A hierarchical Bayesian-based model for hazard analysis of climate effect on failures of railway turnout components

Serdar Dindar[b], Sakdirat Kaewunruen[a,*], Min An[c]

[a]*Department of Civil Engineering, Izmir Katip Celebi University, Izmir, 35620 ,Turkey*
[b]*School of Civil Engineering, University of Birmingham, Birmingham, B15 2TT, UK*
[c]*School of Built Environment, University of Salford, Manchester, M5 4WT, UK*

Abstract

There has been a considerable increase in derailment investigations, in particular at railway turnouts (RTs), as the majority of derailments lead to lengthy disruptions to the appropriate rail operation and catastrophic consequences, being potentially severely hazardous to human safety and health, as well as rail equipment. This paper investigates the impact of climates with different features across the US on the derailments to light up a scientific way for understanding importance of climatic impact. To achieve this, ocial derailment reports over the last five years are examined in detail. By means of geographic segmentation associated with spatial analysis, different exposure levels of various regions have been identified and implemented into a Bayesian hierarchical model using samples by the M-H algorithm. As a result, the paper reaches interesting scientific findings of climate behaviour on turnout-related component failures resulting in derailments. The findings show extreme climate patterns impact considerably the component failures of rail turnouts. Therefore, it is indicated that turnout-related failure estimates on a large-scale region with extreme cold and hot zones could be investigated when the suggested methodology of this paper is considered.

Keywords: Climate effect, Railway operation, Derailment, Bayesian network

*Corresponding author
Email address: s.kaewunruen@bham.ac.uk (Sakdirat Kaewunruen)

Nomenclature

As scientific texts use different notations for constant values, random variables, dependencies with respect to distributions, Table 1 provides the convention applied in this research paper.

Table 1: The list of symbols and notations used in the methodology of this paper.

Notation	Description
y	The number of observed derailment
j	Climate region
e	Exposure
$\widetilde{y_j}$	The number of expected derailment for region j with exposure e_j in a future sample.
ρ_j	The volume of rail traffic in j^{th} region
ϕ_j	The number of turnout in j^{th} region
λ	The unknown hyper-parameter vector
θ	Random sample by using the "Metropolis within Gibbs"
$g(\lambda \mid y_j, e_j)$	The posterior density of λ
α_1	A parameter of gamma density at the first level
μ	A parameter of gamma density at the first level
α_2	A parameter of an inverse gamma density at the second level
a	The shape parameter of the inverse gamma function
b	The scale parameter of the inverse gamma function
κ	A proportionality constant

1. Introduction

Train derailments result in not only operational shutdown and financial losses but also fatalities and injuries affecting passengers, the workforce and members of the public. Turnout-related derailments are revealed to account for nearly one-third of all major derailment cases in the US, and received a great deal of

attention by UK line operators as a problem to be remedied as almost half of all UK maintenance costs are made to ensure a smooth rail operation [1].

Mechanical failure of various turnout components has been observed to be one of the common primary or contributory factors to train derailments[2, 3, 4]. Interest in the investigation of such failures in railway engineering has recently been gained over the last five years. A real train flow model that excludes turnout related failures is produced to show railway vulnerability under both single and multiple component failures[5]. A nonlinear 3D finite element (FE) model is developed, considering the nonlinearities of materials, in order to evaluate the interaction and behaviour of turnout components [6].

Safety-based maintenance has been discussed to optimise geometry restoration of railway turnout systems [7]. It is pointed out that there is a striking relation between weather patterns and the number of turnout component failures, leading to derailments [8]. A Bayesian network-based failure prediction, which does not give rise to any derailment prediction, is modelled to evaluate the effect of weather on component failures at railway turnouts (RTs) [9]. The study is not applicable to large-scale investigations as climate pattern often varies. Moreover, the methodology of this study does not respond to the consequences of the failures. This study suggests the industry and ocial rail agencies as to how regions can be segmented according to climate patterns, and failures, causing derailments, can be predicted throughout entire rail network within a given time. Multi-level modelling with fixed variables and random coefficients was introduced [10]. Exchangeability of hyperparameters on multi-level hierarchical Bayesian analysis was introduced and improved [11, 12]. Accident-related safety estimations on road sections by the Empirical Bayes method, using Rukhin‘s approaches were analysed [13]. The Gibbs sampler, combined with a Metropolis-Hastings (GSMH) algorithm, was introduced [14]. GSMH has often been used to test the reliability and prediction of various engineering systems. In 2016, a Bayesian network model that considers light signals, speed limit along with various rail infrastructures, was developed to determine safety of a travelling rolling stock in particular rail networks [15]. On the other hand, the railway accident

analysis has been underlined that the preciseness of accident investigation and feedback mechanisms between actors contribute is vital to form a positive safety culture [16]. A bayesian method is also used to evaluate several existing predictive methods dependent on the type of the decision-making problems regarding train accidents [17]. Various environmental effects on intersection-related train accidents were analysed through a hierarchical Bayesian model.[18].

In this study, different exposure levels gathered throughout the process of GIS (geographical information system), has been integrated into a multi-level Bayesian model, and then all possible factors, inuencing the number of derailments in each given region, has been levelled out through the exchangeability model. The significant differences between this study and its predecessors are firstly that a large-scale investigation, including the entire US rail network, has been performed. Secondly, the rail component failures whose consequences result in a turnout infrastructure-related derailment, are considered. Thirdly, as a unique outcome, the study has developed a hypothesis underlying that any safety-risk analysis dealing with derailments caused by component failures on a large scale could be satisfying providing that regional segmentation on the basis of climate patterns is conducted. Thus, the study comes up with a novel methodological contribution by putting forward a bayesian based hierarchical model for probabilistic accident assessment at railway engineering.

2. Description of Failures and Risk

Risk is identified in British Standards No. 4778, 1979 as the combined effect of the magnitude of the event, and the probability of the occurrence of an undesirable event. Consequently, risk (R) might be attributed to a measure of a combined effect of the severity (S), and probability (P) of the hazard, which is the subject of this study. The relationship of both, therefore, might mathematically present as follows:

$$R = S \times P$$

Table 2: Reported failures of Frogs, Switches and Track Appliances at RTs.

n	FRA Code	Description of failures
1	T301	Derail, defective
2	T302	Expansion joint failed or malfunctioned
3	T303	Guard rail loose/broken or dislocated
4	T304	Railroad crossing frog, worn or broken
5	T307	Spring/power switch mechanism malfunction
6	T308	Stock rail worn, broken or disconnected
7	T309	Switch (hand operated) stand mechanism broken, loose, or worn
8	T310	Switch connecting or operating rod is broken or defective
9	T311	Switch damaged or out of adjustment
10	T312	Switch lug./crank broken
11	T313	Switch out of adjustment because of insufficient rail anchoring
12	T314	Switch point worn or broken
13	T315	Switch rod worn, bent, broken, or disconnected
14	T316	Turnout frog (rigid) worn, or broken
15	T317	Turnout frog (self-guarded), worn or broken
16	T318	Turnout frog (spring) worn or broken
17	T319	Switch point gapped (between switch point and stock rail)
18	T399	Other frog, switch and track appliance defect

The occurrence of derailment at RTs might lead to one of, or a combination of the following consequences: financial loss, loss of time, damage to functioning railway components, personal injury and even loss of life. This study is limited to the probability (P) of the hazard, as derailments associated with component failures on rail turnouts have similar consequences, namely casualty and financial, and the main focus of the study reveals an exact relationship between derailments and climate. The database of the study utilised from the FRA[1] accident reports is shortened through the following ocial accident codes, shown in Table 2. The FRA codes are selected considering as to whether to associate

[1]The Federal Railroad Administration, which is the US ocial rail agency.

with rail turnouts. In other words, the failure description in this study is only related to various component failures of rail turnout. The FRA discretises RT-related component failures in 18 types of accident, each of which responds to different failures at RTs, as seen in Table 2.

This study utilises from accident reports published by the FRA when any accident exceeds the monetary threshold, results in any causality or both. Reportable damage consists of labour costs and all other costs to replace or repair dam- aged tracks, track structures, on-track equipment, trackbeds or signals. The cost of clearing a wreck, the environmental clean-up costs, or damaged lading etc. are excluded in the calculation of total damages presented in FRA accident re- ports. In addition to reportable damages, the casualties are recorded to chart the magnitude and nature of the injury and death across the USA. The reports also present concrete information on the location of an accident and the environ- mental temperature when a derailment takes place.

Temperature and precipitation have been proven to im- pact considerably on RT component failures [19, 9]. In order to determine the gathering of values of both, the locations of accidents are used. NOAA[2] provides temperature and precipitation values by way of daily summary observations. The gathered data by NOAA has been checked through the suggested map (Figure 1). It is seen that both values comply with each other. Therefore, it is decided to use the map as a climate zone map. The numbers 1 to 7 in the figure rep- resent from the hottest to coldest temperature-based climate regions, while A, B and C are used to describe various humidity levels of the climate.

3. Data selection

The United States of America consists of 50 states, and the FRA collects derailment-related data from various local railway operators who run a service in

[2]The official US climate authority providing a source of timely and authoritative scientific data and information about climate.

these states. Each state has different variables including precipitation, temperature and trac density, and an intersectional variable, track class. This research disregards two states, namely Hawaii and Alaska, due to the non-existence of railway lines in Hawaii and extreme cold weather patterns, albeit a low volume of trac, in Alaska, which would give rise to divergence from the expected estimate values.

As part of the FRA's jurisdiction, all track is categorized into six classes, which indicates the quality of track, and are segregated by maximum speed limits.

This research will concentrate on derailment estimates on entire networks state by state. It is assumed that the condition of the turnouts, and their maintenance quality, is homogenously distributed through the states. On the other hand, exposure to derailments in the states is said to be relevant both to the length of the railway network and the density of trac (rail ton-miles per track mile per year[3]). Therefore, the selected data can be confirmed to offer information on how much turnouts on the entire network are exposed.

3.1. Region selection

As aforementioned in Sections 1 and 2, it is well-known that weather patterns have a considerable impact on derailment cases on RTs. As a result of this, climate regions composed by some or entire coverage of the states are selected considering the annual amount of precipitation and the yearly average temperature.

As the climate is inuenced by a number of factors including proximity to the ocean, altitude attitude, latitude, topography etc., the variability in climate pattern of the US is plenty. Therefore, climate classification can be performed to provide significant insight into the potential vulnerability of RT components to various weather patterns into the analysis. Regionalisation of the United

[3]The product of annual total weight (including the weight of locomotives and loaded/unloaded wagons) and the distance moved by a rolling stock.

States is accomplished using hierarchical cluster analysis on precipitation and temperature data, which appears to yield a set of candidate clustering levels

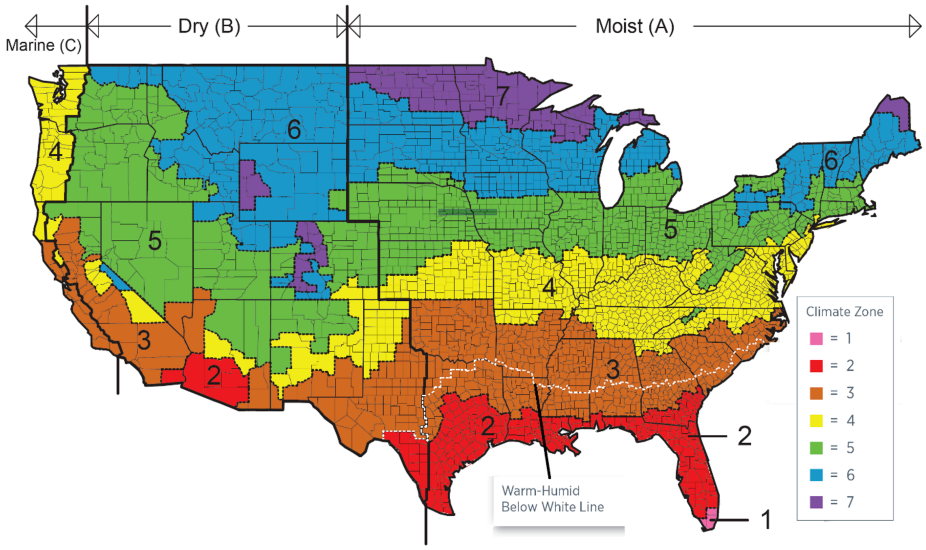

Figure 1: Seven US climate regions

On the other hand, recent research allows the identification of the appropriate climate designation for the US counties, describing the climate zone designations used by the US Department of Energy Building America Program [20].

Figure 1 shows the regional designation based on average temperatures (TBCZs)[4], and precipitation (PBCZs)[5]. The map is separated from left to right to point out the degree of precipitation (A to C), while climate zones (1 to 7) are assigned to express, zone by zone, the degree of the monthly average temperature. To be more precise, climate zone 4, lightened up with yellow on the map, might be addressed to areas where moderate temperature along with either a marine, dry or moisture regime is present.

To illustrate different climate patterns in the US, seven states are decided

[4]Temperature based climate zones.
[5]Precipitation based climate zones.

upon. These states, and their various climate-related patterns are shown in Table 3. The states altogether are assumed to almost correspond to the average regime of the US climate, considering the size of each of the seven climate zones and the three moisture regimes in the US. Zone 1 is not included as being covered as it is a negligible, small area.

Therefore, it is clearly seen that the research needs to deal with analysing derailments zone by zone rather than state by state, due to the fact that the states might be composed of different climate zones.

Table 3: Reported failures of Frogs, Switches and Track Appliances at RTs.

The Name of State	Climate Zones	Moisture Regime	Average Annual Temperature	Average Annual Precipitation
Illinois (IL)	5&6	A	10.8°C	991 mm
Kansas (KA)	4&5	A	13.0°C	992 mm
Nebraska (NE)	5	A	10.4°C	768 mm
North Dakota (ND)	6&7	A	6.0°C	453 mm
Oregon (OR)	4&5	B&C	11.7°C	1006 mm
Texas (TX)	2, 3&4	A& B	17.7°C	623 mm
Utah (UT)	5&6	B	12.7°C	472 mm

4. Anatomy of Turnout Use in the US Rail Network

Even though there is still vigorous ongoing research interest regarding rail turnout-related investigation to ensure a proper rail operation, this kind of study seems to lag considerably behind the other kinds of transportation study. This might be considered to stem fundamentally from a lack of previous research and comprehensive data acquisition. The study also aims to offer plenty of data for prospective researchers in the section.

4.1. Analytics

It is observed that the US passenger and freight users of rail services do not scatter as humongous as the study could assume. On the other hand, the number

of turnouts and the volume of rail trac could not also be assumed in accordance with the size of the climate regions. Therefore, ArcGIS, a tool based on the geographic information system (GIS) for working with layers offering various geographic information, is used. The layers can be expressed to be reliable as they form a comprehensive database of the nation's railway system and are provided by official authorities[6]

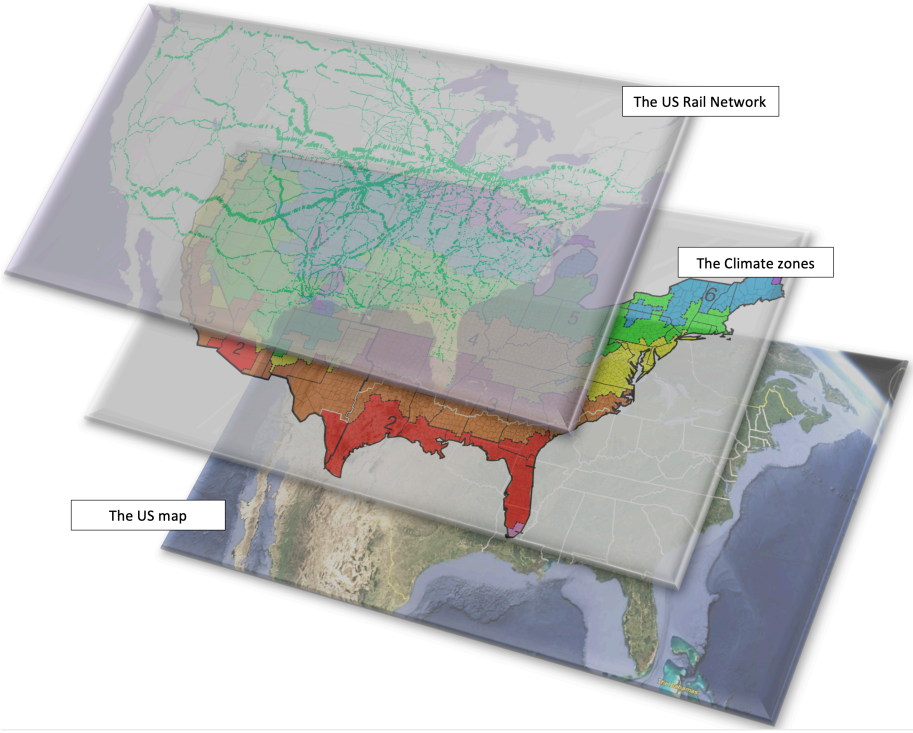

Figure 2: 2 Stacked layers for spatial analysis on ArcGIS

As a foundational concept of spatial analysis, layers that contain different kinds of data are used to find out regional exposures to train derailments at RTs. The bottom layer in Figure 2 includes information as to each county within the geographic coordinate systems of the chosen states. The middle layer (the climate zones) provides regional borderlines of each climate pattern, while the

[6]https://www.arcgis.com/home/item.html?id=088a858aa479444fae9d3bade2d457e5.

upmost layer (the US rail network[7]) trac and the number of rail turnouts where each county is located on the US map. ArcGIS was only used to determine turnouts and derailment locations within the regions. It was first identified that the turnout counts and available traffic on each identified turnouts (around 80,000). Then, we started to determine all turnouts within the segmented region to able to extract proper data.

The number of derailments appears to be determined by some metric of rail trac, such as car-miles, train-miles or million gross tons (MGT) [21, 22, 23]. Therefore, it is underlined herein that the research expresses trac density in MGT.

4.2. Identification of Rail-Turnout Characteristics by the States and the Climate Regions

In order to find out a mathematical explanation as to whether climate has an impact on turnout component failures causing derailments, chosen regions and state-based numeric values forming exposures[8], are illustrated in this section. The calculation of exposure, by using these numeric values, is shown in Section 6.

It has been stressed that the research is based not on exposure by states, but on exposure by climate to first find out the impact of climate on derailments[9] and then true derailment estimates considering this impact. Therefore, the distribution of the indicators by climate zones and regions, see 3.3, is shown in Figure 3. It is identified that there are just under 80,000 turnouts in use through the US rail network.

On the other hand, Figure 3 represents the first investigation in literature, which shows turnout counts and total track density over turnouts. Thus, the

[7] https://www.arcgis.com/home/item.html?id=96ec03e4fc8546bd8a864e39a2c3fc41

[8] being in a situation which has some contributory risk of involvement in turnout-related rail derailments. An exposure composes of turnout counts (ρ) and rail trac (ϕ).

[9] The research is limited to turnout component failures causing derailments. The investigation of climate impact on other kind of failures is left for the future studies.

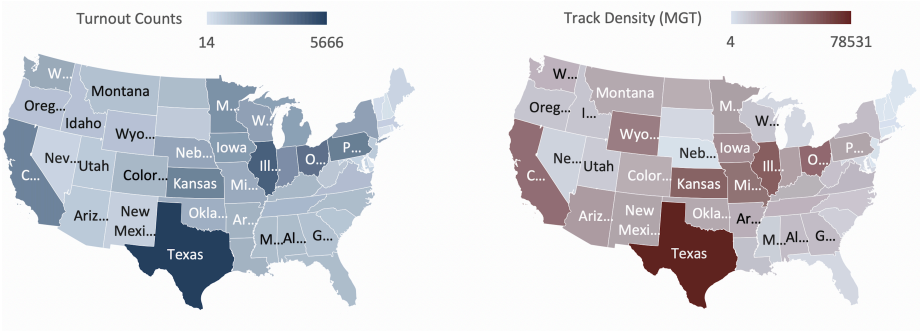

Figure 3: The number of usable turnouts and the volume of trac over them across US rail network

distribution of these two exposure indicators is illustrated state by state throughout the US. It is now quite apparent that Texas might be expected to be where the largest number of derailments are seen as the state has the largest number of turnouts and track density. Texas is followed by Illinois, California, Ohio and Kansas, respectively.

As a result of real observations, Figure 4 illustrates the proportions of turnout counts and rail traffic throughout the climate regions. The yearly average volume of trac over these turnouts is observed to be around 700,000 MGT.

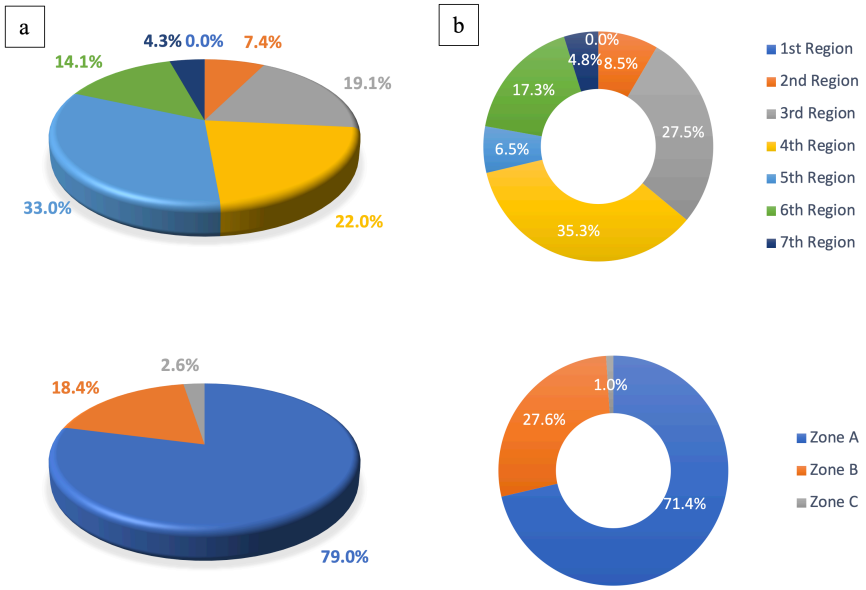

Figure 4: Proportions of turnout counts (ρ) and proportions of rail traffic (ϕ) across the 7 regions and 3 zones: (a) Turnout counts (ρ) (b) Rail Traffic (ϕ)

5. Model Establishment

As discussed in previous sections, it is necessary to conduct an analysis on failure rates and severity to reveal the risk of derailment at RTs. Therefore, a structure, capable of estimating the rates of the derailment accidents within the zones of a particular climate region, is modelled mainly for the derailment cases. Estimates for the probability of derailments within each zone need to be dealt with separately through the same approaches embedded in mathematical formulae, albeit different statistical inputs as the varied statistical realisation of the same sample space is constructed and many different independent random outcomes present. Therefore, modelling consists of nine precipitation and temperature elements, each of which handles its unique count data (y_j).

5.1. Hierarchical Modelling Through Unique Exposure Levels

This study is established on a Bayesian-hierarchical model to identify environmental factors on derailment cases, particularly at RTs. The model gives the simultaneous mathematical estimation of the true failure rates from each climate region and each US state. As all regions often comprise of different states, causing unique volumes of rail trac, a two-stage Bayesian hierarchical model[10] , a mixture of gamma distributions with different hyper-parameters and parameters, is used to obtain more accurate estimates. The choice of gamma function for methedology allows predicting how many derailment events take place in the given period of time. Additionally, it has two different parameterization sets (scale and shape), both of which can be used for the rate of events and the number of derailments.

The proposed model also illustrates exchangeability of prior knowledge by which the true component failure rates of RTs are assigned on a multi-layer Bayesian structure. The posterior distribution of the hyperparameters is composed of a simulation from nine unique gamma distributions corresponding to the regions.

The study relies on derailment counts over a specific period of time. Thus, data distribution, which is statistically a function illustrating all the possible values of such given derailment data, can be obtained as follows:

$$y \sim f(y \mid \theta) \tag{1}$$

where y denotes derailment counts. The observations of derailment cases across the US are given distributions conditional on a parameter, which is θ in Eq.1. On the other hand, the parameter is, in turn, assumed to be of distributions conditional on other parameters, called hyperparameters, as shown in the following

[10]It is calculated through the LearnBayes Pack in Software R.

Eq. 2:

$$\theta \sim g_1(\theta \mid \lambda) \tag{2}$$

where λ is generic unknown hypermeters and relates the number of derailment. On the other hand, Eq. 2 might be pronounced, as the prior vector θ follows the dependent function g_1. Hierarchical establishment, therefore, could be conducted by virtue of the distribution of λ.

$$\lambda \sim g_2(\lambda) \tag{3}$$

As indicated before, λ changes through nine regions since derailment rates within each region is different.

5.2. Hierarchical Prior Choices

Even though six temperature and three precipitation-based climate zones (TBCZs & PBCZs) were pronounced for segmentation of climate characteristics prevailing in the US, the pink-coloured region, see Figures 1 and 2, will not be included due to several reasons underlined through Section 4. Therefore, it is considered that nine climate regions could give the response of the most reliable quantitative to the search objectives. Although such limitation will not impact overall aim of the study, the investigation of these extremely hot regions is left for future studies.

Let J be the generic symbol representing the chosen regions, and each J thus denotes a particular climate region. In order to show the hierarchical Bayesian model of derailment causing component failures for railway turnouts, a directed acyclic graph (DAG)[11] is illustrated in Figure 5.

As seen, each climate region has unique random variable and derailment observations, and are symbolised as J1 to 6, which are represented in an orange colour and address differences in temperature, while the last three, coloured as

[11] A directed acyclic graph (DAG) is a graph that is directed and without cycles connecting the other edges.

purple in the figure, deal with precipitation across the US. Climate zones 2 to 7 in Figure 1 are named as 1 to 6 to ensure the elimination of ambiguity in the research.

All regions have been processed through three successive model layers. The data model, which is the first layer, presents a number of derailments within the climate regions. The process model corresponds to known parameters of the distribution, which is discussed in 5.3. The last model, the parameter model, regards a probabilistic distribution on hyperparameters of the known parameters. The arrows in the Figure 5 illustrates the direction of calculation.

Figure 5: DAG of hierarchical modelling for temperature and precipitation-focused climate zones

5.3. Structural Definition of the Model

The identification of the impact of environmental factors, see Sec. 3.3, on derailment counts is one of the main concerns. Thus, the problem of learning about the rate of derailments is modelled region by region.

Let y_j refer to the number of derailment variables in future in one of the given nine regions j 'and $j = 1, \ldots, J, J = 9$. That is, the counts of events within

a set unit of time are observed. In such situations, a probability model for the distribution of counts of derailments at RTs is considered to be a Poisson distribution with mean π. Due to the nature of Poisson distribution, a parameter with a high value yields a high frequency of occurrence. However, it has been observed that a large number of U.S states with extreme weather conditions behave in a different way, see Figure 7. As a result, the methodology was established on a hierarchical model rather than only Poison distributions. However, the parameter exposure is still used as below;

$$\pi_j = e_j \cdot \lambda_j \tag{4}$$

where e_j is the exposure (per given period of time); see section 4.2 for introductory information. The mathematical formula of the exposure is shown in Eq.5.

$$e_j = \rho_j \cdot \varphi_j \tag{5}$$

where ρ_j and φ_j denote the total volume of rail traffic which passes every turnout and the number of rail turnouts within the jth region of the assigned climates, respectively. This research only uses a Poisson distribution based probability model with mean π to compare the suggested hierarchical model. The results of Pois(π_j) are shown in Figures 8 and 9.

$$\lambda_j = \sum_{j=1}^{9} \sum_{s=1}^{s=48} y_{js} \tag{6}$$

where y_{js} represents the number of derailment within the j-th region of the assigned climate and the s-th contiguous states with $\mathrm{S} = 1, \ldots, \mathrm{s}, \mathrm{s} = 48$. Therefore, the distribution of possible unobserved derailments is conditional on the observed derailments and is given by

$$f\left(\widetilde{y_j} \mid e_j, y\right) = \int f_P\left(y_j \mid \pi_j\right) g\left(\lambda_j \mid y\right) d\lambda \tag{7}$$

where $g\left(\lambda_j \mid y\right)$ and $f_P\left(y_j \mid \pi_j\right)$ denote posterior density and Poisson sampling density, respectively. Eq. 7 is used to check if the consistency of the observed

derailments within state 's' in one of the given climate regions 'j', with s = 1, ..., s, S = 48, and j = 1, . . . , j, J = 9 is present.

The target, therefore, is set to the estimate of the derailment rate at RTs per unit exposure (e). As expected, the total number of derailments throughout the regions is quite low. As a natural result, maximum likelihood estimate (MLE) cannot satisfy the estimate of $\tilde{\lambda}$ under any circumstances, as the denominator, $y_{i,j}$, of MLE, $\tilde{\lambda}_{i,j} = y_{i,j}/e$, the equation will be pronounced with low discrete values, which gives rise to a poor estimate.

It, thus, is desired to benefit from Bayesian estimates having prior belief about the sizes of the derailment rates for the regions, which leads to opting for a gamma function with (y_j, e_j) density.

$$g\left(\lambda \mid y_j, e_j\right) = \frac{1}{e_j^{\alpha}\Gamma\left(y_j\right)}\lambda^{y_j-1}\exp\left(-e_j\lambda\right) \tag{8}$$

On the other hand, Eq.7 is also modelled through a noninformative Bayesian prior, which is $p(\lambda) \propto \lambda^{-1}$. This is thought to give rise to a broadening discussion along with the intended model. The posterior density of λ might be obtained[12]

$$g_1\left(\lambda \mid y_j, e_j\right) \sim \sum_{j=1}^{6}\left[\lambda^{y_j-1}\exp\left(-e_j\lambda\right)\right] \tag{9}$$

Using Eq. 9, Eq.2 and Eq.3 might be modelled through the following equation ($g_1\left(\lambda \mid \propto_1, \mu\right)$):

$$\begin{aligned} &= \tfrac{1}{\alpha_1\Gamma(\propto_1)}\left(\tfrac{\propto_1}{\mu}\right)^{\alpha_1}\exp\left(-\alpha_1\lambda/\mu\right), \\ &\qquad\qquad \lambda \in [0, +\infty) \end{aligned} \tag{10}$$

where $g_1\left(\lambda \mid \propto_1, \mu\right)$ represents a gamma function used to generate samples of λ at the first level of the hierarchical structure. α and μ (parameters) are assumed

[12]The equation 9 is designed for TBCZs as the index of the summation starts at 1 and goes to 6. The last value of the index is replaced with 3 for PBCZs. The same replacement is performed for Eq. 13 too.

to be a priori independent and follow inverse-gamma function [12, 24, 25]:

$$\begin{aligned} \alpha_1 &\sim \mathcal{JG}\left(\propto_2\right) \\ \mu &\sim \mathcal{JG}(a, b) \end{aligned} \tag{11}$$

As each region has unique posterior distribution considering independent values of π j, the PDF is given by:

$$\pi_j \sim g\left(y_j + \alpha, e_i + \alpha/\pi\right) \tag{12}$$

The marginal posterior density of the log hyper-parameters $(\log(\alpha), \log(\mu)\ )$ is as the following equation:

$$\kappa \frac{z}{\Gamma^6(\propto)(\propto + z)^2 \mu} \sum_{j=1}^{6} \left[\frac{\propto^{\propto} \mu^{-\alpha} \Gamma\left(\alpha + \prod_{i=1}^{48} y_{ij}\right)}{\left(\alpha/\mu + \pi_j\right)^{\left(\alpha + \prod_{i=1}^{48} y_{ij}\right)}} \right] \tag{13}$$

where κ is the constant of proportionality, z denotes the median of α.

5.4. Metropolis-Hastings (H-M) Algorithm

It is suggested that the posterior distribution function is sampled using the H-M algorithm [26]. This algorithm is expressed to bring out the Gibbs sampler, which is a Markov chain Monte Carlo (MCMC) algorithm, as a special case[25].

The algorithm is used to approximate the output of Eq. 4 in Sec. 5.1, acquiring a sequence of random walk proposals from the Metropolis-Hastings algorithm itself.

The reason for the choice of Metropolis-Hastings for this research over Gibbs sampling is firstly that it is highly unlikely or practical to obtain the conditional distributions for each of the random variables in the suggested model, even within an environment with the full posterior joint density function (FPJDF) [27]. Secondly, it is also unlikely that the posterior conditionals for each variable have a known form. As a result, samples from these conditionals cannot be drawn in an uncomplicated way [28].

6. Results

6.1. Fundamental Data Analysis Findings of Suggested

To deliver usable and useful information, a further aim is followed by this research, aside from estimating the significant characteristics of regional impacts on derailments, such as absolute numbers in various categories of the suggested framework. The distribution of categorised data and initial results of the suggested workow are preferred to be firstly discussed. The fundamental statistical data obtained through the framework suggested in Figure 6 may, therefore, be presented.

Responses of the selected states to Eq. 5 are illustrated in Figure 6-a. Many states, such as Texas (green) and California (yellow), might be underlined to overpower the distribution of e total, whereas the others seem to be impacted by a low number of turnouts, a low number of rail trac, or a combination of both, due to the mathematical nature of the Eq.5. It is worth noting that the 3D bar chart is not directly used to respond to the research question. Instead, region and state based exposures are illustrated. Figure 6-b illustrates the distribution of e through six regions from Region 2 (red) to Region 7 (purple), see 3.3 for a colour match.

It may be identified that total exposure in Region 2 might be pronounced to be higher than the other five regions in a deterministic way. Therefore, more derailment in Region 2 could be expected. Additionally, Regions 4 and 5, which are coloured yellow and green respectively, can be asserted to have more turnout related derailments than Regions 6 and 7 have together.

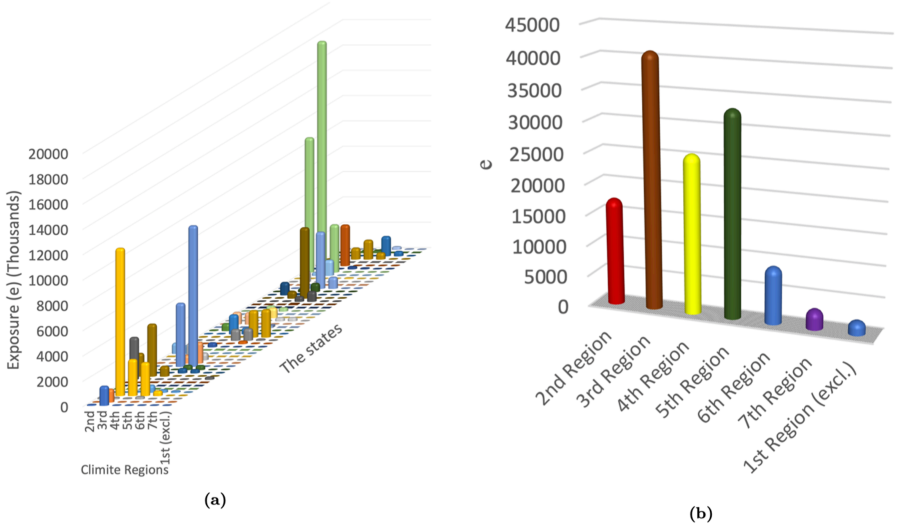

Figure 6: Exposure distribution by the States (a) and Regions (b)

Nevertheless, Figure 7[13] illustrates explicit inconsistency with this claim considering the derailment-based behaviour of states in Region 2. For instance, Florida, shown as FL in Figure 7, induces considerable unexpected risk (see Sec. 2) at turnout-related derailments, although Figure 6-b shows a low rate of exposure[14]. Therefore, this phenomenon might be elucidated to require a better mathematical algorithm, which includes a polled database not by states, but climate, as well as more detailed consideration, such as hierarchical establishment.

[13]To deliver the main point, and enhance visual quality, the figure is plotted excluding a few states with extreme exposure, i.e. Texas, since they have relatively quite high values of y_j/π_j, which results in rounding up many values on a small area at a logarithmic scale. Therefore, it could be said that Poisson distribution alone will not match observations as some states with less exposure experience relatively more accidents or vice versa.

[14]It is worth noting that 1st region is almost composed of FL.

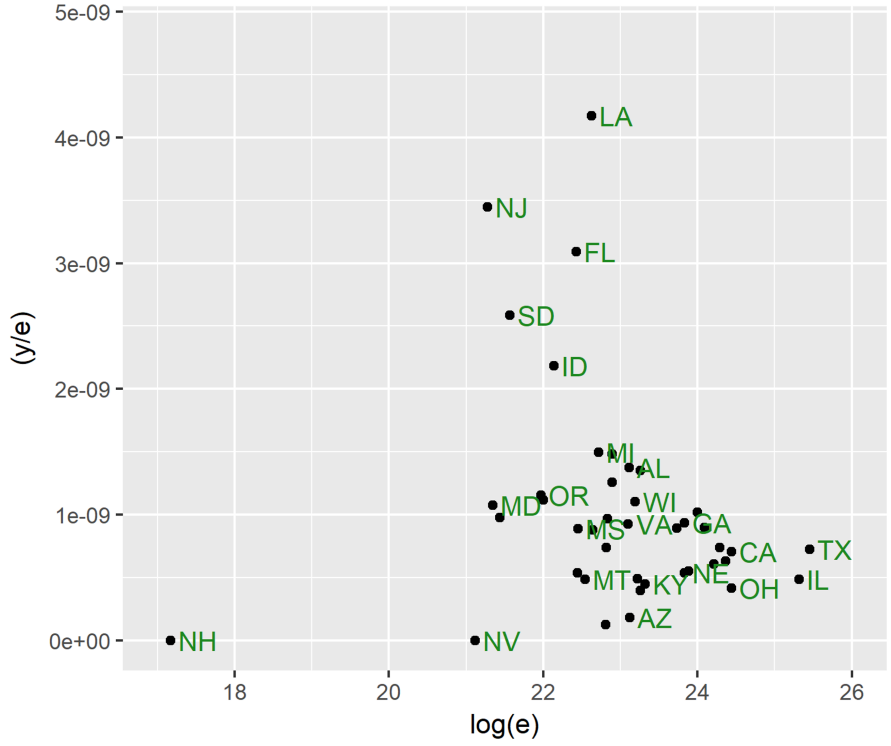

Figure 7: Derailment rates against log exposure for majority of the states

However, it is still intended to find a more robust understanding in a stochastic way, as the output of the deterministic model is fully determined by the overall parameter values, i.e. $^{y_j}/\pi_j$, which responds to a limited scenario. That is, the output of the stochastic model possessing some inherent randomness is needed to solve the degree of the impact of climate conditions on derailments in particular turnout systems, which are specifically chosen for this research due to the fact that they are of many engineering systems.

Figure 8[15] shows the approximation of a posterior predictive distribution (PPD) with an equal derailment parameter through regions, i.e. $\lambda_2 = \lambda_j, \{\mathrm{j} =$

[15]The abscissa and ordinate of the Figure 8 represent the number of derailment (y) and generated sample counts, respectively. The same abscissa and ordinate are also used in the Figure 9.

380 $(3, 7), j \in \mathbb{N}\}$. In other words, the results of Eq. 8 and Eq. 9 are presented region by region in the Figure. The PPDs might be useful to identify to what degree the estimates are appropriate, as the numbers of turnout-related derailment observations in all regions are also illustrated as red dotted vertical lines on the same plots.

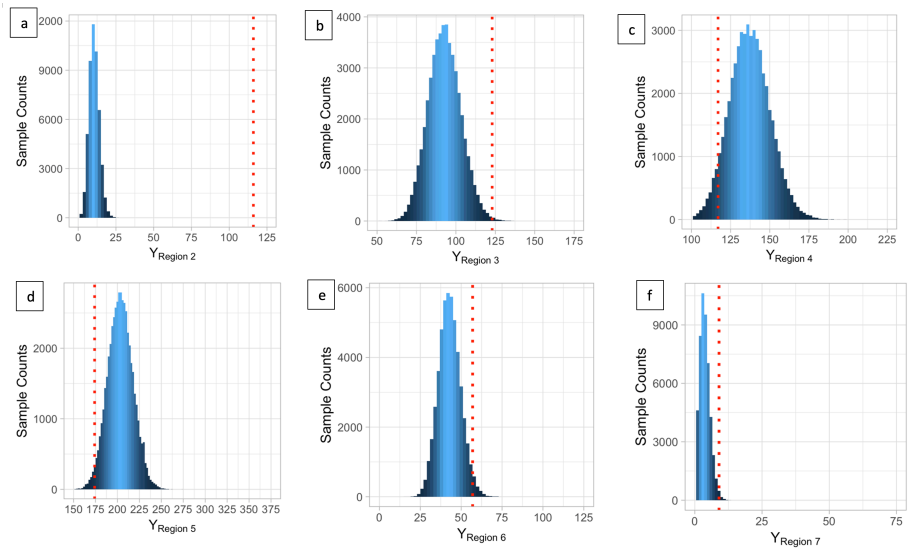

Figure 8: Histograms of TBCZs by 50,000 draws simulated from the posterior predictive distribution of Region 2 (a), Region 3 (b), Region 4 (c), Region 5 (d), Region 6 (e) and Region 7 (f)

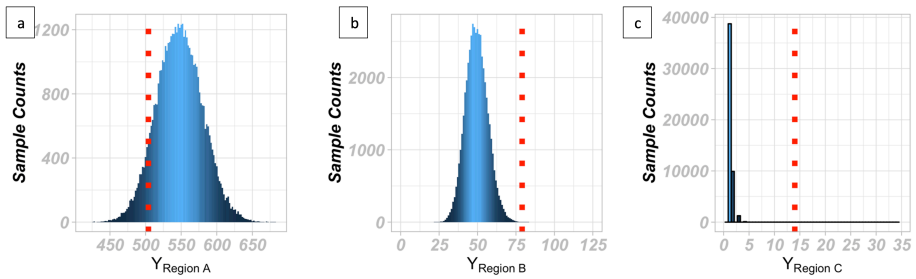

Figure 9: Histograms of PBCZs by 50,000 draws simulated from the posterior predictive distribution of Region A (a), Region B (b) and Region C (c)

385 *6.2. Results of the Stochastic Model: Response of the Climates*

Moderate climate regions, namely 3 to 6, seem to almost respond to the related assumptions and the function. For instance, the observation of Region 4 is placed near the middle of the predictive distributions. However, the majority of the other observations are in the tail portion of the distribution. Furthermore,
390 the estimate for Region 2 is proved to fail rather evidently, which reveals that there is no absolute agreement of these observations with the fitted model. Therefore, regarding the histogram associated with the second region, it might be pronounced that the distribution of the samples needs to be focused on a higher range of observations (x axis) than simulated through the Eq. 7, based
395 on an agreed value of λ. In section 3, the US has been said to be also separated through three regions, considering the characteristics of precipitation that the regions have. The results of these three regions are illustrated in Figure 9.

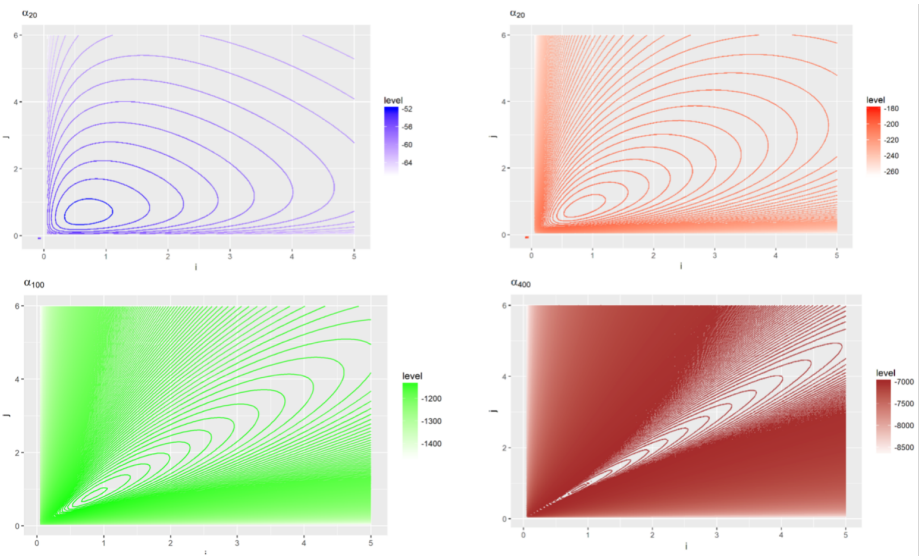

Figure 10: Behaviours of the exchangeable prior on (λ1(i), λ2(j)) against inverse gamma (alpha = 10, beta = 10)[16]

Similar to the harsh climate conditions in Figure 9, the estimate-distribution of the latter does not show the reality. On the other hand, although the number
400 of derailments within Climate A is almost seven times higher than the other

two zones, the observed number is placed on the left tail. Similar to Region 2 and Region 7 in Figure 8, it might, therefore, be concluded that the derailment counts of the three regions cannot be estimated properly and needs to be a hierarchical structure.

It is proposed by Eq. 9 that the most approximated rate parameter π of derailments through the regions could be derived from $g(\lambda \mid \propto_1, \mu)$. In order to achieve this , a random sample of 50,000 is generated, which results in understanding the behaviour of the relationship between shape ($\propto_1$) and the mean (μ) parameters of the gamma function. As this study is designed on a priori hierarchical structure (see 5.2), both the generated sample and therelationship of parameters can only be reached in the event that the behaviours of hyperparameters, namely; ($\propto_2$), a and b, which are nested at the second layer of the hierarchy, are revealed. In this circumstance, as $\propto_2$ is the only available hypermeter in the function, assigning random values on $\propto_2$ rather than the other two hyperparameters, has been considered to require less workload.

On the other hand, (μ), one of the two parameters at the first prior stage, is assumed to follow an inverse gamma function with hyperparameters; namely, a and b equal to 10 [29]. Thus, π_1 and π_2 can be fixed at a certain value, which leads to understanding behaviours of their exchangeable prior while $\propto_1$ changes. Figure 10. illustrates such behaviour when the unknown hyperparameter of an assigned inverse gamma function, $\propto_{1)}$, is at 4, 20, 100 and 500.

In addition to these discussions, it might be stressed that each line of the distributions show an increment in the concentration of $\propto$. Therefore, the more $\propto_2$ values are assigned, the more λ_1 approximates to λ_2, and vice versa, which makes λ_2 equal to infinity. There seems to be a centre around (0.95, 0.95)[17].

[17]There are two conditional samplings as seen in Figure 12. Acceptance rates of log and log µ are observed to be around 0.93 and 0.50, respectively. This highly likely arises from small σ^2 of the proposal distributions. Iterations at the beginning of the MCMC run were not thrown away, as the limited number of samples (around 10) has been identified to be out of the sample average.

Prior density of the form is suggested [12] to be $g(\propto_2) = z/(\propto_2 + z)^2$, providing that $\propto_2$ must be bigger than 0[18]. In this case, Z is assigned as 0.98[19]. In addition to these discussions, it might be stressed that each line of the distributions show an increment in the concentration of $\propto$. Therefore, the more $\propto_2$ values are assigned, the more λ_1 approximates to λ_2, and vice versa, which makes $\propto_2$ equal to infinity.

These parameters and hyperparameters are also assigned with a sample of 50,000 randomly simulated from the joint posterior. Log(μ) and log($\propto$) are illustrated in accordance with the level of the posterior density in Figure 11.

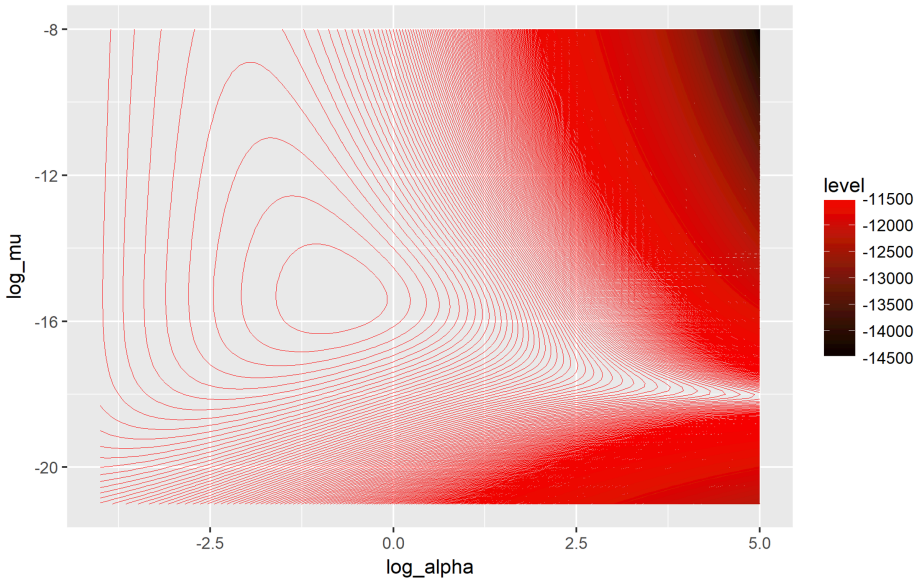

Figure 11: Behaviours of log(μ) and log($\propto$) against US Climate Patterns

The samples of log(μ) and log($\propto$) are obtained through Eq.12 with a value of z, which has previously been found to be 0.98. To find out the centre of the lines, a plot with a long range of log(μ) and log($\propto$) is performed, and then

[18]This is not because $g(\propto)$ but because of the behaviour of $\lambda 1$ and $\lambda 2$, as those would be undefined.

[19]The centre of alpha contours in Figure 10 presents z value.

the plot between log(μ)[3, 5] and log($\propto$)[21, 8] is selected to show the core of distribution. The Figure 11 also illustrates the modal values through curving-alike red contour lines, each of which represents an interval of the logarithmic scale.

For instance, the 2 nd , 4 th and 6 th lines from the core of distribution represent 0.1%, 1% and 10% of the modal value, respectively to assign a sample into the distribution, the function GIBBS[20], which is available in the LearnBayes package in R , is used. The function enables researchers to define an arbitrary real-valued posterior density into a Metropolis-Hastings merged with a Gibbs algorithm. Assignment of "start", one of the arguments under the GIBBS, might be seen as the starting value of the parameter vector in Figure 12. As noticed at first glance, the practice called burn-in is not used in this research since the nearby-real starting point is found to be earlier, which makes the technique an unnecessary part of this execution.Log(μ)[5] and log($\propto$)[22] is selected as starting points, which are shown as the first pair on the left side of the Figure 12. As clearly seen, the range of log(μ) and log($\propto$) values starts to fluctuate in a way that ranges roughly from 1.5 to -1.5 for log($\propto$) and from -15.5 to -17.5 for log(μ). In other words, the trace of the parameter vector seems to be well-distributed throughout the 50,000 different iterations of the chain.

The coordinates of generated samples, composing the parameter vector and assigned by the MH algorithm, throughout log(μ) and log($\propto$) are plotted out in Figure 13-a. The dog-tooth pattern indicates that the chain converged almost immediately, considering the distribution of log values of parameters in Figure 12. Figure 13-b shows the density of these samples, which indicates that most samples are generated within the same value range, randomly near where log($\propto$) equals 1. There might seem to be some samples dispersed out of the dog-tooth. This is because of the nature of the MH algorithm. A set.seed function[21]n R is

[20]Usage: Gibbs (logpost,start,m,scale,...). Logpost, start, m, scale are assigned as Eq.12, start = c (2, -22), 50,000, c (1.00, 0.25), respectively. For further details see; https://www.rdocumentation.org/packages/LearnBayes/versions/2.15/topics/gibbs.

[21]The seed of R's random number generator, which is used to create random objects that

assigned as 737327(the letter of the name of the first author on dial pad). It has been observed that different set.seed values do not give a different sample population, which presents a sample size of 50,000 is enough for such a study. In other words, the undesired samples, however, might be deduced to not impact on results as the longer tail of its density function, see Figure 13-b, places it at a relatively low level due to the large size of the generated sample.

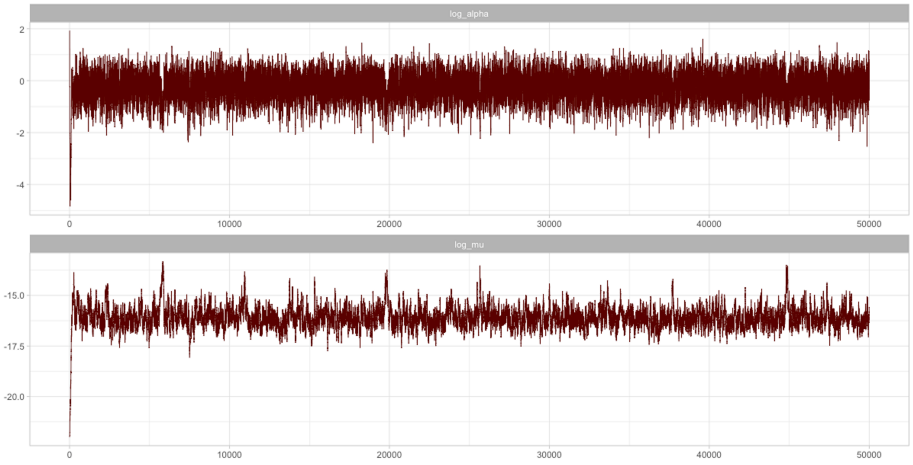

Figure 12: A representative trace of $\log(\propto)$ and $\log(\mu)$ iterations of 50,000 cycles through M-H algorithm

can be reproduced.

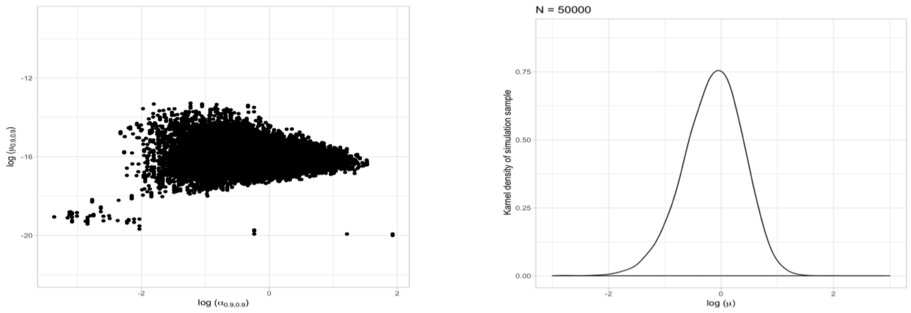

Figure 13: A scatterplot of simulated values based upon the samples in the chain (a) and its density function (b) by Metropolis -Hasting integrated in Gibbs Sampling

The main point of this scientific investigation by real cases is to provide the required response to the question as to whether characteristics of a climate regime manipulate annual derailment counts at railway turnouts. It has been attempted to level out all covariates, such as π, e, λ throughout the chosen regions. After some assumptions, see Sec.4, this theoretically brings out the ultimate outcome of the entire endeavour, shown in Figure 14.

Essentially, Figure 14 illustrates the density functions of nine chosen climate regions as the derivative of the Eq. 6 throughout a continuous distribution of 50,000 samples per each region. The areas that are covered under these curves, produced by the density functions and the x- axis, are equal to 1. On the other hand, the Y axis points out not absolute counts, but relative frequencies of the curves.

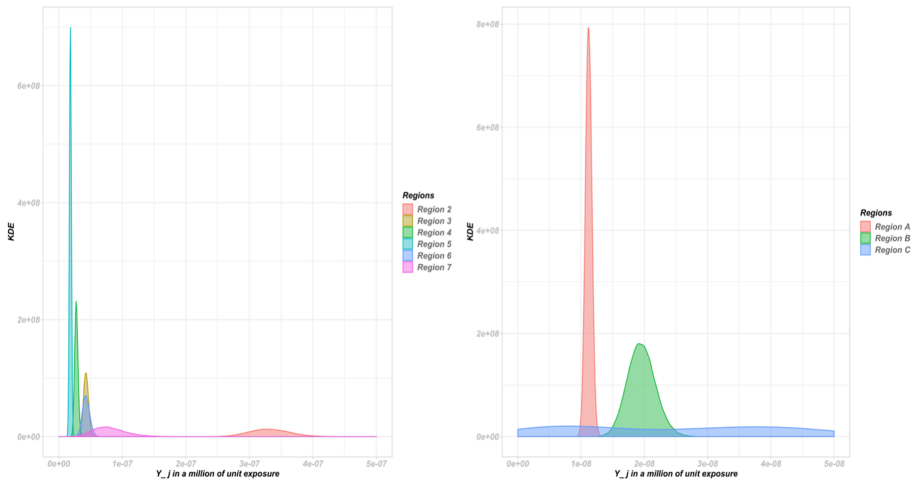

Figure 14: Regional posterior density functions of derailment at railway turnouts in a year for TBCZs (a) and PBCZs (b)

In the Figure 14-a, three distributions, associated with Regions 3 to 6, are observed to behave in quite similar manner to each other. Considering that the first region is left out of the research for several reasons, Regions 3, 4 and 5 have, in fact, more moderate climatic conditions through the year than the others. As a result, this common behaviour could be expected. As for Region 6, covering the large northern US, it is observed that this climate tends to approach Region 7 by diverging the three moderate regions. Even as the climatic characteristic of Region 6 is already known (see 3.3), it is now revealed that a light impact of these characteristics on derailment counts, Y_6, presents. However, this hypothesis could be open to discussion, considering that this study has not dealt with the distribution of operational error. For instance, maintenance, as an operational error, has recently been found to have been involved in 20% of all turnout-related derailments in the UK[30]. The majority of operational errors on rail turnouts in the US is still an unknown phenomenon [31, 32]. This research can be extended by working on operational error. The authors believe that the distribution of such errors might be considered to disperse quite uniformly across the US rail network at the rate of relative areas that are covered by regions. It

should not be forgotten that these assumptions might result in outcomes that are not able to deduce an absolute meaning. Therefore, nothing can be said directly about Region 6. Instead, the general characteristics of a cold region are likely to impact the derailment counts in the region, considering that the region 6 starts to resemble the region 7 in Figure 14-a.

On the other hand, the regions in which extreme weather conditions prevail, namely, Region 2 (red line) and Region 7 (pink line), show explicitly different patterns from the others. The first striking feature is that both regions take a place on the right side in a discrete way. As a certain result of this, it can now be pronounced that the extreme weather conditions[22] impact on component failures for railway turnouts, causing derailment, even though it does not present verbally on ocial reports [23].

The kernel density function (KDF) of Region 7, covering cold regions, is seen to be placed in a long range of Y_7 values. This behaviour of the density function stands for a lack of preciseness in the estimate, as the density function reacts intrinsically to fulfil a real Y7 value, due to insucient knowledge, compared to the others. The study investigates 596 derailments across the US over the five years from 2010 to 2015. It is observed that a range of roughly 70 to 90 derailments is distributed throughout five climate regions, while Region 7 has a population of under 50 derailments. Therefore, the KDF function of the coldest region seems to be dispersed. However, the median value of the KDF is seen to take a place in-between the values of moderate regions and Region 2.

The sample distribution of Region 2 is seen to have the most distinguishing

[22]The phrase is, in general, used to express weather events that are significantly different from the usual or average weather pattern taking place over a period of time. The research refers to seasonal adverse conditions of the US‘s average weather pattern.

[23]Ocial accident reports, in fact, mention environmental reasons, covering codes beginning with “M”, such as M102, Extreme environmental condition - TORNADO. This paper has not attempted to discuss such deterministic explicit reasons, but investigation of climate impacts on component failures for rail turnouts on a stochastic methodology has so far been a milestone for rail researchers to understand the real climate impact beyond observable facts.

features. As an abundance of data provided by the region is enough, a distribution with a higher mode and median than Region 7 is observed. Moreover, the KDF is the highest median of all. In light of these findings, it is mathematically possible to claim that the more extreme environmental conditions a region has, the more component-related failures railway turnouts of this region possess.

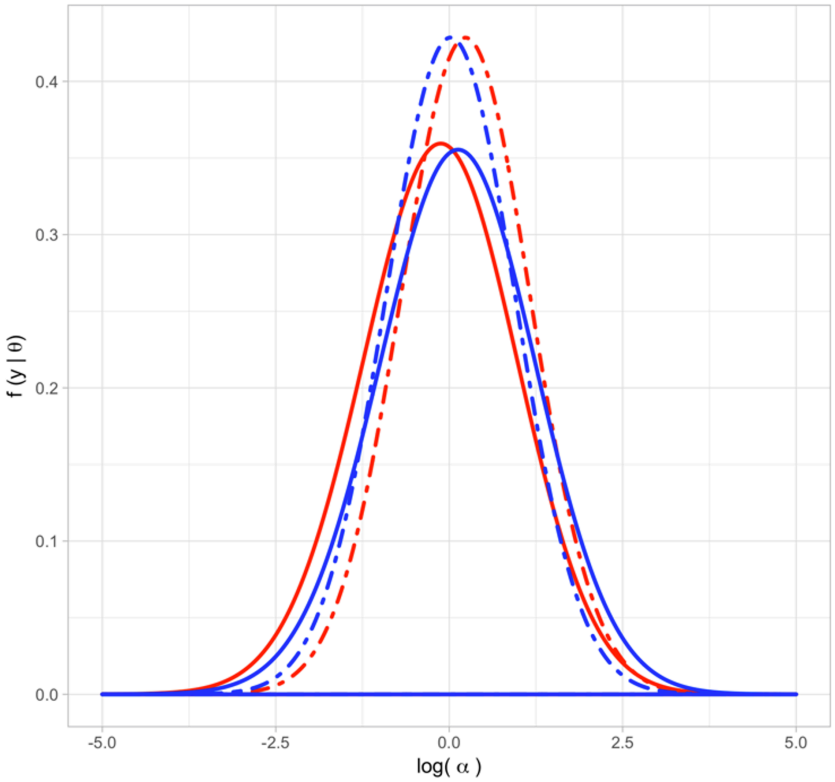

Figure 15: Posterior (solid lines) and Prior (dot-dashed lines) Density Functions for TBCZs

The KDF of precipitation-based climate zones, on the other hand, is shown in Figure 14-b. Similar to the results of the KDF of TBCZs, the more derailment observations, a better estimate can be conducted. Fourteen cases are observed to have occurred across Zone C, which is shown with blue colour. Although the peak of KDF is around 3 x 10^6, it is almost impossible to make a comment

on the precipitation regime of Zone C. This is highly likely to be due to a low number of derailment cases and low exposure compared to those of the other two regions. The green-coloured distribution, Zone B, which is moist, visualises the distribution of derailment y_j probability over a million of (e), which is lower than Zone A, where a dry climate prevails.

6.3. Robustness of the Model for Component Failures at RTs

It is assumed that the derailment rates were sampled from the gamma function, see Eq-9, with two parameters (μ and $\propto$). While a non-informative prior proportional to $1/\mu$ was assigned to μ, $z_0/(\propto + z_0)$ was assigned to $\propto$. Z[24] is assessed as 0.98 for this research. Figure 13 shows prior and posterior results by the chosen value of z (0.98) and the testing value of (98).

These prior (red lines) and posterior (blue lines) functions are shown in Figure 15. The y-axis of the figure illustrates the density of samples while the x-axis of the figure gives an idea on $\log(\propto)$. Dotted lines denote posterior functions, while blue lines are obtained through an increased z value. It is seen that the assumption 'z' does not impact largely upon the prior and posterior functions, even though the z value is increased by 100 times. In other words, the z value does not seem to be largely effective, which shows the robustness of the model.

As seen in Figure 8, Y_i values of climate zones 2 and 7, which are relatively extreme hot and cold respectively, have been shown to reveal how effective is one layer of the derailment estimate. Moreover, the estimates (bar chart) have been observed to not satisfy, considering a real number of derailments in a given time (red dot line) in this Figure 8. As a result of hierarchical modelling and the preciseness of exchangeability of π_j, ultimate estimates are expected to be better than found previously.

[24] Z value is assumed considering the centre of contour lines shown in Figure 10. The model is conducted in a way that further parameters such as Z impact negligibly the result. To test the robustness of the model, Z value is changed largely. It has been understanding that the proposed model immunes considerably the changes at Z values.

In contrast, the estimates for climate zone 2 and 7 (of TBCZs), illustrated in Figure 16-a & Figure 16-b, respectively, yields better estimate. Taking into account that the distribution of the derailment estimate for zone 7 on one-layer structure in Figure 8 did not meet the real count (shown as red dotted line), the distribution in the hierarchical structure seems to be placed in a desired way. On the other hand, the worst estimate has been seen on the second zone.

The estimate of the hierarchical model for PBCZs is illustrated in Figure 17. In order to visualise whether or not the distribution of samples by the suggested hierarchical model better estimates, Figure 17 might be compared to Figure 9. One layer modelling has given the histogram as placed on the far-left side of y_A (the number of real observed cases). However, $\widetilde{y}_A$ is seen to be placed properly at the middle of the distribution in the Figure 16. As for Region B, 79 derailment cases, y_B is observed over the observed period. The distribution through MLE is revealed to miss y_B in Figure 9. The suggested hierarchical model for turnout-related derailment estimates is more precise, placing y_B at around the middle of $\widetilde{y}_B$.

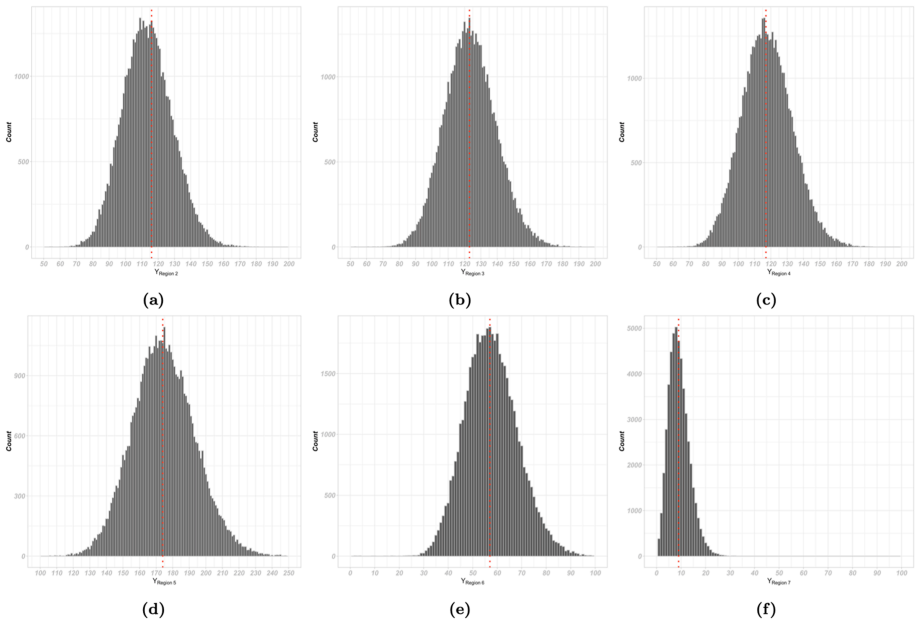

Figure 16: Histograms of TBCZs by the posterior predictive distributions of Region 2 (a), Region 3 (b), Region 4 (c), Region 5 (d), Region 6 (e) and Region 7 (f) from the hierarchical model

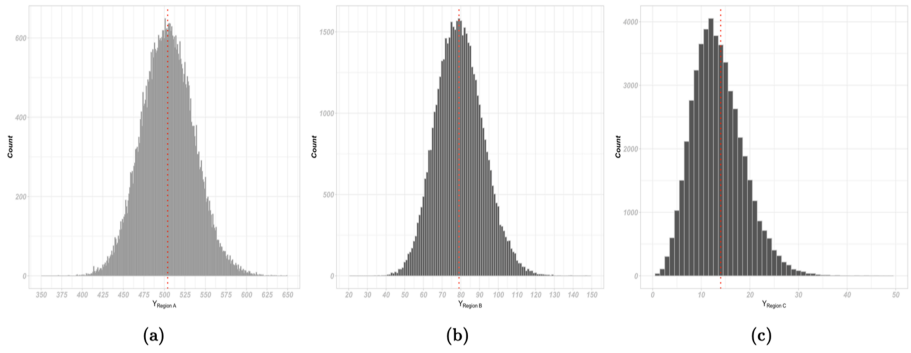

Figure 17: Histograms of PBCZs by the posterior predictive distributions of Region A (a), Region B (b) and Region C (c) from the hierarchical model

7. Discussion

Train safety and risk analysis are dependent on the accurate estimation of the derailment rate as well as understanding its contributory factors. This research is aimed at better identifying both, for turnout components-related derailments, dealing with a large amount of ocial data. In order to do this, a hypothesis regarding environmental factors, i.e. precipitation and temperature, and their impact on derailment rates regarding turnout components is asserted. Therefore, the entire US, except Alaska and Hawaii, is divided into seven different temperature-based regions (climate zones 1 to 7) and three different precipitation-based regions (climate zones A to B).

Unique exposures of the zones are exclusively researched and identified, which are defined as rail trac density (MGT) and the number of turnouts across the zones. A hierarchical Bayesian model is used to argue the hypothesis, as a single layer Bayesian model is suspected to not respond properly due to the environmental factors, which also gives an opportunity to test them against each other. This is accredited to reveal the preciseness of the research. It is observed that 596 derailments, which are due to various turnout component failures, have occurred across the US over the last five years. To be able to interpret the derailments, the unique exposures of the climate zones, which are affected by rail trac over each turnout, along with the number of rail turnouts, are found out, individually dealing with the nine climate zones. The results are firstly exhibited through Poisson distribution, which is, in general, preferred by the rail industry[33], and then a hierarchical Bayesian model is used. Therefore, the results of the two different statistical approaches are compared with the number of reported derailments (real counts) within the climate zones to determine the best fitting estimate, and how both techniques are close to the reality.

It is identified that a hierarchical model yields better results, in particular within a scarce data environment, i.e. low derailment rate or low exposure. The estimate distributions of Zone B and C of PBCZs, relatively within such

an environment as compared to Zone A, are not considered to be a proper estimation without a hierarchical model, as the real observations do not place on the predictive distributions. In contrast, the distributions are updated through the hierarchical model and revealed that the numbers of observed derailments are placed within the updated distributions.

That is, the number of observed derailments for extreme climate zones is indicated to have an agreement of these derailment cases with the fitted hierarchical Bayesian model. On the other hand, the posterior predictive distributions (PPD) of PBCZs and TBCZs examine the appropriateness of Albert‘s exchangeable hierarchical Bayesian model with unique exposures in order to identify true derailment rates for each assigned climate region. As a concrete result, the behaviour of this phenomenon is found to vary across the region. More specifically, the regions of extreme weather conditions almost yield similar rates Y_j. For instance, PPDs, belonging to climate zones 3 to 5 of TBCZs, are observed to place nearby each other on the axis, which shows $\widetilde{y}$ values in million exposure units. In contrast, the other TBCZs, which have been called extreme climate regions in this study, are likely to be elaborated as regions posing much risk at turnout-related derailments. It is observed that these regions possess derailments up to a few times higher compared to regions where moderate climate regimes prevail. As regional exposures, utilised from the total volume of trac and the number of turnouts within the regions, are assigned to the same unit, it could be assumed that there are other contributory factors, including a different maintenance regime, track quality and different approaches in reporting derailment accidents.

However, all derailments are chosen in accordance with the same track class reported by the FRA, which has ocial jurisdiction over track quality with a responsibility gained by the Railroad Safety Act of 1970. Thus, in the chosen turnout-related accidents it might be said that the quality and design of turnouts are similar across the US. Even if this statement can be debatable, the distribution of turnout characteristics would be made homogeneously, considering that the US rail network has a large number of turnouts in operation. As

for track inspection, the FRA as well as the railroads themselves are required to perform this. Although some carriers possess in-house rail detection vehicles, the majority of the Class 1 systems have their own engineering trains, used to routinely monitor a wide range of factors critical to safe operation, creating a profile of various components, alignment and gauge. The frequency of turnout inspection falls to two different categories, namely "main track" or "other than main track". It is observed that turnout derailments scarcely occur on the main track. Therefore, the turnouts associated with such derailments, with which this research deals, must have been inspected at similar intervals.

Therefore, the characteristics of separated regions seem to be the main driver for strikingly different derailment rates. The study shows there is an important relationship between climate zones and derailment rates. As suspected, some weather conditions, including extreme cold and hot, as well as a moist climate, are observed to impact largely on the rates. Turnout component related-derailment estimates of the 2nd and 7th climate regions of TBCZs are identified to appear two to three-fold compared to those of the other regions, respectively.

Moreover, sample distributions of regions, except for 2nd and 7th climate regions, are determined to correspond similarly to the hierarchical model, gathering together. In other words, it is identified that mild climate zones, i.e. the 3rd to 6th climate regions of TBCZs, are not as impactive on turnout component-related derailments as the extreme zones.

It can be noted that the proposed model requires considerably resource consumption. It is observed that a computer with I7-8559U processor, 16GB of ram and macOS Mojave operating system needs around 20 days to compute the equations and plot the results. There is no arguing that the hierarchical model takes more processing time than a single layer Bayesian model. However, a multi-layer structure is found out to result in more precise estimates on a large scale rail network.

8. Conclusion

This study might be considered to start a novel discussion field, which is associated with rail safety research, particularly derailment estimates. A rail accident analysis is presented through a stochastic process, which responds to a large number of environmental conditions for turnout operations. The previous studies are often of deterministic approaches, which are limited to conditions on which the studies were conducted. Therefore, the study has suggested that rail researchers can adopt a stochastic based methodology, which associates with component failures of rail infrastructures. The paper comes up with a hypothesis asserting that climate considerably impacts the rate of component failures, depending on its characteristics. Concrete results, which provide accuracy of the asserted statement, are found, which is highly likely to result in helping researchers to better estimate derailment rates. This is constitutive because pooling derailment data is proved to not be the way to precisely estimate the rates for a large-scale geographic region, where climate zones with different characteristics prevail. The proposed hierarchical model can be integrated into the current train safety and risk analysis, mainly associated with accurate estimation of derailment rate.

The study is conducted on the basis of derailment cases, that occurred in the US rail network, and were recorded by the United States Department of Transportation. As the rail network is in operation in many prevailing climate zones, whose characteristics are divergent from each other, the study has been able to identify the striking impact of the climate. Therefore, the researchers conducting derailment estimate might be suggested that the presented methodology is considered to eliminate this impact for obtaining better results.

The other factors such as train speeds, quality of railway tracks and applied maintenance strategies, variations of rail management techniques might be added to the research. Theoretically, if all factors were found out and were nested layer by layer (as many as they are), the exact or more precise results would be observed in Figures 16 and 17. However, this requires quite high

computational power.

Consequently, the authors believe this study can be applied to large-scale rail networks such as China's and the EU's, in order to estimate more accurately turnout component related-derailments and component failures. As they have a long-term plan to expand the network, derailment rates will likely be deviated from previous experiences. On the other hand, a Road Safety Programme which aims to reduce rail accidents in Europe by half in the next decade was adopted by the European Commission a few years ago. Overall statistics across the EU have been recorded and will be recorded by Eurostat to benchmark the countries, often making a deterministic comparison with each other, and investigating any performance increase or decrease with previous experience. It could be highly effective to implement this study into Eurostat's framework. This would reveal the intended performance of the EU countries, considering changes in climate, and the rail networks.

9. Acknowladgement

The authors would also like to thank British Department for Transport (DfT) for Transport - Technology Research Innovations Grant Scheme, Project No. RCS15/0233; and the BRIDGE Grant (provided by University of Birmingham and the University of Illinois at Urbana Champaign). The first author gratefully acknowledges Turkish Ministry of Education for supporting his PhD at University of Birmingham. The authors are sincerely grateful to the European Commission for the financial sponsorship of the H2020-RISE Project No. 691135 "RISEN: Rail Infrastructure Systems Engineering Network", which enables a global research network that tackles the grand challenge of railway infrastructure resilience and advanced sensing in extreme environments (www.risen2rail.eu). The authors wish to thank the members of RAILTEC group for their help throughout the course of this comprehensive research; in particular Christopher P.L. Barkan provided many useful discussion and ocial rail accident database.

[1] S. Dindar, S. Kaewunruen, Assessment of turnout-related derailments by

various causes, in: International Congress and Exhibition" Sustainable Civil Infrastructures: Innovative Infrastructure Geotechnology", Springer, 2017, pp. 27–39.

[2] X. Liu, M. R. Saat, C. P. Barkan, Analysis of causes of major train derailment and their effect on accident rates, Transportation Research Record 2289 (1) (2012) 154–163.

[3] S. Dindar, S. Kaewunruen, M. An, Á. Gigante-Barrera, Derailment-based fault tree analysis on risk management of railway turnout systems, in: IOP Conference Series: Materials Science and Engineering, Vol. 245, IOP Publishing, 2017, p. 042020.

[4] P. Rungskunroch, A. Jack, S. Kaewunruen, Benchmarking on railway safety performance using bayesian inference, decision tree and petri-net techniques based on long-term accidental data sets, Reliability Engineering & System Safety 213 (2021) 107684.

[5] J. Liu, F. Schmid, K. Li, W. Zheng, A knowledge graph-based approach for exploring railway operational accidents, Reliability Engineering & System Safety 207 (2021) 107352.

[6] J. Sae Siew, O. Mirza, S. Kaewunruen, Nonlinear finite element modelling of railway turnout system considering bearer/sleeper-ballast interaction, Journal of Structures 2015.

[7] M. F. Ishak, S. Dindar, S. Kaewunruen, Safety-based maintenance for geometry restoration of railway turnout systems in various operational environments, in: Proceedings of The 21st National Convention on Civil Engineering, Songkhla THAILAND, 2016.

[8] S. Dindar, S. Kaewunruen, M. An, M. H. Osman, Natural hazard risks on railway turnout systems, Procedia engineering 161 (2016) 1254–1259.

[9] G. Wang, T. Xu, T. Tang, T. Yuan, H. Wang, A bayesian network model for prediction of weather-related failures in railway turnout systems, Expert systems with applications 69 (2017) 247–256.

[10] I. G. Kreft, I. Kreft, J. de Leeuw, et al., Introducing multilevel modeling, Sage, 1998.

[11] J. Albert, A mcmc algorithm to fit a general exchangeable model, Communications in Statistics-Simulation and Computation 25 (3) (1996) 573–592.

[12] J. H. Albert, Criticism of a hierarchical model using bayes factors, Statistics in medicine 18 (3) (1999) 287–305.

[13] E. Hauer, Overdispersion in modelling accidents on road sections and in empirical bayes estimation, Accident Analysis & Prevention 33 (6) (2001) 799–808.

[14] C. Robert, G. Casella, Monte Carlo statistical methods, Springer Science & Business Media, 2013.

[15] E. Castillo, Z. Grande, A. Calviño, Bayesian networks-based probabilistic safety analysis for railway lines, Computer-Aided Civil and Infrastructure Engineering 31 (9) (2016) 681–700.

[16] G. J. Read, A. Naweed, P. M. Salmon, Complexity on the rails: A systems-based approach to understanding safety management in rail transport, Reliability Engineering & System Safety 188 (2019) 352–365.

[17] A. Iranitalab, A. Khattak, Probabilistic classification of hazardous materials release events in train incidents and cargo tank truck crashes, Reliability Engineering & System Safety 199 (2020) 106914.

[18] Z. Li, C. Chen, Y. Ci, G. Zhang, Q. Wu, C. Liu, Z. S. Qian, Examining driver injury severity in intersection-related crashes using cluster analysis and hierarchical bayesian models, Accident Analysis & Prevention 120 (2018) 139–151.

[19] S. Dindar, S. Kaewunruen, M. An, Rail accident analysis using large-scale investigations of train derailments on switches and crossings: Comparing the performances of a novel stochastic mathematical prediction and various assumptions, Engineering failure analysis 103 (2019) 203–216.

[20] M. C. Baechler, T. L. Gilbride, P. C. Cole, M. G. Hefty, K. Ruiz, Building america best practices series-high-performance home technologies: guide to determining climate regions by county, Tech. rep., EERE Publication and Product Library, Washington, DC (United States) (2013).

[21] T. Treichel, C. Barkan, Working paper on mainline freight train accident rates, Research and Test Department, Association of American Railroads.

[22] C. P. Barkan, C. T. Dick, R. Anderson, Railroad derailment factors affecting hazardous materials transportation risk, Transportation research record 1825 (1) (2003) 64–74.

[23] R. T. Anderson, C. P. Barkan, Railroad accident rates for use in transportation risk analysis, Transportation Research Record 1863 (1) (2004) 88–98.

[24] P. Gustafson, S. Hossain, Y. C. Macnab, Conservative prior distributions for variance parameters in hierarchical models, Canadian Journal of Statistics 34 (3) (2006) 377–390.

[25] A. Gelman, Prior distributions for variance parameters in hierarchical models (comment on article by browne and draper), Bayesian analysis 1 (3) (2006) 515–534.

[26] W. K. Hastings, Monte carlo sampling methods using markov chains and their applications.

[27] I. Yildirim, Bayesian inference: Metropolis-hastings sampling, Dept. of Brain and Cognitive Sciences, Univ. of Rochester, Rochester, NY.

[28] S. M. Lynch, Introduction to applied Bayesian statistics and estimation for social scientists, Springer Science & Business Media, 2007.

[29] J. Albert, Bayesian computation with R, Springer Science & Business Media, 2009.

[30] S. Dindar, S. Kaewunruen, M. An, J. M. Sussman, Bayesian network-based probability analysis of train derailments caused by various extreme weather patterns on railway turnouts, Safety science 110 (2018) 20–30.

[31] S. Dindar, S. Kaewunruen, M. An, Bayesian network-based human error reliability assessment of derailments, Reliability Engineering & System Safety 197 (2020) 106825.

[32] M. Catelani, L. Ciani, G. Guidi, G. Patrizi, An enhanced sherpa (e-sherpa) method for human reliability analysis in railway engineering, Reliability Engineering & System Safety (2021) 107866.

[33] X. Liu, Optimizing rail defect inspection frequency to reduce the risk of hazardous materials transportation by rail, Journal of Loss Prevention in the Process Industries 48 (2017) 151–161.