EVALUATION OF SEMIVARIOGRAM MODELS IN THE STUDY OF SPATIAL INTERPOLATION OF SOIL SALINITY

A CASE STUDY OF CHAO PHRAYA AND THA CHIN LOWER BASINS, THAILAND

Kanoksuk Chankon[1,]*, Sanwit Iabchoon[1],Pariwate Varnakovida[1], Amnat Chidthaisong[2]
[1] KMUTT Geospatial Engineering and Innovation Center,King Mongkut's University of Technology Thonburi, Thailand
[2] The Joint Graduate School of Energy and Environment, King Mongkut's University of Technology Thonburi, Thailand
sming.gt@gmail.com

***Abstract*— Sea level rise accelerated by climate change is one of the concerns in causing sea water intrusionand salt contamination in surface and ground water as well as soil.Data compilation and synthesis in the past have indicated that the sea levels in the Gulf of Thailand have been rising at 3-5 mm/year. This study evaluated several semivariogram models for the study of spatial distribution of soil salinity in the main central basins of Thailand. Soil salinity tests were carried out in Chao Phraya and Tha Chin lower basins alongside of the Gulf of Thailand. Study area is approximately 4,900 km^2. Soil salinity measurements were performed in 2018 at two levels of depth at 30 and 100 cm. Line transect sampling method was used to determine the location of measuring points. Each of measuring points was about 10 km apart. Totally 40 sampling points across the basins were selected. Suitable semivariogram model was selected from a comparison result of geostatistical analysis. Four models of semivariogram were compared: circular, spherical, exponential and Gaussian. Kriging interpolation method was used to estimating soil salinity values of unknow point. Soil salinity as measured by electrical conductivity (EC) in these areas ranges between 0.02-3.83 dS/m at 30 cm depth and 0.08-4.00 dS/m at 100 cm depth, respectively. There is a clear gradient of soil salinity at both depths; higher along the shore (>2.00 dS/m) and gradually decreases towards inland (0.08-2.01 dS/m at 50-100 km from the coastline). Evaluation of semivariogram model performance reveals that the Gaussian model and exponential model were the best models for estimating the value of soil salinity at 30 cm and 100 cm with root mean square error (RMSE) is 0.92 and 1.10, respectively.**

Keywords—soil salinity; kriging; spatial interpolation; Chao Phraya and Tha Chin river basin; sea level rise

I. INTRODUCTION

Climate change is an urgent issue that affected many aspects of human life such as public health, diseases, fresh water scarcity, and especially agricultural areas. Importantly, climate change causes an increase of global temperature resulting extreme weather and sea level rise [1]. Sea level rise is one of theconcerns that causing seawater intrusionin land and salt contamination in surface and groundwater as well as soil.

Thailand faces salinity problems in many regions including the northeastern and the central plain, especially salinity in the soil. Soil salinity is the salt content in the soil. Soil salination in Thailand causes by natural processes of the environment such as sea level rise and saltwater encroachment as well as human activity such as salinity intrusion in land from agriculture activities in upstream areas and unsuitable irrigation. Saltwater encroachment is most common in coastal areas, where the freshwater is displaced by the inland movement of saltwater from the sea [2], which also happens in coastal areas along the Gulf of Thailand. The increased of salinity in soil and water (both surface and ground water) would impacts the quality of water for drinking and irrigation with serious economic, social and environmental consequences for both rural and urban communities. [3] [4]

The study of soil salinity in the central regions of Thailand has been limited. Two studies were performed by Arunin (1996) and Takaya (1987) [5][6]. Additionally, there is not enough information about the phenomenon of temporal and spatial change of salinity in the land. To be able to closely monitor a real-time situation of soil salinity, it is necessary to have accurate, standardized and up-to-date supported data. Thus, there is the urgent need for the studies of soil salinity to guide a more efficient soil and water management. One of the common methods used for assessing spatial distribution of soil salinity is geostatistical method. Geostatistics is a class of statistics used to analyze and predict the values associated with spatial or spatiotemporal phenomena. This method describes spatial patterns and interpolates values for locations where samples were not collected [7]. The major key behind geostatistical method is called semivariogram. It is an important process, but also a difficult step of geostatistical analysis. Semivariogram used to determine a suitable model for experimental values and useful to identify the prediction errors after the process of spatial interpolation for locations where unmeasured.

The propose of this study is to evaluate the appropriate semivariogram model for spatial interpolation of soil salinity in the lower basinof Chao Phraya and Tha Chin river basins where located in the central of Thailand.

II. MATERIAL AND METHODS

A. Study Area

The study was performed in the lower part of Chao Phraya and Tha Chin river basins, located alongside of the Gulf of Thailand.This study area covering eight provinces including Bangkok, Nonthaburi, Nakhon Pathom, PathumThani, Phra Nakhon Si Ayutthaya,SamutPrakan, SamutSakhon and SuphanBuri. It is extended between 13°30'N-14°25'N and 99°40'E-101°5'E covering an area of approximately 4,900 km^2 (Fig.1) andthe distance from shoreline across to the north is around 100 km.The average height in the area is about 9.4 m with a relative height difference is 19 m. This is an important economics area in the central of country, which is dominated by agricultural land and various industries.The location of study areais illustrated in Fig. 1.

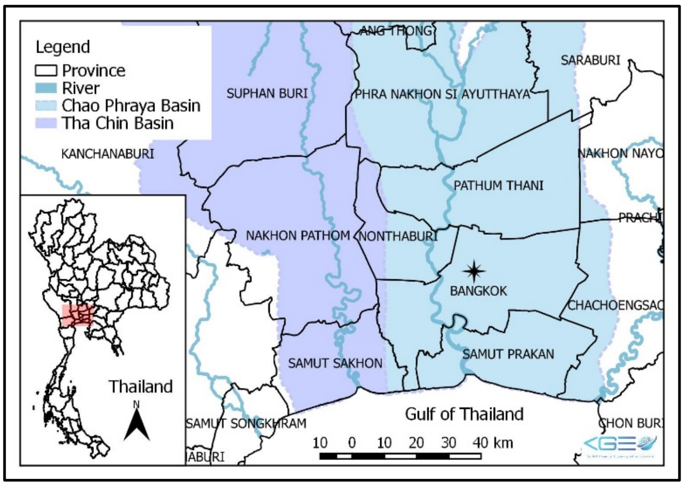

Fig. 1. Location of study area.

B. Data Collection

Soil salinity was assessed by measuring electrical conductivity (EC) in which a high EC value indicates a high salinity level [8] [9]. The EC data were collected in June 2018 at the two levels of depth which are 30 and 100 cm. Reconnaissance soil survey was carried out using LULC data during 2010-2012 and road network as base map. The sampling point were determined using Line transect sampling method [10][11]. Totally 40 soil sampling sites were selected. Each of points is located along the transect line with 10 km apart spreading over the basins as shown in Fig. 2. A handheld global positioning system (GPS) was used to record the location of each sample site. Undisturbed soilsamples at two depths (30 and 100 cm) were collected with 4 soil cores from each site and mixed well into a composite soil sample. Soil electrical conductivity values were measured by using EC meter type Hanna HI98331 and expressed as deci-Siemens per meter (dS/m).

Two series of soil test were carried out in two level of depth. For each series,statistical analyses were preceded by the calculation of basically statistical: mean, standard deviation, kurtosis, skewness, minimum value, median and maximum value.

C. Spatial Interpolation

There are many types of spatial interpolation method, for example, inverse distance weighting (IDW), spline, kriging and trend. However, in the field of soil science and geology, the often and widely used method is kriging interpolation. There are various comparative studies about surface modeling of soil properties, especially electrical conductivity. The results showed clearly that the kriging method is the most efficient to EC estimating [12] [13] [14]. Kriging is a geostatistical interpolation technique that considers both the distance and the degree of variation between known data points when estimating values in unknown areas.Spatial variability in soilsalinity was assessedusing semivariogram model, which is the basic and widely used geostatisticalfunction. This model is given by theformula [15] [16] [17]:

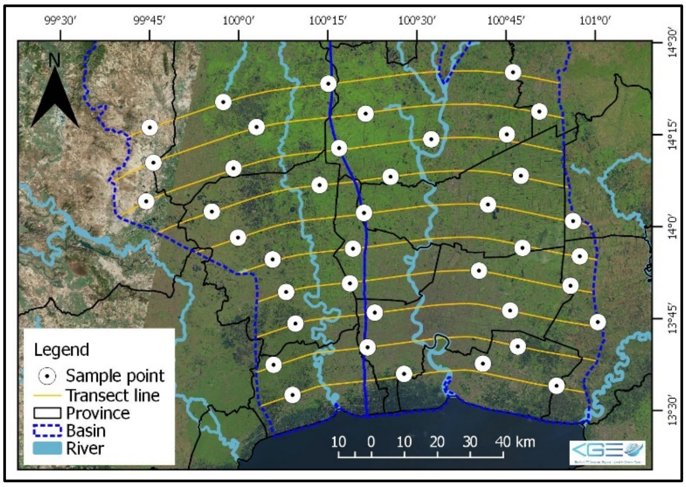

Fig. 2. Transect linesand sample points

$$\gamma(h) = \frac{1}{2N_h}\sum_{i=1}^{N_h}[z(x_i) - z(x_i + h)]^2 \quad (1)$$

where: $\gamma(h)$ is the empirical semivariogram; $z(x_i)$, $z(x_i + h)$ = soil moisture values at sample points x_i and $x_i + h$, spaced apart at distance h; N_h = number of pairs (x_i, $x_i + h$) of soil salinity values at points spaced at distance.These were used for calculating the semivariogram function.

In this study, four functions were compared to model the empirical semivariogram as follows: circular, exponential, spherical and Gaussian. Because some good results have appeared in several kinds of soil salinity research [17]. Next step, mathematical models which given the best fit for each empirical semivariograms were selected and compared with one another.

The prediction of an unknown point was interpolated using the kriging method[10]. Kriging is a spatial interpolation method of good statistical properties. The initial equation of kriging is given by the formula [18]:

$$\hat{z} = \sum_{i=1}^{n} w_i z_i \quad (2)$$

where: $\hat{z}$ = predicted value;z_i = sample value at location i; w_i = weight; n = number of sample data.

Assessment of prediction error is an important step used to evaluate interpolation efficiency. The following prediction errors were analyzed: mean, root mean square, mean standardized, root mean square standardized, average standard error. These parameters will help to distinguish the suitable model for this study area.

III. RESULTS AND DISCUSSION

Soil salinity were assessed using an electrical conductivity or EC value, unit in this study is deci-Siemens per meter or dS/m. A high EC value indicates a high salinity level. Results of basic descriptive statistics of the EC measurements are shown in Table 1.

TABLE I. STATISTIC OF ELECTRICAL CONDUCTIVITY FOR EACH SERIES OF TESTS.

Parameter	**Series / Depth (cm)**	
	1st Series / 30	***2nd Series / 100***
Count	40	40
Mean	1.01	1.49
Standard Error	0.16	0.19
Median	0.70	1.17
Mode	2.62	4.00
Standard Deviation	0.99	1.19
Sample Variance	0.99	1.42
Kurtosis	1.31	-0.08
Skewness	1.38	0.92
Minimum	0.02	0.08
Maximum	3.83	4.00
Sum	40.38	59.73

Statistical analysis presented that the EC values in the first series (30 cmdepth) of soil tests were low salinity than those values found in the second series (100 cmdepth). In addition to the first series of measurements, EC ranged from 0.02 to 3.83 dS/m. The mean of EC values for the first series was 1.01 dS/m. In the second series of soil tests, EC ranged between 0.08 to 4.00 dS/m. Also, the mean value was shifted to 1.49 dS/m higher than the first series of soil tests.The standard deviation calculated for the second series was 1.19 which is higher than the values obtained in the first series (0.99). Accordingly, the sample variance of second series were more varied than the first series of measurements, value is 1.42 and 0.99 respectively. The median of the calculated parameter of both series was similar lower than the arithmetic meanof each series, which is represents that EC values measured were asymmetrically distributed. This is related to the coefficient of skewness. Furthermore, both series of EC measurements had positively skewed data distribution (skewness > 0).

According to the study of semivariogram, spatial relationshipswere studied. This tool describesthe variationbetween measured pointsdepending on the distance between them. Prior thesemivariogram was generated, the trends which appeared in both EC data series were removed. The followingfunctions were used to model the semivariograms: circular exponential, spherical and Gaussian.

The semivariograms obtained for the first series of EC measurements and the estimate of semivariograms using theoretical models were illustrated in Fig.3. Table 2 presented the semivariogram properties of each model.

In the first series, analysis of the semivariogram graphs showed that the range of spatial autocorrelation was similar for all models. On the other hand, nugget and sill are different. The nugget was lowest in the exponential model but highest in the Gaussian model. The sill was lowest in the exponential model but highest in the circular model.

In the second series,the semivariogram were shown in Fig. 4 and the properties of the semivariogram models were shown in table 2. The properties of the spatial autocorrelation range were between 40,383 and 47,437.04 m. The lowest range was found in the circular model,but the highest range was found in the exponential model. The lowest nuggetappeared in the exponential model, whereas the highest nugget appeared in the Gaussian model. The sill was lowest in the Gaussian model but highest in the exponential model.

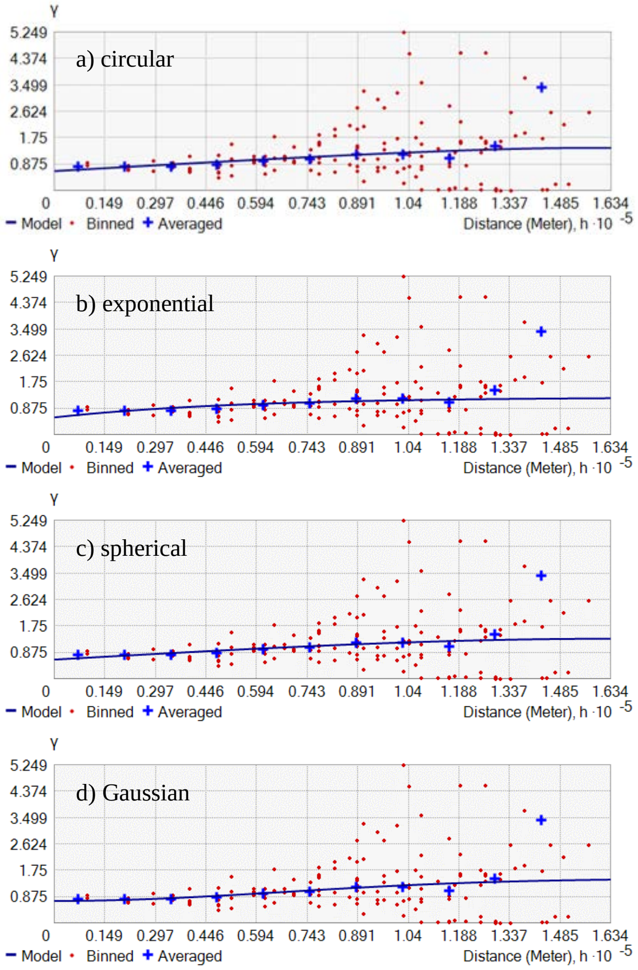

Fig. 3. Semivariogram of electrical conductivity assigned for the first series (depth, 30 cm) of soil tests and their approximate theoretical models: a) circular, b) exponential, c) spherical and d) Gaussian.

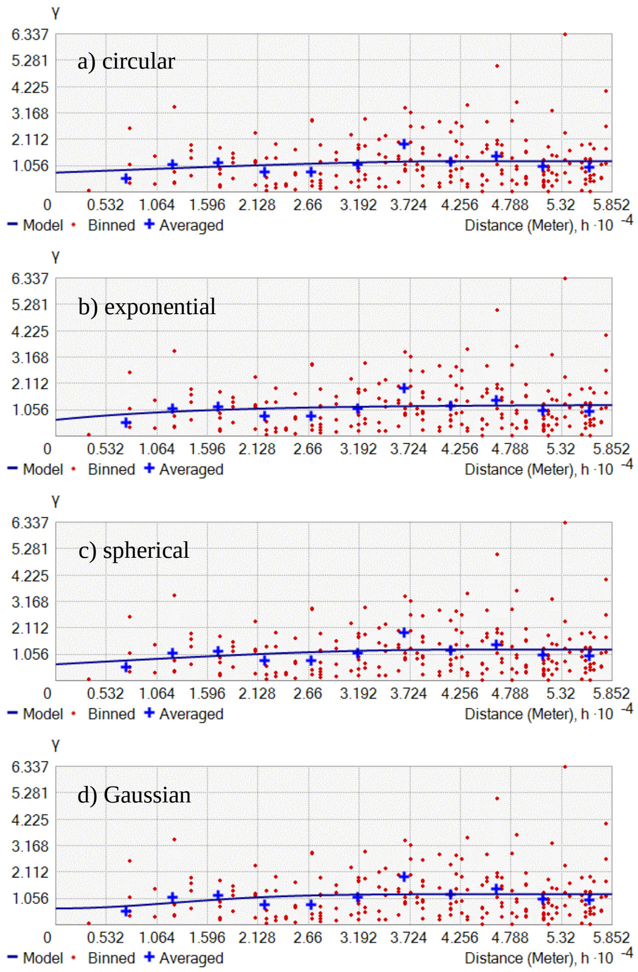

Fig. 4. Semivariogram of electrical conductivityassigned for the second series (depth, 100 cm) of soil tests and their approximate theoretical models: a) circular, b) exponential,c) spherical and d) Gaussian.

The calculation results of the prediction error for those two series of EC measurements were shown in Table 3.In the first series, the distinguishes of smallest values of root mean square error (RMSE) and root mean square standardized error (RMSD) at the measurement points were observed for the Gaussian model (RMSE = 0.917, RMSD= 1.008).The biggest of RMSE was found in the exponential model (0.955). These results suggested that the Gaussian model is a suitable choice for data interpolation in this case.

For the second series, the exponential model provided the lowest values of RMSE and RMSD which were 1.099 and 1.013,respectively. These results suggested that the exponential model is the most suitable model for studying soil salinity study at the 100 cm depth.

The spatial distribution of salinity level in soil was studied by EC data as explained before. Electrical conductivity in the unmeasured area was estimated using the widely employed geostatistical method called kriging interpolation [7][13]. Those two measured data series were used to generate the distribution of EC for each selected semivariogram model.

TABLE II. PROPERTIES OF THE SEMIVARIOGRAM MODELS OF EC WITHBOTH LEVELS OF DEPTH.

Series / Depth	Model	Nugget	Sill	Range
1st series / 30 cm	Circular	0.64	0.76	163,371.19
	Exponential	0.57	0.67	163,371.19
	Spherical	0.63	0.69	163,371.19
	Gaussian	0.73	0.74	163,371.19
2nd series / 100 cm	Circular	0.78	0.46	40,383.17
	Exponential	0.66	0.60	47,437.04
	Spherical	0.77	0.47	43,921.68
	Gaussian	0.89	0.35	43,008.99

TABLE III. PREDICTION ERRORS FOR SELECTED SEMIVARIOGRAMS MODELS.

Depth	Prediction error	Semivariogram Model			
		Cir	*Sph*	*Exp*	*Gau*
30 cm.	Mean	-0.008	-0.008	-0.007	-0.010
	Root Mean Square	0.935	0.937	0.955	0.917
	Mean Standardized	-0.007	-0.007	-0.006	-0.009
	RMS Standardized	1.029	1.032	1.044	1.008
	Average Standard Error	0.907	0.907	0.913	0.908
100 cm.	Mean	-0.004	-0.006	-0.001	0.002
	Root Mean Square	1.118	1.108	1.099	1.101
	Mean Standardized	0.001	-0.001	0.002	0.005
	RMS Standardized	1.037	1.026	1.013	1.017
	Average Standard Error	1.074	1.076	1.083	1.079

[a] Cir = circular, Sph = spherical, Exp = exponential, Gau = Gaussian

The spatial interpolation results of both series are illustrated in Fig. 5 and Fig. 6.There is a clear gradient of soil salinity at both depths; higher along the shore (EC >2.00 dS/m) and gradually decreases towards inland (EC 0.08-2.01 dS/m at 50-100 km from the coastline).When the salinity distributions from those two series of measurements are compared (Fig. 7),some similar spread trends of salinity were observed.

The salinity in soil at 30 cm depth gradually decreases from south to north (coastline to land).In 30 cm depth, the EC value in Tha Chin river mount is higher than Chao Phraya river mount. This value is the same trend withsalinity in the river from the report of water quality by Pollution Control Department that measured at the same time [15]. And of course, it might be the effect by the location of field survey points. Because of, this study focus on collect the soil samples on crop areas. That means no information from the sample location in Bangkok city center.

As shown in Fig. 2 (study area and sample points),those space that has no sample points and either the influence of longer distance between sample points in the left side and right side of Chao Phraya river. On the other hand, there are extremely increases of salinity intrusion towards in the land at 100 cm depth level.Especially in the area between the two rivers. The result shows hotspot in the border between Nonthaburi and Pathum Thani provinces, which is the rice paddy field growing areas. As shown in Fig.7, kriging result for each series which is lowest errors were compared.

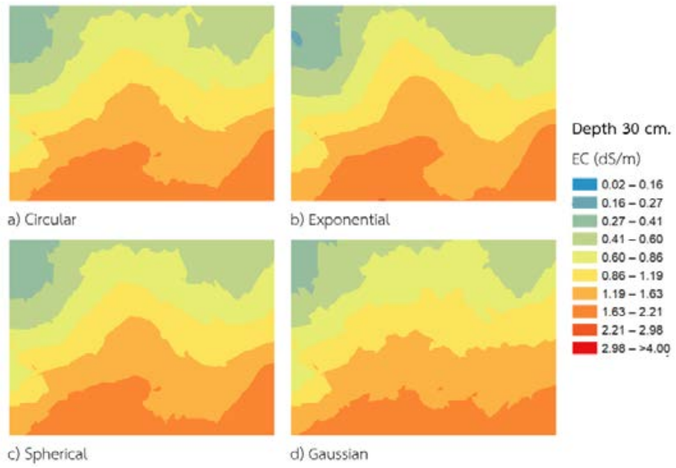

Fig. 5. The distributions of electrical conductivity in the first series of measurements (30 cm - depth), interpolated by ordinary kriging base on semivariogram models: a) circular, b) exponential, c) spherical and d) Gaussian

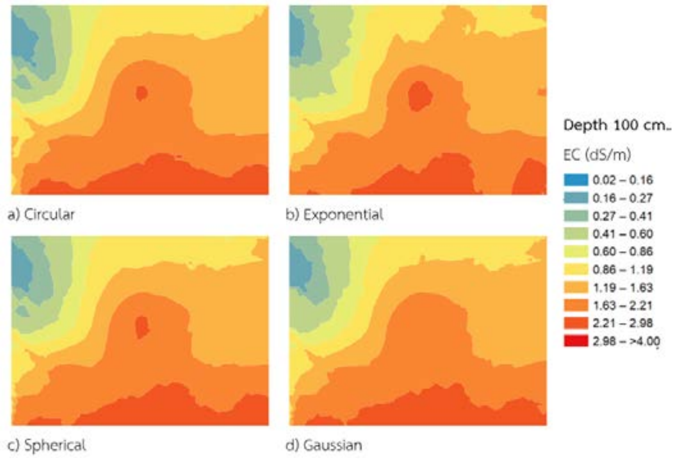

Fig. 6. The distributions of electrical conductivity in the second series of measurements (100 cm - depth), interpolated by ordinary kriging base on semivariogram models: a) circular, b) exponential, c) spherical and d) Gaussian

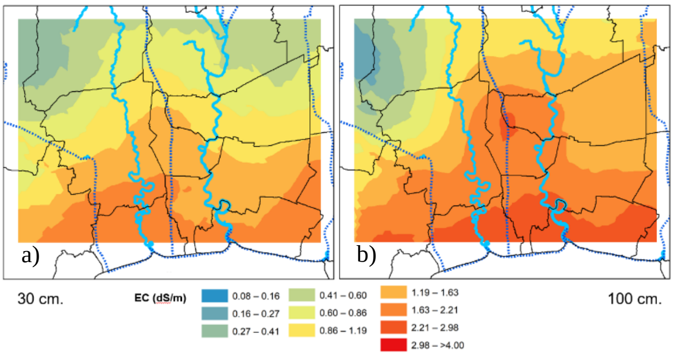

Fig. 7. A comparison of kriging results with the lowest errors model per series of EC measurements: a) Gaussian model at 30 cm. depth, and b) exponential model at 100 cm. depth.

IV. CONCLUSIONS

This study is one part of our soil salinity research project that purpose to assessing the spatial and temporal changes of soil salinity in Chao Phraya and Tha Chin lower basins using geostatistical method. The results in thisstudy are the preliminary process to study the spatial variation of soil salinity in this study area. Aims to select the appropriate equation or model for spatial interpolation for all datasets.

Among the four models of semivariogram analyzed in this study, Gaussian and exponential models provided the lowest errors in estimating the unknown values of soil salinity relative to the empirical data. In should be noted that soil is a non-homogeneous object and diverse mixture with many particles. In the fieldwork, we found different EC values from nearby points, which highlight the important of selected location for data collection. These criteria could influence the samples data variance. Taking together, there is no perfect model for any datasets. The suitable model should have minimized differences in residuals between experimental and theoretical semivariograms.

ACKNOWLEDGMENT

This research was undertaken as part of the research programentitled “Evaluating salinity in river water and soil associated with climate change in Chao Phraya and Tha Chin river basins.” Funding support for this study was provided by Biodiversity-Based Economy Development Office (Public Organization) (BEDO) and National Research Council of Thailand(NRCT). We are grateful to the Faculty of Science, King Mongkut’s University of Technology Thonburi (KMUTT)for their place and equipmentsupport during the realization of this project.The opinions expressed here belong to the authors, and do not necessarily reflect those of BEDO, NRCT or KMUTT.

REFERENCES

[1] S. Naruchaikusol, “Climate change and its impact in thailand. a short overview on actual and potential impacts of the changing climate in southeast asia,” TransRe Fact Sheet, no. 2, Department of Geography, University of Bonn, Bonn, June 2016.

[2] B. D. Hoyle, “Saltwater Encroachment.” Climate Change: In Context, Encyclopedia.com, November 3, 2018. [https://www.encyclopedia.com/environment/energy-government-and-defense-magazines/saltwater-encroachment]

[3] Department of Groundwater Resources, Project report of the groundwater observation network to groundwater monitoring, Department of Groundwater Resources, Bangkok, Thailand, 2015.

[4] Queensland Government, “Impacts of salinity,” December 8, 2016. [https://www.qld.gov.au/environment/land/management/soil/salinity/impacts]

[5] S. Arunin, Salt affected soil in Thailand, Land Development Department (LDD), Bangkok, 1996.

[6] Y. Takaya, Agricultural Development of a Tropical Delta: A Study of Chao Phraya Delta, Univ. of Hawaii Press, Honolulu, 1987.

[7] ESRI, “Introduction to Geostatistical Analyst,” Pro.arcgis.com, November 3, 2018. [http://pro.arcgis.com/en/pro-app/help/analysis/geostatistical-analyst/what-is-geostatistics-.htm]

[8] Y. Shirokova, I. Forkutsa and N. Sharafutdinova, “Use of electrical conductivity instead of soluble salts for soil salinity monitoring in Central Asia.” Irrigation and Drainage Systems, 14, 2000, pp.199–205.

[9] J.D. Rhoades, F. Chanduvi and S. Lesch, Soil Salinity Assessment: Methods and interpretation of electrical conductivity measurements, Food and Agriculture Organaization of the United Nations (FAO), Rome, Italy, 1999.

[10] Zheng, Zhong, F. Zhang, F. Ma, X. Chai, Z. Zhu, J. Shi, and S. Zhang, “Spatiotemporal Changes in Soil Salinity in a Drip-Irrigated Field,” Geoderma, 2009. [Doi:10.1016/j.geoderma.2008.12.002]

[11] S. Miyamoto and A. Chacon, “Soil Salinity of Urban Turf Areas Irrigated with Saline Water II Soil Factors,” Landscape and Urban Planning, 2006. [Doi:10.1016/j.landurbplan.2004.12.011]

[12] G. S. Bhunia, P. K. Shit and R. Maiti, “Comparison of gis-based interpolation methods for spatial distribution of soil organic carbon (soc),” Journal of the Saudi Society of Agricultural Sciences, 17, pp.114-126, 2018.

[13] W.J. Shi, T.X. Yue and Z.M. Fan, “An overview of high accuracy surface modeling of soil,” Procedia Environmental Sciences, 13, pp.2304-2309. 2012.

[14] H. K. Poshtmasari, Z. T. Sarvestani, B. Kamkar, S. Shataei and S. Sadeghi, “Comparison of interpolation methods for estimating pH and EC in agricultural fields of Golestan province,” Intl J Agri Crop Sci., vol. 4 (4), pp.157-167, 2012.

[15] R.A. Olea, Geostatistics for engineers and earth scientists, Boston. Springer, 1999, pp. 303. [ISBN 978-0-7923-8523-3]

[16] R. Webster , M. A. Oliver, Geostatistics for environmental scientists, 2nd Ed., John Wiley & Sons, 2007, pp. 271. [ISBN 978-0-470-02858-2]

[17] R. Obroślak and O. Dorozhynskyy, “Selection of a semivariogram model
in the study of spatial distribution of soil moisture,” Journal of Water and Land Development, no. 35, pp. 161-166, 2017. [DOI: 10.1515/jwld-2017-0080].

[18] P. J. Diggle, J. A. Tawn and R. A. Moyeed, “Model-based geostatistics,” Appl. Statist., 47, Part 3, 1998, pp. 299-350.

[19] Polution Control Department, Weekly Water Quality Report (June 2018),Ministry of Natural Resources and Environment, Bangkok, Thailand, 2018.