



Development of Predictive Model for 3-D Distribution of Key Soil Parameters

Dr. Yu Wayne Wang

Associate Professor in Civil & Environmental Engineering

School of Science, Engineering & Environment University of Salford, Manchester UK





Contents

- 1. Geostatistical interpolation techniques
- 2. Modelling and prediction for sampling plots
- 3. Software development
- 4. Software installation and operation





1. Geostatistical interpolation techniques

- Geostatistics is used to analyze and predict the values associated with spatial phenomena.
- Geostatistical methods not only provide interpolated values, but also the uncertainty for those values.
- Geostatistics is widely used in environmental science and engineering.







The spatial distributions of geo-physical properties, such as elevations, soil contamination, etc. can be approximated using spatial interpolation.







An example of the deterministic interpolation the Inverse Distance Weighting (IDW)

- IDW is simple and popular in practice.
- It assumes that the value of a geo-property at a spatial position relates to the data at other surrounding points.
- Their relation can be approximated in terms of their distance between.







$$Z(\boldsymbol{X}_p) = \sum_{i=1}^n w_i Z(\boldsymbol{X}_i)$$

$$w_{i} = \frac{\frac{1}{(x_{p} - x_{i})^{m}}}{\sum_{j=1}^{n} \frac{1}{(x_{p} - x_{j})^{m}}}$$

 w_i : the weight factor at each sampling point i.

 $X_p - X_i$: the distance between point p and i.

m : an order number for approximation.





The Geostatistical Analysis

Geostatistics introduces in a concept which considers random variation of geo-properties and represents them in terms of their spatial covariance.







For a spatially continuous geo-property, Z(X), where X indicates the spatial coordinates (x, y), It has a deterministic mean value, M(Z(X)). As a result, the geo-property can be expressed in a form of:

$$Z(\boldsymbol{X}) = M(Z(\boldsymbol{X})) + \Delta(\boldsymbol{X})$$

where Δ is called stochastic spatial residual – variation against the mean value, *M*.





Two Assumptions

- I. the value, M, over a certain spatial region is the same as that within the whole area, X + h,
 - h all the distances of all the location X to all other points in the area. i.e.:

 $E\{M(Z(X)) - M(Z(X+h))\} = 0$

E – an error function stands for error or difference.





II. The variance the geo-property over the spatial area is auto-correlated to the distance, and the correlation can be evaluated using the semi-variance :

$$\gamma(\boldsymbol{X},h) = \frac{1}{2} \left[Z(\boldsymbol{X}) - Z(\boldsymbol{X}+h) \right]^2$$

 $\gamma(X, h)$ – the semi-variance at distance, h, referring to the value at position, X.





For certain number of collection of data (samples), an average semi-variance (also simply called the **semi-variance**) for all is also defined as:

$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^{n} [Z(X_i) - Z(X_i - h)]^2$$

i – the individual sampling points n – the total number of sampling points

Now $\gamma(h)$ is not relevant to a specific position but for all points in the area.





The relationship of the **semi-variance**, $\gamma(h)$, against distance, *h*, is called the variogram. Mathematical functions are employed to represent the relationship, i.e.,

 $\gamma(h)\approx f(h)$







The more the number of sample points the higher the accuracy of the variogram model representation.







Using the **semi-variance** function to represent the stochastic spatial residual value, $\Delta(X)$, we have:

$$Z(\boldsymbol{X}) = M(Z(\boldsymbol{X})) + \gamma(h)$$

The equation states that a geo-property (Z) at a specific site, X, equals to the mean value of the geo-property, M(Z(X)), in the region plus the semi-variance, $\gamma(h)$, of the geo-property in the whole region against site, X





The Geostatistical interpolation – Kriging

In term of the interpolation principle:

and the Geostatistical theory:

 $Z(X_p) = M(Z(X)) + \gamma(X_p - X_j), (j = 1, 2, \dots, n).$ Eq. (2)







$$M(Z(\mathbf{X})) + \gamma(\mathbf{X}_p - \mathbf{X}_j) = \sum_{i=1}^n w_i Z(\mathbf{X}_i)$$

where, *j* = 1,2, …, *n*.

At last, the weighting coefficient, w_i , for each sampling points can be determined solving the produced equation system:

$$[\mathbf{A}] \begin{cases} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_n \end{cases} = \begin{cases} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{cases}$$





2. Modelling and prediction two for plots







The semi-variance of TPH







The semi-variance of Aliphatic C8-C16







The semi-variance of Aliphatic C16-C35





The predicted areas in the two plots



Antopia agente Alexa (e Alexaño) en portes A de anto y el estat fonciar de portes







The Modelling results





Plot 1



Hazardous Daily Soil Ingestion (I_s) g/day



Plot 2









3. Software development







- KOC Soil Contamination Information Tool (KOC-SCIT) is a database software for two specific oil lake areas in Kuwait.
- The database contains all the analysed soil contamination properties for two rectangular areas in the two plots.
- The data extraction uses both curve fitting and linear interpolation methods within the database.
- KOC-SCIT runs on the MS Excel. A technician with basic knowledge and skills using MS Excel can operate the software.



The software structure









KOC-SCIT has three operating worksheet in front, they are **Home**, **File Installed Path**, **Data Found**, and numbers of data worksheets in background.







Each of the data worksheet contains the geo-statistical modelling results computed using an in-house Kriging solver at 40,401 locations of a mesh for a sampling layer.









In depth direction, KOC-SCIT adopts curve fitting to predict local vertical contamination profile.







Each modelling task result has been recorded in mesh map format

File	Hor	ne In	isert Draw	Page Layout	Formulas	Data Rev	iew View	Autom	nate D	eveloper	Help	D					ç	□ Com	iments d	含 Share
Past V	¦	Arial B		10 - A^ A [~] ~ <u>A</u> ~ <u>A</u> ~		≫~~ et	Wrap Text Merge & Center	•	General 📧 ~ %	• 9 (~ .00 →.0	Conditional Formatting ~	Format Table	as Cell Styles *	Insert → Delete → Format →	$ \begin{array}{ccc} \Sigma & \bullet & A \\ \hline & & Z \\ \hline & & & \\ \hline \hline & & & \\ \hline \hline & & & \\ \hline \hline \\ \hline & & & \\ \hline \hline & & & \\ \hline \hline & & & \\ \hline \hline \\ \hline & & & \\ \hline \hline \\ \hline \hline \\ \hline \\$	Find & Ar Gelect ~	alyze Data	Sensitivity	
Clipt	board Is	ā	Font		2	Alignment		۲ _N	Nu	mber	Гы		Styles		Cells	Editing	Ar	alysis	Sensitivity	_
012		x :	XZ	f 1 166600	10967925															~
015				1.100035	10807825															
	A	0005	В	C	D	E	F		G	H				J	K	L	M		N	•
1	1.16388	9605	1.164416394	1.164938523	1.165455525	1.16596692	2 1.1664/222	2 1.1	66970924	1 1.1674	462517	1.167946	648 1	.168422282	1.168889385	1.169347245	1.169795	031 1	1.170233028	1.1
2	1.		3D distr	ibution, Plot	1 - TPH 0-1	Dcm			(Contour	- map	. Plot 1 - T	PH 0-	10cm		1.109101012	1.1090029	75 4	1.170043300	1.17
3	1.			,								,				1.100923942	1 1601205	1.109300073 1		2 1 1
4	1.														Series201	1.100071329	1.1091290	12 1	1.109074723	1 1 16
6	1.															1.100309273	1 1685450	0/ 1	1.16800560/	1 1 16
7	1 35														Series181	1 167745196	1 1682140	104 1	1 168667701	1.10
8	1							<u>.</u>							-	1 167385259	1 1678585	23 1	1 168315825	1.16
9	1							1							Series161	1 167001215	1 1674794	91 1	1 167941086	1 16
10	1 3							. N								1 166593831	1 1670778	36 1	1 167544496	1 16
11	1.														Series141	1.166163765	1.1666543	328	1.16712695	5 1.16
12	1. 25														5enes141	1.165711549	1.1662096	602 1	1.166689198	3 1.16
13	1.															1.165237568	1.1657441	39 1	1.166231823	1.16
14	1.											0			Series121	1.164742045	1.165258	324	1.16575522	2 1.1
15	1. 2															1.164225024	1.1647520	018 1	1.165259577	1.16
16	1.								, P						Series101	1.163686357	1.1642253	375 1	1.164744856	5 1.16
17	1. 1.5												SE MARSHINA SE			1.163125692	1.1636779	92 1	1.164210775	i 1.16
18	1.														Series81	1.162542464	1.1631093	816 1	1.163656804	1.16
19	1.															1.16193589	1.1625185	61 1	1.163082149) 1.
20	1. 1														Series61	1.161304968	1.1619046	94 1	1.162485755	i 1.1
21	1.															1.16064848	1.1612664	48 1	1.161866299) 1.16
22	1. 0.5	5				Se	eries175								Series/1	1.159964995	1.1606023	819 1	1.161222206	1.16
23	1					Series	117								361/6341	1.159252879	1.159910	58 1	1.160551647	1.16
24	1.											N. L.				1.158510313	1.1591892	96 1	1.159852566	i 1.16
25	1.	13 13		TURNING		Series59									Series21	1.157735308	1.1584363	343 1	1.159122694	1.15
26	1.		85 85 85	6 6 5 8 9		Forior1										1.156925727	1.1576494	33 1	1.158359578	3 1.15
27	1.			12 13 145	157 169 181 93	-C+-C31			8 2 9	10 4 m	1	0 6 8 2	0 0 4		Series1	1.156079312	1.1568261	38 1	1.157560607	1.15
28	1.							1 1	4 m 5	1000	00 0	10: 11: 12: 12:	14, 13, 13, 13, 13, 13, 13, 13, 13, 13, 13	16. 177 187 190	19	1.155193713	1.1559639	928 1	1.156723054	1.15
29	1.															1.154266521	1.15506	502 1	1.155844106	6 1.15
30	1.		0.5-	1 ■ 1-1.5 ■ 1.5-	2 2-2.5 2.5-	3 3-3.5					0-1	1-2 2-3	3-4			1.153295298	1.1541123	818 1	1.154920913	1.15
31	1.14202	0232	1.143009139	1.144001000	1.140410191	1.14027912	4 1.14/142/9	1 1.1	40003313	1.1400	000347	1.149723	110 1	. 130300403	1.131431342	1.152277616	1.1531176	53 1	1.153950623	1.15
30	1 1/158	Sheet1	1 1/2/55068	1 1/3331258	1 1//208303	1 1/508652	2 1 1/506530	3 11	46844000	1 1/17	721038	1 1/18508	302 1	1/0/72611	1 1503/3702	1 151211003	1 1520736	1 1	152030/3/	1 15
Ready	Ē	Access	ability: Unavailable													H	U – U –			+ 100%





A user interface has been developed to facility to extract data by the request and definition of users.

	KOC-SCIT-2023-v1							×	
ste 🗈									
- 4 7	Location						1		
pboard	Fast		North		Dept	h	Find		
	784422		209734		10				
A	704422		200754	_	10		Reset		
L								N 3209631	= d
X	Property distribution ·	Area 1 -				1		E 785105	A
7844	Depth (cm)	10.0	0-10	10-20	20-30				.44
7844	TPH (HEM) %	10	12	0.7	0.3		In.	A	.44
7044	Trift (neil), so	1.0	1.2	0.7	0.5		13		.44
	Aliphatic C8-C16, mg/kg	200.1	246.8	162.2			2 124 8	187	
	Aliphatic C16-C35, mg/kg	780.7	950.2	641.5			Je is	and a start of the	
	Aromatic C8-C16, mg/kg	276.4	346.7	220.4					
_	Aromatic C16-C35, mg/k	703.2	756.1	654.1			*19-	1110	-
			2020	1202			11.***	~	
	рН	8.5	8.3	8.6	8.7			204	
	EC, uS/kg	614.8	691.7	565.4	549.0			350 33	
							A	194 24	
	HZ Daily Rate, g/day	2.5	2.3	2.9			31.36	1 N 1	-
	PAHs (20%) ma/ka	0.0	0.0	0.0			1 1 A A	10	
_		0.0	0.0	0.0					
	- Depth reached the s	earched c	ontent -						
		curcifica c	oncent	_		10-30100	New Mean	and and and	
	1.0 TPH (H	EM)		- Se	earch	N 3208734		Area 1 👻	
_			-		()	E 784422			-
_		Depth (cm)	Tota	Volum	e (m3)				-
	ТРН	8.86					Save	Close	
	Linu .								
-		Υ							





All local contaminating information is listed out in a single table







Find the local depth given the contaminating degree of properties

ne Insert	Draw Page I	ayout F	ormulas	Data	Review	View	Automate	Developer	Help	2	Comme
DC-SCIT-2023	-v1								-		×
E L	ocation —										
	F 1							Find			
	East		North		Dept	.n			-		
	784885	3	210324		10			Reset	1		
									1	N 321	0434
Property	distribution - /	Area 2 —				7				E 785	282
Depth (cn	ו)	10.0	0-10	10-20	20-30						
TPH (HEM	1), %	1.4	2.0	0.9	0.0						
						and the second second			and the second		-
Aliphatic (C8-C16, mg/kg	318.2	401.4	252.3							N
Aliphace	210°C33, mg/kg	1140.5	1250.5	1014.0		de					A
Aromatic	C8-C16, mg/kg	188.7	291.7	122.1							1
Aromatic	C16-C35, mg/kg	958.8	1055.7	870.6			43	46 48	51.	54	-
рH		8.0	7.8	8.2	0.0	38	44	45 10	52		200
					10.000	1	44		0		20
EC, µS/kg		1236.7	1382.7	1106.1	0.0	C. A.	Alter	- and	20		18
HZ Daily F	Rate, g/day	1.9	1.8	2.1		100	- Ann				
	, 5,,			1		STREET.	0 50 100) 200	300	400	24
PAHs (20	%), mg/kg	5.9	6.3	5.6		100	And in case of	a start of the	11000	Me	ter
						_					
- Depth r	eached the se	arched co	ontent -								
1.0	TPH (HEN	1)		5	arch	N 32	10324			Dist 2	
1 A		N N			aich	E 78	1885			Plot 2	-
T T	Aliphatic	1)	4	/olume	e (m3)	E 70	4000				
	Aliphatic	C16-C35						Carro	1	Class	1
TPH	Aromatic	C8-C16						Save		Close	-
	Aromatic	C16-C35		-		-					_
	pH										
Home D	ata F. EC		-	-						• []	





University of Salford

Record and output all the search history







Find the total contaminated volume given a criterion – function has not finished yet.

KOC-SCIT-2023-v1						×		
Location East	Nor	th	Depth			Find data f		
784422	3208734		10			Find data f	or Depth	
Property distribution - Pla	ot 1 —							N 3209631 F 785105
Depth (cm)	10.00	0-10	10-20	20-30				-
TPH (HEM), %	1.0	1.2	0.7	0.3		21	N.	A
Aliphatic C8-C16, mg/kg	200.1	246.8	162.2				1	
Aliphatic C16-C35, mg/kg	780.7	950.2	641.5			1	AS	
Aromatic C8-C16, mg/kg	276.4	346.7	220.4				7 10	
Aromatic C16-C35, mg/kg	703.2	756.1	654.1				15	
рН	8.5	8.3	8.6	8.7			1 10 20a 14 ²¹ 35a	2
EC, μS/kg	614.8	691.7	565.4	549.0		13	27 31 22 ²⁵ 196	2 <mark>3</mark> 2 1
HZ Daily Rate, g/day	2.5	2.3	2.9				30.	34
PAHs (20%), mg/kg	0.0	0.0	0.0			0 50100 200 300	100	
Depth reached the sear	rched co	ntent –			Meter			
1.0 TPH (HEM))	•	Se	earch	N 3208			Plot 1 💌
Dep	oth (cm)	Total	Volume	e (m3)	E /844			
TPH 8.80	6				Total	Volumn	Save	Close





4. Software installation and operation

Information is provided by the software document

