higher-order eigenvalues located in the higher subspaces where the noise components typically arise.

7.6.3 Experiment 3: The SSA for Separating Sine Wave and Wind Noise

This experiment is to demonstrate the developed SSA as a wind noise reduction method in this system verification phase. Also, this experiment is to examine to what extent the separability can be achieved by analysing a generated time series composed of a sine wave and wind noise added. This is mainly what this study is aiming at as an interesting case for further investigation and development of the SSA algorithms with realistic samples in the system validation phase.

Aims:

- Developing an algorithm based on the aspects of the SSA for separating wind noise out when added to a generated deterministic signal (e.g., Sine wave), and apply the main concepts of the developed method including the SVD to produce the eigentriples.

- To investigate and develop the grouping criterion and apply several constrains to be included in the algorithm following the systematic approach developed for the method in this study as discussed in Chapters 4 and 5.

- To demonstrate the capabilities of the developed SSA system for separating signals mixed with wind noise as a particular type of environmental noise in this thesis.

Theory:

The SSA is a statistical and mathematical approach based on linear algebra. The SSA main concepts from the general SSA theory and different algorithm steps for separating signals out from noise are considered in writhing the algorithm for this experiment. Furthermore, it is important to adopt the developed systematic approach that includes all the mathematical formulations formulated and explained in the previous chapters. Also, in the system verification phase, it important to bring these mathematical formulations to practice for developing SSA algorithms for wind noise separation and verify the requirements and specifications. The key elements discussed in Chapters 4 and 5 which are grouping and reconstruction techniques along with the proper selection of window length and the dominated eigenvalues are among the aspects considered for developing the SSA particularly for wind noise separation.

Procedure and description:

The SSA generates a trajectory matrix $\mathbf{Y}$ from the original time series $X(t)$ by sliding a window of length m. The trajectory matrix is approximated using Singular Value

Decomposition method to produce the additive components. Constructing the covariance matrix used for computing the eigenvalues and eigenvectors by performing matrix operations is an important step in SSA algorithm. The eigenvectors are needed to compute the principle components as they represent the axes of projection and this can be simply accomplished using the covariance matrix. The eigenvectors are presented in a square matrix that corresponds to a number of eigenvalues of matrix **C** presented along the main diagonal of the eigenvalue's matrix. Determining the eigenvalues in this way is known as Singular Value Decomposition.

The amplitude of each eigenvector will be comparable to the amplitudes of all other eigenvectors with the normalisation used in the construction of the trajectory matrix. The first eigenvectors represent a high-frequency oscillation when considering successive elements of the eigenvectors over time, while the rest capture the lower-frequency components of the time series. The components to be used to reconstruct the series must be properly chosen based on the singular spectrum appearance as explained in Chapter 5.

The last step, however, is to reconstruct the series from the approximated trajectory matrix after computing the principle components as they are produced by the projection of **Y** onto the eigenvectors. The principle components are incomparable to the original series because they are Eigen domain representation. However, the principle components should be selected in identified groups to compute the reconstructed components. As previously mentioned in the first experiment, the first *RCs* are commonly defining an oscillatory signal due to the phase quadrature of the corresponding *PCs*. This means the *PCs* that correspond to the eigenvectors associated with the dominant eigenvalues which are the lower-order ones.

The interesting output of this experiment is to see how the developed SSA method functioning with regards to eliminating wind noise from a signal contaminated of this noise in most likely source separation approach. Also, to finally obtain from the singular spectrum of $X(t)$ a column vector represents a reconstructed one-dimensional series after applying the diagonal average method as explained in Chapter 4.

Results and discussion:

The strategy adopted in this system verification phase was by conducting an empirical study when using testing signals mixed together along with different types of noise added. Developing and implementing the SSA technique in such a way helped to lead to better understanding the concepts of the method regarding time series analysis, decomposition and reconstructions, etc. Importantly, this strategy paved the way towards improving the method

with regards to certain key elements to make a substantial contribution to knowledge as in Chapters 4 and 5. In fact, as mentioned previously, the systematic procedure developed for the SSA in the framework of this thesis was adopted in all the experiments in this verification phase.

This illustration typically shows a complex case as wind noise separation is presented in this experiment when considering the original signal is a sine wave. The main aim is to verify the specification and requirements in the developed SSA method. Among the aims of this experiment is to demonstrate the capabilities of this developed system for reducing wind noise. Also, this experiment is aiming at evaluating the performance of the developed algorithm regarding the main concern of this study which is wind noise separation and improve separability of the decomposed components of the given noisy signals.

In this discussion, Figure 7.15 shows the eigenvalues spectra that represent the Eigen mode in the horizontal axis versus the eigenvalues number in the vertical axis. Figure 7.15 also illustrates typical eigenvalues spectra; however, the number of the dominant eigenvalues located in the lower-subspace might differ from one case to another as previously mentioned.
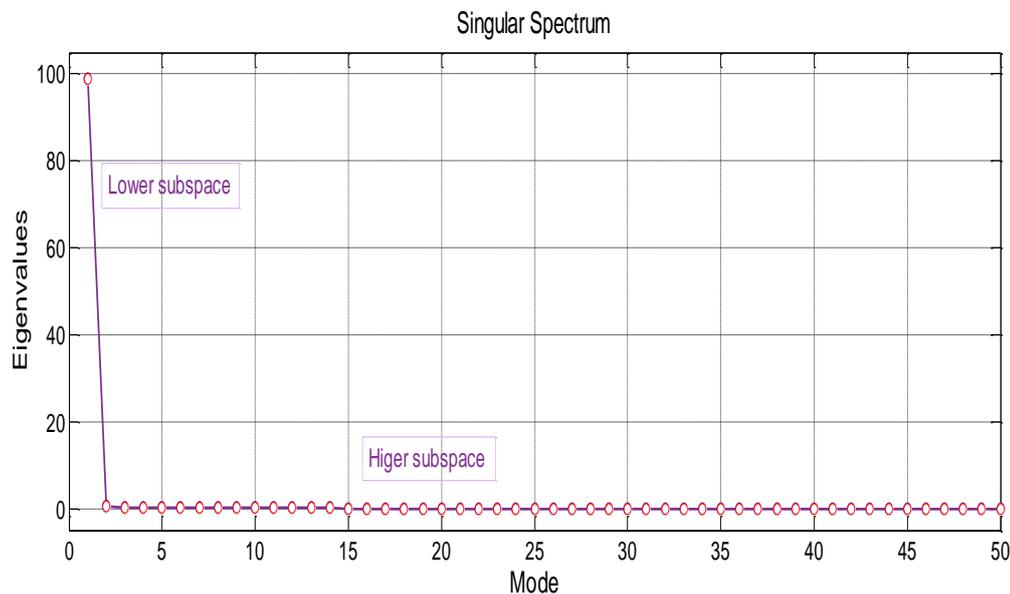
Figure 7.15. A graphical representation of the eigenvalues

Figure 7.16 illustrates a combination of the different signals in the experiment which are the original sine wave, the mixed wave (sine wave and wind noise), and the reconstructed series. The top part of Figure 7.16 shows the original signal which is the sine wave contaminated with wind noise. Accordingly, the positive result that can be emerged from Figure 7.16 is that the reconstructed series is perfectly resembles the original signal with no

185

reconstruction errors. However, these results of this verification phase helped in considering the developed SSA a promising technique for wind noise separation.
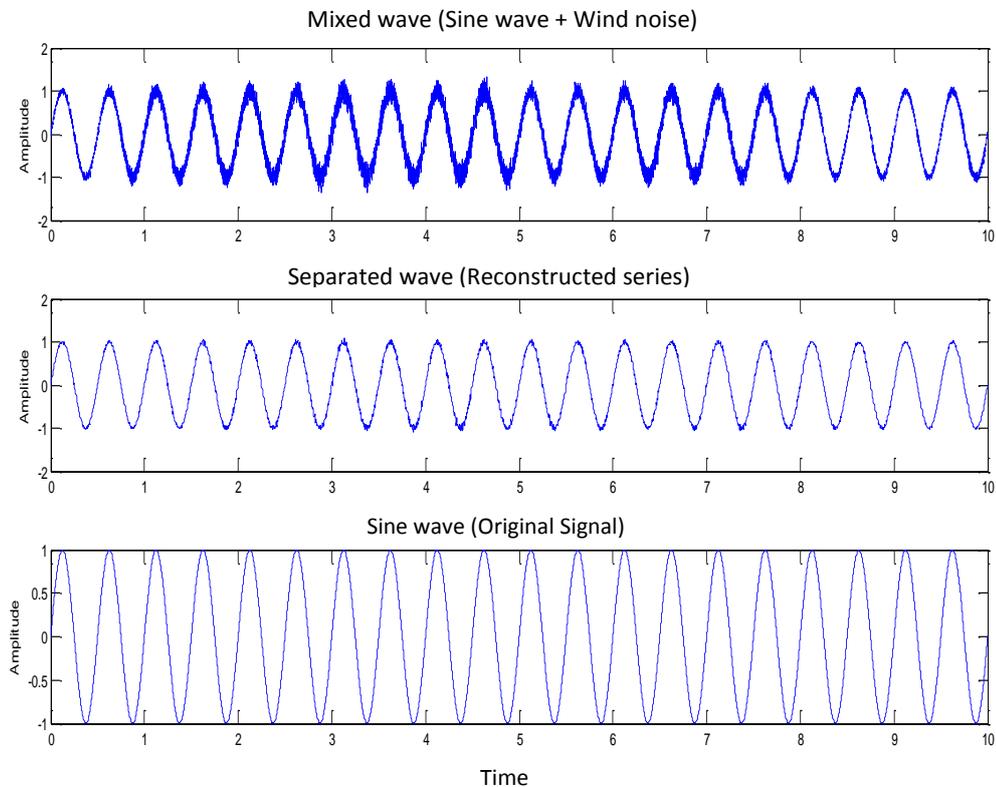
Figure 7.16. Comparison between the different signals in the experiment

The results of this experiment increase the effectiveness of the developed SSA method for distinguishing wind noise components in the Eigen domain. The results show that reconstructing the decomposed components based on gathering wind noise components in a separate cluster as residual noise and the desired components in another group is possible. This experiment produced a positive outcome that shows the capability of the developed SSA regarding the separability particularly for the separation of wind noise component. Furthermore, from this experiment, it can be concluded that defining a subspace for the desired components and another subspace that corresponds to unwanted components is possible using the developed SSA based on developing groping and reconstruction techniques and using standard SSA tools.

Figure 7.17 illustrates the comparison between mixed (noisy) signal and reconstructed series (de-noised signal). The differences between reconstructed series and original signal are highlighted in the top part of Figure 7.17. The findings revealed that the reconstructed signal got resembles the original signal when comparing to the original with no reconstruction errors.

The findings also provide evidence that considerable noise reduction was obtained as it can be seen in the bottom of Figure 7.17. A considerable amount of wind noise was separated out from the signal and denoted as residual series although this remains to be validated when using real-world sounds and critically evaluating the developed SSA using different objective measures.
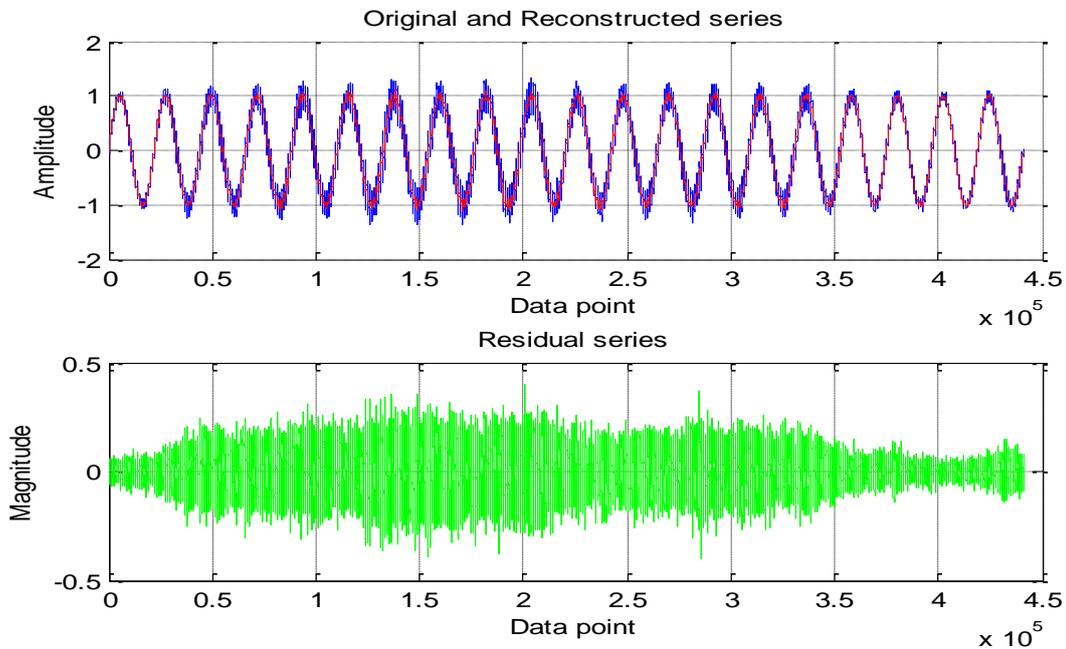
Figure 7.17. Reconstructed series vs original signal and residual series

The single most interesting observation to emerge from the data comparison regarding applying the signal-to-noise ratio (SNR) as an objective measure to evaluate the developed algorithm in this experiment was the considerable improvement for an average of about 8 dB as it is apparent from Table 7.1 for a step of 3 that is covering the range from 0 dB to -6dB.

Table 7.1. SNR measure applied for evaluating the SSA for wind noise reduction

SNR (Before) dB	SNR (After) dB	Difference dB
-6	1	7
-3	5	8
0	7	7

As demonstrated by this experiment, the first *RCs* are generally defining an oscillatory signal due to the phase quadrature of the corresponding *PCs*. These principle components correspond to the eigenvectors associated with the lower-order eigenvalues which are considered as the most dominant. It can be concluded from this experiment that a considerable amount of wind noise has been removed proved by the obtained results which demonstrate the capabilities of the developed SSA in separating signals contaminated with wind noise.

7.7 Summary

The experiments performed during the SSA system verification phase presented in this chapter are for clarifying the principle of the developed SSA method and test the developed functionalities and algorithm capabilities in recovering the original time series and separating the noise out in noisy signals. Furthermore, performing such experiments helped in understanding multiple aspects like generating signals, adding noise, etc., and most importantly, writing and developing SSA algorithms in a step by step manner.

In this chapter, a principle framework has been established considering many important key elements in developing the SSA algorithm throughout the different experimental phases of the empirical studies in this thesis. The procedure of drawing a map to perform a wide range of experiments has been established in the system verification phase using typical testing signals and noise including wind noise. This chapter also introduced the proposed algorithm framework as the algorithm has been written in different versions to meet the specifications laid down according to the needs and expectations of each experimental phase. This chapter also introduced the arrangement of the practical work phases of the study.

The system verification phase using typical testing signals and noise has been carried out and presented with only three experiments that have been selected as examples among the wide range of the experiments performed in this phase. These experiments have been presented in such a way to include the aims, the theory formulated for each case based on the general theory of the SSA and SSA recommendation and standards. Also, the experiments have been reported to include procedure and description along with results, discussion, and conclusion. The straightforward procedure that has been established from the outset was performed by first decomposing the time series into several components and reconstructing the desired signal by grouping desired components based on the eigentriples. This procedure aimed at separating the decomposed components after grouping similar components together.

To make a comprehensive evaluation of the performance of the developed algorithms and measure the effectiveness of the SSA as a method for wind noise separation, the criteria used to evaluate the method will be with real-world sounds. Therefore, the system validation and critical evaluations of the results are carried out in the next experimental phase presented in the next chapter.

8 CHAPTER EIGHT   SYSTEM VALIDATION: EMPIRICAL STUDY WITH REAL-WORLD SOUNDS AND DISCUSSIONS

8

System Validation: Empirical Study with Real-World Sounds and Discussion

8.1 Overview

A wide range of experiments have been carried out to identify the potential and capability of the developed SSA in microphone wind noise reduction. The experimental investigation procedure and practical work were divided into several phases based on developing an incremental methodology as explained in the preceding chapters. Recall from Chapter 7, for the verification of the developed SSA, several experiments have been performed for the separation of wind noise from many deterministic signals along with the separation of such signals from each other. However, in the system validation phase presented in this chapter, real-world sounds have been used in the experiments.

In the system validation phase, it is to deal with realistic samples; therefore, some interesting environmental sounds such as birds' chirps have been used as wanted signals for testing and validation. Also, for validation purposes, different samples have been selected which are alarms and car series as mentioned in Chapter 6 (section 6.4.1). It is worth mentioning that during conducting this research, the developed SSA algorithm was also implemented to mitigate wind noise in corrupted speech signals in specific experiments whose results have been previously published in an IEEE conference (Eldwaik and Li, 2017).

To examine the developed SSA method and determine significant oscillations, noisy audio recordings composed of wanted signals and wind noise have been used. However, this testing and validation phase is based on establishing a dataset of audio files containing different environmental sounds such as birds' calls, alarms, car sirens, etc. The soundtracks used in the experiments performed in this phase contain a mixture of such signals at different mixing ratios to enlarge the sample size and study the effect of the content of wind noise on the output of the SSA algorithms. To extract information by specifying the oscillation that represents the signal of interest while excluding the wind noise, a complete SSA algorithm has been written considering all the key aspects in the method and following the systematic approach developed for the SSA within the framework of this study as explained in the previous chapters.

The main aim is to validate the developed SSA system and investigate the capability of the developed algorithms to decompose such mixed soundtracks that include wind noise and wanted signal into different components. That is to group the desired signal in one group and omit the other which is the unwanted wind noise content in a different group. Reconstructing the decomposed components into time domain representation has been made after applying the developed grouping criterion in the algorithm. However, two groups are defined as one represents the signal of interest and the residual noise presented in the other group represents the wind noise component with typically non-over-lapping content.

Although in all cases, notable wind noise reduction was observed, results are slightly different from case to another due to the complexity of the environmental sounds. Using the dataset specified for testing, results from the separation of outdoor wind noise from birds' chirps are detailed in the current chapter. The results of the validation and evaluation using the other real-world sounds as specified in the dataset are presented in this chapter as well.

8.2 Theory and Practice

The SSA has been showed as a valid method for time series decomposition using spectral decomposition (Vautard and Ghil, 1989; Elsner and Tsonis, 2013). As mentioned previously, from the SSA theory, the spectral (eigenvalue) decomposition of a given matrix into a set (spectrum) of eigenvalues and eigenvectors gave us a clear theoretical and scientific understanding of the term "singular spectrum". As stated in (Elsner and Tsonis, 2013), the covariance matrix plays a vital role for the spectral decomposition. However, when it is estimated from short records in particular, it can make a difference to the results. To clarify, for instance, when the SSA is used to decompose a time record with oscillations of varying frequencies and similar amplitudes, the ordering of the eigenvalues might be affected by the selected estimated matrix. Consequently, the main objective of the development process of the SSA in the context of this research is to decompose mixed soundtracks into a number of components which are wanted signal and wind noise.

In principal, as known, the SSA embeds a given time record into multi-dimensional Euclidean space and finds a subspace corresponding to the sought-for component as a first stage (decomposition stage). In other words, the SSA decomposes the time series into a number of elementary matrices by mapping the time record under test into a trajectory matrix and then processing it using the SVD method. Basically, the SSA decomposition depends on the approximate separability of the different components of the time record (Harmouche *et al.*,

2017; Traore *et al.*, 2017). The window length should be properly identified to obtain the lag k required for constructing a new matrix (trajectory matrix) according to delay coordinates.

At the decomposition stage and after computing the covariance matrix from constructed trajectory matrix, the time record is decomposed into mutually orthogonal components. The trajectory matrix is constructed from the real observations of the time series and decomposed into additive components or elementary matrices by means of SVD (Chu, Lin and Wang, 2013; Golyandina and Shlemov, 2015; Harmouche *et al.*, 2017). As explained in the preceding chapters, the SVD method is also used to determine the principal components of a multi-dimensional signal. Once the SSA decomposes mixed signals in the eigen-subspaces, it selects and groups the principle components according to their contributions to wind noise and desired signal in the singular spectrum domain.

Eventually, the second complementary stage involves applying the grouping criterion and the diagonal averaging to reconstruct a time series component corresponding to its specific subspace. In the grouping step, different rank-one matrices obtained from the SVD can be selected and grouped to be processed separately. The groups will reflect different components of the original time record if they are properly partitioned (Chu, Lin and Wang, 2013; Golyandina and Shlemov, 2015). Therefore, the elementary matrices can be grouped into smaller groups by applying the grouping criterion. After arranging such matrices in different groups, however, it is to sum-up matrices that show glimpses of a specific signature of oscillation patterns while separating out the wind noise based on many tools such as the eigenvalues spectra and the selection of most dominant principle components. Indeed, generating the elementary matrices can be perceived as a convolution of the principle components with their corresponding eigenvectors and weighted by the eigenvalues.

The SSA reconstructs the wanted components back to the time domain resulting in the separation of wind noise and wanted signal. The reconstruction of the one-dimensional series is accomplished by using estimated trend and harmonic components. The reconstruction of the time record is by selecting those components that reduce the noise in the time series (Ghil *et al.*, 2002; Patterson *et al.*, 2011).

In the SSA, there are some useful approaches that can be used in the separation of the wanted signal from the noise. In general, a harmonic component produces two eigentriples with close singular values. Using visual SSA tool by checking breaks in the eigenvalue spectra is considered as useful insight. It is worth noting that a slowly decreasing sequence of singular values is typically produced by a pure noise time series (Golyandina and Lomtev, 2016).

8.3 Testing Criteria and Experimental Procedure

After examining the developed SSA algorithms in the system verification phase with general testing signals as presented in Chapter 7, it is to apply suitable criteria for the testing and validation of the SSA algorithms with detailed specifications and requirements when using real audio recordings containing wind noise. These testing criteria include; the effect of the environment type, SNR ratio, as well as the optimisation and proper selection of the sole parameter of the embedding stage in the SSA algorithm which is the window length. Further development of grouping and reconstruction techniques when using real-world sounds are also included in these criteria.

To implement the methodology of this research in the testing and validation stages of this final experimental phase presented in this chapter, the experimental procedure of the method has been established as shown in Figure 8.1. The experiments were carried out on MATLAB platform. There are five main stages in the experimental procedure as shown in Figure 8.1.
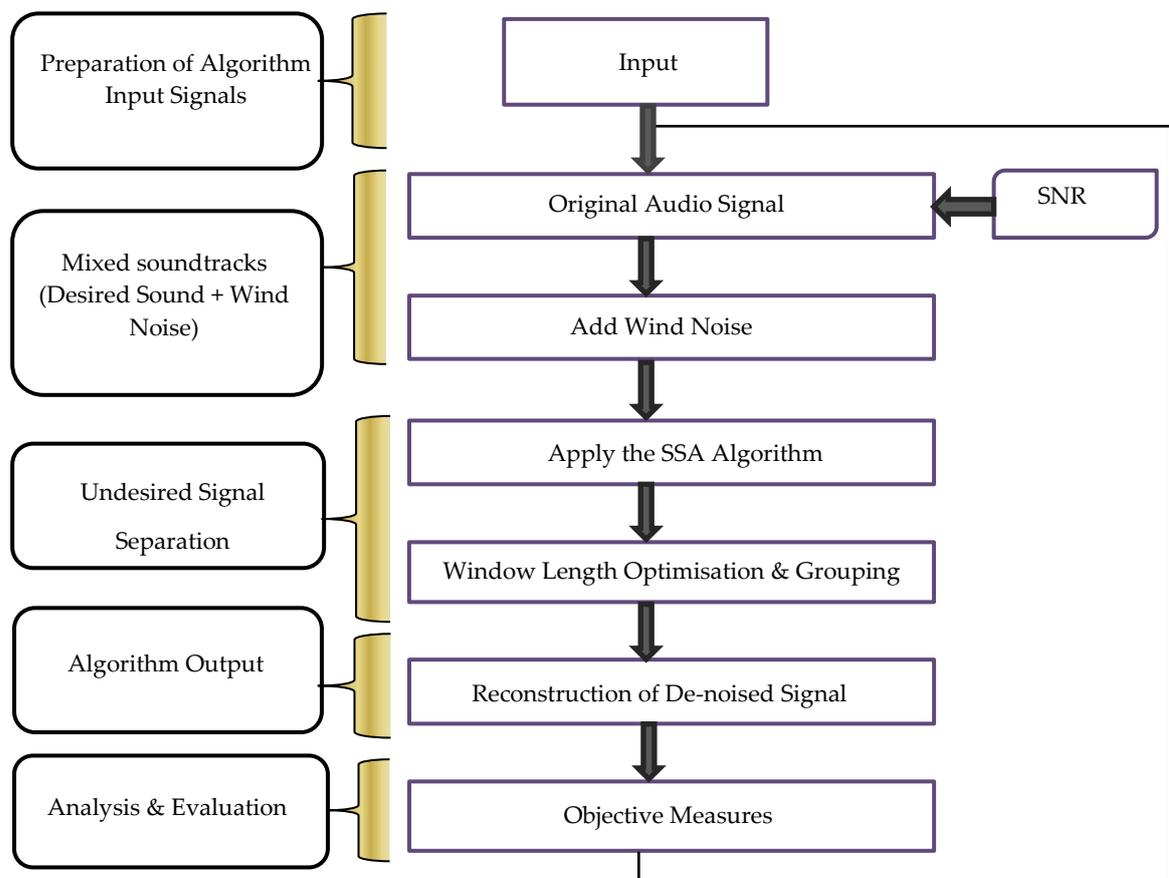

Figure 8.1. A flowchart of the experimental procedure of testing and validation phase

A complete SSA algorithm has been developed based on adopting the systematic approach in which all the key aspects have been considered. Based on the results of the verification phase and empirical study conducted in early stages, the first stage is to prepare the dataset and make it suitable for the experiments. Also, many factors have been considered as explained in Chapter 6, such as the length of the soundtracks the sampling rate, etc.

The second stage is the mixing stage, which is shown as one step here, but it is mainly an audio mixer model that mixes pre-recorded samples according to their signal intensity. In the third, it is to apply the developed grouping and reconstruction techniques followed by analysing the outputs of the algorithm, reporting the results, and eventually evaluating the method by applying objective measures as fourth and fifth stages. As explained in Chapter 6, evaluating the developed method for noise reduction capabilities, measuring the performance of the algorithm, the effectiveness, and efficiency will be through applying objective measures including sound analysis methods, DSP measurements and *w*-correlation matrix.

8.4 Testing Platform

The testing criteria introduced in this chapter have been developed to meet the requirements of the testing and validation phase. At this stage, these testing criteria have been further improved compared to the criteria used in the system verification phase. However, this improvement is based on the new requirements and specifications as mixed soundtracks of audio recordings of wanted sounds and wind noise have been used for the validation and critical evaluation of the developed system. It is worth mentioning that some of the results presented in this chapter are from related articles that have been previously published in the context of this research (Eldwaik and FF Li, 2017; Eldwaik and Li, 2017; Eldwaik and F. Li, 2018).

The system architecture of implementing the developed SSA system in the testing and validation phase for the empirical study with real-world sounds is presented in Figure 8.2. The testing and validation system architecture of the developed SSA has been designed to include several stages. Figure 8.2 shows the adequate preparation of the sounds used in the system validation phase and the configuration of the testing platform. The audio files database has been originally established to contain audio files of 10 s excerpts of desired sounds and wind noise. In the processing stage, it is to apply specific algorithms to detect and remove silence. Specific software tools have been used to split the audio files in the database to 1 s length which is fully representative to reduce the computational load of the SSA algorithms.

The configuration process of the testing platform also includes preparing mixed soundtracks using a mixing model and averaging of the produced samples using MATLAB codes.   The mixing model works based on the principle of mixing the samples of clean signals, which are the output of the splitting stage located in the database, with the wind noise average sample (the output of the averaging model) at desired SNR. The output of the mixing model is mixed soundtracks; however, this procedure is repeated at different SNR. The framing and averaging methods are used to process the mixed soundtracks to produce the SSA algorithm input noisy signals. The averaging model also produces an audio file represents the clean signal. For all the case studies when different sounds are used, clean signals are required for comparing the results. Therefore, framing and averaging methods are also used to process the clean signals. When running the SSA algorithm to process input noisy signals, all the SSA processing and algorithm steps are implemented. Eventually, this procedure results in the separation of wind noise and the reconstruction of desired signals as shown in Figure 8.2.
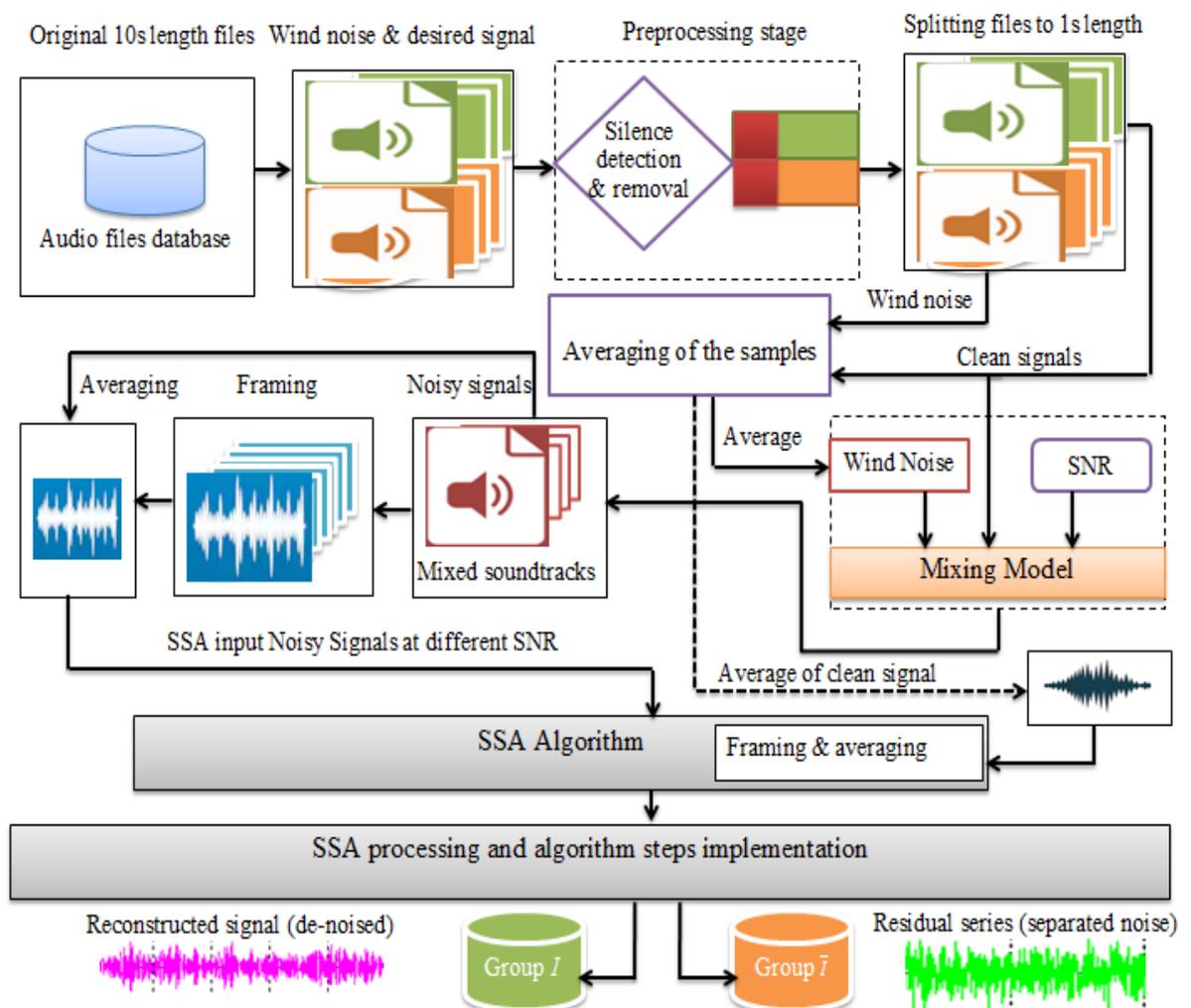

Figure 8.2. Testing and validation system architecture of the SSA method

8.4.1   Testing Requirements

A benchmark database consists of the signals of interest is needed to conduct the case studies in this system validation phase. A freefield 1010, which is a dataset of standardised 10 seconds excerpts from Freesound field recording, has been selected. However, some effort has been made in order to make the samples suitable to the case through using automatic methods where possible to ensure that all the samples included in the datasets for testing and validation purposes are pure desired sounds such as birds' chirps, alarms, car sirens, etc. In addition, all silent gaps have been removed from the samples.

The dataset also includes the samples of wind noise as the samples from this dataset have been mixed to generate mixed soundtracks as shown in Figure 8.2. The testing requirements for developing and implementing the SSA algorithms for wind noise separation are indicated as follows:

- In case of a single class of environmental noise (wind noise).
- Using different types of environmental sounds based on selecting multiple recordings from different locations in the city.
- Using different values of SNR ratio to represent different noise content.
- A careful preparation of the dataset as explained in Chapter 6.
- Using framing and averaging methods along with window length optimisation method.

8.4.2   Framing Method

Since the processing time of the SSA depends on the length of the given time series and the selected window length, therefore, to reduce the processing time of the SSA, the frame by frame processing method has been applied. This method helps in processing the soundtrack samples instead of processing the whole audio file directly. Therefore, it is not practical to use embedding dimension of the same length of the given time series as the SSA works with a large number of matrices and their size will be increased with bigger window lengths, and as a consequence the processing time of the SSA will be very long. Since the frame will be mapped into a two-dimensional trajectory matrix which will be used to compute principle components vectors before applying grouping criterion using the eigentriple, both the window length and frame size might affect the distribution of the eigenvalues in the singular spectra. As a consequence, the separation of the undesired noise components and the reconstruction of the desired signal might be affected.

In the experiments conducted in this phase, samples from the datasets of audio recordings of the multiple desired signals have been selected for testing and validation using an average of 100 ms frame size of five thousand frames. It is a common practice to use such size or even more to be fully representative likewise for selecting the sliding window to contain sufficient information. However, smaller frame sizes are not recommended as they might not contain enough information as required. Even though the average method has been used in all cases, a balance between the window length and the frame size is required. Using sufficient frame size as recommended and the method of computing the optimal value of the window length based on reviewing the literature as discussed in the next sections might lead to achieve an adequate balance and therefore satisfactory results.

8.5 Key Objectives for the SSA Testing and Validation

8.5.1 Window Length and Singular Spectra Method of Calculation

The calculation of the singular spectra for the given time records using seven different window lengths is the method used to optimise the window length and obtain its optimal value in the system validation phase. The selected range of m is between 15 ms and 90 ms based on the size of the frames for the reasons explained above. The size of the frames is 100 ms; however, this range indicates in different steps a percentage from 15% to 90% of the new calculated N_t. Therefore, to obtain the optimal value of the window length, a comparison of the eigenvalues of the lagged-covariance matrix was made for these window lengths which used to generate the trajectory matrix from the frame.

To perform the calculation of N_t and the multiple suggested window lengths to be entered to the algorithm for the optimisation purpose, Equation (8.1) has been formulated. This is the first study to undertake this new equation in the optimisation method of the window length. This equation has been particularly developed in this research and can be considered as one of the major contributions. Also, it is among some other equations formulated within the framework of this study which can all make a significant contribution to knowledge.

$$m_i = p_i N_t = p_i L F_s, \tag{8.1}$$

where m_i is the window length, p_i is the desired percentage of the length for $i = 1, ..., n$, for n number of required steps of the desired percentage with the maximum percentage that gives value of $m_i \leq N_t$, N_t is the new calculated length, L is the frame size in seconds, F_s is the sampling rate in Hertz.

8.5.2 The Detection of Eigenvalues Pairs in the Singular Spectra

Since the construction of the trajectory matrix by means of mapping the soundtrack vector is one of the main steps in the SSA algorithm, the selection of the window length is always seen as an initial and key step towards implementing the SSA method. Therefore, the window length must be optimised to make the best selection. As previously mentioned in Chapter 5, the window length is highly related to spectral information or the frequency width that corresponds to each principle component.

Many assumptions have been made by several authors and researchers about how the window length should be selected, however, such selection is greatly dependant on some criteria such as the aim of the analysis and the complexity of the data. Importantly, achieving appropriate separability of the components can be considered as a central aim of selecting the window length as the case of this study. The formulated concept that introduces each pair of eigenvalues with nearly equal values in the singular spectrum has also been reported as a key aspect in relation to significant oscillations and dominant frequencies in the signal. The eigenvalue ordering along the singular spectrum may be affected by the window length (Vautard and Ghil, 1989; Elsner and Tsonis, 2013; Mohammadi *et al.*, 2016).

Regarding the detection of eigenvalues pairs, it has been showed that greater value of the window length m can enhance such detection (Vautard, Yiou and Ghil, 1992). Various types of noise with several generated time records were tested in (Vautard, Yiou and Ghil, 1992) to observe the impact of the on the dominant signal. The recommendation given in (Vautard, Yiou and Ghil, 1992) is possible for their case study which is trend forecasting for weather time record where the frequency is very low; however, this is not the case for audio data.

In some studies such as the one conducted by (Penland, Ghil and Weickmann, 1991), the authors argued that as long as $m < N_t$, the window length did not significantly affect the results of the SSA. However, the findings of many other studies including this study contradict this assumption to a certain extent. Although the optimisation and justification of the selection of the window length is always required, it can be assumed, as a common practice, that the ratio $N_t/2$ is the optimal window length as reported by many researchers.

Since the lagged-covariance matrix is seen as symmetric matrix of Toeplitz structure, however, with values of m that are nearly equal to N_t, the construction of the trajectory matrix will be affected, and in turn the covariance matrix. As previously mentioned in Chapter 3, the Toeplitz structure of the covariance matrix can be determined by the length of the time series

and the window length. Also, according to (Elsner and Tsonis, 2013), that the lagged-covariance matrix is used for the spectral decomposition and when it is estimated from short records in particular, it can make a difference to the results. In fact, the expected Toeplitz structure which is indicated by the ratio N_t/m cannot always be obtained in practice because of the shortness of the time series or high values of m.

As the SSA was not previously implemented for wind noise reduction and since there is no benchmark for the window length, the optimal window length can be determined using a heuristic method. In the optimisation of the window length, smaller variances between the lower-order subspaces of the eigenvalue's spectra should be highlighted. Consequently, the window length that enhances the detection of pairs of the eigenvalue's spectra can be seen as the optimal value.

8.6 Testing Phase, a Case Study Using Birds Chirps as Wanted Signal

In the third and last phase which is the system validation, it is to bring together the developed SSA algorithms presented in the systematic approach for testing, validation, and critical evaluation. As previously mentioned, this systematic approach includes developing grouping and reconstruction techniques to ensure further improvement with regards to the separability. In this phase, a suitable dataset that links up to the application area of the study using real-world sounds has been adopted.

The standard procedure that has been established in the context of this study will be considered in this final phase for testing and validation of the developed system. Therefore, it is to follow the steps of this procedure in a similar way regardless the sound of the wanted signal. However, the dataset used in the testing and validation phase share same characteristics as explained in Chapter 6. At this stage, birds' chirps have been selected as a wanted signal. The SSA window length optimisation method has been examined for determining the optimal value which might differ upon the dataset.

8.6.1 Window Length Optimisation

For its optimisation, different SSA window lengths will be computed using Equation (8.1). To clarify, when using 100 ms frame size then $L=$ 100 ms and as defined $F_s=$ 44100 Hertz, assuming $n=$ 7 which denotes the required different steps that gives $i = 1, ... , 7$, therefore the suggested window lengths can be calculated to be tested with the SSA algorithm for computing the optimal value. The optimisation method has been developed to fulfil the

conditions set out for the selection of the window length that are concordant with the observation in the other studies as discussed in Chapter 5 and mentioned in the above section. Any number of steps equal or different based on the desired percentages can be selected. Therefore, Equation (8.1) which is one of the significant contributions of this study is used to calculate m_i for the selected values of p_i for $i = 1, ... ,7$, and fill Table 8.1 as below.

Table 8.1.   The calculation of different SSA window lengths in optimisation method

p_i	m_i	Optimal value
15%	662	Testing all these values to select the optimal one or a suitable range
30%	1323	
40%	1764	
50%	2205	
60%	2646	
75%	3308	
90%	2969	

The processing has to be performed for soundtrack samples of the above-mentioned size considering applying the average method. Figure 8.3 shows the average of the eigenvalues measured using thousands of samples of the mixed soundtracks in the dataset that represent the time records of mixed birds' chirps with wind noise for the seven above-mentioned lengths. The optimisation of the window length is based on using Equation (8.1) that has been formulated and applied for the first time in this thesis to develop the optimisation method. Due to its vital role in the embedding stage to generate the trajectory matrix, the eigenvalues ordering along the singular spectrum, and the way that the grouping and reconstruction can be performed, window length optimisation in this study is a major finding. Consequently, this optimisation process has a significant effect on the results particularly with regards to the separability improvement as proved later in this chapter.

The audio files in the database are originally 10s length with a sampling rate of 44100 samples per second. Figure 8.3 also shows different dominant pairs of nearly equal eigenvalues for almost values of m in the range from $\frac{1}{4}N_t$ to $\frac{3}{4}N_t$. This size has too much significance in maintaining most of the variance of the time record by generating the required symmetric covariance matrix. The eigenvalues in the singular spectrum which can be defined as variance peaks can be affected by the window length. To clarify, the adjacent eigenvalues in the singular spectrum in case of small window length might be merged together and represented by one eigenvalue. However, large window length defines a high-resolution case and the variance

peaks might be split into a number of consecutive frequency components. The window length value that enhances the detection of pairs of nearly equal values in the singular spectra and monitors most of the variance has been experimentally obtained. Therefore, as previously mentioned, window length optimisation is a major finding in this thesis.
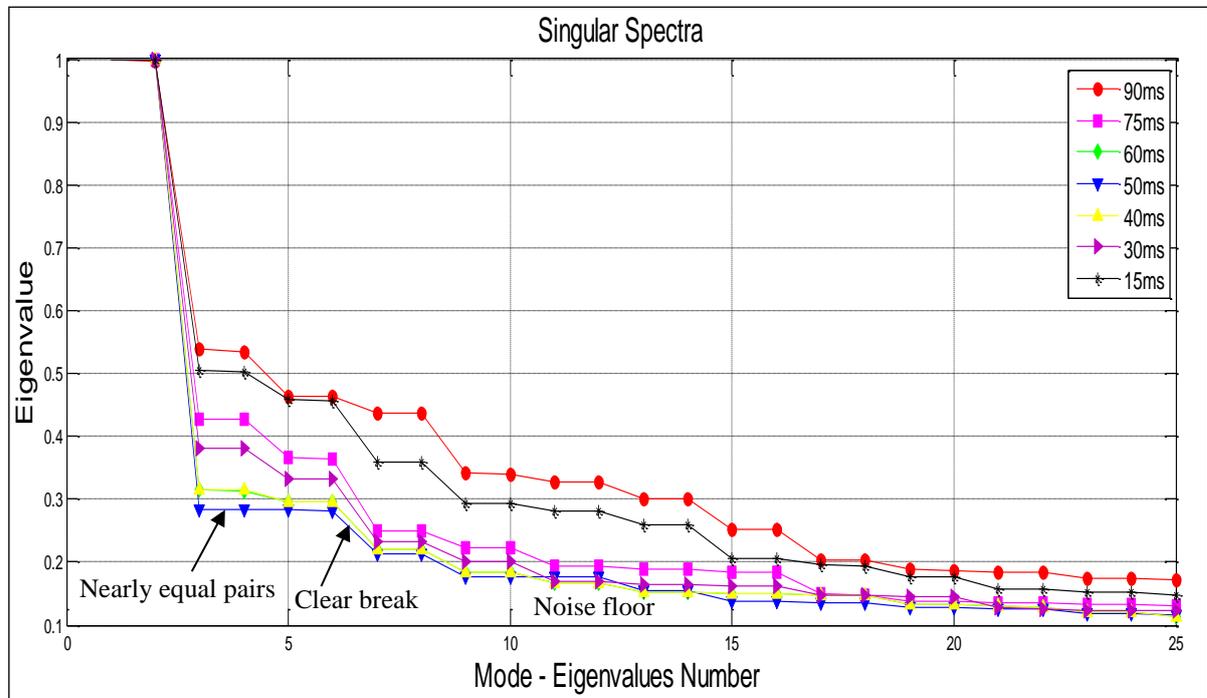
Figure 8.3. Window length optimisation using seven different values for the input signal of the SSA algorithm

More precisely, the eigenvalues spectrum for the window lengths in the range from 40 ms to 60 ms gave apparently concordant results as the first two pairs are considerably satisfy the assumption of obtaining pairs of eigenvalues that have nearly equal values before the noise floor in the higher subspace. This information is valuable when applying the grouping criterion as each pair of the nearly equal values corresponds to an important frequency while the noise is defined in the higher subspace of the eigenvalue's spectrum. In contrast, the other examined window lengths could not satisfy the assumption of determining pairs of nearly equal values. In addition, a definite threshold cannot be specified for the pairs obtained above or below the aforementioned optimal range of window lengths which mainly becomes indifferent when moving away from the optimal range in both sides.

A clear break between these singular values and the others which spread out in a nearly flat noise floor can also be realised. For window lengths above and below this band, indifferent eigenvalues are obtained. It is not strange to find the optimal window around $N_t/2$, however,

there was a general consensus of opinion as most other researchers and authors have noticed. Therefore, based on the above-mentioned observation, reasonable results can be obtained from the indicated range as still the length $N_t/2$ produces two dominant pairs of nearly equal eigenvalues as illustrated in Figure 8.3. Thus, the statistical dimension $d_s = 4$ seems to be the most dominant value in this case since the record is the superposition of oscillations perturbed by wind noise. It is worth mentioning that dominant eigenvalues in the singular spectrum correspond to an important oscillation of the system for each pair of nearly equal as stated in (Elsner and Tsonis, 2013; Yang *et al.*, 2016)

Using values of the window length from the above-mentioned range, approximately same results will be achieved. Selecting higher window length when there is a choice such as in the indicated range is not recommended because this selection is likely to affect the processing time of the SSA algorithms. In fact, the computational load of the SSA is one of the challenges when developing and implementing the SSA particularly with long time records and large dataset. The SSA uses a massive number of matrix multiplications whose complexity increases with the size of the dataset along with the size of the generated matrices that mainly depend on the window length. Therefore, processing the soundtrack samples instead of processing the whole audio file directly by using framing and averaging methods along with the optimisation of the window length helped in reducing the processing time of the SSA algorithms.

8.6.2   Description and Implementation

The applied SSA algorithm in the final phase has been written and developed following the systematic approach developed for the method including all the technical and practical key aspects explained in Chapters from 3 to 5 along with the procedures and dataset explained in Chapter 6 and evaluated according to the measures explained in Chapter 6 as well. Decomposing noisy signals into numbers of oscillations that correspond to desired signals and unwanted wind noise components is a key to success in order to improve the performance of microphone wind induced removal methodologies. In fact, exploiting the SVD in the SSA method and wind noise features as explained in Chapter 2 along with wind noise spectrum of mixed samples to generate groups of oscillatory and wind noise components lead to important observations towards achieving the separability of the components of the given noisy signals.

As mentioned previously, the main objective to make a significant contribution to knowledge is based on developing the method with regards to grouping the SVD components' matrices to transform back to time series expansion from the expansion of grouped matrices towards achieving a proper separability of these components. As stated in (Golyandina and

Shlemov, 2015; Golyandina and Lomtev, 2016; Hansen and Noguchi, 2017), the ability of allocating these components from an observed sum when suitable grouping criterion is applied defines the separability of the components of the time series. However, this approach has been developed and adopted as a central objective to separate wind noise out from the signal.

In the performed experiments in the system testing and validation phase, the SSA is developed and applied for clustering the desired signal oscillation patterns and wind noise components comprised within the mixed recordings of the sample into a number of spaces indicated by the elementary matrices as explained in the previous chapters. These elementary matrices have to be divided into two groups, desired signal and wind noise through the specified grouping criterion developed in the algorithm. The elementary matrices presented within each group are summed up to generate a single matrix; however, in one group such matrix reflects significant oscillations while the matrix in the other group reflects wind noise component.

In the reconstruction stage, each of the final matrices that correspond to a specific group is recovered back to the time domain from the Eigen domain. Accordingly, the mixed soundtracks will be represented by more than one time series as reconstructed wanted signal and residual wind noise, results the separation of wind noise out from the desired signal. Based on the results of the optimisation method, window length $m = N_t/2$ is the optimal value for the given record. Since $N_t = 4410$ for the 100 ms frames, the window length is therefore equal to 2205. However, this selection can generally be made based on SSA recommendations in the case of harmonic or oscillatory components (Traore *et al.*, 2017). Consequently, after performing the SVD of the trajectory matrix, the most dominant eigentriples ordered by their contribution in the decomposition can be obtained.

From the plot of logarithms of the singular values shown in Figure 8.3, a significant drop in values can be seen around component 6 which indicates the start of the noise floor. Therefore, three obvious pairs can be considered as with almost leading singular values. Basically, to perfume the experiments, the mixed soundtracks have been produced from mixing wind noise with the clean birds' chirps at different SNR to obtain noisy singles. The normalisation has been considered to the mean at scale 1 in the singular spectra for all the calculated averages to make a valid and fair comparison between the two signals. Figure 8.4 shows the eigenvalues arranged in descending order in the eigenvalue's spectra of a decomposed signal of birds' chirps corrupted with wind noise and a clean one.
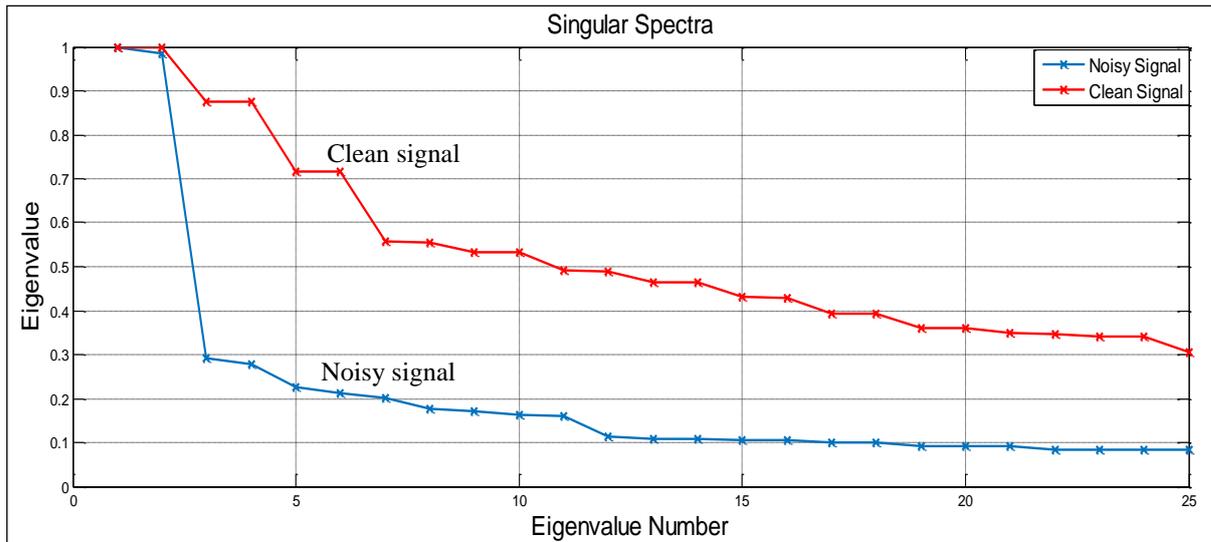
Figure 8.4. Singular spectra of birds' chirps clean and noisy records (wind noise added)

The eigenvalues spectra shown in Figure 8.4 represent the average of both clean and noisy signals. In the corrupted signal, only first few of the eigenvalues carry large amount of energy. The first pairs of eigenvalues, however, are the ones with less correlation. The high correlation ones are those which left behind in the higher subspace and generally represent noise. As illustrated in Figure 8.4, the distribution of the eigenvalues of nearly equal pairs along with equal location of the eigenvalues within the pairs themselves over a specified threshold can be clearly seen in the clean signal. However, this indicates a normal ordering of the eigenvalues in the absence of noise. On the other hand, different ordering with only few of nearly equal eigenvalues can be seen in the noisy signal. The grouping technique is developed based on defining such aspects for the best selection of the most dominant pairs of eigenvalues.

8.6.3   Results and Discussion

To separate wind noise component out and retain the signal of interest, fewest numbers of eigenvalues in the lower subspace before the noise floor will be examined to establish the separation boundaries as explained in complete details in Chapter 5. This procedure is based on the contribution of the principle components to the wanted signal and wind noise according to admissible orthonormal bases of eigenvectors. Basically, from the calculation of the singular spectra for the given noisy signal which is a superposition of oscillations of birds' chirps perturbed by wind noise in this case study, the statistical dimension $d_s = 4$ has been determined as the most dominant value. As previously mentioned, the optimal value of the window length has been determined as $m = N_t/2$ from the optimisation method.

The leading four principle components have been used for grouping and reconstruction of the desired components that represent birds' chirps form a noisy signal with additive wind noise as shown in Figure 8.5. From this time domain representation shown in Figure 8.5, it can be clearly seen that the first two pairs of eigenvectors correspond to important oscillations represented by their corresponding principle components. However, birds' chirps signal can be reconstructed based on the selection of these pairs in the grouping step. It is worth mentioning that the experiment was performed using the mixed soundtracks that have been produced from five thousand of samples in the dataset and the input of the SSA algorithms represents the average so as the output which is the reconstructed signal. The separability cannot be clearly indicated with time domain representations, but when using different evaluation measures, such as *w*-correlation matrix, as explained later in this chapter.
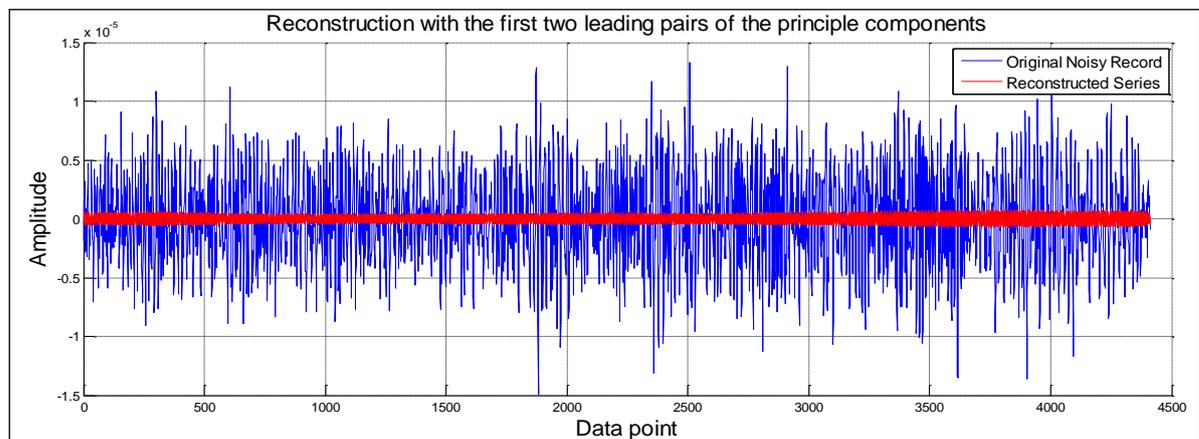
(**a**)
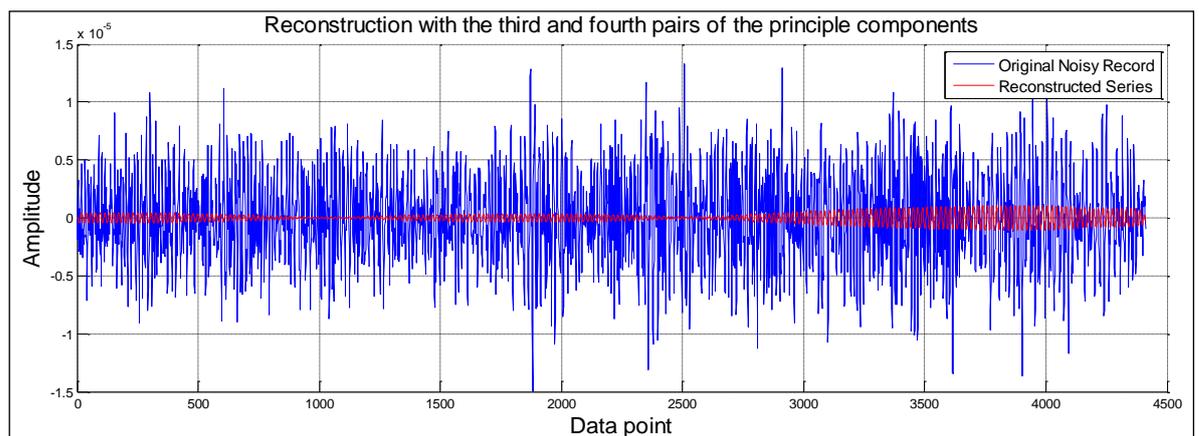
(**b**)

Figure 8.5. The leading four principle components used for grouping and reconstruction of birds' chirps record with additive wind noise added: (a) Reconstruction with the first two leading pairs; (b) Reconstruction with the third and fourth pairs

The phases are in quadrature and regular changes in amplitude are obviously present for the reconstruction with the first two leading pairs of principle components as shown in the top part of Figure 8.5. In contrast, for the third and fourth pairs shown in the bottom part of Figure 8.5, there is a slight coherent phase relationship between their two eigenvectors. To reconstruct the desired signal while separating the wind noise out, the pairs of nearly equal eigenvalues are considered dominant while the others which spread out in a nearly flat noise floor are not of interest, or otherwise they correspond to wind noise components. However, the eigenvalues associated with the ones represent the higher-order values are mostly located in the noise floor or the higher subspace of the singular spectra as they are of low variance. Figure 8.5 also shows the original noisy signal of the give record in this case study.

The principal components can be computed using the eigenvectors by projecting the embedded time series onto the individual eigenvectors. In the principle components matrix; each column vector represents a separate *PC*. The principal component pairs correspond to the most dominant eigenvectors consist of clean structures of the signal are in marked contrast to the next pairs of principal components, which are noisy with low amplitudes. The grouping therefore has been performed when clearly setting the separation boundaries between the subspaces in the spectra and defining a specified threshold as explained in Chapter 5.

As known, a pure noisy time record typically produces a slowly decreasing sequence of singular values which is the start of the noise floor; however, the selection of such higher-order eigenvalues will only produce noisy signals. Therefore, it has been experimentally found that selecting eigenvalues beyond a prescribed limit that precisely indicates the boundaries in the singular spectra based on setting defined constrains and imposing certain conditions on such allocation as explained earlier, will lead to add portion of the noise to the signal. Eventually, to produce the reconstructed components of the time record, the principle components are then projected onto the orthogonal matrix of the eigenvectors.

The developed grouping technique applied in the SSA algorithms in the experiments performed in this system testing and validation phase is to group pairs of the resultant elementary matrices in two groups I and $\bar{I}$ as explained in Chapter 5 (section 5.4). Recall from Chapter 5, dividing the resultant elementary matrices in two groups is mathematically expressed in Equation 5.3 for specifying the threshold and Equation 5.4 for identifying the groups. Group I consists of the desired elementary matrices after aggregating them together to implement each group in a single matrix of the same dimension as of the trajectory matrix.

However, group $\bar{I}$ contains unwanted components which reflect the wind noise. Restoring to the time domain is now possible by applying the diagonal averaging as described in Chapter 4.

Figure 8.6 illustrates a comparison of the noisy record (bird chirps and wind noise), with the reconstructed one. The differences between reconstructed series and original noisy signal are highlighted in the top part of Figure 8.6. It is clear from the bottom part of the figure, a considerable amount of noise (denoted as residual series) was separated out. However, the findings revealed that the de-noised signal resembles the clean one. As seen above, the SSA can readily extract and reconstruct periodic components from noisy time series.
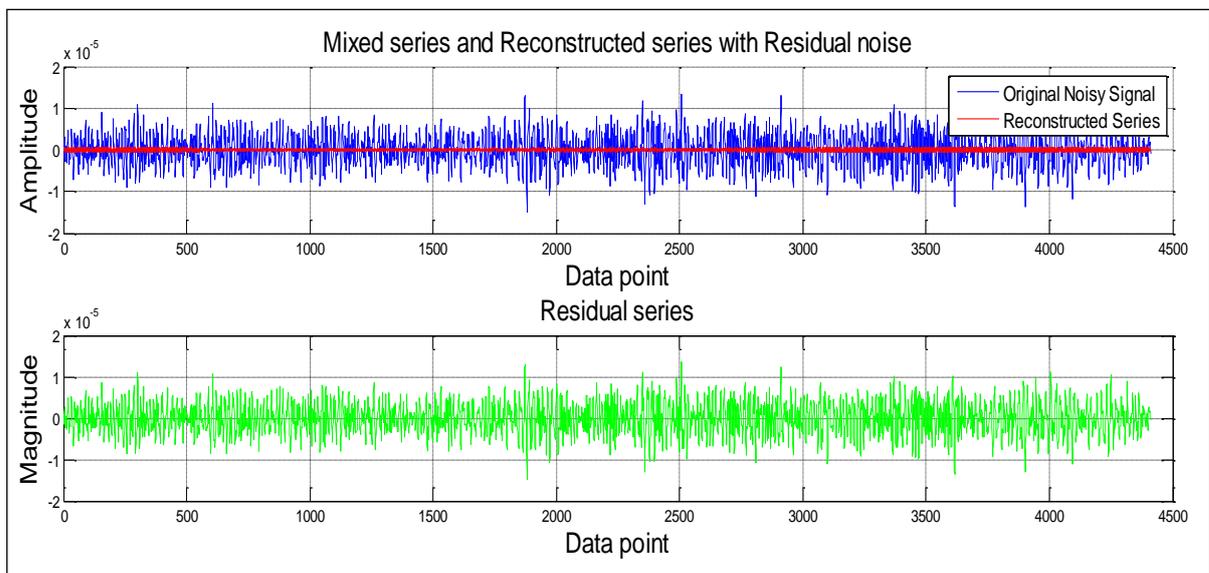
Figure 8.6. Reconstructed series vs original noisy signal and residual series

A combination of the signals used in the experiment and presented in the time domain is shown in Figure 8.7. These signals are the original signal (birds' chirps) which is denoted as clean signal shown in the tope part Figure 8.7 (a), the mixed signal which represents bird chirps and wind noise shown in Figure 8.7 (b), and the reconstructed signal shown in Figure 8.7 (b). Sound pressure level (SPL) measurements are also shown in Figure 8.7. The single most significant observation to emerge from the data comparison was that the SSA can separate wind noise out and reconstruct the desired signal which is birds' chirps in this case. However, these results will be proved when evaluating the performance of the developed SSA system using sound analysis methods, DSP measures, and *w*-correlation matrix as the time domain representation cannot clearly indicate to what extent the separability can be achieved and the desired signals can be retained with no reconstruction errors.

(**a**)

(**b**)

(**c**)
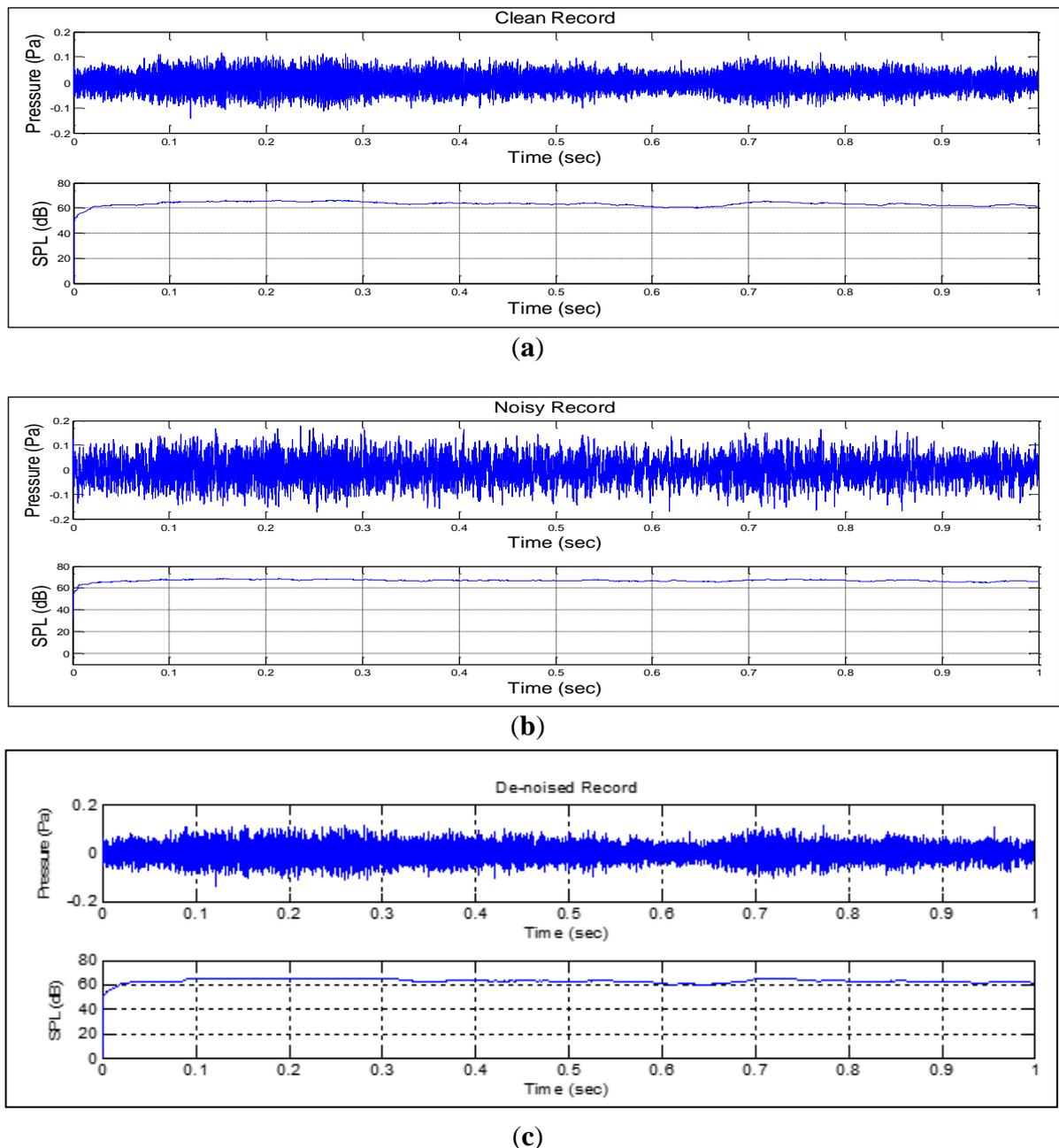
Figure 8.7. A combination of the signals used in the experiments: (a) Clean record of birds' chirps; (b) Noisy record (birds' chirps and wind noise); (c) The de-noised record separated by retaining only the leading two pairs of the eigenvalues

To assess the effect of noise on a signal, the signal-to-noise ratio (SNR), which is an objective measure, is generally used. This measure can be used jointly with some other criteria to evaluate the output of a noise reduction algorithm. It was applied to evaluate the developed SSA system for wind noise separation in this study. The SNR ratios are calculated using standard definition before the removal process. After the separation process, signal levels and noise levels have been estimated. The SNR ratios are calculated using Equation (8.2) (Crocker, 2007; Vaseghi, 2008; Brandstein and Ward, 2013; Arul Elango, Sudha and Francis, 2017).

$$SNR = 10\log\frac{\sum_{n}\hat{s}(n)^2}{\sum_{n}\hat{w}(n)^2}, \tag{8.2}$$

where $\hat{s}(n)$ is the estimated signal, $\hat{w}(n)$ the estimated noise.

Results of using the SNR measure are given in Table 8.2 for comparison. A notable improvement for an average of about 9dB as it is apparent from Table 8.2 has been achieved. The result shown in Table 8.2 is a reported average of many cases.

Table 8.2. SNR measure applied for evaluating the developed SSA method for wind noise reduction, measurement cases and difference as a reported average.

The Objective Measure for Evaluating the Method	Before	After	Difference
SNR in dB	0dB	9.47dB	9.47dB

8.7 Validation Phase, a Case Study Using Alarms as Wanted Signals

As mentioned previously, in the system validation phase, however, the developed systematic approach with regards to many key elements in the method has been adopted in the empirical study using real-world sounds. Therefore, the algorithm has reached its final stage after regular updates and verification to meet the final requirements and specification set based on deep experimental investigations and testing. Regarding the dataset used in this phase, environmental sounds, have been used as specific signals of interest. In addition to the sound of birds' chirps, some other interesting signals such as car sirens (e.g., police car siren), sound of alarms (e.g., fire alarm) are given as examples in the validation phase.

As previously mentioned in Chapter 6, the dataset used in this study has been prepared for testing and validation purposes according to clear rules. The samples in the dataset share same characteristics and the mixed soundtracks have been prepared using the same model and under the same conditions using the same wind noise dataset. The defined testing criteria and the experimental procedure along with all the above defined technical approaches have been used in this validation phase. The developed SSA system has been validated and evaluated using different real-world sounds and desired signals.

8.7.1 Defining the Embedding Dimension in the Algorithm

The embedding dimension parameter or the window length has to be defined in the SSA algorithm by its optimal value. Window length optimisation is always seen as a preliminary

step to select the optimal value to be defined in the SSA algorithm instead of trial and error method, however, this gives more accurate results although as discussed earlier the value $N_t/2$ can mostly be considered as a common practice.

To plot the singular spectra for comparison, Equation (8.1) has also been used to calculate different window lengths similar to those calculated in the testing phase as provided in Table 8.1 in (section 8.6.1). These window lengths will be used to generate trajectory matrices for each one from a frame of the same size as explained in the previous sections. However, the experiments in the system validation phase have to be carried out under strictly controlled conditions to validate the method and make a fair and valid comparison.

As the aim is to map the soundtrack vector in the embedding step into a trajectory matrix using a specified window length that might differ depends on the used dataset. However, to find this specific embedding dimension, the way of distributing the eigenvalues in the eigenvalue's spectra considering the lower subspace has to be compared. These eigenvalues are produced from the SVD of the covariance matrices for each proposed embedding dimension for $m \le N_t$. The distribution of the eigenvalues in this subspace depends on the type of the dataset itself and the selected embedding dimension from which the statistical dimension can be defined. The soundtrack samples have been processed using the framing method with a frame size as mentioned earlier and then applying the average method. The average of the eigenvalues has been measured using thousands of samples of the mixed soundtracks in the dataset that represent the time records of mixed wanted signals with wind noise for the seven window lengths as in the testing phase.

Figure 8.8 shows the values proposed for the window lengths presented in the lower subspace of the eigenvalues spectra and taken as a percentage of the length of the time records N_t measured from the applied framing method. Two cases have been selected to be presented in this validation phase which are police car siren and fire alarm as wanted signals. The optimisation has to be accomplished with SSA input signal which is a mixed noisy signal; however, it can be at different considerable SNR such as 0dB or -10dB for example.

For the eigenvalue's spectra shown for both cases at different embedding dimensions as in Figure 8.8, it is to focus on dominant pairs of nearly equal eigenvalues which can be seen for almost values of m in the range from $0.4N_t$ to $0.6N_t$ for both cases. Consequently, the embedding dimension that enhances the detection of pairs of nearly equal eigenvalues is the optimal window length as well as this is one of the key aspects in grouping technique. This

specific embedding dimension should help in maintaining most of the variance of the time record by generating the required symmetric covariance matrix.

(**a**)

(**b**)
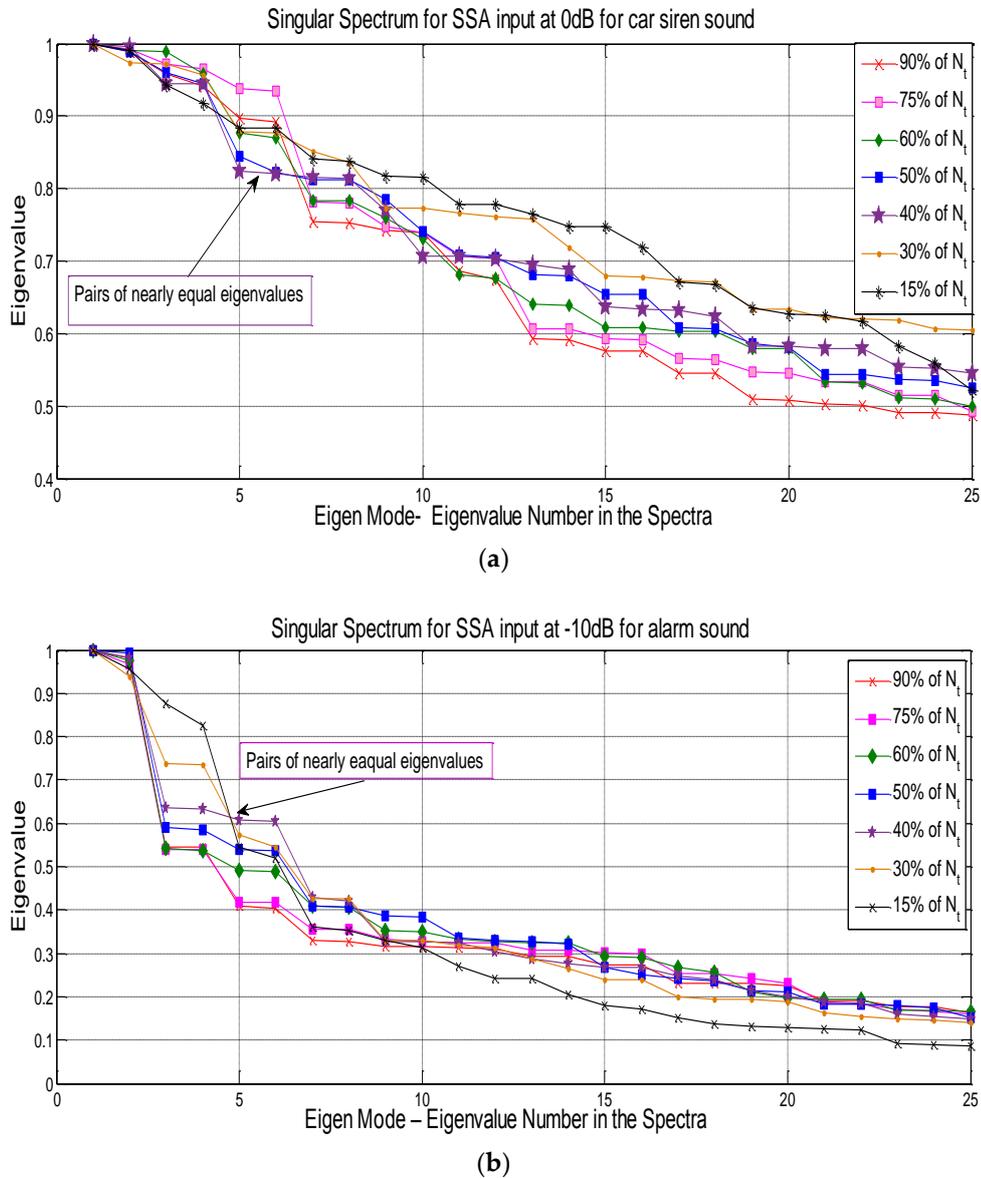
Figure 8.8. Comparison between different values of *m* for SSA input signal at 0dB for case (a) and −10dB for case (b) to obtain and define the optimal value in the algorithm: (a) Police car siren; (b) Alarm sound

From the lower subspace of the eigenvalues spectra shown in Figure 8.8 for the selected seven window lengths, window lengths in the range from 40% to 60% of N_t gave apparently concordant results as the first pairs are considerably satisfy the assumption of obtaining pairs of eigenvalues that have nearly equal values. As mentioned previously, dominant eigenvalues in the singular spectra correspond to an important oscillation for each pair of nearly equal values. Each pair of the nearly equal values therefore correspond to an important frequency; however, this information is valuable in the grouping step. In contrast, the assumption of

determining pairs of nearly equal values could not be fulfilled with the other examined window lengths. For example, in case of considering 15% of the length of the time record, even a single pair cannot be identified as shown in bottom part of Figure 8.8 for the case of the alarm sound. Furthermore, a definite threshold cannot be specified for the pairs obtained above or below the above-mentioned range of window lengths. In fact, moving away in both sides from this range, which can be seen as an optimal range, defining specific threshold becomes indifferent.

Checking clear breaks between the singular values presented in the lower subspace and the others presented in the higher subspace that represent the noise floor is another important aspect. However, clear breaks can be realised in the indicated range, whereas, indifferent eigenvalues that distributed randomly are obtained above and below this range. Based on such calculations, it can be observed that the embedding dimension from the above indicated range can help in producing dominant pairs of nearly equal eigenvalues from which positive results can be obtained. Accordingly, the statistical dimension, which is one of the constrains required for defining boundaries in the singular spectra in the grouping criterion as explained in Chapter 5, can be computed based on the above-mentioned observation. To sum up, the optimal selection of the embedding dimension can also lead to better applying the grouping technique.

8.7.2 Description and Implementation

This phase is aiming at validating the developed system by means of implementing the SSA algorithm which has been written following the systematic approach developed for the method. The experimental procedure has been established for conducting multiple experiments and case studies. The main objective to achieve certain separability is to distinguish between wind noise component and other components that represent numbers of oscillations correspond to the wanted signals. This can be accomplished by decomposing the SSA input noisy signal into additive components after mapping the soundtrack vector into a trajectory matrix using a defined length of the sliding window. Therefore, it is to apply the SVD for obtaining the eigentriples required for grouping along with using wind noise features and wind noise spectrum of mixed samples.

As mentioned previously, the adopted systematic approach developed for the SSA method within in the context of this study was mainly based on deeply reviewing the literature concerning all the key aspects in the method. In this approach, it has been defined that each principle component vector which is corresponding to eigenvalues of nearly equal values has a unique oscillation pattern. In fact, the principle components are produced by projecting the

two-dimensional matrix onto a new set of axes which are the eigenvectors associated to those eigenvalues. The *PCs* are similar to the embedded time series; however, they are presented in a different coordinate system.

In all cases, the mixed soundtracks used in the experiments are composed of wind noise and wanted signal. When considering the wind noise spectrum and harmonics that exponentially spaced in terms of frequency, the *PCs* can be determined according to their contribution to wind noise and wanted signal. The oscillation patterns of the wanted signal are presented in the elementary matrices and as a consequence in the principle components in a different way to the wind noise components. Therefore, summing up each corresponding number of elementary matrices together leads to generate the required groups. Errors arising from adding a portion of the desired signal to the noise or part form the noise to the wanted signal, which might happen in some cases, can be avoided and then reducing any undesirable effect on the separability when applying all the constrains described in Chapter 5. As previously mentioned, the separability might differ based on the dataset.

Considering all these aspects can lead to generate groups of oscillatory and wind noise components to make important observations towards achieving the required separability. After going through the whole processes and stages defined in the SSA algorithm for processing mixed soundtracks, the system output when returning from Eigen domain to time domain will be represented in two time series which are reconstructed desired signal and residual wind noise. Detailed description of the method implementation regarding some other aspects as described in the algorithm has been given in section (8.6.2).

As indicated by the above mentioned range from $0.4N_t$ to $0.6N_t$, the window length $m = 0.4N_t$ has been selected as an optimal value for the given record based on the optimisation explained above for both cases. As mentioned previously, this value is nearly concordant with the SSA recommendations which is $m = 0.5N_t$. However, approximately same results can be achieved when using values of the window length from the above-mentioned range. Since $N_t = 4410$ for the 100 ms frames, the window length is therefore equal to 1764 as calculated using Equation (8.1). However, this value has been defined in the algorithm to map the frames into the trajectory matrix when using the average method.

The significant drop in values indicates the start of the noise floor that can be seen from the plot of logarithms of the singular values. For example, what stands out in the bottom part (b) of Figure 8.8 in the case of alarm sound is the significant drop in values that can be seen

around component 6. The results obtained from the preliminary analysis of the eigenvalues spectra using the SSA visual tools are shown in Figure 8.8, however, three obvious pairs can be considered as with almost leading singular values. For proper grouping, it is important to select the most dominant eigentriples obtained from the SVD method that must be ordered by their contribution. As described in the case study presented in the testing phase, the normalisation to the mean at scale 1 in the singular spectra for all the calculated averages has been considered to make a valid and fair comparison between the different signals used in the experiment. Figure 8.9 shows the eigenvalues (arranged in descending order) of a decomposed signal of police car siren (a) and alarm sound (b) which both corrupted with wind noise along with the original clean signal for each.
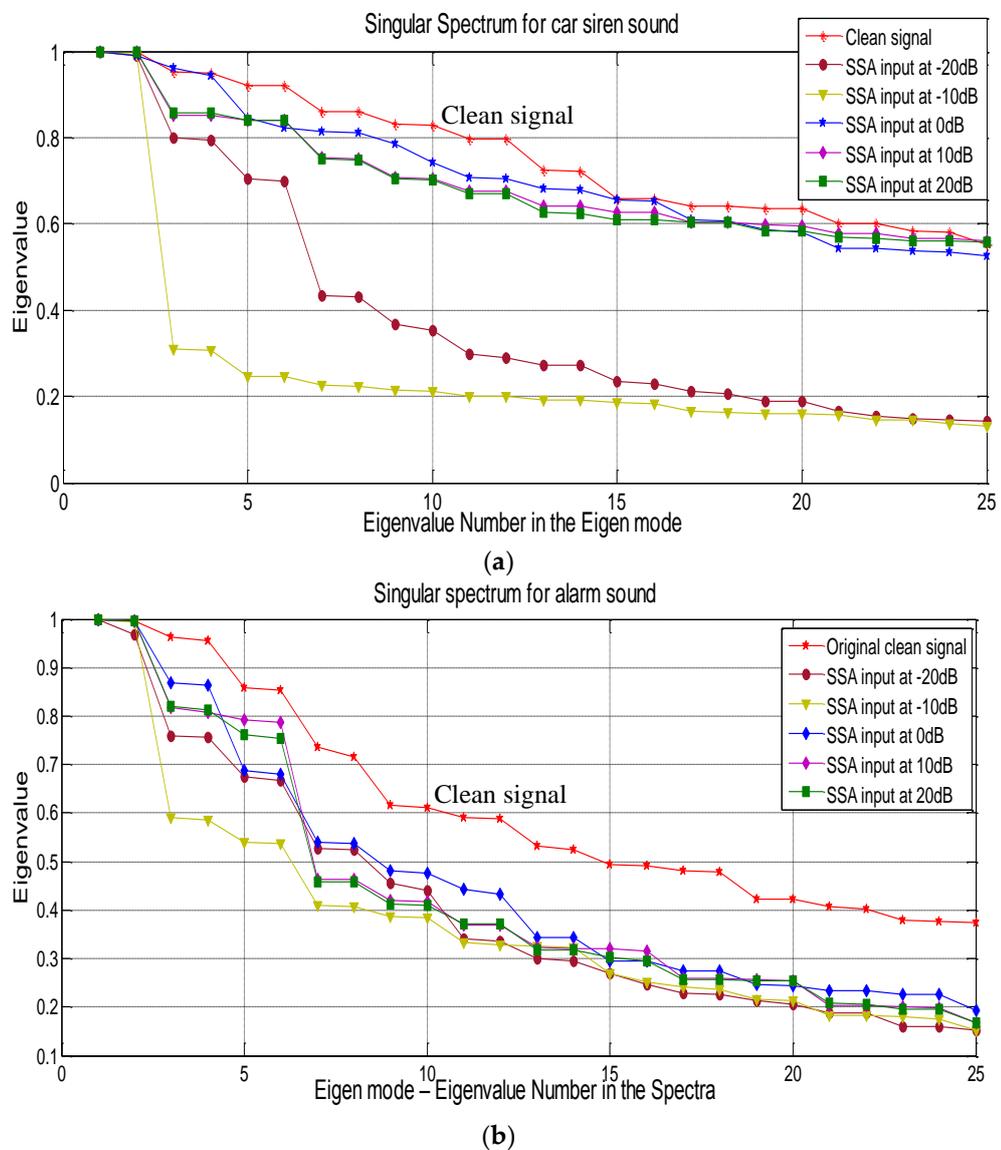

Figure 8.9. Singular spectra of clean and SSA input noisy records of the two selected cases (a) Police car siren; (b) Alarm sound in the validation phase with wind noise added at different SNR for $m = 0.4N_t$

In Figure 8.9, the comparison is made when using $m = 0.4N_t$. The figure compares the eigenvalues spectra for the clean signal and the SSA input noisy signal at different SNR mixing ratios. One of the aims of using different SNR values in producing the mixing soundtracks is to study the effect of wind noise content on the distribution of the eigenvalues in the singular spectra. From Figure 8.9, significant differences between the noisy and clean signals can be noticed. As mentioned previously, the significant drop and slowly decreasing sequence of singular values typically indicates the noise floor.

Generally, concordant with the observations of other studies, the dominance of the first two or three pairs of eigenvalues are noticeably greater than any other values. In the noisy signal, only first few of the eigenvalues carry large amount of energy. The first pairs of eigenvalues in the lower subspace, however, are the ones with less correlation. The high correlation ones are those which left behind in the higher subspace and generally represent noise. Looking at Figure 8.9, it is apparent that the eigenvalues spectra of the clean signal shows multiple eigenvalues with nearly equal values with gradual descending in order which is not the case for the noisy signals. The SNR mixing ratio of the SSA input signals indicates how strong the effect of wind noise on the distribution of the eigenvalues in the singular spectra.

8.7.3 Results and Discussion

As the main aim of this study is to separate wind noise components out and retaining the desired signal, however, the statistical dimension that represents fewest numbers of dominant eigenvalues before the noise floor will be examined to establish the separation boundaries as clarified in Chapter 5. This is based on the contribution of the principle components to the wanted signal and wind noise according to admissible orthonormal bases of eigenvectors. When fixing the boundary in the singular spectra to define the starting point of the noise floor by means of determining the appropriate statistical dimension, the principle components in the higher subspace should contain no much information. Therefore, the objective is to separate all the principle components that correspond to eigenvalues indexes greater than the defined statistical dimension. The projection of the trajectory matrix onto the eigenvectors can be interpreted as a transformation process of the trajectory matrix with reference to the selected eigenvectors in a way that the trajectory matrix is multiplied by one representative eigenvector at a time.

Basically, for the value of window length $m = 0.4N_t$ defined in the algorithm for calculating the singular spectra of the record used in this case study as shown in Figure 8.8, it

is apparent that the first pairs of eigenvectors correspond to important oscillations. The desired signal can be reconstructed in the grouping step based on the selection of these pairs to produce dominant pairs of principle components *PCs*. Dominant pairs selected according to admissible orthonormal bases of eigenvectors can show regular changes in the amplitude of the oscillations as well as quadrature phase. In contrast, for other pairs, there is slight coherent phase relationship between their two eigenvectors.

Figure 8.10 shows the reconstruction of the wanted signal using the dominant leading *PCs* and the original SSA input noisy signal of the produced soundtracks (fire alarm sound and wind noise) mixed at 0dB and -10dB as examples.

(a)

(b)

Figure 8.10. The dominant leading principle components used for grouping and reconstruction of the alarm sound record with additive wind noise: (a) Reconstruction for SSA noisy input signal at 0dB; (b) Reconstruction for SSA noisy input signal at −10dB

As mentioned previously, based on the defined aspects and grouping constrains, pairs of nearly equal eigenvalues in the eigenvalue's spectra are always considered dominant. However, others particularly those which spread out in a nearly flat noise floor do not contain much

information, or otherwise they are corresponding to wind noise components. The higher subspace mostly represents the eigenvalues of low variance, whereas a slowly decreasing sequence of singular values is typically produced by a pure noise time series (Golyandina and Lomtev, 2016).

Similarly, Figure 8.11 shows the case of selecting police car siren as a wanted signal. As known, the calculation of the principal components is by projecting the embedded time series onto the individual eigenvectors. However, each column vector represents a separate *PC* in the principle components matrix. It is worth mentioning that the principal component pairs correspond to the most dominant eigenvectors consist of clean structures of the signal are in marked contrast to the other pairs of principal components, which are noisy with low amplitudes. Clustering the elementary matrices produced from the SVD in different groups can be performed by clearly setting the separation boundaries between the subspaces in the singular spectra and defining a specified threshold as explained in Chapter 5.
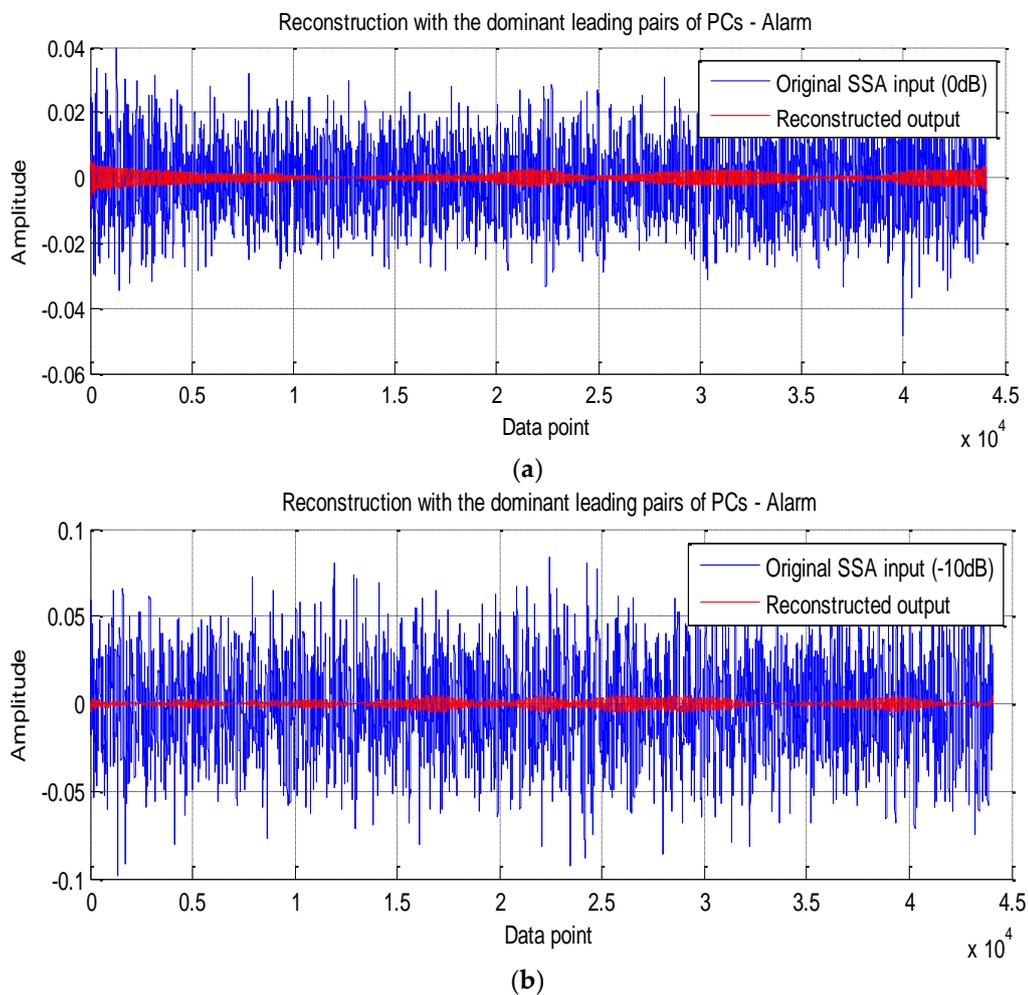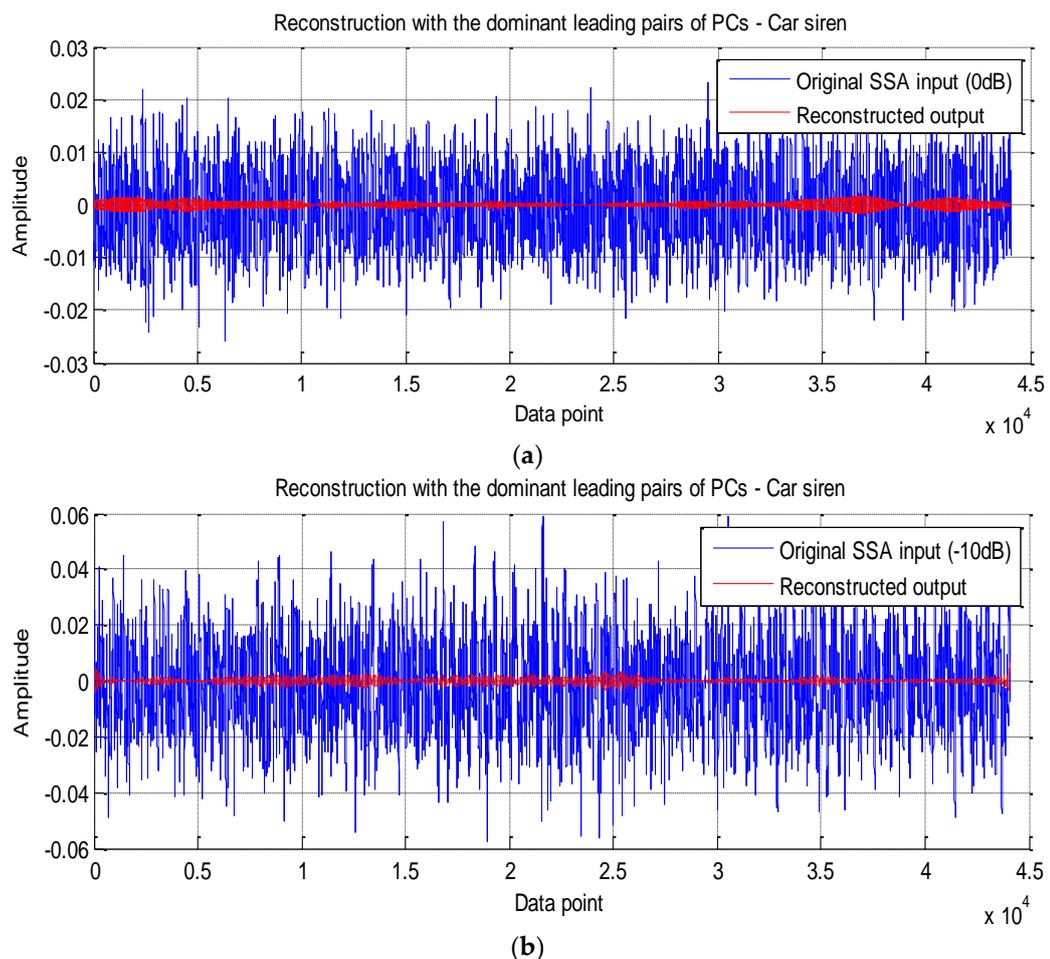

Figure 8.11. The dominant leading principle components used for grouping and reconstruction of the record of police car siren with additive wind noise: (a) Reconstruction for SSA noisy input signal at 0dB; (b) Reconstruction for SSA noisy input signal at -10dB

The results obtained as an output of the SSA algorithm are shown in Figure 8.12 which compares between the original SSA input noisy record and the reconstructed output signal for the alarm sound used in this case study. It can be seen from Figure 8.12 that the original noisy input signal is plotted in time domain with the reconstructed desired signal for comparison purposes. However, the reconstructed series is obtained by using the most dominant principle components as clarified above, whereas the SSA input signal represents the standard case used to report the results which is at 0dB SNR along with −10dB SNR for comparison.

Figure 8.12. Mixed SSA input at 0dB and −10dB and reconstructed output series with the residual unwanted series for alarm sound: (a) SSA input signal at 0dB; (b) SSA input signal at −10dB

The output audio signals of the SSA algorithm are hearable which can also be subjectively evaluated, however, the SSA shows better output sound and performance in reconstructing the desired signal for positive SNR mixing ratios compared to the negative ones with a considerable amount of wind noise removal in all cases.

Similar to the case of fire alarm sounds, Figure 8.13 shows the results of the case of police car siren. The differences between reconstructed series and original noisy signal are highlighted in the top part (a) of Figure 8.13. Looking at the bottom part (b) of Figure 8.13, a considerable amount of wind noise (denoted as residual series) was clearly separated out. However, the findings revealed that the de-noised signal resembles the clean one. As can be seen from Figure 8.13, the SSA can readily extract and reconstruct periodic components from noisy signals contaminated with wind noise.
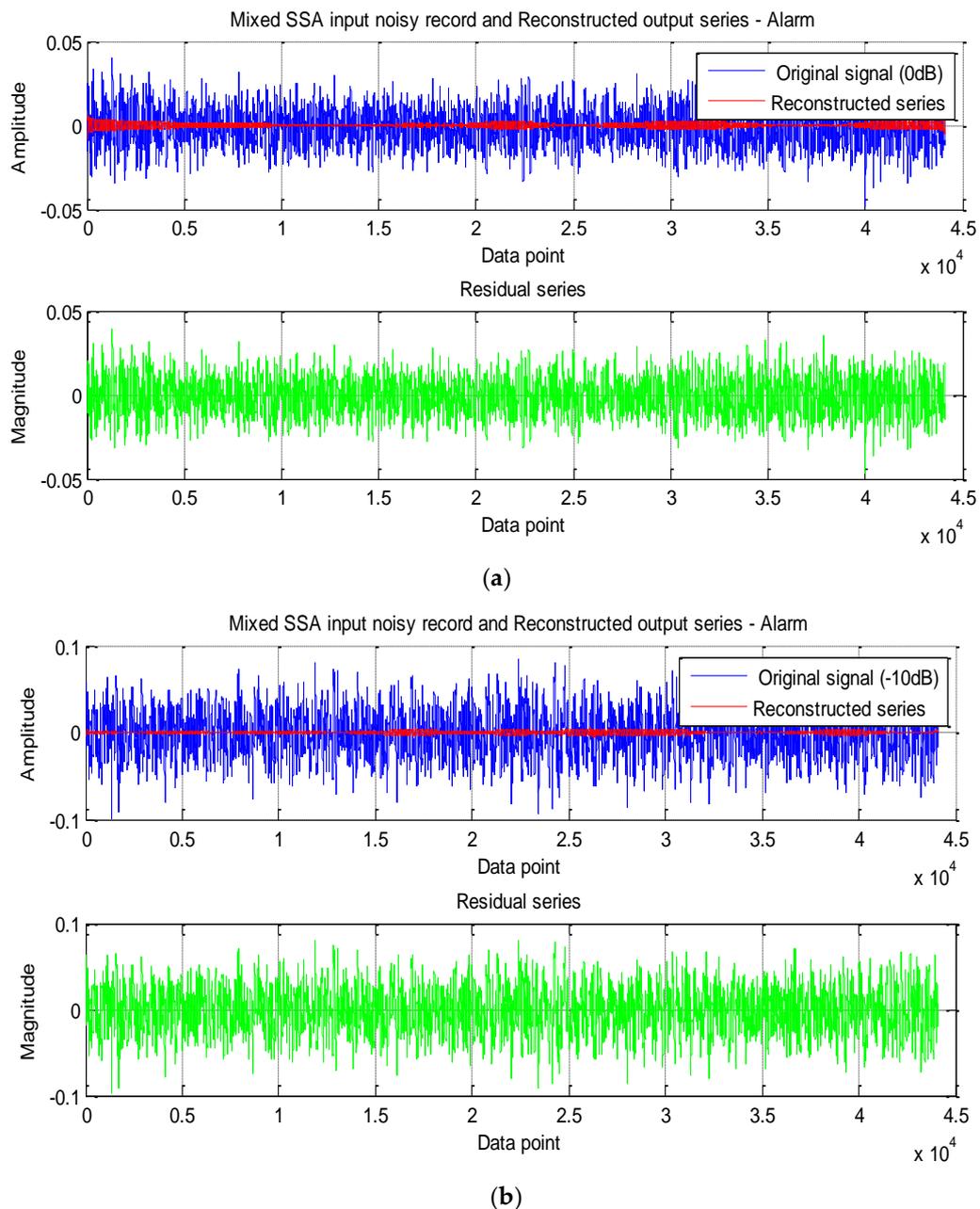

(**a**)

(**b**)
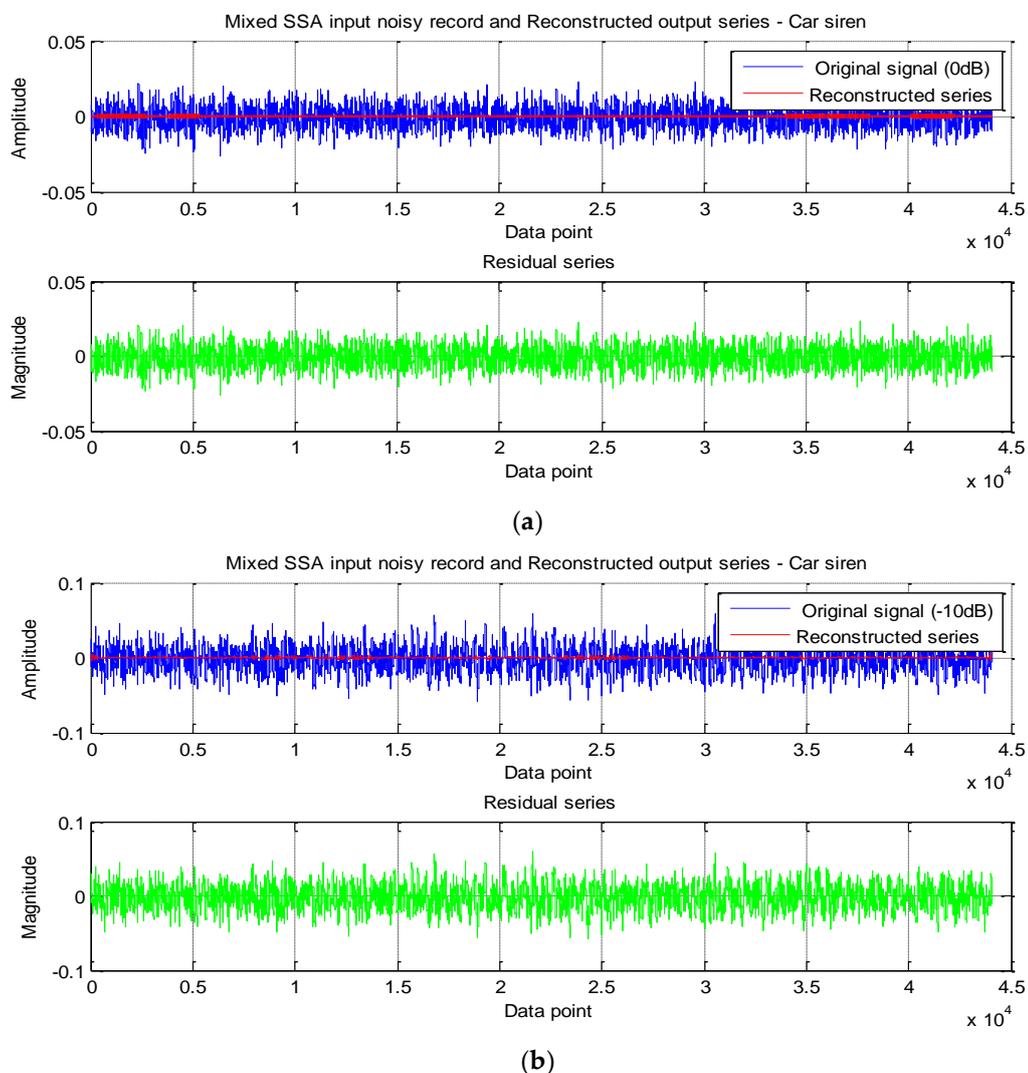
Figure 8.13. Mixed SSA input at 0dB and −10dB and reconstructed output series with the residual unwanted series for car siren sound: (a) SSA input signal at 0dB; (b) SSA input signal at −10dB

A combination of the main types of the signals used in the experiments conducted in the validation phase presented in the time domain with the SPL measurement and given as examples for both case studies are shown in Figure 8.14
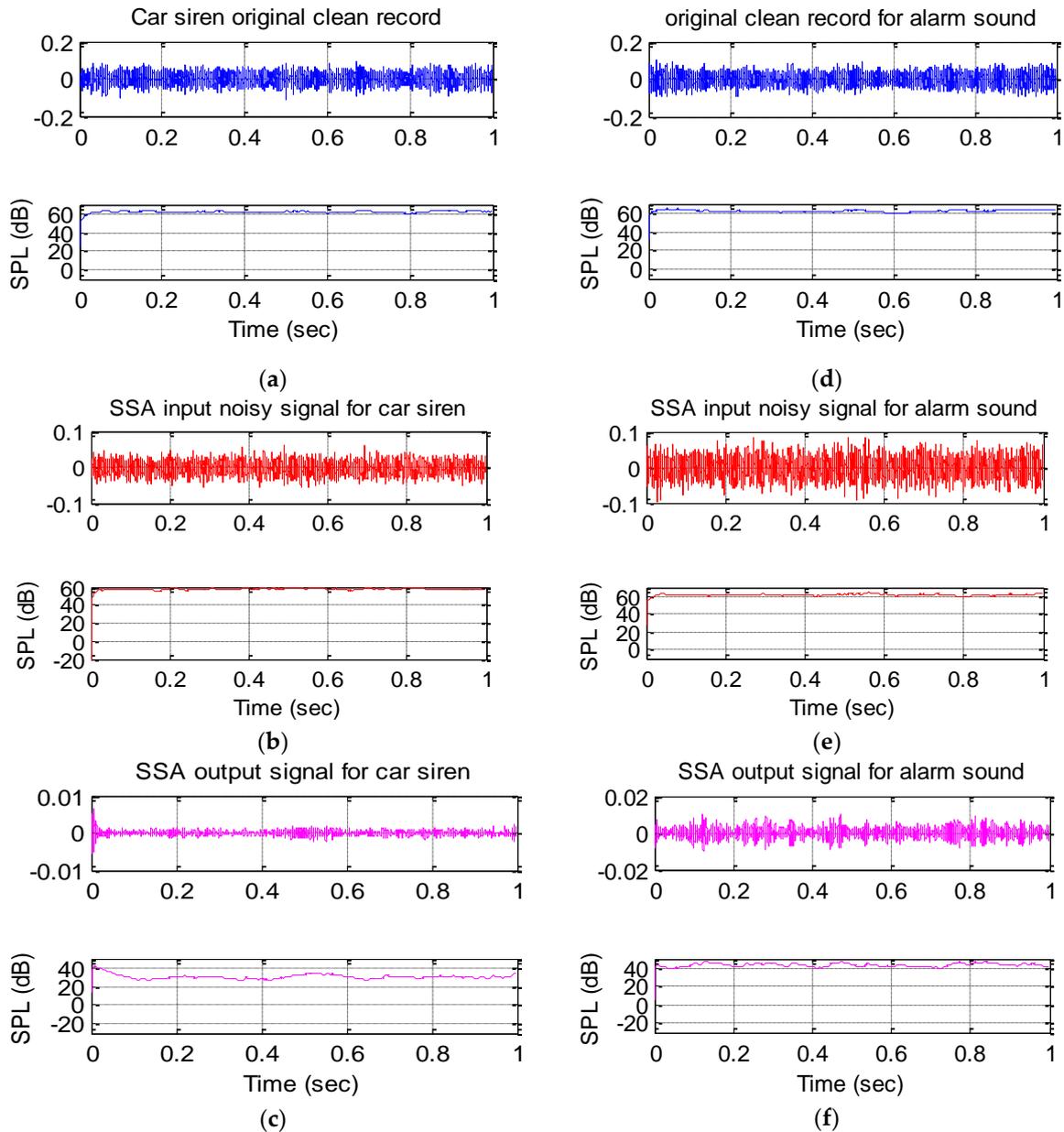
Figure 8.14. A combination of the signals used in the experiments: (a), (b), (c) for the case car siren; (d), (e), (f); for the alarm signal

It is worth mentioning that the grouping technique has been developed to generate two groups I and $\bar{I}$ for gathering the resultant elementary matrices produced from the SVD as detailed in Chapter 5 (section 5.4) and mathematically expressed in Equation (5.3) for specifying the threshold and Equation (5.4) for identifying the groups. The oscillation pattern of the desired signal is presented in the elementary matrices and consequently in the principle

components in a different way to the wind noise components. Hence, in the grouping step, it is to sum up each corresponding number of elementary matrices together to generate the required groups. Desired elementary matrices have been grouped in group I after aggregating them together to implement each group in a single matrix of the same dimension as of the trajectory matrix. Whereas, undesired components corresponding to wind noise have been grouped in group $\overline{I}$. Eventually, the projection of the principle components onto the orthogonal matrix of the eigenvectors produced the reconstructed components of the given time record.

Strong evidence was found when the SSA can separate a significant amount of wind noise out from the SSA input noisy signals and reconstruct the desired signals in all the cases. What stands out from the results presented in the figures shown above is that the SSA can work towards a satisfactory solution of the problem of wind induced microphone signals addressed in this study. As the focus of this study is on the separation of wind noise, results of the critical evaluation phase will demonstrate the improvement of the separability of the SVD components. In-depth analysis will be given in the next section when critically evaluating the developed SSA system.

8.8 Critical Evaluation

The third phase of implementing the developed SSA system was performed with bigger dataset of real-world sounds for testing and validation process as explained earlier in this chapter. The objective measures explained in Chapter 6, which are signal processing measurements and sound analysis methods, were mainly used for evaluating the performance of the developed SSA method and measuring its effectiveness with regards to wind noise separation. Importantly, the *w*-correlation matrix which is an important measure to indicate the separability was used in the critical evaluation process.

8.8.1 Objective Evaluation

To make a comprehensive evaluation of the performance of the algorithm and measure the effectiveness of the developed SSA as a method of wind noise separation, at this stage, an example from the large number of case studies conducted in the system testing and validation phase using realistic samples has been selected. In this section, the results and discussions of the critical evaluation process of the developed SSA method using common approaches as explained in Chapter 6 are presented. For instance, sound analysis method and different signal processing measures have been used to present experimental findings to provide practical

assistance to the evaluation process. Therefore, this section consists of some tables and figures along with comments on the significant data shown in these.

It seems impossible to present all the results of all the experiments and case studies and even for a single case due to the different experimental conditions required for performing the experiments in a convenient and effective way. For example, mixing the soundtracks with different SNR values, the reconstruction with different principle components to give evidence in analysing and evaluating the results, and the variety of the desired signals are some of the experimental conditions and testing requirements. Therefore, the systematic approach that has been established based on the given data is followed. Also, the indispensable available tools are used to cover every single aspect in the evaluation process such as *w*-correlation method which will be presented later in a separate section.

To assess the evaluation process, some common DSP measurements for sound analysis method have been used. It is important to choose the most representative cases based on the procedure of the testing and validation and the results obtained. To clarify, Table 8.3 shows an overview of the sound analysis method when using repeated-measures which are DSP measurements. Table 8.3 presents the experimental data for the reconstructed signals with different principle components when using these measures. Meanwhile, it also compares between the different signals produced from the mixed soundtracks that have been prepared at a selected range of the SNR which is from -20 to 20 dB in a step of 10 when performing the reconstruction process using the first four leading principle components.

Moreover, Table 8.3 presents the original clean signal for comparing with the reconstructed signals for the four leading *PCs* with regards to the indicated measures. In fact, these signals obtained as an output of the SSA algorithm and represented in audio files as de-noised signals. However, it is worth noting that the SNR was mainly the variable that has been changed to produce a number of soundtracks for testing purposes in the proposed dataset along with the type of the signals of interest that indicate different cases. As mentioned previously, the audio files are not treated individually; however, the average method was used for preparing the data in the experiments.

The first aim of presenting the experimental data in this way as in Table 8.3 is to examine the reconstruction technique with regards to possible selected *PCs* when considering different SNR used for preparing the soundtracks that represent the noisy signals. The second aim is to compare between the reconstructions using the selected *PCs* with the original signal that

represents the average of the clean soundtrack signals. Table 8.3 also helps in analysing the effect of SNR. In other words, presenting the experimental data in this way somehow indicates the effect of the content of wind noise mixed with the original clean signals on the performance of the algorithm. Basically, due to the wind noise content presented in the mixed signals, the singular spectra might be affected. Consequently, the grouping and reconstruction techniques introduced in the developed systematic approach followed in the experiments have to be modified in case of any changes in the singular spectra. To sum up, it is to examine and compare the developed SSA algorithm output with the original clean signal with regards to the indicated measures in the selected SNR range when always considering the grouping and reconstruction techniques defined in the algorithm to achieve the most possible separability.

Table 8.3. Sound analysis method for the reconstructed signals with the leading four *PCs* and other types of signals

Type of signal	DSP measurements			
	RMS value	Dynamic range D	Crest factor Q	Autocorrelation time
Clean signal	0.030315	71.6413	11.6991	0.58231
SNR (dB) = -20				
PC1	0.046354	71.3711	7.7403	0.88569
PC2	0.045924	71.2435	7.6936	0.88544
PC3	0.027786	69.5453	10.3599	0.95875
PC4	0.027434	68.6754	9.6007	0.95868
SNR (dB) = -10				
PC1	0.014653	61.3414	7.7139	0.94549
PC2	0.014379	61.2064	7.743	0.94501
PC3	0.010014	59.3697	9.0488	0.95234
PC4	0.010526	59.7354	8.9812	0.95231
SNR (dB) = 0				
PC1	0.041064	51.3405	10.7644	0.87478
PC2	0.040829	54.253	10.7212	0.87451
PC3	0.027281	49.2779	10.2518	0.95866
PC4	0.026857	48.2995	9.4094	0.95429
SNR (dB) = 10				
PC1	0.03293	51.8213	11.1606	0.80077
PC2	0.032912	51.8213	11.1651	0.80077
PC3	0.025598	49.4843	11.0112	0.75714
PC4	0.025589	49.4843	11.0143	0.75848
SNR (dB) = 20				
PC1	0.029199	50.9061	11.2899	0.63766
PC2	0.029193	50.9061	11.2917	0.63766
PC3	0.02387	49.8831	12.0169	0.62372
PC4	0.023863	49.8831	12.0197	0.62372

The indicated DSP measurements can provide with a prevailing view in such comparison as it is required to consider how the reconstruction in selecting pairs of principle components might differ. Also, it is to evaluate to what extent the performance of the SSA algorithm becomes affected with the change of the SNR. It has been previously mentioned that the RMS value is a mathematical way of expressing the amount of energy in a signal and its measurement is time-dependent. Although RMS calculation for audio signals which are made of multiple frequencies is far more complex as indicated in Chapter 6, however, it is particularly useful as it allows comparing signals in equal terms.

Crest factor is a measure of a waveform that indicates the difference between the peak value and the effective value. It represents the ratio of peak value to the root mean square value. Generally, sound waves tend to have high crest factors in a considerable range between 10 to 20 dB for most typical sounds. Accordingly, crest factor indicates how extreme the peaks are in a waveform. It can be useful to judge overall perceived loudness. The dynamic range measures the ratio between the maximum strongest un-distorted signal and the minimum discernible signal, and often describes the ratio of the amplitude of the loudest possible to the RMS noise amplitude. The level of a signal can vary broadly from one moment to the next and that indicates the dynamic range of the signal. In other words, it can indicate the variance in an audio signal from optimum recording level into severe distortion.

The autocorrelation for a signal, which is mathematically corresponding to a time delay, can be used for comparing the signal with time-delayed versions of itself. For instance, periodic signals can be considered as perfectly correlated signals with versions of such signals when the time-delay is an integer number of periods. The autocorrelation of a signal is very much associated to the power spectrum of that signal. The autocorrelation has been involved in signal processing for human hearing as an important part. Since audio signals are not necessarily periodic, however, they have autocorrelations that are closed to 1 and therefore they are often interpreted. Without going through further details, in the context of this study the autocorrelation time has been presented in the experimental data for having an overall and clear view of the evaluation process.

It is important to maintain a balance between these measures; however, values that are closed to those of the clean signal can be always considered as the most optimal in this evaluation process. A graphic representation shown in Figure 8.15 has been generated based on the findings presented in Table 8.3.
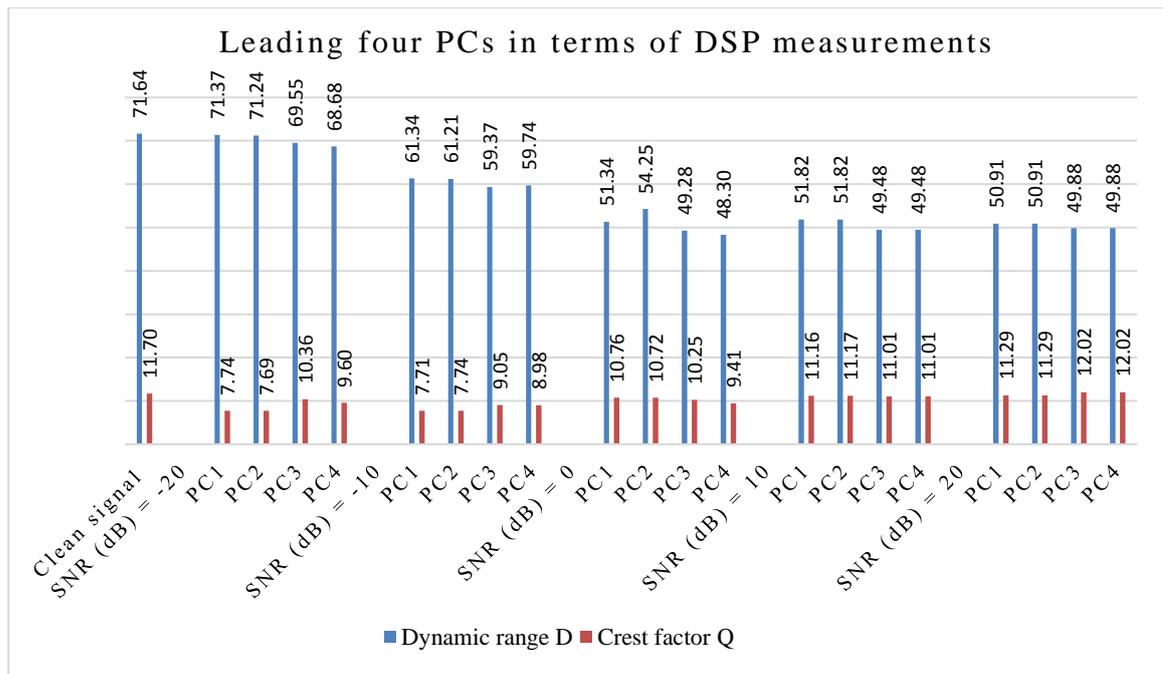
(**a**)
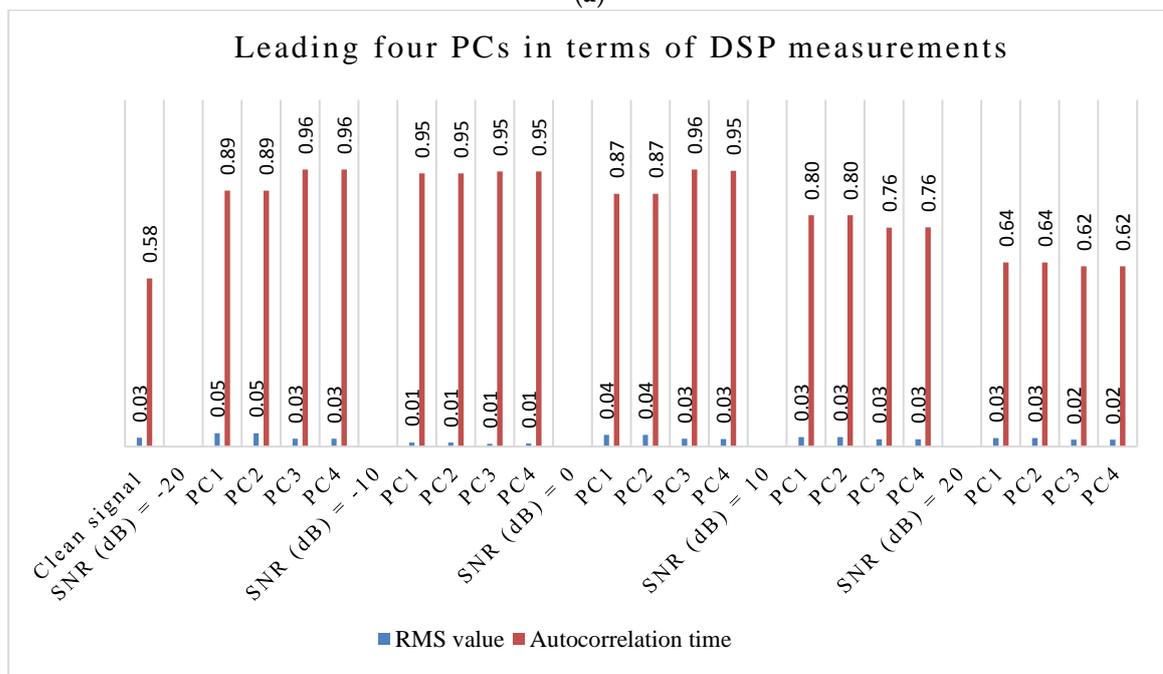
(**b**)

Figure 8.15. Graphic representation to show the leading four *PCs* used for reconstruction with reference to the original clean signal when using noisy signals at different SNR: (a) for D and Q values; (b) for RMS and autocorrelation time

As can be seen from Figure 8.15, the first two categories with the lowest SNR reported significantly more values of D than the other three groups particularly with the first two *PCs* pairs, however, with lower values of Q. What also stands out in Figure 8.15 is that autocorrelation is generally closed to that of the clean signal for all categories although it is more reasonable for the last two categories. Closer inspection of Figure 8.15 shows that the

RMS values are quite stable for mostly of all the five categories particularly for the first two pairs of the *PCs*. It is apparent from Figure 8.15 that the last two levels are more stable with regards to the indicated measures even with lower values of D.

In the graphic representation shown in Figure 8.15, the focus is on presenting the four leading *PCs* calculated for the soundtracks that mixed at different SNR when using the indicated DSP measurements. At lower SNR, with much content of wind noise, the performance of the method is slightly affected. Still, at different *PCs*, the dynamic range is nearly equal to that of the original signal particularly at low SNR, however, this is not always positive as the content of noise is considered in the signal. At higher SNR, with less content of wind noise particularly at 20dB, very much closer results to the original clean signal regarding RMS and autocorrelation time are obtained. With regards to the leading principle components, better results have been obtained with the first two *PCs* and particularly with less wind noise content when considering a balance among all these measurements. Regarding the crest factor, nearly equal values to those of the original signal can be mostly seen with the first two *PCs* and at higher SNR.

On average, the *PCs* were shown to have a fair and meaningful comparison of the reconstructed signals with regards to the indicated measures among each other and with the original clean signal to ensure the developed approach for grouping and reconstruction techniques. In addition to presenting the results in the selected case studies in this chapter which are based on adopting the systematic approach developed for the method in context of this study and used in the algorithm, strong evidence of selecting the first two pairs as dominant ones was found when considering these measures as well.

To conclude, the results, as provided in Table 8.3 and Figure 8.15, indicate that the first two pairs of the principle components are the most dominant ones which confirm and demonstrate the results presented in the previous sections. Therefore, Table 8.4 has been set to provide the results obtained from the first two pairs only showing all the signals in this case for the five options of the SNR and with the same indicated measures. Table 8.4 presents the results obtained from the preliminary analysis of reconstructing the desired signals using only the first two pairs of *PCs*. For each value of the SNR, it is to compare between the noisy signals and the outputs of the SSA algorithm which are the de-noised signals when using the original clean signal as a reference.

Table 8.4. Results for the first two pairs with other signals for the selected SNR range

Type of signal	DSP measurements			
	RMS value	Dynamic range D	Crest factor Q	Autocorrelation time
Clean signal	0.030315	71.6413	11.6991	0.58231
SNR (dB) = -20				
Noisy signal	0.31752	90.3087	9.9644	0.68415
De-noised signal	0.045924	71.2435	7.6936	0.88544
SNR (dB) = -10				
Noisy signal	0.10503	82.7369	12.0018	0.68401
De-noised signal	0.014379	61.2064	7.743	0.94501
SNR (dB) = 0				
Noisy signal	0.044878	75.3134	11.9637	0.18002
De-noised signal	0.040829	54.253	10.7212	0.87451
SNR (dB) = 10				
Noisy signal	0.033208	72.8552	12.1213	0.58188
De-noised signal	0.032912	51.8213	11.1651	0.80077
SNR (dB) = 20				
Noisy signal	0.031769	72.2768	11.9277	0.58202
De-noised signal	0.029193	50.9061	11.2917	0.63766

It is apparent form Table 8.4 that at different SNR values, the value of D for the noisy signals is always higher than the one measured for the clean signal, and particularly much higher at lower values of SNR. What is interesting about the results in Table 8.4 is that practical values for the measures at all the SNR mostly those above 0dB have been obtained except for the slight effect on the values of D. However, this can be considered as a reliable indication of the SSA output mainly for these values of SNR. Figure 8.16 is also a graphic representation that has been generated based on the findings presented in Table 8.4.

After considering the best reconstruction which can be achieved with the first two *PCs*, this graphic representation is therefore generated to focus on the effect of the content of wind noise represented by the SNR used for generating the mixed soundtracks. Consequently, this illustration is to show the performance of the SSA algorithm with respect to the content of wind noise in the noisy signals. In other words, it is to show how the reconstruction using these pairs for the SSA output signals becomes affected by the content of wind which, in turn, indicates the performance of the method and always with regards to the selected measurements. From Figure 8.16, it can be summarised that at extremely low SNR values with high content of wind

noise, the performance of the developed SSA method is slightly affected compared to the case with low wind noise content.
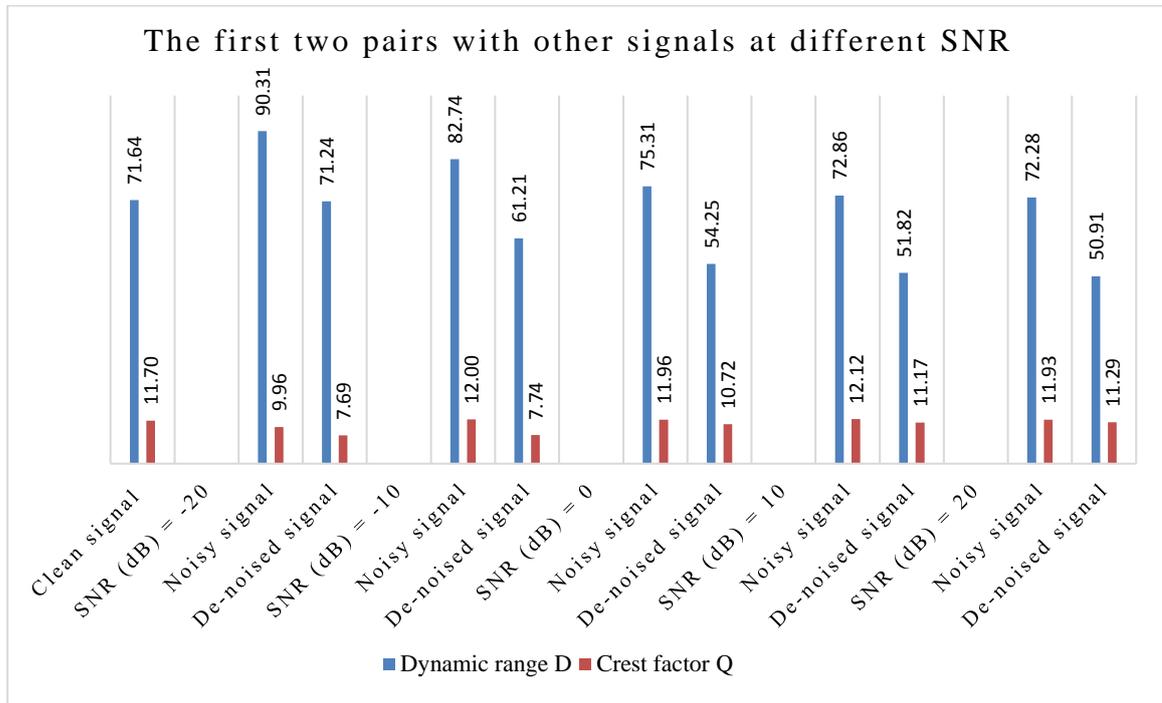
(**a**)
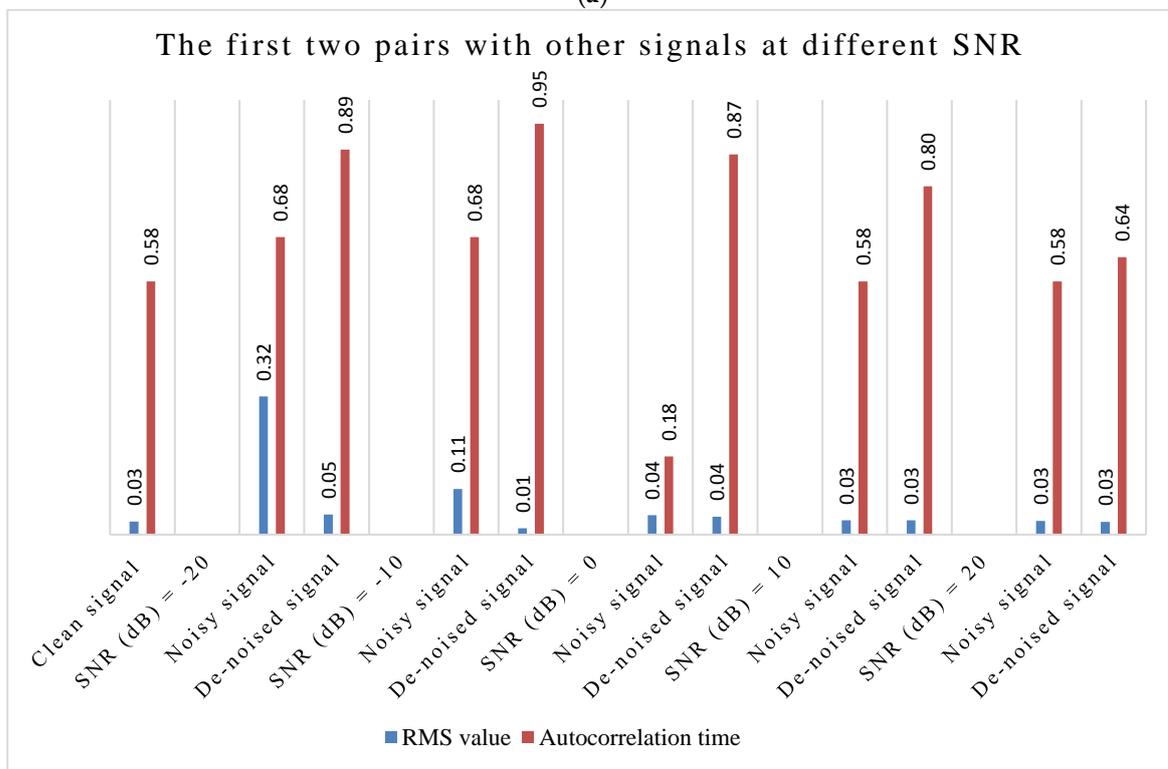
(**b**)

Figure 8.16. Graphic representation to show the first two *PCs* pairs used for reconstruction with reference to the original clean signal and noisy signals at different SNR: (a) for D and Q values; (b) for RMS and autocorrelation time

As mentioned previously in Chapter 6 (section 6.7), there are several DSP analysis techniques that are commonly used in sound analysis along with visual SSA tools that have been used in presenting the results of these empirical studies. Time-domain analysis and spectral analysis are rather important concepts to visually present audio signals. Such common DSP techniques are useful to demonstrate valuable interpretation while analysing any given signals regarding viewing important aspects such as frequency content, noise presented, silence periods, etc. Also, such analysis techniques can be used for observing the changes occurred to the signals used in the experiments during different processes for comparison purposes. Therefore, these techniques can help in interpreting the changes occurred to the signals after the noise separation process when compared to the original signals and noisy signals before applying the noise reduction algorithms as the case for this study. Discussion centred on sound analysis techniques would make to the evaluation process. These techniques can help in critically evaluate noise reduction methods as they can give strong and reasonable indication to the effectiveness and performance level of the developed SSA algorithms.

To summarise the results, a representative case has been selected. However, this case provides a comprehensive overview of the three types of signals used in the experiments conducted in this empirical study which are noisy mixed signals represent SSA input, original clean signals used as reference, and de-noised signals which are the output of the noise separation process of the SSA algorithm. Therefore, as it is in the middle of the five levels selected of the SNR range and interesting results have been obtained, 0dB case has been selected for presenting more results and carrying out further analysis in the evaluation process. The results presented in Table 8.5 have been produced from the above tables to summarise the 0dB case. The results of the reconstruction signals for the first leading *PCs* along with clean and noisy signals with regards to the indicated measurements are summarised in Table 8.5.

Table 8.5. Summary of the results obtained for the reconstructed signals and other signals for the case of 0dB

DSP measurements	Type of signal		Reconstruction with different *PCs*			
	Clean signal	Noisy Signal	*PC1*	*PC2*	*PC3*	*PC4*
RMS value	0.030315	0.044878	0.041064	0.040829	0.027281	0.026857
Dynamic range D (dB)	71.6413	75.3134	51.3405	54.253	49.2779	48.2995
Crest factor Q (dB)	11.6991	11.9637	10.7644	10.7212	10.2518	9.4094
Autocorrelation time (s)	0.58231	0.18002	0.87478	0.87451	0.95866	0.95429

Figure 8.17 is a summarised comparison graph for the leading four *PCs* at 0dB case. This graphic representation shows the leading four *PCs* used for the reconstruction with the reference clean signal and noisy signal at 0dB with regards to the signal processing measurements which are D values (a) and for Q and autocorrelation time values (b).
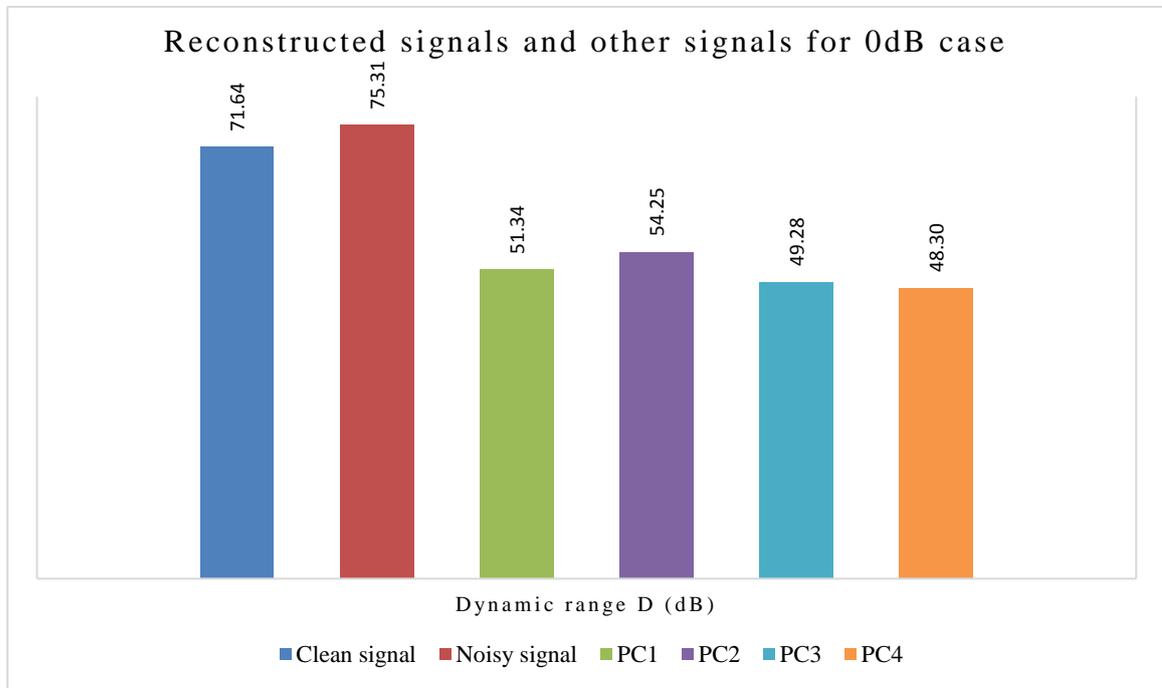
(**a**)
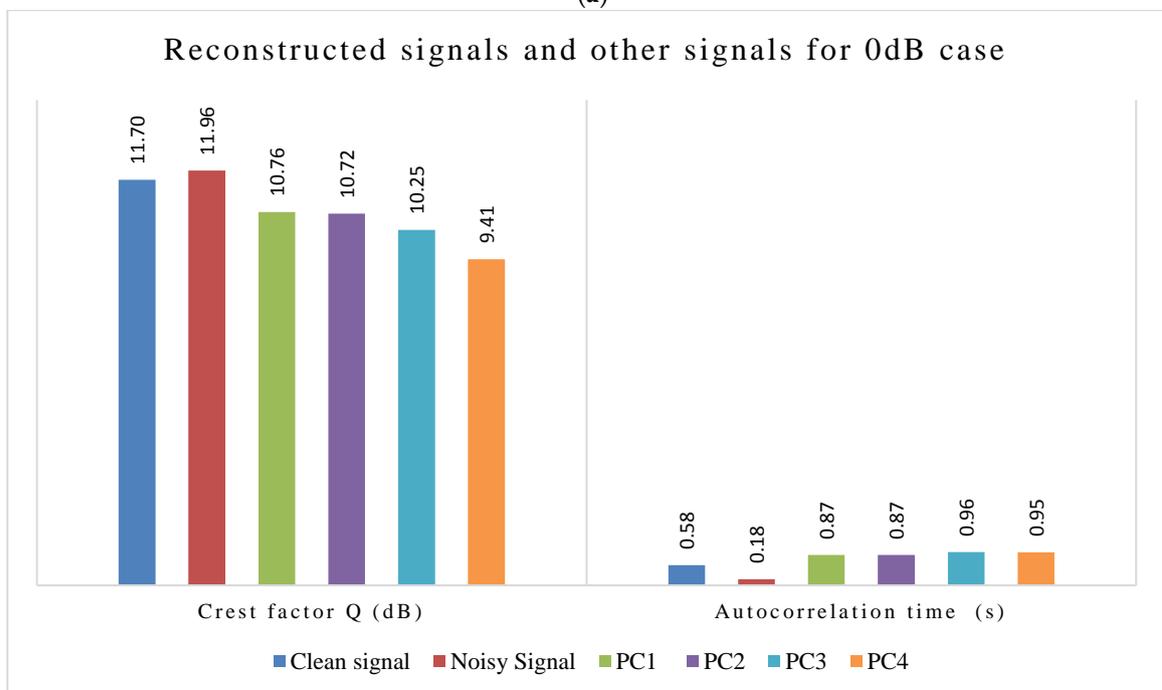
(**b**)

Figure 8.17. Graphic representation to show the leading four *PCs* for reconstruction with reference to the original clean signal and noisy signals at 0dB: (a) for D values; (b) for Q and autocorrelation time

It is apparent from the graphic representation shown in Figure 8.17 that the value of D for the noisy signal is higher than the one measured for the clean signal and the values obtained for the leading four *PCs*. Although there is a slight effect on the values of D for these leading four *PCs* when comparing to the value related to the reference clean signal, however, these values still considered practical. As previously mentioned, it is important to achieve an adequate balance between all the used measures. What is interesting about the results shown in Figure 8.17 is that the most practical values for D has been measured when reconstructing using the first two leading *PCs* which demonstrate the results presented in the previous sections.

As can be seen from Figure 8.17, the first three leading *PCs* reported significantly more values of Q than the fourth *PC* particularly with the first two *PCs* pairs. What also stands out in Figure 8.17 is that the autocorrelation is generally closed to that of the clean signal for all leading *PCs* although it is more reasonable for the first two leading *PCs*. Closer inspection of Figure 8.17 shows that the values of the indicated measures are fairly stable and considered practical for the first two pairs of the *PCs* even with slightly lower values of D. However, this can be considered as a reliable indication of the SSA output mainly for two leading *PCs*. To sum up, signal processing measurements gave a reliable indication to the effectiveness and performance level of the developed SSA algorithms particularly with regards to the two leading *PCs* used for the reconstruction. Clear separation can be further demonstrated using sound analysis methods and particularly *w*-correlation which helps in indicating the separability.

In addition to the importance of analysing the signals used in this evaluation process by using their graphical representations, the results produced by the sound analysis algorithm are presented for comparison purposes. The main aim of this discussion is to critically evaluate the developed SSA method for wind noise separation and measure the algorithm effectiveness and performance particularly with regards to the key aspects involved in the systematic approach developed for the SSA method which are grouping and reconstruction techniques.

Figure 8.18 shows an overview of noisy signal for the case of 0dB which is the case that has been chosen for further discussion as mentioned above. The top left part of Figure 8.18 shows a sample of the noisy signal in the time domain while the amplitude spectrum is shown in the right top part in which the changes in the magnitude with respect to the frequency can be seen. The difference between this part of Figure 8.18 and the corresponding one of Figure 8.19 which is for the clean signal was significant as the effect of wind noise components in the lower frequency range for approximately less than 4000Hz can be clearly seen.

The visual representation in the spectral analysis of the signals is shown by means of the spectrogram which shows the amplitude of the frequency components of the signal over time. The spectrogram also presents interpolated colours of magnitude versus time to allow seeing which frequencies are present in the signal with reference to time and amplitude. As known, a colour scheme is used to show the approximate amplitudes of the components to indicate intensity by varying shades of darkness of the pattern. Therefore, the bottom left part of Figure 8.18 shows the spectrogram of the noisy signal used as an example in this evaluation process.
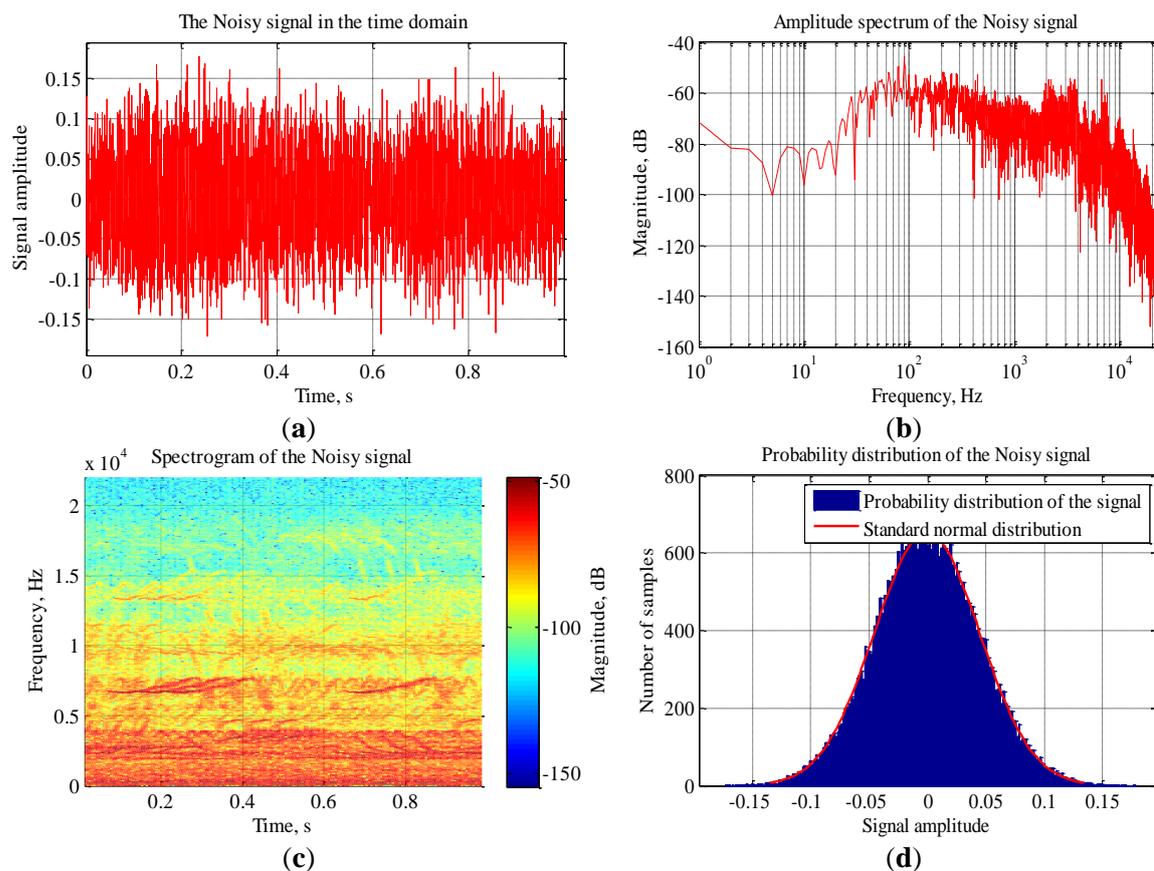
Figure 8.18. An overview of noisy signal for the case of 0dB: (a) Noisy signal in the time domain; (b) Amplitude spectrum; (c) Spectrogram (d) Probability distribution

Based on the shading system used in the spectrogram, components that make up complex signals like the sound signals used in this study do not share the same amplitude value. However, frequency components with the highest amplitude values are shown in darkest region of the colours and components of lower values are shown in ladder shad of light colours. The spectrogram also shows changes in frequency values of the components of the signal over time. However, looking at part c of Figure 8.18 for the noisy signal, it is obvious that wind noise components which indicated by the colour (red) are located in the lower frequency region with frequency range for approximately less than 4000Hz as mentioned above.

The probability distribution of the noisy signal is illustrated in the bottom right part of Figure 8.18. A higher amplitude value can be seen compared to the probability distribution of the other two signals presented in Figure 8.19 and Figure 8.20. However, this represents the content of wind noise components in the signal. Similar to Figure 8.18 while using the same visual DSP techniques, Figure 8.19 graphically shows an overview of the original clean signal and Figure 8.20 for the de-noised output signal of the SSA algorithm reconstructed using the first two pairs of *PCs*.

With regards to the amplitude spectrum shown in the right top part of these three figures, there was a significant difference between both the clean and de-noised signals with the corresponding one which is the noisy signal presented in Figure 8.18. Also, it can be seen that the de-noised signal has similar amplitude spectrum as for the clean signal. However, the effect of wind noise components in the indicated range discussed above cannot be seen in the de-noised signal. Also, similar indication can be found when comparing the de-noised signal shown in Figure 8.20 with the noisy and clean signals with regards to the other measurements.
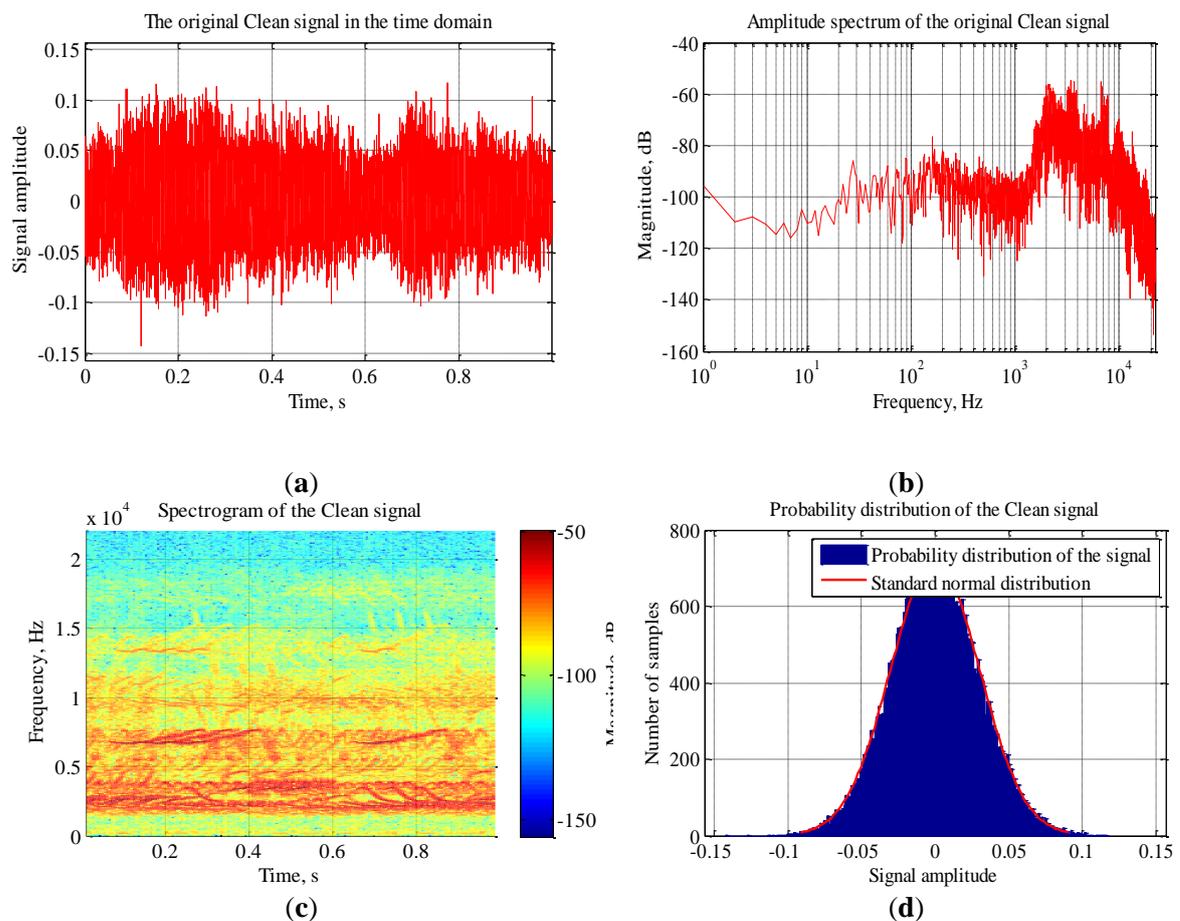
Figure 8.19. An overview of the original clean signal: (a) Clean signal in the time domain; (b) Amplitude spectrum; (c) Spectrogram (d) Probability distribution
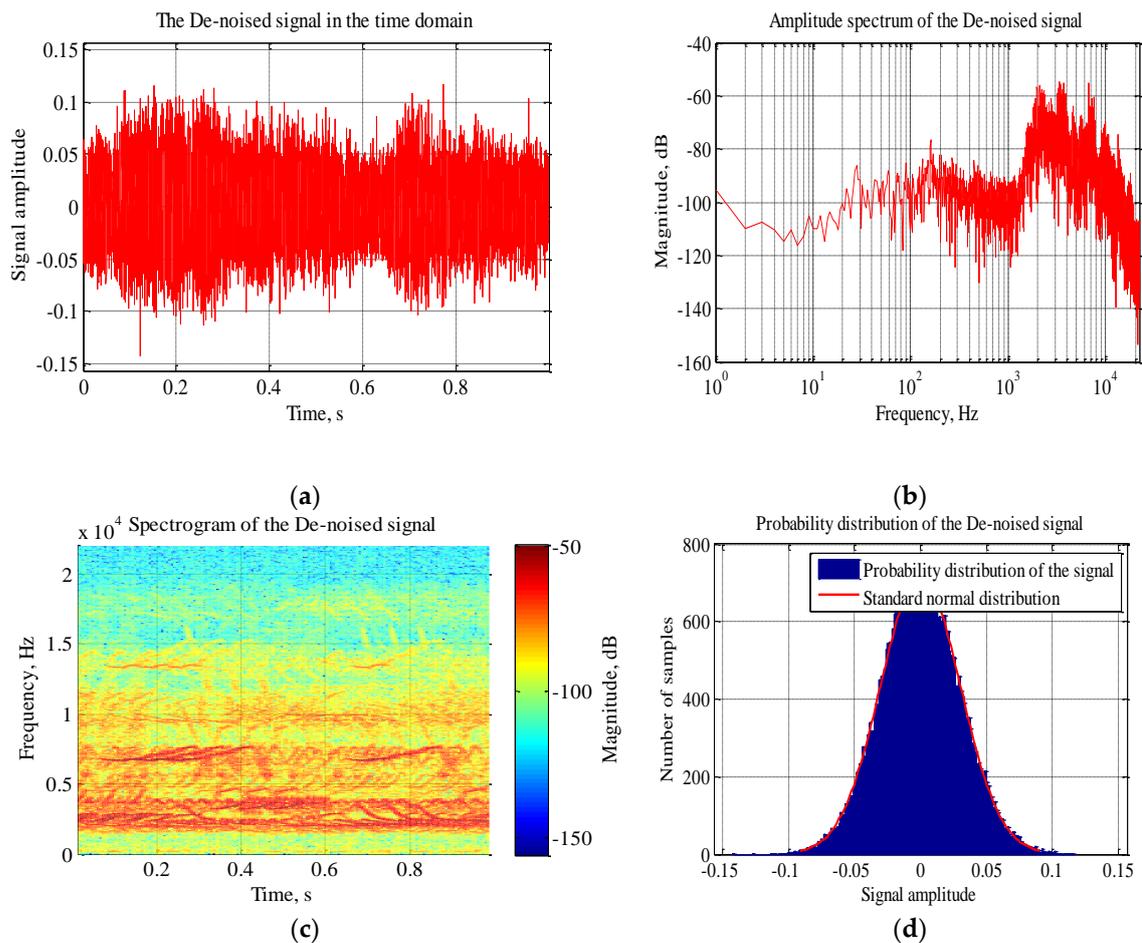
Figure 8.20. An overview of the output signal of the SSA algorithm reconstructed using the first two pairs of *PCs*: (a) representation in the time domain; (b) Amplitude spectrum; (c) Spectrogram (d) Probability distribution

These visual DSP and sound analysis techniques used in the evaluation process yielded significant results as the output signal of the SSA algorithm is very much similar to the original clean signal. A significance difference can be noticed between the de-noised wind noise free signal and the input noisy signal in a way that ensures the performance and effectiveness of the developed SSA system. These results gave a reliable indication that the developed grouping and reconstruction techniques worked in an efficient with regards to separating wind noise components out and reconstructing the desired signals with no reconstruction errors. However, another important measure, which is the *w*-correlation matrix, provides with a strong evidence to approve the separability approach will be discussed in the next sub-section.

8.8.2   W-correlation

Recall from Chapter 5, an important quantity known as weighted correlation (*w*-correlation) has been introduced. This measure is defined as a natural measure of the

dependence between the reconstructed components; however, it can be used to realise the separability and a method for determining the grouping criterion. Reconstructed components that have zero *w*-correlation are the well-separated ones, whereas, large values of *w*-correlation indicate that the reconstructed components are of one group. Thus, in the SSA decomposition, this corresponds to the same component (Harmouche *et al.*, 2017; Rodrigues and Mahmoudvand, 2017; Xu, Zhao and Lin, 2017).

Examining the matrix of the absolute values of the *w*-correlations is a useful measure that can be used to show that microphone wind noise is separable or not in the singular spectrum domain. The matrix of *w*-correlation contains information that can be very helpful for detecting the separability and identifying grouping. It is a standard way used to check the separability between elementary components. Furthermore, *w*-correlations can also be used for checking the grouped decomposition. The *w*-correlation matrix consists of weighted cosines of angles between the reconstructed components of the time series. The number of entries of the time series, which terms into its trajectory matrix, is reflected by the weights (Golyandina and Shlemov, 2013; Hansen and Noguchi, 2017; Rodrigues and Mahmoudvand, 2017).

In the case of this study, the analysis using *w*-correlation matrix is to distinguish between frames containing mostly the energy of the desired signal and wind-only frames presented in the subspace of the higher-order eigenvalues of the singular spectra. Basically, the separability between two components such as $X^{(1)}, X^{(2)}$ of the time record characterises how well these two components can be separated from each other. Using the *w*-correlation can help in evaluating the separability and can be simply shown as in Equation (8.3) (Harmouche *et al.*, 2017; Traore *et al.*, 2017).

$$\rho_{12}^{(w)} = \frac{<X^{(1)}, X^{(2)}>_w}{\left\|X^{(1)}\right\|_w \left\|X^{(2)}\right\|_w},$$ (8.3)

The values of $\rho_{12}^{(w)}$ ensure the concept of separability, however, small absolute values, particularly the ones closed to zero, indicate that the components are well-separated. Whereas, big values show that these components are inseparable and therefore they relate to the same components in the SSA decomposition (Harmouche *et al.*, 2017; Traore *et al.*, 2017). A specific software known as Caterpillar software was used and recommended in many studies such as in (Hassani, Mahmoudvand and Zokaei, 2011; Golyandina and Lomtev, 2016). However, in this study this software has also been used for carrying out the *w*-correlation investigation.

Figure 8.21 shows the moving periodogram matrix of the reconstructed components from which the corresponding eigentriples ordered by their contribution can be seen. Generally, poorly separated components have large correlation while well-separated components have small correlation. Groups of correlated series components can be found while looking at the matrix which indicates the *w*-correlations between elementary reconstructed series. Such information can be used for the consequent grouping. Strongly correlated elementary components should be grouped together in one group. As an important rule, it is recommended that correlated components should not be included into different groups. The matrix of *w*-correlations between the series components is graphically depicted in absolute magnitude in white-black scale. Correlations with moduli close to 1 are displayed in black, whereas small correlations are shown in white (Golyandina and Shlemov, 2013).
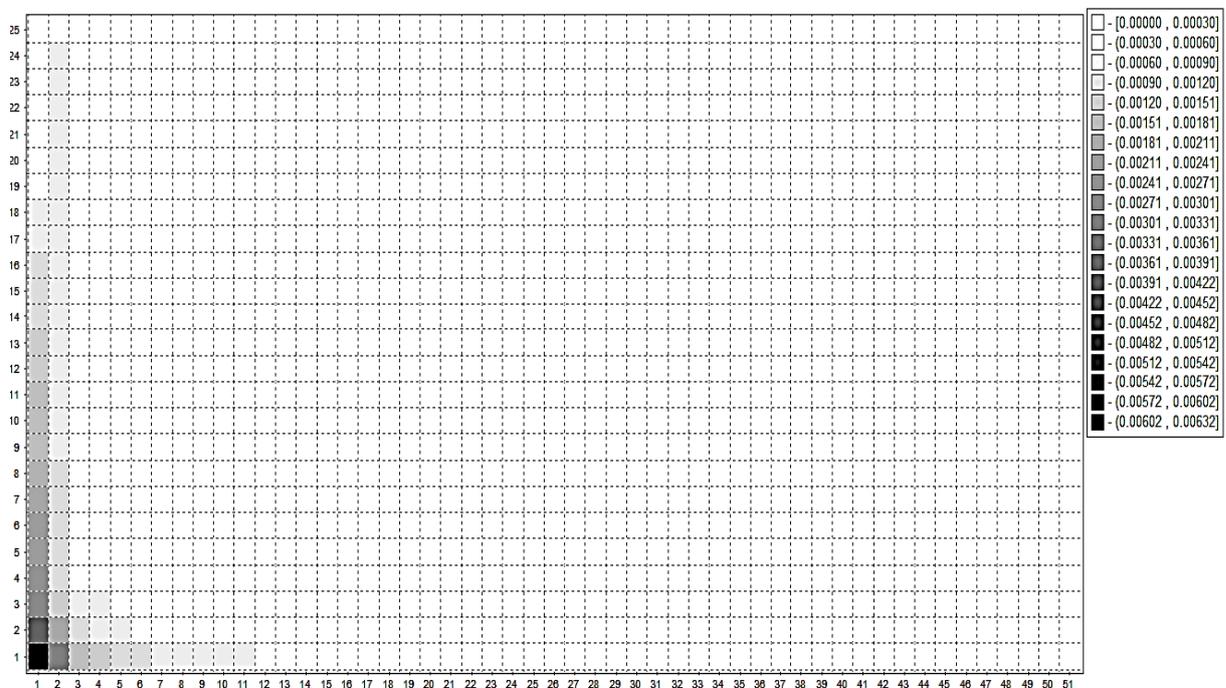
Figure 8.21. Moving periodogram matrix of reconstructed components

To evaluate the developed SSA system with regards to the separability using *w*-correlation matrix, harmonic and broadband noise, mainly, birds' trill and wind noise has been selected as an example. These two signals are of different waveform characteristics such as energy, duration and frequency content. As previously mentioned, only few of the reconstructed components can lead to the best reconstruction. The moving periodogram matrix of the reconstructed components presented in Figure 8.21 shows the contribution of the eigentriples. The first two eigentriples indicate the maximum contribution with regards to the desired signal, whereas the energy presented along the corresponding eigenvectors of the rest is low.

Figure 8.22 shows the *w*-correlations matrix for only the first 50 reconstructed components. The system uses a 20-grade grey scale ranging from white to black which corresponds to the absolute values of correlations from 0 to 1. Components with small correlations are shown in white and indicate well-separated components; however, such components are considered as w-orthogonal which points out that they are highly separable. Whereas big correlation ones shown in black with moduli close to 1 are not w-orthogonal and therefore, this means that they are poorly separated components.
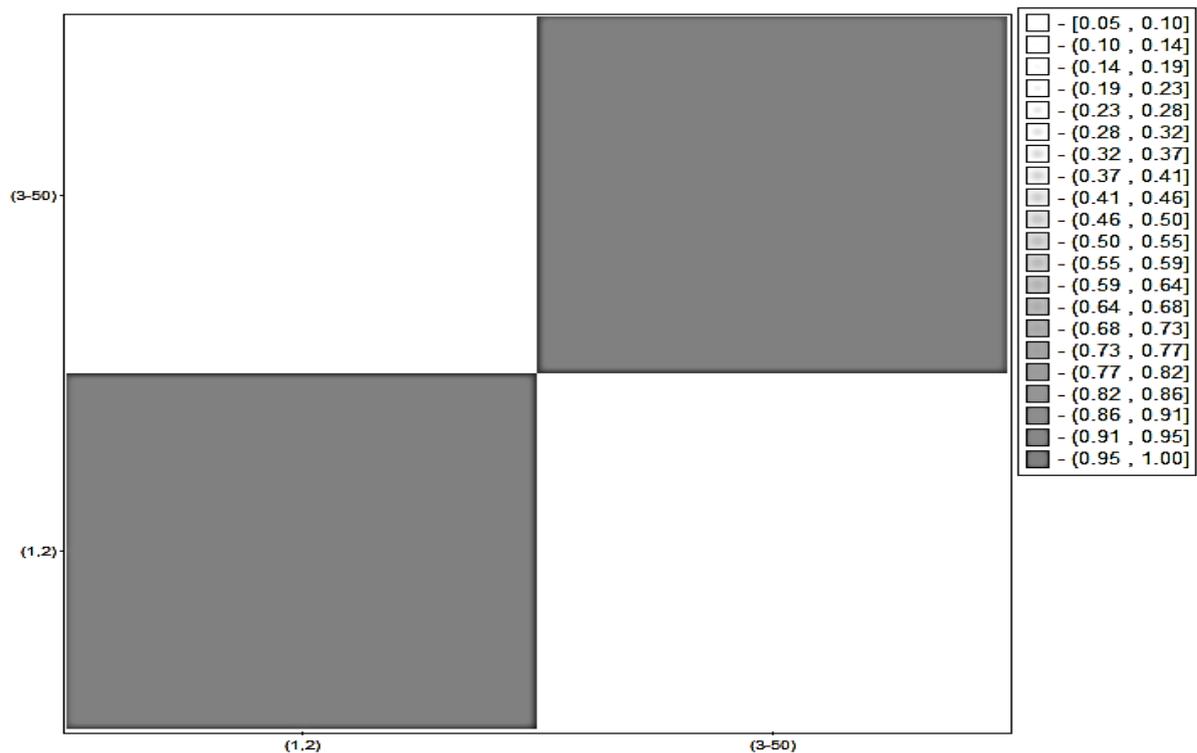
Figure 8.22. Matrix of *w*-correlations of the selected 50 eigenvectors of the SVD of **Y**

Based on the information extracted from the matrix of correlations, the first two pairs of the eigentriples can be used to reconstruct the de-noised signal, while separating out the rest of the matrices presented in the second group as they represent the wind noise components. With regards to the separability, the analysis of this matrix shows that the given record has two independent components. Also, it can be observed that *w*-correlation shown in white is the correlation between pair eigenvector that represents the noise component and nearly equals to zero. The resultant reconstructed record based on the results approved by the *w*-correlation is assumed to be noise free and hence; $\rho^{(w)}_{(S^{(1)},S^{(2)})} = 0$.

Using the eigenvalues spectra helped in efficiently selecting the eigenvectors associated with S1 as a noise free signal and eigenvectors associated with S2 that represent noise. These

two groups of eigenvectors are mainly corresponding to different frequency events which eventually led to their reconstruction. Based on this information; it can be concluded that the dataset used in this study is separable with the selection of the first dominant eigentriples to reconstruct of the original series while wind noise component is represented by the rest.

8.9 Summary

This chapter presented the results of the empirical study using real-world sounds in the system validation and evaluation phases along with comprehensive analyses and detailed discussions. It included the results obtained from these experimental phases using some selected experimental case studies. In the system testing and validation phase, the window length has been optimised, the developed systematic approach which includes the developed grouping and reconstruction techniques has been adopted. Along with SSA visual tools, relevant DSP techniques, sound analysis methods, and *w*-correlation matrix has been used to indicate grouping and separability in the critical evaluation stage.

Regarding the dataset used in these phases, realistic samples of environmental sounds have been used. For testing the SSA algorithms and, as specific signals of interest, birds' chirps sound has been taken as an example. The results of the validation phase followed by the critical evaluation considering and adopting suitable dataset that links up to the application area of the study (e.g., smart city interesting sounds such as police car siren and fire alarm sound) have been reported in this chapter.

The only parameter that can be adjusted in the decomposition stage is the window length. This parameter is known to have significant impact on the performance and effectiveness of the SSA algorithm. Therefore, systematic investigation has been carried out using a mathematical model provided in the optimisation method that has been particularly developed in this study. More advanced grouping and reconstruction techniques have been applied to ensure and improve the separability. The grouping technique reported in this empirical study was mainly based on the eigentriples for separating the decomposed components after grouping similar components together. The current results have been obtained using an optimisation method based on certain important aspects in the grouping technique such as the nearly equal singular values and other constrains as defined in the systematic approach developed within the framework of the developed SSA method.

The *w*-correlation method has also been involved in the grouping technique to ensure the separability. However, positive results and notable wind noise reduction were achieved. The plausible findings from the investigation suggest that microphone wind noise and wanted sounds are separable in the singular spectrum domain by the developed SSA in this study. These findings are evidenced by the *w*-matrix and considerable amount of wind noise separation which is indicated by the applied objective measures and observed after the SSA de-noising in the testing, validation, and critical evaluation phases as discussed in this chapter.

The experimental investigation and findings indicate the potential of the developed SSA as a valid method for wind noise reduction. Indeed, with this developed version of SSA, the *w*-matrix showed that wind noise is separable for all the real-world sounds used in the system testing and validation phase. Based on these findings, wind noise has its features clearly separable in the proposed SSA subspaces indicated by the *w*-matrix. This suggests that the developed SSA method is effective and generally robust in separating wind noise. This further suggests that this developed system should be able to extend to other outdoor audio acquisition or recording. This opens much more extended applications and impact. Therefore, effective deployment in real-time noisy environments will hopefully lead this developed SSA to a universal microphone wind noise separation method.

9 CHAPTER NINE   CONCLUSION AND FUTURE WORK

9

Conclusion and Future Work

9.1 Conclusion

This thesis set out to develop the separation approach of wind noise in microphone signals based on singular spectrum subspace method in the context of outdoor sound acquisition. Data for this thesis was collected from two main sources: freefield1010 dataset, internet based Freesound recordings as described in Chapter 6. Suitable testing criteria have been followed for the justification of the developed method at the very early stage. Several tests have been designed in the justification stage before further developing and examining new functionalities according to the requirements of each of the next experimental phases. Conducting several experiments as well as following a case-study design with in-depth analysis of the results using SSA visual tools, related signal processing, and sound analysis techniques was the strategy that has been drawn up to meet the aim of this thesis. These experiments and case studies have been carried out in the empirical studies for the verification of the developed system as in Chapter 7 and system validation and evaluation in Chapter 8 based on the systematic approach developed for the method presented in Chapters 3-5. The key concepts and wind noise characteristics and spectrum explained in Chapter 2 have been considered in the development process. Additionally, this procedure was based on using the established measurable objectives to address the research question and achieve the central aim stated in Chapter 1.

In this thesis, a new method to mitigate wind induced noise in microphone signals has been developed. The present thesis developed, for the first time, the Singular Spectrum Analysis method for wind noise separation in the singular spectral subspace based on the systematic approach that has been developed and established considering the development of grouping and reconstruction techniques as key aspects. In this approach, new mathematical models and formulations used in developing SSA algorithms with particular emphasis on investigating the factor that determines and improves the separability and window length optimisation have been introduced. In the context of this study, the SSA has been developed as a two-step point-symmetric noise reduction method of noisy records contaminated with wind noise based on data-adaptive nature of its functions which improved the ability of adding or removing such additive components either low or high frequencies. However, noiseless signals have been obtained with high separability and no reconstruction errors.

In the testing and validation phases with real-world sounds, the window length has been optimised using a new mathematical model particularly developed in the context of this study. Grouping technique has been developed and *w*-correlation matrix has been used to indicate grouping and separability. In the system validation phase, it was to bring the developed method for critical evaluation considering and adopting suitable dataset that links up to the application area of the study (e.g., birds' chirps, smart city interesting sounds such as siren of ambulance cars, police cars, and different alarm sounds). For this purpose, soundtracks generated from a mixing model through mixing desired signals with wind noise at different SNR has been established and examined.

This work provides an exciting opportunity to advance and contribute to existing knowledge of the SSA theory and practice by developing such separation approach. This approach utilises a conceptual framework, has, in its final form, three key objectives; grouping, reconstruction, and separability. Before this study, evidence of improving the understanding of such key elements was purely anecdotal. Besides, new descriptive figures have been introduced which can be seen as a significant part to add to existing literature. In the context of this study, improving separability has therefore been considered as a key element in the developed SSA using non-orthogonal decompositions of time series and independent component analysis. However, this is because the signals obtained by the SSA decomposition are generated from eigenvectors in the Eigen subspace. Wind induced noise is statistically separated from wanted signals in a singular spectral subspace. Thus, the new developed method provides a convenient alternative to existing standards.

Returning to the question posed at the beginning of this thesis, it is now possible to state that the results of the experiments and case studies conducted during the empirical studies increased the effectiveness of the developed SSA for distinguishing wind noise in the Eigen domain. This separation approach is based on reconstructing the decomposed components of the SVD method by gathering wind noise components in a separate cluster as residual noise and reconstructing the desired components in one group. Moreover, the obtained results from both empirical studies for system verification and validation produced a positive outcome that showed the capability of the modified SSA with regards to the separability which has been approved by the *w*-correlation matrix. Also, from the results discussed in Chapters 7 and 8, it can be concluded that defining a subspace for the unwanted wind noise and another subspace that corresponds to desired components was possible with the developed method. This indicated the effectiveness of the developed grouping and reconstruction techniques.

These results further support the idea of noise separation using singular spectral subspace method. The finding of window length optimisation with optimal value around $N_t/2$ is consistent with that of many previous studies. It is worth noting that higher window lengths are likely to affect the performance of the SSA with regards to the processing time. In addition to the size of the generated matrices that depends on the window length, the SSA uses a massive number of matrix multiplications which increases with the size of the dataset. Therefore, the computational load increases due to the mathematical complexity of SSA algorithms. In fact, this is one of the challenges and a major limitation of the SSA particularly with long time records and big dataset.

Several important aspects such as the selection of the *PCs*, the generated *RCs*, and the grouping and reconstruction techniques that allow creating groups for the decomposed components to improve the separability have been introduced in the developed method. It can be concluded from the obtained results that not all the principle components are dominant and accordingly only few of the reconstructed components are required. Each *RC* retranslates its corresponding principal component into the original units of the time series which permits a comparison between the reconstructed components. Therefore, it can also be concluded that to reflect oscillatory modes of interest, the original noisy time series can be de-noised through a convolution when selecting a small number of principle components and their associated eigenvectors based on the eigentriple produced by the SVD. Using less of the total number of the produced reconstructed components can be used as an appropriate manner for reflecting the oscillatory modes of interest when properly defining the grouping constrains. It is a satisfactory outcome indeed as separating wind noise in a different cluster as a residual noise and the reconstruction of one-dimensional time series have been performed in this way when considering these developed key objectives that eventually led to achieve high separability.

In the experiments and case studies conducted within the empirical studies for system verification and validation, it can be obviously seen from the eigenvalues spectra that the eigenvalues located in the lower subspace are meaningfully higher than the others and take more of the energy of the given signal, whereas others in the higher subspace represent the noise floor. Accordingly, the corresponding principle components are significant for reconstructing a noise free series as they provide glimpses of the periodic components. Meanwhile, the random components are represented by the ones that correspond to the other subspace. Notably, all the variance of the given time series can be specified by the dominant principle components which is consistent with the lower-order eigenvalues.

243

For the types of data used in the empirical studies, extremely low reconstructions errors were found as a clear indicative sign of an optimal recovery and separability. A comprehensive evaluation of the performance of the developed system has been made to measure its effectiveness for wind noise separation using suitable evaluation criteria. Therefore, results show that microphone wind noise is separable in the singular spectrum domain after completing the verification and validation of the developed system and the critical evaluation when different objective measures have been applied and finally evidenced by the *w*-correlation. The findings indicate the potential of the developed SSA as a valid method for wind noise separation. The results improve the reliability of outdoor acoustic sensing for soundscapes and environmental noise monitoring. In addition, the developed method should be able to extend to any outdoor audio acquisition and recording. As a result, it may have much more extended applications and impact.

Finally, the new aspect presented in this thesis to wind noise in microphone signals, which is the separation of the decomposed components using the developed SSA, might even lead to better results. It is only with reliable developed SSA method wind noise can be separated and wanted sounds can be retained. Hence, it is possible to improve soundscapes monitoring and construct smart urban environments based on acoustic sensing technology in such a way that their true capabilities may be exploited. The thesis outcomes including main technical achievements and improvements can be outlined as follows:

1. The separation approach has been developed.
2. The findings suggest that microphone wind noise and wanted sounds are separable by the developed SSA.
3. Wind noise can be distinguished and separated in the singular spectral subspace evidenced by *w*-correlation matrix and other measures.
4. Results of reconstructed signals indicate significant improvement in separability and the effectiveness of the developed grouping and reconstruction techniques.
5. Window length has been optimised using new mathematical model with optimal value obtained around $N_t/2$.
6. The developed grouping technique helped in efficiently selecting the eigentriples associated with one group as a noise free signal, and others associated with the second group represent noise.

7. Best reconstruction has been achieved based on the first two *PCs* with high separability proved by the applied measures (*w*-correlation matrix).

8. At extremely low SNR with high content of wind noise, the performance of the method is slightly affected.

9. For both system verification and validation and with all the selected signals of interest, the method shows high performance with no reconstruction errors and clear separability.

10. The developed systematic approach which includes new mathematical models and formulations as well as new descriptive figures can be seen as a significant part to add to existing literature.

11. As a major limitation, mathematical complexity of SSA algorithms increases the computational load and the processing time.

12. The generalisability of the results as to deploy the developed method in real-time noisy environments and with other types of environmental noises or different desired signals might be another limitation.

9.2 Future Work

It is probably hard to conclude whether the developed method will still work efficiently in real- time noisy environments despite the positive results obtained for wind noise separation with multiple real-world sounds even under adverse condition with low SNR values. Therefore, the generalisability of these results is subject to this limitation. Also, whether the method will work in the same way with other types of environmental noises or not. Although the effectiveness of the developed grouping and reconstruction techniques needed to be identified for noise components as for such difficult case is evidenced by the *w*-correlation matrix, however, there is no likely apparent reason why the method should not work.

This procedure of wind noise separation can be considered as complex case for a time-varying broadband noise as previously discussed, however, it might be less complex with other environmental noises. Hence, considering this issue can lead to indicate the generalisability of the developed method with a more convenient way. Also, despite of the positive results obtained from the new derived mathematical model for window length optimisation and its consequence effects on other processes, this optimisation method was implemented manually. Therefore, using machine learning technique could be a practical suggestion in this case. To

summarise, this research has thrown up some questions in need of further investigation. It is recommended to undertake further research in the following areas.

1. To what extent the method can be deployed in real-time noisy environments along with selecting other desired signals to indicate the generalisability.
2. To what extent the developed method will work with other types of environmental noise, such as rain noise, as to indicate the generalisability as well.
3. It would be interesting to assess the effects of the computational load and processing time of the SSA on its real-time implementation for environmental noise separation.
4. Applying machine learning technique could be a practical suggestion to implement window length optimisation method.

Finally, it is worth mentioning that analysis regarding improving soundscapes design in urban smart environment has not been included in the thesis and might be thoroughly recommended for future work. A further study could assess the long-term deployment of acoustic sensing considering acoustic modelling as one of the key initiatives. Furthermore, such important aspect may underpin the vision to probably introduce a model based on a virtual platform that includes characteristics of applications to have a more complete solution which could improve performance and ensures more accurate production and delivery of sensed information in different applications.

APPENDICES

APPENDIX A: A Brief Summary of Existing Wind Noise Reduction Methodologies

Table A.1. Conventional noise reduction schemes applied for wind noise reduction

Conventional noise reduction schemes, (spectral subtraction and statistical-based estimators)	
Approach	- Based on subtracting an estimate of the noise magnitude spectrum from the noisy signal magnitude spectrum. - Spectral subtraction has been used for wind noise reduction (Boll, 1979).
Main advantage	**Main limitations**
Robust, easy to implement, comprehensively studied and generalised through many years	- Show limited effectiveness and cannot effectively attenuate wind noise (Schmidt, Larsen and Hsiao, 2007; King and Atlas, 2008). - Assumes that the noise is stationary, noise estimate is mostly obtained during the silence period (e.g. speech pauses). - Unable to obtain new noise estimates as long as the signal of interest is there (Boll, 1979). - Special detection and processing might be required to better reduce the effect of wind noise when using such methods.
Traditional remedy for wind noise (e.g. Wiener filter method)	
Approach	- Random noise optimal removal filter, used for removing noise from a signal when the signal of interest and the noise have different frequency characteristics. - It works by attenuating frequencies where the noise is expected to be the most dominant.
Main advantage	**Main limitations**
Robust, easy to implement, comprehensively studied and generalised through many years	- It compromises the quality of the sensed acoustic data, shows some but limited performance (Schmidt, Larsen and Hsiao, 2007). - It assumes stationary signals, which might give inappropriate approximation for speech and wind noise (Dietrich and Utschick, 2005; Vaseghi, 2008). - Kalman and Wiener filters are not applicable in case of reducing noise levels in certain experimental time series related to dynamical systems.

Table A.2. The post filters and their extension with methods introduced for the estimation of noise PSD

Post-filters and extension of the conventional post-filters	
Approach	- Post-filters: It is to emphasise the formant frequencies and deemphasise the spectral portion where noise contributes the most to the observed distortion (Nemer and Leblanc, 2009).
	- An extension post-filter: Based on adapting the emphasis parameters.
	- A time-domain adaptive post-filter proposed in (Nemer and Leblanc, 2009) to reduce wind noise in corrupted speech. It is based on tracking the changing envelop spectrum of wind noise similarly to other post-filters except it deemphasises the wind 'resonance'.
Main advantage	**Main limitations**
LPC analysis was used in time-domain adaptive post-filter to distinguish between frames	Used in model-based speech coders (Juin-Hwey Chen and Gersho, 1995) as an attempt to speech quality improvement in the presence of wind noise but not an optimal solution to broadband noise such as wind especially with a non-stationary speech signals.
Single microphone methods introduced for the estimation of noise PSD from noisy speech signals and other methods that dealing with wind noise reduction in single microphone	
Approach	- The estimation of the noise Power Spectral Density (PSD).
	- Exploiting the spectral characteristics of the wanted signal (speech) and noise in order to estimate the wind noise PSD is the applied approach in (Nelke *et al.*, 2014).
	- Considering magnitude spectrum towards higher frequencies of wind noise and the harmonic structure of the wanted speech signal (Nelke *et al.*, 2014).
	- The concept relies on directly modify the noisy input signal (Kuroiwa *et al.*, 2006; King and Atlas, 2008; Hofmann *et al.*, 2012).
Main advantage	**Main limitations**
Focused on wind noise reduction from speech signals	- The assumption of considering noise signal as slower varying over time than the speech signal is not true for wind noise signals.
	- According to (Nelke *et al.*, 2014), due to inaccurate estimates of noise PSD, such conventional algorithms provide insufficient level of noise reduction.
	- Better performance but still there is a remarkable distortion.

Table A.3. The different models and comb filters introduced for wind noise reduction for speech

Based on the approach below, many models have been introduced, such as Gaussian Mixture Models (Ding *et al.*, 2005)**, Hidden Markov Models** (Roweis, 2001)**, Non-negative Sparse Coding** (Schmidt, Larsen and Hsiao, 2007) **and Vector Quantization** (Ellis and Weiss, 2006)	
Approach	- Modelling the sources in noisy signals independently in order to use these models to find the best estimate of the signal of interest and noise signal has been involved in the more recent methods. - Using binary masking to remove individual signals (Cermak *et al.*, 2007).
Main advantage	**Main limitations**
The estimation of the signal of interest and noise	Mainly for speech signals cleaning, this approach often models an individual speaker rather than speaker independent.
Comb filters	
Approach	Based on reinforcing the harmonic nature of speech signals depending on accurate pitch estimation.
Main advantage	**Main limitations**
Easy to use	The difficulty of achieving accurate pitch estimation in noisy environments (King and Atlas, 2008).

Table A.4. The dual microphone and microphone array technologies

Dual microphone technologies	
Approach	For speech understanding (e.g. in hearing aid applications)
Main advantage	**Main limitations**
- Innovative designs brought notable progress in noise reduction - Helped in thriving multi-microphone technology	- Inherent microphone placement is susceptible to wind noise at the input of the microphone which is considered a downside - More than half of the users of this technology (hearing aid users) reported the downside of methods used in their hearing aid devices and the low performance in windy environments (Kochkin, 2010).
Multichannel with more microphones available (microphone array) **Such as Independent Component Analysis**	
Approach	- The correlation between the desired signals in the microphones - Based on exploiting the difference in propagation delay between the acoustic signals and the wind.
Main advantage	**Main limitations**
Efficient and effective for reducing wind noise	Such methods are computationally considered prohibitive as well as their complicated setups are difficult to deploy which limit their usage.

APPENDIX B: Best Paper Award Certificate

The paper was selected based on the contribution. The criteria used are novelty/originality, method, experimentation, inference, presentation, and contribution to the domain.
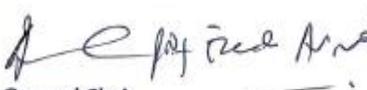
REFERENCES

Aarts, E. and Encarnação, J. (2006) *True Visions: The emergence of ambient intelligence*. Springer Science & Business Media. doi: 10.1007/978-3-540-28974-6.

Abulifa, A. *et al.* (2013) 'A vibration measurement system for deaf people's emergency apparatus', *Proceedings of the 11th International Conference on Manufacturing Research (ICMR2013)*, pp. 101–106.

Adavanne, S. *et al.* (2017) 'Sound Event Detection in Multichannel Audio Using Spatial and Harmonic Features', *arXiv preprint arXiv:1706.02293*.

Ahmed, E. *et al.* (2016) 'Internet-of-things-based smart environments: state of the art, taxonomy, and open research challenges', *IEEE Wireless Communications*, 23(5), pp. 10–16. doi: 10.1109/MWC.2016.7721736.

Alcoverro, B. and Le Pichon, A. (2005) 'Design and optimization of a noise reduction system for infrasonic measurements using elements with low acoustic impedance', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 117(4), pp. 1717–1727. doi: 10.1121/1.1804966.

ALdwaik, M. and Eldwaik, O. (2012) 'Simulation of Tracking Time Delay Algorithm using Mathcad Package', *World Academy of Science, Engineering and Technology, International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering*, 6(7), pp. 659–662.

Alexandrov, T. (2009) 'A Method of Trend Extraction Using Singular Spectrum Analysis', 7(1), pp. 1–22.

Alexandrov, T., Golyandina, N. and Spirov, A. (2008) 'Singular Spectrum Analysis of Gene Expression Profiles of Early Drosophila embryo: Exponential-in-Distance Patterns.', *Research letters in signal processing*, 2008, pp. 1–5. doi: 10.1155/2008/825758.

Alonso, F. J., Castillo, J. M. D. and Pintado, P. (2005) 'Application of singular spectrum analysis to the smoothing of raw kinematic signals', *Journal of Biomechanics*. Elsevier, 38(5), pp. 1085–1092. doi: 10.1016/J.JBIOMECH.2004.05.031.

Arlinger, S. *et al.* (2009) 'The emergence of Cognitive Hearing Science', *Scandinavian Journal of Psychology*. Wiley/Blackwell (10.1111), 50(5), pp. 371–384. doi: 10.1111/j.1467-9450.2009.00753.x.

Arrowsmith, S. J. and Hedlin, M. A. H. (2005) 'Observations of infrasound from surf in southern California', *Geophysical Research Letters*. Wiley-Blackwell, 32(9), p. L09810. doi: 10.1029/2005GL022761.

Arul Elango, G., Sudha, G. F. and Francis, B. (2017) 'Weak signal acquisition enhancement in software GPS receivers – Pre-filtering combined post-correlation detection approach', *Applied Computing and Informatics*, 13(1), pp. 66–78. doi: 10.1016/j.aci.2014.10.002.

Attenborough, K. (1983) 'Acoustical characteristics of rigid fibrous absorbents and granular materials', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 73(3), pp. 785–799. doi: 10.1121/1.389045.

Attenborough, K. *et al.* (1986) 'The acoustic transfer function at the surface of a layered poroelastic soil', *The Journal of the Acoustical Society of America*. Acoustical

Society of America, 79(5), pp. 1353–1358. doi: 10.1121/1.393663.

Balakrichenan, S., Kin-Foo, A. and Souissi, M. (2011) 'Qualitative Evaluation of a Proposed Federated Object Naming Service Architecture', in *2011 International Conference on Internet of Things and 4th International Conference on Cyber, Physical and Social Computing*. IEEE, pp. 726–732. doi: 10.1109/iThings/CPSCom.2011.72.

Ballou, G. (2015) *Handbook for sound engineers*. Focal Press.

Batchelor, G. K. (1951) 'Pressure fluctuations in isotropic turbulence', in *Mathematical Proceedings of the Cambridge Philosophical Society*. Cambridge University Press, pp. 359–374. doi: 10.1017/S0305004100026712.

Benesty, J., Souden, M. and Huang, Y. (2012) 'A Perspective on Differential Microphone Arrays in the Context of Noise Reduction', *IEEE Transactions on Audio, Speech, and Language Processing*, 20(2), pp. 699–704. doi: 10.1109/TASL.2011.2163396.

Bentler, R. *et al.* (2008) 'Digital noise reduction: Outcomes from laboratory and field studies', *International Journal of Audiology*. Taylor & Francis, 47(8), pp. 447–460. doi: 10.1080/14992020802033091.

Berman, S. and Stearns, C. R. (1977) 'Near-earth turbulence and coherence measurements at Aberdeen Proving Ground, Md.', *Boundary-Layer Meteorology*. Kluwer Academic Publishers, 11(4), pp. 485–506. doi: 10.1007/BF02185872.

Bjornsson, H. and Venegas, S. A. (1997) 'A manual for EOF and SVD analyses of climatic data', *CCGCR Report*, 97(1), pp. 112–134.

Boll, S. F. (1979) 'Suppression of Acoustic Noise in Speech Using Spectral Subtraction', *IEEE Trans. on Acoust., Speech, and Signal Proc.*, 27(2), pp. 113–120.

Bonizzi, P. *et al.* (2015) 'Sleep apnea detection directly from unprocessed ECG through singular spectrum decomposition', in *2015 Computing in Cardiology Conference (CinC)*. IEEE, pp. 309–312. doi: 10.1109/CIC.2015.7408648.

Bowman, H. S. and Bedard, A. J. (2010) 'Observations of Infrasound and Subsonic Disturbances Related to Severe Weather', *Geophysical Journal of the Royal Astronomical Society*. Oxford University Press, 26(1–4), pp. 215–242. doi: 10.1111/j.1365-246X.1971.tb03396.x.

Bradley, S. *et al.* (2003) 'The Mechanisms Creating Wind Noise in Microphones', in *Audio Engineering Society Convention 114*. Audio Engineering Society.

Brandstein, M. and Ward, D. (2013) *Microphone arrays: signal processing techniques and applications*. Springer Science & Business Media.

Burridge, R. (2010) 'The Acoustics of Pipe Arrays', *Geophysical Journal of the Royal Astronomical Society*. Wiley/Blackwell (10.1111), 26(1–4), pp. 53–69. doi: 10.1111/j.1365-246X.1971.tb03382.x.

Caesarendra, W. *et al.* (2016) 'Acoustic emission-based condition monitoring methods: Review and application for low speed slew bearing', *Mechanical Systems and Signal Processing*. Elsevier, 72–73, pp. 134–159. doi: 10.1016/j.ymssp.2015.10.020.

Canavier, C. C., Clark, J. W. and Byrne, J. H. (1990) 'Routes to chaos in a model of a bursting neuron.', *Biophysical journal*. Elsevier, 57(6), pp. 1245–51. doi: 10.1016/S0006-3495(90)82643-6.

Cermak, J. *et al.* (2007) 'Blind Source Separation Based on a Beamformer Array

and Time Frequency Binary Masking', in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*. IEEE, pp. I-145-I–148. doi: 10.1109/ICASSP.2007.366637.

Cheeke, J. D. N. (2016) *Fundamentals and Applications of Ultrasonic Waves*. CRC Press. doi: 10.1201/b12260.

Christie, D. R. and Campus, P. (2010) 'The IMS Infrasound Network: Design and Establishment of Infrasound Stations', in *Infrasound Monitoring for Atmospheric Studies*. Dordrecht: Springer Netherlands, pp. 29–75. doi: 10.1007/978-1-4020-9508-5_2.

Chu, M. T., Lin, M. M. and Wang, L. (2013) 'A study of singular spectrum analysis with global optimization techniques', *Journal of Global Optimization*, 60(3), pp. 551–574. doi: 10.1007/s10898-013-0117-3.

Chu, S., Narayanan, S. and Kuo, C.-C. J. (2008) 'Environmental sound recognition using MP-based features', in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, pp. 1–4. doi: 10.1109/ICASSP.2008.4517531.

Chung, K. (2004) 'Challenges and Recent Developments in Hearing Aids: Part I. Speech Understanding in Noise, Microphone Technologies and Noise Reduction Algorithms', *Trends in Amplification*. SAGE PublicationsSage CA: Los Angeles, CA, 8(3), pp. 83–124. doi: 10.1177/108471380400800302.

Cicirelli, F., Guerrieri, A., *et al.* (2017) 'An edge-based approach to develop large-scale smart environments by leveraging SIoT', in *2017 IEEE 14th International Conference on Networking, Sensing and Control (ICNSC)*. IEEE, pp. 738–743. doi: 10.1109/ICNSC.2017.8000182.

Cicirelli, F., Fortino, G., *et al.* (2017) 'Metamodeling of Smart Environments: from design to implementation', *Advanced Engineering Informatics*, 33, pp. 274–284. doi: 10.1016/j.aei.2016.11.005.

Cicirelli, F. *et al.* (2018) 'Edge Computing and Social Internet of Things for large-scale smart environments development', *IEEE Internet of Things Journal*, 5(4), pp. 2557–2571. doi: 10.1109/JIOT.2017.2775739.

Claessen, D. and Groth, D. (2002) 'A beginner's guide to SSA', *CERES-ERTI, Ecole Normale Superieure,*.

Clifford, G. D. (2005) 'Singular Value Decomposition & Independent Component Analysis for Blind Source Separation', *Biomedical Signal and Image Processing*, 44, pp. 489–499.

Connolly, D. M. *et al.* (2013) 'Adolescents' perceptions of their school's acoustic environment: the development of an evidence based questionnaire.', *Noise & health*. Medknow Publications and Media Pvt. Ltd., 15(65), pp. 269–80. doi: 10.4103/1463-1741.113525.

Cook, D. and Das, S. (2004) *Smart environments: Technology, protocols and applications.* Vol. 43. John Wiley & Sons.

Cook, D. J. and Das, S. K. (2007) 'How smart are our environments? An updated look at the state of the art', *Pervasive and Mobile Computing*, 3, pp. 53–73. doi: 10.1016/j.pmcj.2006.12.001.

Crocker, M. J. ed. (2007) *Handbook of noise and vibration control*. John Wiley & Sons.

Daniels, F. B. (1952) 'Acoustical Energy Generated by the Ocean Waves', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 24(1), pp. 83–83. doi: 10.1121/1.1906855.

Daniels, F. B. (1959) 'Noise-Reducing Line Microphone for Frequencies below 1 cps', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 31(4), pp. 529–531. doi: 10.1121/1.1907747.

Das, S. K. and Cook, D. J. (2005) 'Designing Smart Environments: A Paradigm Based on Learning and Prediction.', in. Springer, Berlin, Heidelberg, pp. 80–90. doi: 10.1007/11590316_11.

Delgado-Arredondo, P. A. *et al.* (2017) 'Methodology for fault detection in induction motors via sound and vibration signals', *Mechanical Systems and Signal Processing*. Academic Press, 83, pp. 568–589. doi: 10.1016/J.YMSSP.2016.06.032.

Dey, A. K., Abowd, G. D. and Salber, D. (2000) 'A Context-Based Infrastructure for Smart Environments', in *Managing Interactions in Smart Environments*. London: Springer London, pp. 114–128. doi: 10.1007/978-1-4471-0743-9_11.

Dietrich, F. A. and Utschick, W. (2005) 'Pilot-assisted channel estimation based on second-order statistics', *IEEE Transactions on Signal Processing*, 53(3), pp. 1178–1193. doi: 10.1109/TSP.2004.842176.

Dillion, K., Howard, W. and Shields, F. D. (2007) 'Advances in distributed arrays for detection of infrasonic events', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 122(5), pp. 2960–2960. doi: 10.1121/1.2942549.

Dillon, H. (2008) *Hearing aids*. Hodder Arnold. doi: 10.1037/023990.

Ding, G. H. *et al.* (2005) 'Speech enhancement based on speech spectral complex Gaussian mixture model', in *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.* IEEE, pp. I–165.

Ding, H., Lee, T. and Soon, I. Y. (2012) 'Two objective measures for speech distortion and noise reduction evaluation of enhanced speech signals', in *2012 8th International Symposium on Chinese Spoken Language Processing*. IEEE, pp. 117–121. doi: 10.1109/ISCSLP.2012.6423488.

Dmochowski, J. P. and Goubran, R. A. (2007) 'Decoupled Beamforming and Noise Cancellation', *IEEE Transactions on Instrumentation and Measurement*, 56(1), pp. 80–88. doi: 10.1109/TIM.2006.887196.

Doclo, S. *et al.* (2007) 'Frequency-Domain Criterion for the Speech Distortion Weighted Multichannel Wiener Filter for Robust Noise Reduction', *Speech Communication*, 49(7–8), pp. 636–656. doi: 10.1016/j.specom.2007.02.001.

Douglas Shields, F. (2005) 'Low-frequency wind noise correlation in microphone arrays', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 117(6), pp. 3489–3496. doi: 10.1121/1.1879252.

Dubbelboer, F. and Houtgast, T. (2007) 'A detailed study on the effects of noise on speech intelligibility', *The Journal of the Acoustical Society of America*, 122(5), pp. 2865–2871. doi: 10.1121/1.2783131.

Duncan, P. J., Gaydecki, P. A. and Burdekin, F. M. (1996) 'Ultrasonic NDT Prototype for the Inspection of Ducted Post Stressing Tendons in Concrete Beams', in *Review of Progress in Quantitative Nondestructive Evaluation*. Boston, MA: Springer US, pp. 1799–1806. doi: 10.1007/978-1-4613-0383-1_235.

Eckhoff, D. and Wagner, I. (2018) 'Privacy in the Smart City—Applications, Technologies, Challenges, and Solutions', *IEEE Communications Surveys & Tutorials*, 20(1), pp. 489–516. doi: 10.1109/COMST.2017.2748998.

Eftaxias, K. *et al.* (2015) 'Detection of Parkinson's tremor from EMG signals; a singular spectrum analysis approach', in *2015 IEEE International Conference on Digital Signal Processing (DSP)*. IEEE, pp. 398–402. doi: 10.1109/ICDSP.2015.7251901.

Eldwaik, O. and F. Li, F. (2018) 'Mitigating Wind Induced Noise in Outdoor Microphone Signals Using a Singular Spectral Subspace Method', *Technologies*. Multidisciplinary Digital Publishing Institute, 6(1), p. 19. doi: 10.3390/technologies6010019.

Eldwaik, O. and FF Li (2017) 'Microphone wind noise reduction using singular spectrum analysis techniques', *Proceedings of the Institute of Acoustics*, Vol. 39(Pt. 1.).

Eldwaik, O. and Li, F. F. (2017) 'Mitigating wind noise in outdoor microphone signals using a singular spectral subspace method', in *2017 Seventh International Conference on Innovative Computing Technology (INTECH)*. IEEE, pp. 149–154. doi: 10.1109/INTECH.2017.8102428.

Elko, G. W. (2007) 'MH ACOUSTICS LLC. Reducing noise in audio systems.', *U.S. Patent 7,171,008.*

Ellis, D. and Weiss, R. (2006) 'Model-Based Monaural Source Separation Using a Vector-Quantized Phase-Vocoder Representation.', *In ICASSP*, 5, pp. 957-960.

Elsner, J. and Tsonis, A. (2013) *Singular spectrum analysis: a new tool in time series analysis*. Springer Science & Business Media.

Enshaeifar, S. *et al.* (2016) 'Quaternion Singular Spectrum Analysis of Electroencephalogram With Application in Sleep Analysis', *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 24(1), pp. 57–67. doi: 10.1109/TNSRE.2015.2465177.

Enshaeifar, S., Sanei, S. and Took, C. C. (2014) 'An eigen-based approach for complex-valued Forecasting', in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, pp. 6014–6018. doi: 10.1109/ICASSP.2014.6854758.

Fukuda, K. (2007) 'Noise reduction approach for decision tree construction: A case study of knowledge discovery on climate and air pollution', in *2007 IEEE Symposium on Computational Intelligence and Data Mining*. IEEE, pp. 697–704.

Fysarakis, K. *et al.* (2018) 'XSACd-Cross-domain Resource Sharing & Access Control for Smart Environments', *Future Generation Computer Systems*, 80, pp. 572–582.

Gámiz-Fortis, S. R. *et al.* (2002) 'Spectral characteristics and predictability of the NAO assessed through Singular Spectral Analysis', *Journal of Geophysical Research: Atmospheres*, 107(D23). doi: 10.1029/2001JD001436.

Garcés, M. *et al.* (2004) 'On using ocean swells for continuous infrasonic measurements of winds and temperature in the lower, middle, and upper atmosphere', *Geophysical Research Letters*. Wiley-Blackwell, 31(19), p. L19304. doi: 10.1029/2004GL020696.

García Plaza, E. and Núñez López, P. J. (2017) 'Surface roughness monitoring by singular spectrum analysis of vibration signals', *Mechanical Systems and Signal*

Processing. Elsevier, 84, pp. 516–530. doi: 10.1016/j.ymssp.2016.06.039.

George, W. K., Beuther, P. D. and Arndt, R. E. (1984) 'Pressure spectra in turbulent free shear flows', *Journal of Fluid Mechanics*. Cambridge University Press, 148(1), pp. 155–191. doi: 10.1017/S0022112084002299.

Gerkmann, T. and Hendriks, R. C. (2011) 'Noise power estimation based on the probability of speech presence', in *2011 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*. IEEE, pp. 145–148. doi: 10.1109/ASPAA.2011.6082266.

Ghaderi, F., Mohseni, H. R. and Sanei, S. (2011) 'Localizing Heart Sounds in Respiratory Signals Using Singular Spectrum Analysis', *IEEE Transactions on Biomedical Engineering*, 58(12), pp. 3360–3367. doi: 10.1109/TBME.2011.2162728.

Ghil, M. *et al.* (2002) 'Advanced spectral methods for climatic time series', *Reviews of Geophysics*, 40(1), pp. 3–1. doi: 10.1029/2000RG000092.

Ghodsi, M. *et al.* (2009) 'The use of noise information for detection of temporomandibular disorder', *Biomedical Signal Processing and Control*. Elsevier, 4(2), pp. 79–85. doi: 10.1016/J.BSPC.2008.10.001.

Glowacz, A. (2016) 'Diagnostics of Rotor Damages of Three-Phase Induction Motors Using Acoustic Signals and SMOFS-20-EXPANDED', *Archives of Acoustics*, 41(3), pp. 507–515. doi: 10.1515/aoa-2016-0049.

Glowacz, A. *et al.* (2018) 'Early fault diagnosis of bearing and stator faults of the single-phase induction motor using acoustic signals', *Measurement*. Elsevier, 113(August 2017), pp. 1–9. doi: 10.1016/j.measurement.2017.08.036.

Głowacz, A. and Głowacz, Z. (2017) 'Recognition of rotor damages in a DC motor using acoustic signals', *Bulletin of the Polish Academy of Sciences Technical Sciences*, 65(2), pp. 187–194. doi: 10.1515/bpasts-2017-0023.

Goedecke, G. H. and Auvermann, H. J. (1997) 'Acoustic scattering by atmospheric turbules', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 102(2), pp. 759–771. doi: 10.1121/1.419951.

Golyandina, N. *et al.* (2013) 'Multivariate and 2D Extensions of Singular Spectrum Analysis with the Rssa Package', *arXiv preprint arXiv:1309.5050*.

Golyandina, N. E. and Lomtev, M. A. (2016) 'Improvement of separability of time series in singular spectrum analysis using the method of independent component analysis', *Vestnik St. Petersburg University: Mathematics*, 49(1), pp. 9–17. doi: 10.3103/S1063454116010064.

Golyandina, N. and Korobeynikov, A. (2014) 'Basic Singular Spectrum Analysis and forecasting with R', *Computational Statistics & Data Analysis*. North-Holland, 71, pp. 934–954. doi: 10.1016/J.CSDA.2013.04.009.

Golyandina, N., Nekrutkin, V. and Zhigljavsky, A. . (2001) *Analysis of Time Series Structure: SSA and related techniques*. Chapman and Hall/CRC (C&H/CRC Monographs on Statistics & Applied Probability). doi: 10.1201/9781420035841.

Golyandina, N. and Shlemov, A. (2013) 'Variations of singular spectrum analysis for separability improvement: non-orthogonal decompositions of time series', *arXiv preprint arXiv:1308.4022*. doi: 10.4310/SII.2015.v8.n3.a3.

Golyandina, N. and Shlemov, A. (2015) 'Semi-nonparametric singular spectrum

analysis with projection', pp. 1–14.

Grancharov, V. *et al.* (2008) 'Generalized Postfilter for Speech Quality Enhancement', *IEEE Transactions on Audio, Speech, and Language Processing*, 16(1), pp. 57–64. doi: 10.1109/TASL.2007.909327.

Grill, T. and Schluter, J. (2017) 'Two convolutional neural networks for bird detection in audio signals', in *2017 25th European Signal Processing Conference (EUSIPCO)*. IEEE, pp. 1764–1768. doi: 10.23919/EUSIPCO.2017.8081512.

Grimm, S. and Freudenberger, J. (2018) 'Wind noise reduction for a closely spaced microphone array in a car environment', *EURASIP Journal on Audio, Speech, and Music Processing*. Springer International Publishing, 2018(1), p. 7. doi: 10.1186/s13636-018-0130-z.

Grover, F. H. (2010) 'Experimental Noise Reducers for an Active Microbarograph Array', *Geophysical Journal of the Royal Astronomical Society*. Oxford University Press, 26(1–4), pp. 41–52. doi: 10.1111/j.1365-246X.1971.tb03381.x.

Hancke, G. P., Silva, B. de C. E. and Hancke Jr, G. P. (2013) 'The role of advanced sensing in smart cities.', *Sensors (Basel, Switzerland)*. Multidisciplinary Digital Publishing Institute (MDPI), 13(1), pp. 393–425. doi: 10.3390/s130100393.

Hansen, B. and Noguchi, K. (2017) 'Improved short-term point and interval forecasts of the daily maximum tropospheric ozone levels via singular spectrum analysis', *Environmetrics*, 28(8), p. e2479. doi: 10.1002/env.2479.

Hänsler, E. and Schmidt, G. (2005) *Acoustic echo and noise control: a practical approach (Vol. 40).*, *Communication*. John Wiley & Sons. doi: 10.1002/0471678406.

Harmouche, J. *et al.* (2017) 'The Sliding Singular Spectrum Analysis: a Data-Driven Non-Stationary Signal Decomposition Tool', *IEEE Transactions on Signal Processing*, 66(1), pp. 251–263. doi: 10.1109/TSP.2017.2752720.

Harris, T. J. and Yuan, H. (2010) 'Filtering and frequency interpretations of Singular Spectrum Analysis', *Physica D: Nonlinear Phenomena*. North-Holland, 239(20–22), pp. 1958–1967. doi: 10.1016/J.PHYSD.2010.07.005.

Hassani, H. (2007) 'Singular Spectrum Analysis: Methodology and Comparison', *Journal of Data Science*, 5, pp. 239–257. doi: 10.1227/01.NEU.0000349921.14519.2A.

Hassani, H. *et al.* (2009) 'Does noise reduction matter for curve fitting in growth curve models?', *Computer methods and programs in biomedicine*, 96(3), pp. 173–181.

Hassani, H. (2010) 'A brief introduction to singular spectrum analysis', *Optimal decisions in statistics and data analysis. Paper in pdf version available at: www. ssa. cf. ac. uk/ ...*, pp. 1–11.

Hassani, H., Dionisio, A. and Ghodsi, M. (2010) 'The effect of noise reduction in measuring the linear and nonlinear dependency of financial markets', *Nonlinear Analysis: Real World Applications,* 11(1), pp. 492–502.

Hassani, H., Heravi, S. and Zhigljavsky, A. (2009) 'Forecasting European industrial production with singular spectrum analysis', *International journal of forecasting*, 25(1), pp. 103–118.

Hassani, H. and Mahmoudvand, R. (2013) 'Multivariate Singular Spectrum Analysis: a General View and New Vector Forecasting Approach', *International Journal of Energy and Statistics*, 1(01), pp. 55–83. doi: 10.1142/S2335680413500051.

Hassani, H., Mahmoudvand, R. and Zokaei, M. (2011) 'Separability and window length in singular spectrum analysis', *Comptes rendus mathematique*, 349(17–18), p. p.987-990.

Hassani, H., Soofi, A. S. and Zhigljavsky, A. A. (2010) 'Predicting daily exchange rate with singular spectrum analysis', *Nonlinear Analysis: Real World Applications*. Elsevier Ltd, 11(3), pp. 2023–2034. doi: 10.1016/j.nonrwa.2009.05.008.

Hayes, M. H. (2009) 'Statistical digital signal processing and modeling', *John Wiley & Sons*.

Hedlin, M. A. H. and Raspet, R. (2003) 'Infrasonic wind-noise reduction by barriers and spatial filters', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 114(3), pp. 1379–1386. doi: 10.1121/1.1598198.

Hendriks, R. C., Heusdens, R. and Jensen, J. (2010) 'MMSE based noise PSD tracking with low complexity', in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, pp. 4266–4269. doi: 10.1109/ICASSP.2010.5495680.

Herrin, E. *et al.* (2001) 'Comparative Evaluation of Selected Infrasound Noise Reduction Methods'. SOUTHERN METHODIST UNIV DALLAS TX.

Higgins, C. W. *et al.* (2012) 'The Effect of Scale on the Applicability of Taylor's Frozen Turbulence Hypothesis in the Atmospheric Boundary Layer', *Boundary-layer meteorology*, 143(2), pp. 379–391. doi: 10.1007/s10546-012-9701-1.

Hill, R. J. and Wilczak, J. M. (1995) 'Pressure structure functions and spectra for locally isotropic turbulence', *Journal of Fluid Mechanics*. Cambridge University Press, 296(1), pp. 247–269. doi: 10.1017/S0022112095002126.

Hofmann, C. *et al.* (2012) 'A morphological approach to single-channel wind-noise suppresion', in *IWAENC 2012; International Workshop on Acoustic Signal Enhancement*, pp. 1-4. VDE.

Hollosi, D. *et al.* (2013) 'Enhancing Wireless Sensor Networks with Acoustic Sensing Technology: Use Cases, Applications & Experiments', in *2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*. IEEE, pp. 335–342. doi: 10.1109/GreenCom-iThings-CPSCom.2013.75.

Howard, W., Dillion, K. and Shields, F. D. (2007) 'Acoustical properties of porous hose wind noise filters', *The Journal of the Acoustical Society of America*, 122(5), pp. 2985–2985. doi: 10.1121/1.2942646.

Hu, H. *et al.* (2017) 'An adaptive singular spectrum analysis method for extracting brain rhythms of electroencephalography', *PeerJ*, 5, p. e3474. doi: 10.7717/peerj.3474.

Hu, Y. and Loizou, P. C. (2007) 'A comparative intelligibility study of single-microphone noise reduction algorithms', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 122(3), pp. 1777–1786. doi: 10.1121/1.2766778.

Huber, D. M. and Runstein, R. E. (2013) 'Modern Recording Techniques', *Focal press*. Routledge. doi: 10.4324/9780240824642.

Jarchi, D. and Yang, G.-Z. (2013) 'Singular spectrum analysis for gait patterns', in *2013 IEEE International Conference on Body Sensor Networks*. IEEE, pp. 1–6. doi: 10.1109/BSN.2013.6575492.

Jena, D. P. and Panigrahi, S. N. (2015) 'Automatic gear and bearing fault localization using vibration and acoustic signals', *Applied Acoustics*. Elsevier Ltd, 98, pp. 20–33. doi: 10.1016/j.apacoust.2015.04.016.

Jiang, J. and Xie, H. B. (2016) 'Denoising Nonlinear Time Series Using Singular Spectrum Analysis and Fuzzy Entropy', *Chinese Physics Letters*. IOP Publishing, 33(10), p. 100501. doi: 10.1088/0256-307X/33/10/100501.

Jingdong Chen *et al.* (2006) 'New insights into the noise reduction Wiener filter', *IEEE Transactions on Audio, Speech and Language Processing*, 14(4), pp. 1218–1234. doi: 10.1109/TSA.2005.860851.

Johnson, K. (2011) *Acoustic and auditory phonetics*. John Wiley & Sons.

Józwik, J. (2016) 'Identification and Monitoring of Noise Sources of Cnc Machine Tools By Acoustic Holography Methods', *Advances in Science and Technology Research Journal*, 10(30), pp. 127–137. doi: 10.12913/22998624/63386.

Juin-Hwey Chen and Gersho, A. (1995) 'Adaptive postfiltering for quality enhancement of coded speech', *IEEE Transactions on Speech and Audio Processing*, 3(1), pp. 59–71. doi: 10.1109/89.365380.

Kaimal, J. C. *et al.* (1982) 'Spectral Characteristics of the Convective Boundary Layer Over Unever Terrain', *Journal of the Atmospheric Sciences*, pp. 1098–1114. doi: 10.1175/1520-0469(1982)039<1098:SCOTCB>2.0.CO;2.

Karbasi, M., Ahadi, S. M. and Bahmanian, M. (2011) 'Environmental sound classification using spectral dynamic features', in *2011 8th International Conference on Information, Communications & Signal Processing*. IEEE, pp. 1–5. doi: 10.1109/ICICS.2011.6173513.

Keith Wilson, David, Keith Wilson, D and White, M. J. (2010) 'Discrimination of Wind Noise and Sound Waves by Their Contrasting Spatial and Temporal Properties', *Acta Acustica united with Acustica*, 96(6), pp. 991–1002. doi: 10.3813/AAA.918362.

Kelly, B. *et al.* (2014) 'Application of Acoustic Sensing Technology for Improving Building Energy Efficiency', *Procedia Computer Science*, 32, pp. 661–664. doi: 10.1016/j.procs.2014.05.474.

Kendrick, P. *et al.* (2015) 'Perceived audio quality of sounds degraded by non-linear distortions and single-ended assessment using HASQI', *Journal of the Audio Engineering Society*, 63(9), pp. 698–712.

Keshavarzi, M. *et al.* (2018) 'Use of a Deep Recurrent Neural Network to Reduce Wind Noise: Effects on Judged Speech Intelligibility and Sound Quality', *Trends in hearing*, 22, p. .2331216518770964. doi: 10.1177/2331216518770964.

King, B. and Atlas, L. (2008) 'Coherent modulation comb filtering for enhancing speech in wind noise', *In International Workshop on Acoustice Echo and Noise Control*, pp. 14–17.

Kinsler, L. . *et al.* (1999) *Fundamentals of Acoustics. Fundamentals of Acoustics, 4th Edition, by Lawrence E. Kinsler, Austin R. Frey, Alan B. Coppens, James V. Sanders,*. pp. 560. ISBN 0-471-84789-5. Wiley-VCH, December 1999., p.560.

Knaak, M., Araki, S. and Makino, S. (2007) 'Geometrically Constrained Independent Component Analysis', *IEEE Transactions on Audio, Speech and Language Processing*, 15(2), pp. 715–726. doi: 10.1109/TASL.2006.876730.

Kochkin, S. (2005) 'MarkeTrak VII: Customer satisfaction with hearing instruments in the digital age', *The Hearing Journal*, 58(9), pp. 30-32.

Kochkin, S. (2010) 'MarkeTrak VIII: Consumer satisfaction with hearing aids is slowly increasing', *The Hearing Journal*, 63(1), pp. 19–20. doi: 10.1097/01.HJ.0000366912.40173.76.

Komninos, N. (2009) 'Intelligent cities: Towards interactive and global innovation environments', *International Journal of Innovation and Regional Development*, 1(4), pp. 337–355. doi: 10.1504/IJIRD.2009.022726.

Koracin, D. and Berkowicz, R. (1988) 'Nocturnal boundary-layer height: Observations by acoustic sounders and predictions in terms of surface-layer parameters', *Boundary-Layer Meteorology*. Kluwer Academic Publishers, 43(1–2), pp. 65–83. doi: 10.1007/BF00153969.

Korhonen, P. *et al.* (2017) 'Evaluation of a Wind Noise Attenuation Algorithm on Subjective Annoyance and Speech-in-Wind Performance', *Journal of the American Academy of Audiology*, 28(1), pp. 46–57. doi: 10.3766/jaaa.15135.

Kostelich, E. J. and Yorke, J. A. (1988) 'Noise reduction in dynamical systems', *Physical Review A*, 38(3), pp. 1649–1652. doi: 10.1103/PhysRevA.38.1649.

Kotus, J., Lopatka, K. and Czyzewski, A. (2014) 'Detection and localization of selected acoustic events in acoustic field for smart surveillance applications', *Multimedia Tools and Applications*, 68(1), pp. 5–21. doi: 10.1007/s11042-012-1183-0.

Kuroiwa, S. *et al.* (2006) 'Wind noise reduction method for speech recording using multiple noise templates and observed spectrum fine structure', in *2006 International Conference on Communication Technology*. IEEE, pp. 1–5. doi: 10.1109/ICCT.2006.341933.

Kuttruff, H. (2006) *Acoustics: An introduction*. CRC Press.

Lakshmi, K., Rao, A. R. M. and Gopalakrishnan, N. (2017) 'Singular spectrum analysis combined with ARMAX model for structural damage detection', *Structural Control and Health Monitoring*, 24(9), pp. 1–21. doi: 10.1002/stc.1960.

Larson, R. J. *et al.* (2010) 'Correlation of Winds and Geographic Features with Production of Certain Infrasonic Signals in the Atmosphere', *Geophysical Journal of the Royal Astronomical Society*. Oxford University Press, 26(1–4), pp. 201–214. doi: 10.1111/j.1365-246X.1971.tb03395.x.

Launonen, I. and Holmström, L. (2017) 'Multivariate posterior singular spectrum analysis', *Statistical Methods and Applications*. Springer Berlin Heidelberg, 26(3), pp. 361–382. doi: 10.1007/s10260-016-0372-9.

Lee, T. K. *et al.* (2013) 'Assessing rehabilitative reach and grasp movements with singular spectrum analysis', in *21st European Signal Processing Conference (EUSIPCO 2013)*. IEEE, pp. 1–5.

Lenth, R. V (2008) 'Some Practical Guidelines for Effective Sample Size Determination', *The American Statistician*, 55(3), pp. 187–193. doi: 10.1198/000313001317098149.

Li, F. F. and Cox, T. J. (2007) 'A neural network model for speech intelligibility quantification', *Applied Soft Computing*. Elsevier, 7(1), pp. 145–155. doi: 10.1016/J.ASOC.2005.05.002.

Loizou, P. (2007) *Speech enhancement: theory and practice*. CRC press.

Luther, D. and Gentry, K. (2013) 'Sources of Background Noise and Their Influence on Vertebrate Acoustic Communication', *Behaviour*, 150(9–10), pp. 1045–1068. doi: 10.1163/1568539X-00003054.

Luts, H. *et al.* (2010) 'Multicenter evaluation of signal enhancement algorithms for hearing aids', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 127(3), pp. 1491–1505. doi: 10.1121/1.3299168.

Luzzi, S. *et al.* (2013) 'Acoustics for smart cities', in *Proceedings of the Annual Conference on Acoustics (AIA-DAGA), Merano, Italy*, pp. 18–21.

Luzzi, S. (2013) 'Urban noise management and its practical implementation', in *Proceedings 20th International Congress on Sound and Vibration (ICSV20'13)*.

Ma, H. G. *et al.* (2012) 'Determine a proper window length for singular spectrum analysis', in *IET International Conference on Radar Systems (Radar 2012)*. Institution of Engineering and Technology, pp. 1–6. doi: 10.1049/cp.2012.1722.

Ma, L., Milner, B. and Smith, D. (2006) 'Acoustic environment classification', *ACM Transactions on Speech and Language Processing*. ACM, 3(2), pp. 1–22. doi: 10.1145/1149290.1149292.

Maddirala, A. and Shaik, R. A. (2016) 'Removal of EOG Artifacts from Single Channel EEG Signals using Combined Singular Spectrum Analysis and Adaptive Noise Canceler', *IEEE Sensors Journal*, 16(23), pp. 8279–8287. doi: 10.1109/JSEN.2016.2560219.

Mahmoudvand, R., Konstantinides, D. and Rodrigues, P. C. (2017) 'Forecasting mortality rate by multivariate singular spectrum analysis', *Applied Stochastic Models in Business and Industry*, 33(6), pp. 717–732. doi: 10.1002/asmb.2274.

Mamou, J. and Feleppa, E. J. (2007) 'Singular spectrum analysis applied to ultrasonic detection and imaging of brachytherapy seeds', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 121(3), pp. 1790–1801. doi: 10.1121/1.2436713.

Manwell, J. F., McGowan, J. G. and Rogers, A. L. (2010) *Wind energy explained: theory, design and application*. John Wiley & Sons. doi: 10.1007/978-1-4419-7991-9_34.

Mao, G., Fidan, B. and Anderson, B. D. O. (2007) 'Wireless sensor network localization techniques', *Computer Networks*. Elsevier, 51(10), pp. 2529–2553. doi: 10.1016/J.COMNET.2006.11.018.

Maropoulos, P. G. and Ceglarek, D. (2010) 'Design verification and validation in product lifecycle', *CIRP Annals*. Elsevier, 59(2), pp. 740–759. doi: 10.1016/J.CIRP.2010.05.005.

Marshall, D. F. (1984) 'Wind noise reduction using a two microphone array', *The Journal of the Acoustical Society of America*, 76(S1), pp. S70–S70. doi: 10.1121/1.2021992.

Martin, R. (2001) 'Noise power spectral density estimation based on optimal smoothing and minimum statistics', *IEEE Transactions on Speech and Audio Processing*, 9(5), pp. 504–512. doi: 10.1109/89.928915.

De Marziani, C. *et al.* (2009) 'Acoustic Sensor Network for Relative Positioning of Nodes', *Sensors*. Molecular Diversity Preservation International, 9(11), pp. 8490–8507.

doi: 10.3390/s91108490.

McBride, W. E. *et al.* (1992) 'Scattering of sound by atmospheric turbulence: Predictions in a refractive shadow zone', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 91(3), pp. 1336–1340. doi: 10.1121/1.402462.

McDonald, J. A., Douze, E. J. and Herrin, E. (2010) 'The Structure of Atmospheric Turbulence and its Application to the Design of Pipe Arrays', *Geophysical Journal of the Royal Astronomical Society*. Oxford University Press, 26(1–4), pp. 99–109. doi: 10.1111/j.1365-246X.1971.tb03385.x.

Meister, H. *et al.* (2013) 'Examining Speech Perception in Noise and Cognitive Functions in the Elderly', *American Journal of Audiology*. American Speech-Language-Hearing Association, 22(2), p. 310. doi: 10.1044/1059-0889(2012/12-0067).

Mert, C. and Milnikov, A. (2011) 'Singular Spectrum Analysis Method as a universal filter', in *2011 5th International Conference on Application of Information and Communication Technologies (AICT)*. IEEE, pp. 1–5. doi: 10.1109/ICAICT.2011.6111004.

Mesaros, A. *et al.* (2010) 'Acoustic event detection in real-life recordings', in *2010 18th European Signal Processing Conference*. IEEE., pp. 1267–1271.

Mika, D. and Józwik, J. (2016) 'Normative Measurements of Noise At Cnc Machines Work Stations', *Advances in Science and Technology Research Journal*, 10(30), pp. 138–143. doi: 10.12913/22998624/63387.

Miles, N. L., Wyngaard, J. C. and Otte, M. J. (2004) 'Turbulent Pressure Statistics in the Atmospheric Boundary Layer from Large-Eddy Simulation', *Boundary-Layer Meteorology*. Kluwer Academic Publishers, 113(2), pp. 161–185. doi: 10.1023/B:BOUN.0000039377.36809.1d.

Misra, S., Reisslein, M. and Xue, G. (2008) 'A survey of multimedia streaming in wireless sensor networks', *IEEE Communications Surveys & tutorials*, 10(4), pp. 18–39.

Mohammadi, S. M. *et al.* (2016) 'Improving time frequency domain sleep EEG classification via singular spectrum analysis', *Journal of Neuroscience Methods*, 273, pp. 96–106. doi: 10.1016/j.jneumeth.2016.08.008.

Monin, A. S. and Obukhov, A. M. (1954) 'Basic laws of turbulent mixing in the surface layer of the atmosphere', in *Contrib. Geophys. Inst. Acad. Sci. USSR*, pp. 163–187.

Moore, J. C. and Grinsted, A. (2006) 'Singular spectrum analysis and envelope detection: methods of enhancing the utility of ground-penetrating radar data', *Journal of Glaciology*, 52(176), pp. 159–163. doi: 10.3189/172756506781828863.

Moreno, M. V, Zamora, M. A. and Skarmeta, A. F. (2014) 'User-centric smart buildings for energy sustainable smart cities', *Transactions on emerging telecommunications technologies*, 25(1), pp. 41–55.

Morgan, S. and Raspet, R. (1992) 'Investigation of the mechanisms of low-frequency wind noise generation outdoors', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 92(2), pp. 1180–1183. doi: 10.1121/1.404049.

Moskvina, V. and Zhigljavsky, A. (2003) 'An Algorithm Based on Singular Spectrum Analysis for Change-Point Detection', *Communications in Statistics-Simulation and Computation*, 32(2), pp. 319–352. doi: 10.1081/SAC-120017494.

Müller, G. and Möser, M. (2012) *Handbook of engineering acoustics, Handbook of Engineering Acoustics.* Springer Science & Business Media. doi: 10.1007/978-3-540-69460-1.

Nelke, C. M. *et al.* (2014) 'Single microphone wind noise PSD estimation using signal centroids', in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* IEEE, pp. 7063–7067. doi: 10.1109/ICASSP.2014.6854970.

Nemer, E. *et al.* (2013) 'Single microphone wind noise suppression', *U.S. Patent 8,515,097.*

Nemer, E. and Leblanc, W. (2009) 'Single-microphone wind noise reduction by adaptive postfiltering', in *2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics.* IEEE, pp. 177–180. doi: 10.1109/ASPAA.2009.5346518.

Nishiyama, R. T. and Bedard, A. J. (1991) 'A "'Quad-Disc'" static pressure probe for measurement in adverse atmospheres: With a comparative review of static pressure probe designs', *Review of Scientific Instruments.* American Institute of Physics, 62(9), pp. 2193–2204. doi: 10.1063/1.1142337.

O'Brien, R. J. *et al.* (2017) 'A pattern recognition system based on acoustic signals for fault detection on composite materials', *European Journal of Mechanics, A/Solids*, 64, pp. 1–10. doi: 10.1016/j.euromechsol.2017.01.007.

Oropeza, V. and Sacchi, M. (2011) 'Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis', *Geophysics*, 76(3), pp. V25–V32.

Panofsky, H. A. and Mazzola, C. (1971) 'Variances and spectra of vertical velocity just above the surface layer', *Boundary-Layer Meteorology.* Kluwer Academic Publishers, 2(1), pp. 30–37. doi: 10.1007/BF00718086.

Panofsky, H. and Dutton, J. (1984) *Atmospheric turbulence: models and methods for engineering applications.* John Wiley and Sons, New York, 397 pp.

Patterson, K. *et al.* (2011) 'Multivariate singular spectrum analysis for forecasting revisions to real-time data', *Journal of Applied Statistics*, 38(10), pp. 2183–2211.

Penland, C., Ghil, M. and Weickmann, K. M. (1991) 'Adaptive filtering and maximum entropy spectra with application to changes in atmospheric angular momentum', *Journal of Geophysical Research: Atmospheres*, 96(D12), pp. 22659–22671. doi: 10.1029/91JD02107.

Petersen, K. S. *et al.* (2008) 'Device and method for detecting wind noise', *U.S. Patent 7,340,068.*

Petroulakis, N. . *et al.* (2013) 'A lightweight framework for secure life-logging in smart environments', *information security technical report*, 17(3), pp. 58–70.

Petroulakis, N. E., Askoxylakis, I. G. and Tryfonas, T. (2012) 'Life-logging in smart environments: Challenges and security threats', in *2012 IEEE International Conference on Communications (ICC)*, pp. 5680–5684.

Pham, C. and Cousin, P. (2013) 'Streaming the sound of smart cities: Experimentations on the SmartSantander test-bed', in *2013 IEEE international conference on green computing and communications and IEEE internet of things and IEEE cyber, physical and social computing*, pp. 611–618.

Pichora-Fuller, M. K., Schneider, B. A. and Daneman, M. (1995) 'How young and old adults listen to and remember speech in noise', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 97(1), pp. 593–608. doi: 10.1121/1.412282.

Del Pozo, S. M. and Standaert, F. X. (2015) 'Blind source separation from single measurements using singular spectrum analysis', in *International Workshop on Cryptographic Hardware and Embedded Systems*. Springer, Berlin, Heidelberg, pp. 42–59. doi: 10.1007/978-3-662-48324-4_3.

Qiao, T. *et al.* (2017) 'Effective Denoising and Classification of Hyperspectral Images Using Curvelet Transform and Singular Spectrum Analysis', *IEEE Transactions on Geoscience and Remote Sensing*, 55(1), pp. 119–133. doi: 10.1109/TGRS.2016.2598065.

Qiongfeng, P. and Aboulnasr, T. (2007) 'Efficient blind speech signal separation combining independent component analysis and beamforming', *Canadian Acoustics*, 35(3), pp. 118–119.

Raspet, R., Webster, J. and Dillion, K. (2006) 'Framework for wind noise studies', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 119(2), pp. 834–843. doi: 10.1121/1.2146113.

Raspet, R., Yu, J. and Webster, J. (2008) 'Low frequency wind noise contributions in measurement microphones', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 123(3), pp. 1260–1269. doi: 10.1121/1.2832329.

Rienstra, S. W. and Hirschberg, A. (2003) *An Introduction to Acoustics*. Eindhoven University of Technology.

Riffat, S., Powell, R. and Aydin, D. (2016) 'Future cities and environmental sustainability', *Future Cities and Environment*, 2(1), p. 1. doi: 10.1186/s40984-016-0014-2.

Rodrigues, P. C. and Mahmoudvand, R. (2017) 'The benefits of multivariate singular spectrum analysis over the univariate version', *Journal of the Franklin Institute*. Elsevier Ltd, 355(1), pp. 544–564. doi: 10.1016/j.jfranklin.2017.09.008.

Rohdenburg, T., Hohmann, V. and Kollmeier, B. (2005) 'Objective perceptual quality measures for the evaluation of noise reduction schemes', in *9th international workshop on acoustic echo and noise control*, pp. 169–172.

Rönnberg, J. *et al.* (2013) 'The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances', *Frontiers in Systems Neuroscience*. Frontiers, 7, p. 31. doi: 10.3389/fnsys.2013.00031.

Roweis, S. T. (2001) 'One microphone source separation', *Advances in neural information processing systems*, pp. 793–799.

Rudner, M. and Lunner, T. (2014) 'Cognitive spare capacity and speech communication: a narrative overview.', *BioMed research international*. Hindawi, 2014, p. 869726. doi: 10.1155/2014/869726.

Rukhin, A. L. (2002) 'Analysis of Time Series Structure SSA and Related Techniques', *Technometrics*, 44(3), pp. 290–290. doi: 10.1198/004017002320256477.

Sabatier, J. M. *et al.* (1986) 'The interaction of airborne sound with the porous ground: The theoretical formulation', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 79(5), pp. 1345–1352. doi: 10.1121/1.393662.

Saito, S. *et al.* (2011) 'Signal discrimination of ULF electromagnetic data with using singular spectrum analysis – an attempt to detect train noise', *Natural Hazards and Earth System Science*. Copernicus GmbH, 11(7), pp. 1863–1874. doi: 10.5194/nhess-11-1863-2011.

Sanchez, L. *et al.* (2014) 'SmartSantander: IoT experimentation over a smart city testbed', *Computer Networks journal*. Elsevier, 61, pp. 217–238. doi: 10.1016/J.BJP.2013.12.020.

Sanei, S., Ghodsi, M. and Hassani, H. (2011) 'An adaptive singular spectrum analysis approach to murmur detection from heart sounds', *Medical Engineering & Physics*. Elsevier, 33(3), pp. 362–367. doi: 10.1016/J.MEDENGPHY.2010.11.004.

Sanei, S. and Hassani, H. (2015) *Singular spectrum analysis of biomedical signals*. CRC Press.

Sanei, S. and Hosseini-Yazdi, A. (2011) 'Extraction of ECG from single channel EMG signal using constrained singular spectrum analysis', in *2011 17th International Conference on Digital Signal Processing (DSP)*. IEEE, pp. 1–4. doi: 10.1109/ICDSP.2011.6004876.

Sarah, M. (2014) *The Importance and Effect of Sample Size - Select Statistical Consultants*. Available at: https://select-statistics.co.uk/blog/importance-effect-sample-size/ (Accessed: 10 February 2017).

Sarode, M. and Deshmukh, P. (2011) 'Performance Evaluation of Noise Reduction Algorithm in Magnetic Resonance Images', *International Journal of Computer Science Issues (IJCSI)*, 8(2), pp. 198–202.

Schaffers, H. *et al.* (2011) 'Smart Cities and the Future Internet: Towards Cooperation Frameworks for Open Innovation', in. Springer, Berlin, Heidelberg, pp. 431–446. doi: 10.1007/978-3-642-20898-0_31.

Schlipf, D.; *et al.* (2010) 'Testing of Frozen Turbulence Hypothesis for Wind Turbine Applications with a Scanning LIDAR System', in *15th International symposium for the advancement of boundary layer remote sensing*. ISARS.

Schmidt, M. N., Larsen, J. and Hsiao, F. T. (2007) 'Wind Noise Reduction using Non-Negative Sparse Coding', in *2007 IEEE workshop on machine learning for signal processing*. IEEE., pp. 431–436. doi: 10.1109/MLSP.2007.4414345.

Schoellhamer, D. H. (2001) 'Singular spectrum analysis for time series with missing data', *Geophysical Research Letters*, 28(16), pp. 3187–3190. doi: 10.1029/2000GL012698.

Sebastian, S. and Rathnakara, S. (2013) 'Separation and localisation of heart sound artefacts from respiratory data by adaptive selection of Eigen triples in singular spectrum analysis', in *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*. IEEE, pp. 1–5. doi: 10.1109/ICCCNT.2013.6726542.

Shafik, M. and Abdalla, H. (2013) 'A new servo control drive for electro discharge texturing system industrial applications using ultrasonic technology', *International Journal of Engineering and Technology Innovation*, 3(3), pp. 180–199.

Shafik, M., Makombe, L. and Mills, B. (2013) 'Computer Simulation and Modelling of Standing Wave Piezoelectric Ultrasonic Motor Using a Single Flexural Vibration Ring Transducer', in *ASME 2013 International Mechanical Engineering*

Congress and Exposition. American Society of Mechanical Engineers., p. p.V012T13A051-V012T13A051. doi: 10.1115/IMECE2013-62925.

Shams, Q. A., Zuckerwar, A. J. and Sealey, B. S. (2005) 'Compact nonporous windscreen for infrasonic measurements', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 118(3), pp. 1335–1340. doi: 10.1121/1.1992707.

Shih, T. H. and Lumley, J. L. (1993) 'Kolmogorov Behavior of Near-Wall Turbulence and Its Application in Turbulence Modeling', *International Journal of Computational Fluid Dynamics*, 1(1), pp. 43–56.

Sivapragasam, C., Liong, S. Y. and Pasha, M. F. K. (2001) 'Rainfall and runoff forecasting with SSA-SVM approach', *Journal of Hydroinformatics*, 3(3), pp. 141–152. doi: 10.2166/hydro.2001.0014.

Slabbekoorn, H. (2013) 'Songs of the city: noise-dependent spectral plasticity in the acoustic phenotype of urban birds', *Animal Behaviour*, 85(5), pp. 1089–1099.

Sovijarvi, A. R. A. *et al.* (2000) 'Definition of Terms for Applications of Respiratory Sounds', *European Respiratory Review.*, 10(77), pp. 597–610.

Srinivasan, S., Samuelsson, J. and and Kleijn, W. B. (2003) 'Speech enhancement using a-priori information', *In Eighth European Conference on Speech Communication and Technology*.

Stenzel, S. and Freudenberger, J. (2012) 'Blind-Matched Filtering for Speech Enhancement with Distributed Microphones', *Journal of Electrical and Computer Engineering*, 2012, pp. 1–15. doi: 10.1155/2012/169853.

Stowell, D. and Plumbley, M. D. (2013) 'An open dataset for research on audio field recording archives: freefield1010', *arXiv preprint arXiv:1309.5275*.

Tao, Y., Lam, E. C. M. and Tang, Y. Y. (2001) 'Feature extraction using wavelet and fractal', *Pattern Recognition Letters*. North-Holland, 22(3–4), pp. 271–287. doi: 10.1016/S0167-8655(01)00003-4.

Temko, A. *et al.* (2006) 'Acoustic event detection and classification in smart-room environments: Evaluation of CHIL project systems', *Cough*, 65(48), p. 5.

Teutsch, H. and Elko, G. W. (2001) 'First-and second-order adaptive differential microphone arrays', in *Proc. IWAENC (Vol. 1)*.

Thuillier, R. H. *et al.* (1964) 'Wind and Temperature Profile Characteristics from Observations on a 1400 ft Tower', *Journal of Applied Meteorology*, 3(3), pp. 299–306. doi: 10.1175/1520-0450(1964)003<0299:WATPCF>2.0.CO;2.

Tomé, A. M. *et al.* (2010) 'SSA of biomedical signals: A linear invariant systems approach', *Statistics and its Interface*, 3(3), pp. 345–355.

Traore, O. I. *et al.* (2017) 'Structure analysis and denoising using Singular Spectrum Analysis: Application to acoustic emission signals from nuclear safety experiments', *Measurement*. Elsevier Ltd, 104, pp. 78–88. doi: 10.1016/j.measurement.2017.02.019.

Tzagkarakis, G., Papadopouli, M. and Tsakalides, P. (2009) 'Trend forecasting based on Singular Spectrum Analysis of traffic workload in a large-scale wireless LAN', *Performance Evaluation*, 66(3–5), pp. 173–190. doi: 10.1016/j.peva.2008.10.010.

Vary, P. and Martin, R. (2006) *Digital speech transmission: Enhancement, coding and error concealment.* John Wiley & Sons.

Vaseghi, S. V. (2008) *Advanced digital signal processing and noise reduction*. John Wiley & Sons.

Vautard, R. and Ghil, M. (1989) 'Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series', *Physica D: Nonlinear Phenomena*. North-Holland, 35(3), pp. 395–424. doi: 10.1016/0167-2789(89)90077-8.

Vautard, R. and Ghil, M. (1991) 'Interdecadal oscillations and the warming trend in global temperature time series', *Nature*, 350(6316), p. 324.

Vautard, R., Yiou, P. and Ghil, M. (1992) 'Singular-spectrum analysis: A toolkit for short, noisy chaotic signals', *Physica D: Nonlinear Phenomena*. North-Holland, 58(1–4), pp. 95–126. doi: 10.1016/0167-2789(92)90103-T.

Villanueva-Luna, A. E. *et al.* (2011) 'De-Noising Audio Signals Using MATLAB Wavelets Toolbox', in *Engineering Education and Research Using MATLAB*. IntechOpen, pp. 25–54.

Wagenmakers, E.-J., Farrell, S. and Ratcliff, R. (2004) 'Estimation and interpretation of 1/fα noise in human cognition', *Psychonomic Bulletin & Review*. Springer-Verlag, 11(4), pp. 579–615. doi: 10.3758/BF03196615.

Walker, K. T. *et al.* (2008) 'Methods for determining infrasound phase velocity direction with an array of line sensors', *The Journal of the Acoustical Society of America*. Acoustical Society of America, 124(4), pp. 2090–2099. doi: 10.1121/1.2968675.

Walker, K. T. and Hedlin, M. A. H. (2010) 'A Review of Wind-Noise Reduction Methodologies', in *Infrasound Monitoring for Atmospheric Studies*. Dordrecht: Springer Netherlands, pp. 141–182. doi: 10.1007/978-1-4020-9508-5_5.

Wang, Y., Liu, Z. and Dong, B. (2016) 'Heart rate monitoring from wrist-type PPG based on singular spectrum analysis with motion decision', in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, pp. 3511–3514. doi: 10.1109/EMBC.2016.7591485.

Weiser, M. (1991) 'The Computer for the 21 st Century', *Scientific American*, 265(3), pp. 94–105. doi: 10.2307/24938718.

De Wolf, D. A. (1983) 'A random-motion model of fluctuations in a nearly transparent medium', *Radio Science*, 18(2), pp. 138–142. doi: 10.1029/RS018i002p00138.

Wu, C. L., Chau, K. W. and Li, Y. S. (2009) 'Methods to improve neural network performance in daily flows prediction', *Journal of Hydrology*, 372(1–4), pp. 80–93.

Wyngaard, J. C., Siegel, A. and Wilczak, J. M. (1994) 'On the response of a turbulent-pressure probe and the measurement of pressure transport', *Boundary-Layer Meteorology*. Kluwer Academic Publishers, 69(4), pp. 379–396. doi: 10.1007/BF00718126.

Xu, X., Zhao, M. and Lin, J. (2017) 'Detecting weak position fluctuations from encoder signal using singular spectrum analysis', *ISA Transactions*. Elsevier Ltd, 71, pp. 440–447. doi: 10.1016/j.isatra.2017.09.001.

Yang, B. *et al.* (2016) 'Singular Spectrum Analysis Window Length Selection in Processing Capacitive Captured Biopotential Signals', *IEEE Sensors Journal*, 16(19), pp. 7183–7193. doi: 10.1109/JSEN.2016.2594189.

Zabalza, J. *et al.* (2014) 'Singular Spectrum Analysis for Effective Feature

Extraction in Hyperspectral Imaging', *IEEE Geoscience and Remote Sensing Letters*, 11(11), pp. 1886–1890. doi: 10.1109/LGRS.2014.2312754.

Zakis, J. A. (2011) 'Wind noise at microphones within and across hearing aids at wind speeds below and above microphone saturation', *The Journal of the Acoustical Society of America*, 129(6), pp. 3897–3907. doi: 10.1121/1.3578453.

Zakis, J. A. and Tan, C. M. (2014) 'Robust wind noise detection', in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, pp. 3655–3659. doi: 10.1109/ICASSP.2014.6854283.

Zeng, T., Ma, J. and Dong, M. (2014) 'Environmental noise elimination of heart sound based on singular spectrum analysis', in *2014 Cairo International Biomedical Engineering Conference (CIBEC)*. IEEE, pp. 158–161. doi: 10.1109/CIBEC.2014.7020943.

Zhuang, X. *et al.* (2010) 'Real-world acoustic event detection', *Pattern Recognition Letters*, 31(12), pp. 1543–1551. doi: 10.1016/J.PATREC.2010.02.005.

Zumberge, M. A. *et al.* (2003) 'An optical fiber infrasound sensor: A new lower limit on atmospheric pressure noise between 1 and 10 Hz', *The Journal of the Acoustical Society of America*, 113(5), pp. 2474–2479. doi: 10.1121/1.1566978.