Int. J. Advance Soft Compu. Appl, Vol.11, No. 2, July 2019
ISSN 2074-8523

Optimization of the Spatial Interpolation Based on the Sliding Neighborhood Operation Method by using K-Mean Clustering for Predicting the Topographic Shape of the Ground Surface

Achmad Fanany Onnilita Gaffar[1], Ibayasid[2], Rheo Malani[1], Arief Bramanto Wicaksono Putra[1],*, Agusma Wajiansyah[1]

[1]Department of Information Technology, Politeknik Negeri Samarinda, East Kalimantan, Indonesia
e-mail: onnygaffar212@gmail.com, rheo@polnes.ac.id , ariefbram@gmail.com, agusma.wajiansyah@gmail.com
[2] Department of Civil Engineering, Politeknik Negeri Samarinda, East Kalimantan, Indonesia
e-mail: yasid@polnes.ac.id

Abstract

The spatial interpolation method estimates the value of the unobserved locations in geographical space based on the spatial relationship between the position of the unknown points to be expected and the emplacement of the known points about it. Image-based spatial interpolation applies adaptive spatial image processing approach. This method considers the image context locally by involving aspects of geometry, morphology or radiometric. The operational window (can be an array or square), likewise called a pixel neighborhood operation, is an area that locally covers a lot of neighboring pixels around the observed pixel. By sliding the window (as a kernel operator) to the entire image, it will generate the new value of a center pixel. This new value determination using the neighborhood relationship between the center pixel and its neighboring pixels. For a large number of data points, the relationship between these data approximated by the clustering concept, where the entire cluster center considered as new points generated. This study applies the neighboring pixels concept for the spatial interpolation through a sliding neighborhood operation. This operation uses the IDW concept to determine the new value of center pixels based on its neighboring pixels. To improve the result, firstly K-Mean clustering used to create new points based on the cluster

center. Once more, the value of the cluster center also determined by using the IDW concept based on all cluster members. The purpose of this study is to optimize the spatial interpolation based on the Sliding Neighborhood Operation method by using K-Mean clustering for predicting the topographic pattern of the ground surface. These results of this study showed that by applying K-Mean clustering to spatial interpolation based on the sliding neighborhood operation, there was a performance improvement from the success rate of 87.58% (MAPE of 12.42%, without K-Mean) to 90.73% (MAPE of 9.27%, with K-Mean).

Keywords: *Spatial interpolation, Sliding neighborhood operation, Neighboring pixels, IDW concept, K-Mean clustering*

1 Introduction

Topographic modeling of a particular area requires the support of data and information that is not always available. One of the simplest ways to obtain the necessary data is to conduct survey activities. The problem that often inhibits survey activities is wide coverage area with various physiographic conditions and other related factors. For effectiveness and efficiency purposes, the use of sample data becomes very significant. The availability of sufficient sample data will create a sound exemplar. This problem has led to a wide range of interest in the study area. Geo-statistics is part of a statistic that deals specifically with the analysis and interpretation of geographically referenced data through a set of statistical tools to examine the spatial variability and spatial interpolation. One of its roles is to bring about a predictive surface. Geo-statistic mapping defined as the production of analytical maps using a set of known data pints obtained from field observation and additional information to calculate certain values in particular ways at interest locations with computer program support [1]. One of the uses of geo-statistic is to predict the value of the unobserved location in the selected region of geographical space widely used in civil engineering, agriculture, military, etc.

A number of data points, obtained by sampling or other acquisition techniques, may represent the values of a function of the independent variable. It is often required to estimate the new values of that function based on the independent variable around the data points. Interpolation is a method of constructing new data points within a specified range other than a known data point. Spatial interpolation methods estimate variables in unobserved locations in geographical space based on known values in the observed locations. In general, there are two types of spatial interpolation which are geo-statistical and deterministic interpolation. Geo-statistical interpolation creates a surface by predicting new values among known values using a statistical approach. Deterministic interpolation creates a surface based on some observed points without taking into account any spatial processes that occur within it [1]. There are many spatial interpolation methods that have been studied, used, even compared and improved in their performance, such as IDW (Inverse Distance Weighting), Average IDW, Modified IDW, NN (Nearest Neighbor), Spline, Thin Plate Spline, Kriging, Triangular Prism, Linear Regression, MDW (Minkowsky Distance Weighting), MARS (Multivariate Adaptive Regression Spline), etc [3-8].

A digital map is a geographical representation of digital-based geospatial information on terrestrial coordinate systems called geodata. One character of digital geodata is the DEM (Digital Elevation Model) which is a theatrical performance of the bare earth surface, without vegetation and urban features [9]. The higher the spatial resolution of the DEM is increasing due to various application needs. Various studies have been undertaken to create DEM, one of which has been to discuss various interpolation methods in [10]. The data used to generate DEM is usually obtained from LIDAR (Light Detection And Ranging) which is a laser scanning technology that can provide very accurate distance and intensity information quickly. LIDAR data is usually a 3D image and spaced irregularly. This means that the points do not have any particular symmetry over the area of interest so may lose data. To obtain a regular grid, a grid-based interpolation method is required [11]. In this case various approaches of image interpolation can be used, such as bi-linear, bi-cubic, natural neighbor, and nearest neighbor interpolation.

Image-based spatial interpolation applies adaptive spatial image processing approach. This approach implies that operators must vary across the image, which considers the image context locally by involving aspects of geometry, morphology or radiometric. This approach requires extrinsic information of image for efficient processing and analysis. The context-dependent intrinsic operational window is one of the techniques to implement this approach which enables the development of multi-scale and adaptive image transformation spatially [12]. The operational window (can be an array or square) is an area that locally covers a set of neighboring pixels around the observation pixel. This operation is also called a pixel neighborhood operation where window as an operator is called the kernel. By sliding the kernel operator to the entire image, it will get context-dependent intrinsic information as a result of autocorrelation between a set of locally-captured pixel neighbors by the kernel [13]. By using the sliding square window will be able to obtain even more spectral-spatially integrated information. Similar to those approaches, the sliding neighborhood operation is a process performed on a pixel called a center pixel or an observation pixel. The output of this process is the new value of a center pixel determined by the neighboring relationship between the central pixel and its neighboring pixels by applying a particular algorithm [14]. This operation can be applied to various image processing purposes, such as image filtering, enhancement, segmentation, etc [15-17].

In principle, spatial interpolation builds a new point in a new location based on the spatial relationship between the location of the value to be estimated and the location of the known points around it. For a large number of data points, the relationship between these data can be approximated by the clustering concept. The purpose of clustering is to group a set of data into the clusters based on the relationship between the data and their cluster center. One of the most common types of relationships is the distance between the data and their cluster center. There are two types of grouping algorithms: hierarchical grouping and partitioning. Hierarchical clustering algorithms have been widely used in various studies [18-20]. There are two types of partitional clustering: soft and hard clustering. In soft clustering, each pair of data in a set of train data has a certain degree of membership toward each cluster expressed in the interval {0...1}. One type of soft clustering algorithm is FCM (Fuzzy C-Mean) which has been used widely in various studies; include its combinations with other algorithms [21-26].

K-Mean Clustering is the most primitive, simple and effective method in which the relationship between the dataset and the cluster center is the shortest distance. This algorithm is classified as hard clustering, where each pair of data in a set of train data occupies only one cluster. The membership of cluster denoted by 1, and vice versa by 0. This algorithm has been used in various studies either independently or in combination with various other algorithms and its various improved performance [2-5].

This study applies the neighboring pixels concept for the spatial interpolation through a sliding neighborhood operation. This operation uses the IDW concept to determine the new value of center pixels based on its neighboring pixels. To improve the result, firstly K-Mean clustering used to create new points based on the cluster center. Again, the value of the cluster center also determined by using the IDW concept based on all cluster members. The aim of this survey is to optimize the spatial interpolation based on the sliding neighborhood operation method by using K-Mean clustering for predicting the topographic pattern of the land surface.

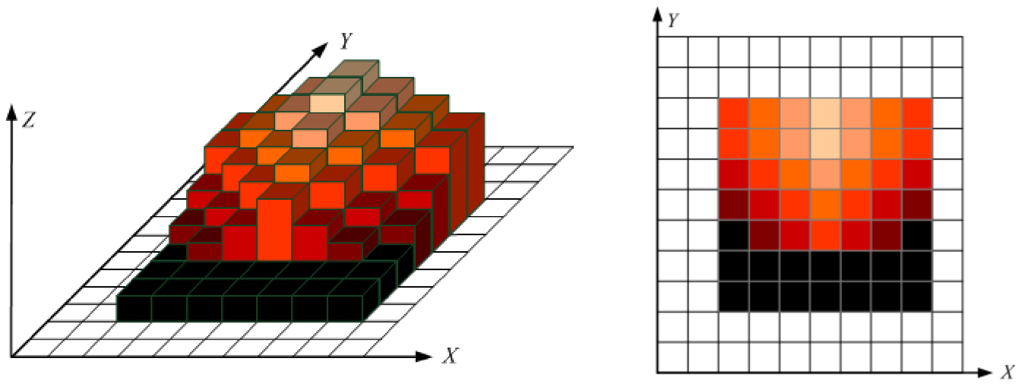
2 Materials and Method

This part reports the observed data samples represented in the regular data grid, spatial interpolation using the sliding neighborhood operation, K-Mean clustering, IDW concept, and their use to predict the topographic pattern of the land surface.

2.1 The map data representation

First, the observed sample data is represented into a regular data grid is apply the concept of image-based interpolation. In this case the map data represented by a 2-D matrix. Each pair of row-column elements represents the position of the tip along the rectangular pattern of the observed topographic area. This pattern is called raster data. The data grid is a raster data used to represent the earth's surface. DEM is one type of map using raster geodata.

By using the concept of a regular data grid, the geographical position of a location represented in a geographic coordinate system. Each position of a spot is represented by (X, Y), whereas Z is a value at that location. In a rectangular grid , an example of a geodata representation is shown in Fig. 1.



(a). 3-D image (X, Y,Z) (b). 2-D image (X, Y)

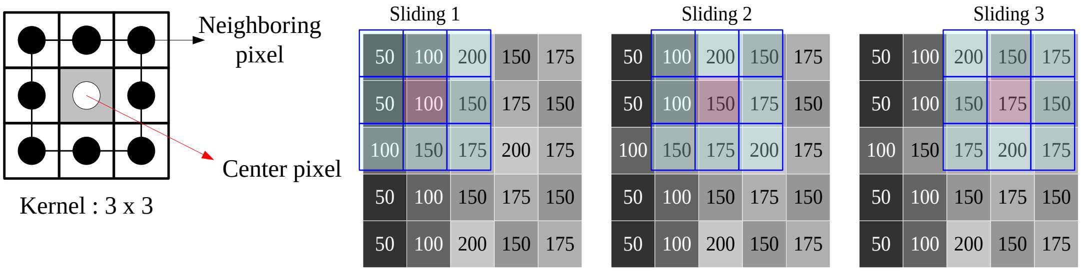
Fig. 1. An example of rectangular grid

Each color representation indicates a value of a point. Several varieties of color map are widely applied to interpret a certain value, such as surface height, elevation, mineral content, the density of occupancy in urban areas, distribution of land prices, etc.

2.2 Sliding neighborhood operation

Sliding neighborhood operation is one of the techniques employed for image filtering. Each pixel of the output image is the result of using a special algorithm to the center pixel of the input image based on its neighboring pixels [31]. The neighboring pixels are a set of pixels around a center pixel defined by a certain kernel. A kernel is an operator of a window that can be a rectangular or square shape. Center pixel is in the center of the window where all pixels other than center pixel are inside the window as neighboring pixels. A 3 x 3 kernel has 8 neighbors' pixels as shown in Fig. 2 (a) whereas a sliding neighborhood operation is illustrated in Fig. 2 (b).

Each sliding operation will generate a new value of the center pixel depending on the algorithm used. The sliding operation is applied to all pixels of the image to produce a new image. This concept can be used to perform image-based spatial interpolation whereby the determination of new values of center pixels uses spatial interpolation algorithms.



(a). A 3 x 3 kernel (b). sliding neighborhood operation

Fig. 2. An illustration of the sliding neighborhood operation

2.3 Spatial interpolation using the IDW concept

The IDW method is a deterministic interpolation method that estimates the value of the observation point by using a linear combination of known sample points around it. This linear combination is the inverse distance weight between the observation point and each sample point [6]. A sample point closer to the observation point would have a heavier weight. The distance function used is the Euclidean distance with the interpolator expressed by:

$$Z_p = \sum_{i=1}^{n} w_i . Z_i \quad A_i = \frac{1}{\sqrt{(X_i - X_p)^2 + (Y_i - Y_p)^2}} \quad w_i = \frac{A_i}{\sum_{i=1}^{n} A_i} \tag{1}$$

Where (X_p, Y_p) and (X_i, Y_i) are the location of the observation point, and the location of the ith sampled point, respectively, Z_p and Z_i are the value to be estimated, and the value of the ith sampled point, respectively.

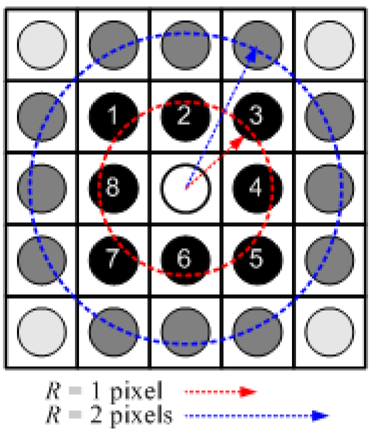


Fig. 3. The application of the IDW concept in image-based interpolation

Using the IDW concept, the estimated value depends on the distance of the location between the observation point and the entire sample point. If applied to a spatial interpolation by a radial distance implementation, there will be different estimates depending on their distance. This is also called spatial variability of the estimated value of the observation point. The application of this concept in image-based interpolation is illustrated in Fig. 3. It shows that for *R* = 1 pixel there will be 8 neighboring pixels as the sample points, whereas for *R* = 2 pixels there will be 12 pixels of the neighborhood. In this study is using the 3x3 kernel to obtain full neighbor pixels.

2.4 K-Mean clustering

K-Mean clustering classified as hard clustering. It is an iterative process that partitions a set of data into a number of *K* clusters. The initial center clusters generated in a certain way are used as the starting point of the partitioning process. The data partitioning process is based on the shortest distance between each data and the cluster center. The partition process is done continuously until the cluster membership does not change anymore. *K*-Mean clustering can use a variety of distance functions, though usually using Euclidean distance.

Suppose a number of *K* initial cluster center denoted by $C(c_1, c_2, \dots, c_K)$ where c_k denoted by $c_k\left(c_k(x), c_k(y)\right)$, and a number of *N* data in a dataset denoted by $D(d_1, d_2, \dots, d_N)$ where d_N denoted by $d_n\left(d_n(x), d_n(y)\right)$. The variables *x* and *y* are the data attributes. The distance between each data and cluster center expressed by:

$$R_{(k,n)} = \sqrt{\left(c_k(x) - d_n(x)\right)^2 + \left(c_k(y) - d_n(y)\right)^2} \quad k = 1 \dots K \quad n = 1 \dots N \qquad (2)$$

The variable *R* is the $K \times N$-size distance matrix.

Suppose *G* is a cluster membership matrix that has a size equal to *R*. Each column of *R* is evaluated. The row position of the smallest value represents the corresponding cluster membership. In the same row position of *G* matrix is denoted by 1, and vice versa denoted by 0.

$$c_k(x) = \frac{1}{m}\sum_{i=1}^{N} d_i(x)_{G(k,i)=1} \quad c_k(y) = \frac{1}{m}\sum_{i=1}^{N} d_i(y)_{G(k,i)=1} \quad k = 1 \dots K \qquad (3)$$

Where *m* is the number of cluster member of each cluster center.

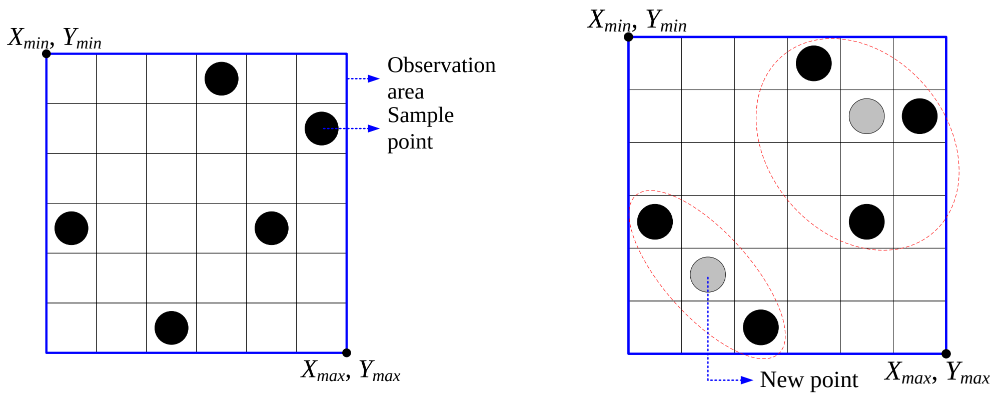
In this study, firstly *K*-Mean clustering used to create new points based on the cluster center. The value of each cluster center is calculated using Eq. (1) where the cluster center as an observation point and all its members as the sample points.

2.5 The proposed method

In this work, the combined K-Mean clustering and sliding neighborhood operations will be utilized to predict topographic shapes of the ground surface based on a set of sample data points. The results will be compared to predicted results of the sliding neighborhood operation to demonstrate its performance. The steps used are as follows:

1) Build sample data points from the selected observation area.
2) Create the boundary of the observation area as a data grid (X, Y) in a Cartesian coordinate system based on the sampled data points obtained.
3) Map the sample data points into the data grid.
4) Apply the *K*-Mean clustering by a number of *K* cluster to generate new points iteratively.
5) Apply the sliding neighborhood operation iteratively using a 3x3 kernel to generate new points until all empty locations within the boundary of observation area have a value.

The illustrations of step 2 - 5 are shown in Fig. 4 and 5.



(a). The illustration of step 2 and 3 (b). The illustration of step 4

Fig. 4. The illustration of step 2 – 4

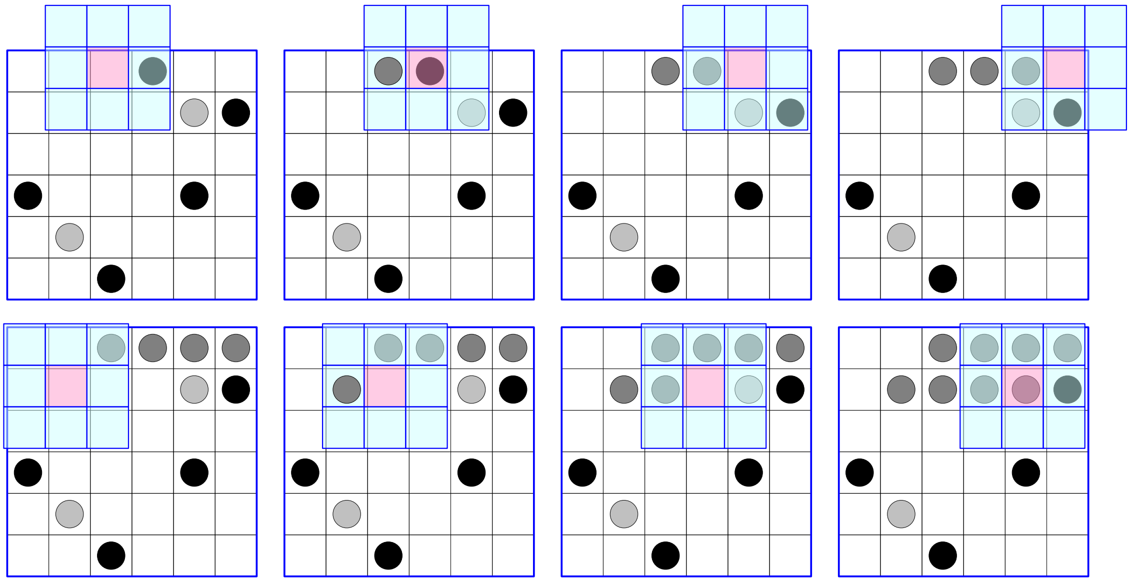


Fig. 5. The illustration of step 5

The number of empty locations to be estimated depends on the size of the observation area boundary and the number of sample data points. Suppose $P(X,Y)$ is the actual geographic position, and N is the number of sample data point. If the boundary of the observation area in the Cartesian coordinate system has $M \times M$ -size then the number of empty locations to be estimated are $(M \times M) - N$. Mapping of sample data points within the boundary of the observation area is done using the following formula:

$$X_n(i) = 1 + (M-1) * \frac{X(i) - X_{min}}{X_{max} - X_{min}}$$
$$Y_n(i) = 1 + (M-1) * \frac{Y(i) - Y_{min}}{Y_{max} - Y_{min}} \quad (4)$$

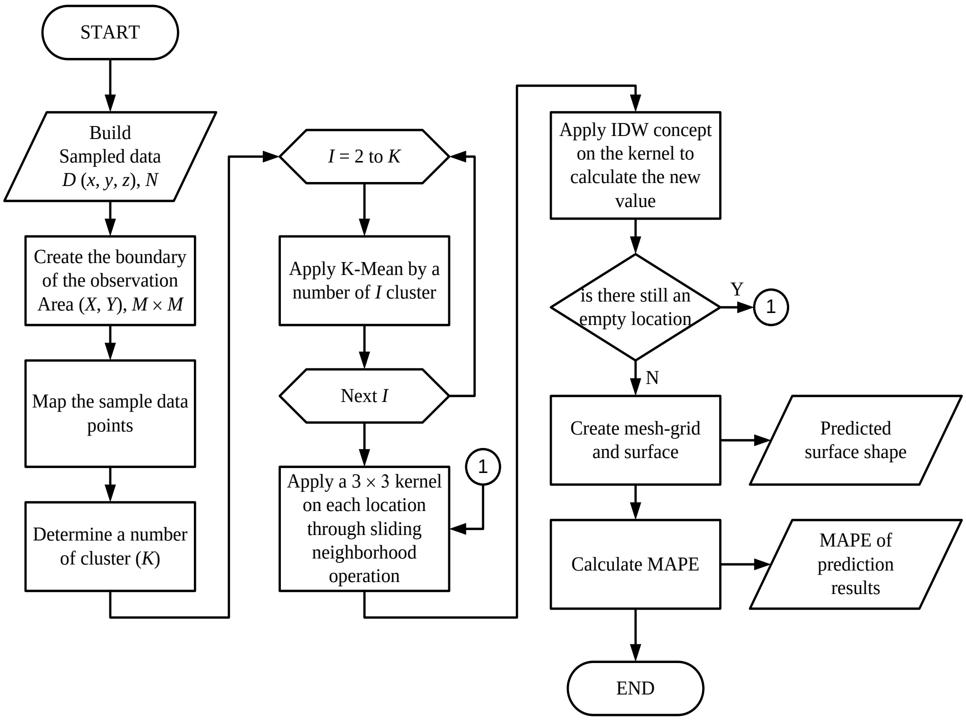


Fig. 6. The proposed method algorithm

MAPE (Mean Absolute Percentage Error) used for the performance of the predicted result expressed by:

$$APE(i) = \frac{|Z_{actual}(i) - Z_{predicted}(i)|}{Z_{actual}(i)} \times 100$$
$$MAPE = \frac{1}{N}\sum_{i=1}^{N} APE(i) \quad (5)$$

The algorithm of the proposed method is shown in Fig. 6.

2.6 Datasets

In this work, the sample data points are the dry land surface elevations obtained from the survey results in selected areas of Samarinda, East Kalimantan, Indonesia. The area of Samarinda is in a geographic location as follows:

Latitude : $0.3123° - 0.7071°S$
Longitude : $117.0471° - 117.3039°E$

The Universal Transverse Mercator System projected into the X-Y Coordinate System for Samarinda's geographic position lies in the 50M zone as follows:

Easting : $5.0524 \times 10^5 - 5.3382 \times 10^5$

Northing: $9.9219 \times 10^5 - 9.9655 \times 10^5$

Observations were conducted at 50 different locations with the results as shown in Table 1 (Step 1). The boundary of the observation area in the Cartesian coordinate system is set to the size of 150×150. The proposed method done by using programs created using MATLAB tool software.

Table 1. Sample Data Observation Results Obtained From 50 Different Locations In Samarinda – East Kalimantan, Indonesia

No.	X	Y	Elevation	No.	X	Y	Elevation
1	517968	9937230	3.437255	26	518753	9942857	23.1344
2	518009	9937561	3.989665	27	518756	9942670	31.55961
3	518073	9947980	3.563305	28	518812	9940775	4.531952
4	518121	9956987	15.45811	29	518821	9941025	2.235591
5	518139	9942725	25.33522	30	518893	9940727	4.030788
6	518169	9942785	26.32243	31	518893	9943217	13.66156
7	518199	9941615	11.31903	32	518911	9953042	8.360943
8	518210	9942849	18.63606	33	518963	9947314	7.565265
9	518228	9947691	5.173994	34	519032	9936442	4.000000
10	518271	9941540	7.702066	35	519172	9955909	8.079804
11	518285	9937410	3.000000	36	519219	9957262	12.89894
12	518305	9937433	3.185289	37	519240	9957333	13.34059
13	518312	9937390	3.276924	38	519255	9943988	29.49929
14	518362	9952889	5.263197	39	519383	9949452	15.55828
15	518378	9941502	4.943363	40	519404	9943019	17.41111
16	518408	9954631	7.416409	41	519447	9944206	42.10487
17	518485	9941849	14.90583	42	519447	9957654	18.02503
18	518567	9943469	16.03800	43	519494	9943278	19.45698
19	518568	9955738	11.83472	44	519520	9943126	20.19802
20	518569	9956053	13.20923	45	519522	9943140	20.77379
21	518585	9953283	13.19609	46	519528	9943212	16.59531
22	518591	9953294	13.24302	47	519549	9943171	16.41016
23	518640	9956007	9.820753	48	519568	9942447	11.63611
24	518690	9941947	6.514025	49	519641	9946292	17.44017
25	518723	9946735	10.32624	50	519670	9951080	14.82371
				Min	517968	9936442	
				Max	519670	9957654	

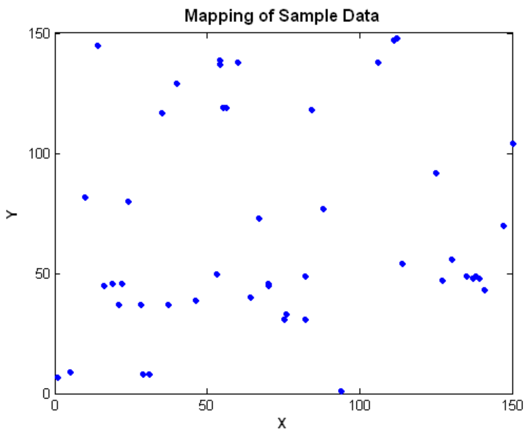
In this section, author should provide the latest related work of the subject matter and critical analyze them. Substantial literatures are expected in this section to ensure the novelty of the proposed work.

3 Results and Discussion

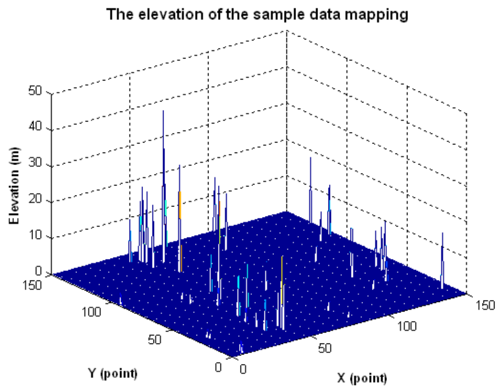
The number of points to be predicted is 150×150=22500 points, using only 50 points as seeds. A minor act of sample data details are intentionally applied to test one of the powers of the proposed method to achieve certain prediction performance. The sample data mapping is performed by using Eq. (4) With the results as presented in Table 2 and Fig. 7 (Step 2 and 3). The minimum number of clusters was 2. By determining the number of clusters K = 3 then we have obtained 2 + 3 = 5 new points as shown in Fig. 8 (Step 4).

Table 2. Normalizaed Sample Data Observation Results

No.	X	Y	Elevation	No.	X	Y	Elevation	No.	X	Y	Elevation
1	1	7	3.437255	19	54	137	11.83472	37	112	148	13.34059
2	5	9	3.989665	20	54	139	13.20923	38	114	54	29.49929
3	10	82	3.563305	21	55	119	13.19609	39	125	92	15.55828
4	14	145	15.45811	22	56	119	13.24302	40	127	47	17.41111
5	16	45	25.33522	23	60	138	9.820753	41	130	56	42.10487
6	19	46	26.32243	24	64	40	6.514025	42	130	150	18.02503
7	21	37	11.31903	25	67	73	10.32624	43	135	49	19.45698
8	22	46	18.63606	26	70	46	23.1344	44	137	48	20.19802
9	24	80	5.173994	27	70	45	31.55961	45	137	48	20.77379
10	28	37	7.702066	28	75	31	4.531952	46	138	49	16.59531
11	29	8	3.000000	29	76	33	2.235591	47	139	48	16.41016
12	31	8	3.185289	30	82	31	4.030788	48	141	43	11.63611
13	31	8	3.276924	31	82	49	13.66156	49	147	70	17.44017
14	35	117	5.263197	32	84	118	8.360943	50	150	104	14.82371
15	37	37	4.943363	33	88	77	7.565265				
16	40	129	7.416409	34	94	1	4.000000				
17	46	39	14.90583	35	106	138	8.079804				
18	53	50	16.03800	36	111	147	12.89894				



(a). The mapping results visualization in the *X-Y* area



(b). The mapping results visualization in the *X-Y-Z* area

Fig. 7. Visualization of sample data mapping

By applying step 5, we have obtained 22500 predicted points that visualized by the mesh-grid and the results of the ground surface shape prediction as shown in Fig. 9. MAPE calculated from 50 predicted points by its actual values. To test the performance of the proposed method, we have used a various number of clusters where the MAPE of predicted results were shown in Table 3. It was showed that the best predicted result obtained by K=4 with MAPE of 9.27%.

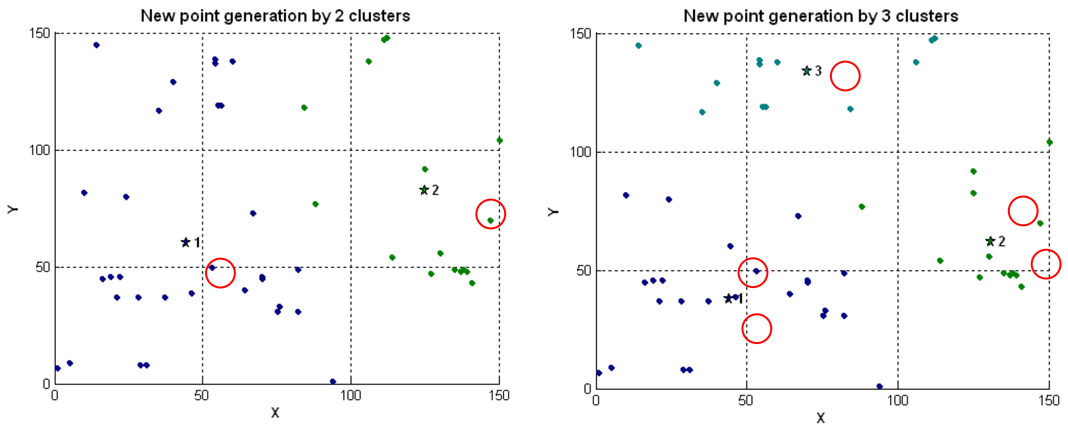


Fig. 8. The new points generated by K-Mean clustering with $K = 3$

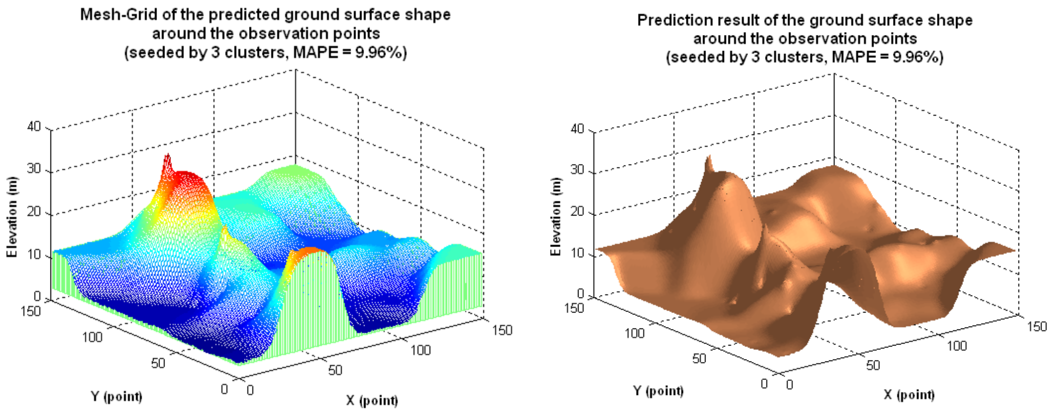
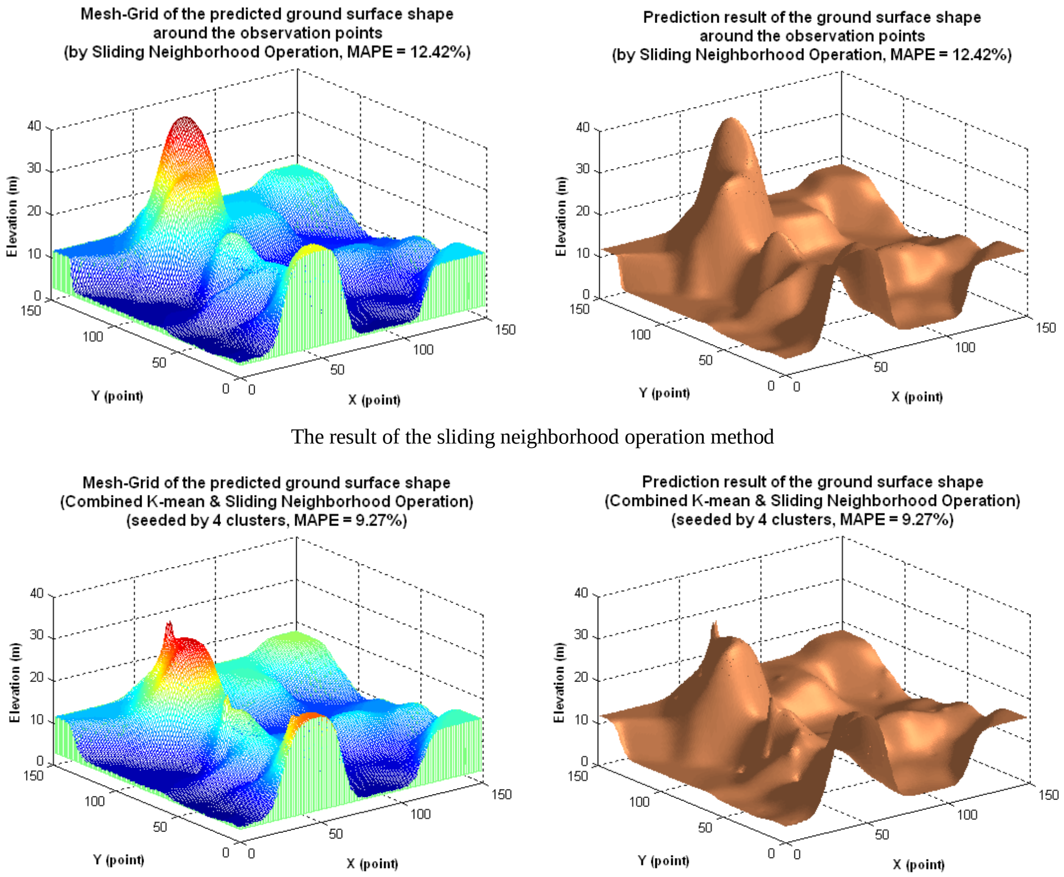


Fig. 9. The Mesh-Grid Visualization And The Result Of The Ground Surface Shape Prediction

Table 3. Prediction Results By The Various Number Of Clusters

No.	No. cluster	New point generation by cluster	No. Iteration of sliding neighborhood operation	MAPE
1	2	2	10	9.90%
2	3	5	11	9.96%
3	4	9	11	9.27%
4	5	14	11	10.21%
5	6	20	11	10.21%
6	7	27	12	12.13%
7	8	35	13	12.73%
8	9	44	13	13.09%
9	10	54	14	13.54%
10	11	65	14	13.45%

If without *K*-Mean clustering, only the sample data points act as seeds to generate new points iteratively by using Sliding Neighborhood Operations. By applying step 5, we have obtained the predicted results with MAPE of 12.42%. Fig. 10 showed the comparison of predicted results between the sliding neighborhood operation method and the proposed method. These results showed that by applying *K*-Mean clustering to spatial interpolation based on the sliding neighborhood operation, there was a performance improvement from success rate of 87.58% (MAPE of 12.42%) to 90.73% (MAPE of 9.27%). By applying the *K*-Mean clustering on the sample data points, the cluster centers obtained were the new points which also acts as the seeds. Those seeds addition have improved the predictive performance. However, this did not apply to all *K* cluster as shown in Table 3. Predictive performance improvements have occurred at *K* = 2 to 7 where the best predictive result was at *K* = 4.

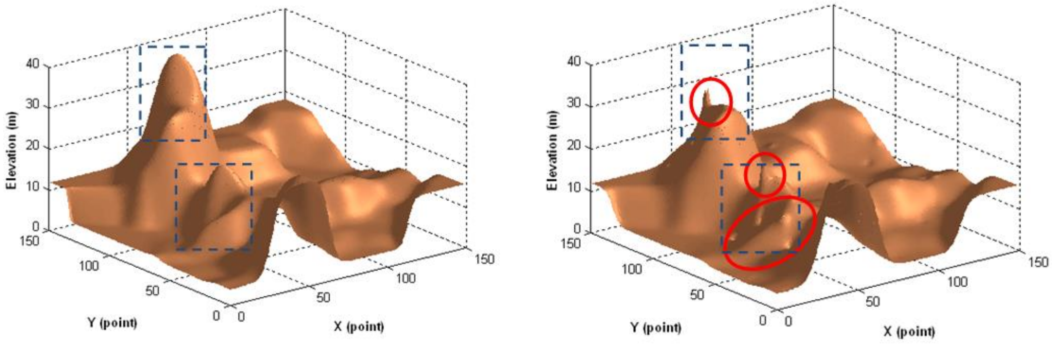


The result of the sliding neighborhood operation method

The result of the proposed method

Fig. 10. The comparison of predicted results between the sliding neighborhood operation method and the proposed method

Compared to the predicted results by Sliding Neighborhood Operations without K-Mean, the predicted results with K-Mean have several rough spots as shown in Fig. 11. The rough points that appear indicated by the surface area bounded by the rectangular box as a result of the optimization of the predicted results. It was possible because the new points generated by the K-Mean clustering give a small distortion during the process of sliding neighborhood operations which carried out iteratively. For further study, smoothing some rough spots is one of the goals of improvement.



(a). without *K*-Mean (b). with *K*-Mean

Fig. 11. The several rough spots caused by new points generated by the *K*-Mean clustering

4 Conclusion

This subject area has applied image-based spatial interpolation by using the Sliding Neighborhood Operation to predict the topographic pattern of the dry land surface based on a set of sample data points. The number of points to be predicted is $150 \times 150 = 22500$ points, using only 50 points as seeds. A minor act of sample data points were deliberately employed to test one of the powers of the proposed method to achieve certain prediction performance. MAPE has used to measure the performance of predicted results that calculated from 50 predicted points by its actual values. If without *K*-Mean clustering, only the sample data points act as seeds to generate new points iteratively using the Sliding Neighborhood Operations. The prediction results that have obtained have a MAPE of 12.42%. By applying the *K*-Mean clustering on the sample data points, the cluster centers obtained were the new points which also acts as the seeds. The prediction results that have obtained have a MAPE of 9.27% at *K*=4. Those seeds addition have improved the predictive performance from the success rate of 87.58% (MAPE of 12.42%) to 90.73% (MAPE of 9.27%).

For further study, the same activity will be carried out by applying the Sliding Neighborhood Operation on selected clusters by *K*-Mean clustering for seeding the set of the new points firstly. It is necessary to smooth out some rough spots that might appear without reducing the performance of the predicted results.

References

[1] A. F. O. Gaffar, R. Malani, A. B. W. Putra, Mulyanto, and Ibayasid, "Prediction of The Topographic Shape of The Ground Surface Using IDW Method through The Rectangular-Neighborhood Approach," in *2018 International Symposium on Advanced Intelligent Informatics (SAIN)*, 2018, pp. 107-111.

[2] M. Curtarelli, J. Leão, I. Ogashawara, J. Lorenzzetti, and J. Stech, "Assessment of Spatial Interpolation Methods to Map the Bathymetry of an Amazonian Hydroelectric Reservoir to Aid in Decision Making for Water Management," *ISPRS International Journal of Geo-Information*, vol. 4, pp. 220-235, (2015).

[3] A. S. Boke, "Comparative Evaluation of Spatial Interpolation Methods for Estimation of Missing Meteorological Variables over Ethiopia," *Journal of Water Resource and Protection*, vol. 09, pp. 945-959, (2017).

[4] Bustani, S. Pradana, Mulyanto, and Nurjanana, "Prediction of electricity sales using neural based inverse distance weighting method," *International Journal of Engineering & Technology (IJET)*, vol. 7, pp. 65-69, (2018).

[5] M. E. Hiswati, A. F. O. Gaffar, Rihartanto, and Haviluddin, "Minimum wage prediction based on K-Mean clustering using neural based optimized Minkowski Distance Weighting," *International Journal of Engineering & Technology (IJET)*, vol. 7, pp. 90-93, (2018).

[6] M. N. Ikechukwu, E. Ebinne, U. Idorenyin, and N. I. Raphael, "Accuracy Assessment and Comparative Analysis of IDW, Spline and Kriging in Spatial Interpolation of Landform (Topography): An Experimental Study," *Journal of Geographic Information System*, vol. 09, pp. 354-371, (2017).

[7] E. Oktavia, Widyawan, and I. W. Mustika, "Inverse Distance Weighting and Kriging Spatial Interpolation for Data Center Thermal Monitoring," presented at the 2016 1st International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, Indonesia, (2016).

[8] M. Wang, G. He, Z. Zhang, G. Wang, Z. Zhang, X. Cao, Z. Wu, and X. Liu, "Comparison of Spatial Interpolation and Regression Analysis Models for an Estimation of Monthly Near Surface Air Temperature in China," *Remote Sensing*, vol. 9, p. 1278, (2017).

[9] Z. Ismail, M. F. Abdul Khanan, F. Z. Omar, M. Z. Abdul Rahman, and M. R. Mohd Salleh, "Evaluating Error of Lidar Derived Dem Interpolation for Vegetation Area," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-4/W1, pp. 141-150, (2016).

[10] B. Szypuła, "Geomorphometric comparison of DEMs built by different interpolation methods," *Landform Analysis*, vol. 32, pp. 45-58, (2017).

[11] I. Ashraf, S. Hur, and Y. Park, "An Investigation of Interpolation Techniques to Generate 2D Intensity Image From LIDAR Data," *IEEEAccess,* vol. 5, pp. 1-11, (2017).

[12] J. Debayle and J.-C. Pinoli, "General Adaptive Neighborhood Image Processing for Biomedical Applications," in *Applied Biomedical Engineering*, G. Gargiulo, Ed., ed: InTech, (2011), p. 500.

[13] L. Zhao, W. Lin, Y. Wang, and X. Li, "Recursive Local Summation of RX Detection for Hyperspectral Image Using Sliding Windows," *Remote Sensing,* vol. 10, p. 103, (2018).

[14] D. Han, "Particle Image Segmentation Based on Bhattacharyya Distance," Degree Master of Science, Arizona State University, (2015).

[15] S. N. S. I. Isa, M. F. Abdullah, N. M. Tahir, M. Mustapha, and N. K. A. Karim, "New Image Enhancement Technique for WMH Segmentation of MRI FLAIR Image," *©2016 IEEE,* p. 34, (2016).

[16] A. Jagadeesan, T.Thillaikkarasi, and K.Duraiswamy, "Cryptographic Key Generation from Multiple Biometric Modalities: Fusing Minutiae with Iris Feature," *International Journal of Computer Applications,* vol. 2, pp. 1-11, (2010).

[17] R. Narayanam, "Translation Invariance (TI) based Novel Approach for better De-noising of Digital Images," *International Research Journal of Engineering and Technology (IRJET)* vol. 4, pp. 1-15, (2017).

[18] S. Kumar and D. Toshniwal, "Analysis of hourly road accident counts using hierarchical clustering and cophenetic correlation coefficient (CPCC)," *Journal of Big Data,* vol. 3, (2016).

[19] A. Salighehdar, Y. Liu, D. Bozdog, and a. I. Florescu, "Cluster Analysis of Liquidity Measures in a Stock Market using High Frequency Data," *Journal of Management Science and Business Intelligence,* pp. 1-8, (2017).

[20] A. Sharma, Y. López, and T. Tsunoda, "Divisive hierarchical maximum likelihood clustering," *BMC Bioinformatics,* vol. 18, (2017).

[21] C. S. Chin, X. Ji, W. L. Woo, T. J. Kwee, and W. Yang, "Modified multiple generalized regression neural network models using fuzzy C-means with principal component analysis for noise prediction of offshore platform," *Neural Computing and Applications,* (2017).

[22] A. Gupta and D. Kumar, "Fuzzy clustering-based feature extraction method for mental task classification," *Brain Informatics,* vol. 4, pp. 135-145, (2016).

[23] A. Khan, D. Katanic, and J. Thakar, "Meta-analysis of cell-specific transcriptomic data using fuzzy c-means clustering discovers versatile viral responsive genes," *BMC Bioinformatics,* vol. 18, (2017).

[24] H.-Y. Li, W.-J. Hwang, and C.-Y. Chang, "Efficient Fuzzy C-Means Architecture for Image Segmentation," *Sensors,* vol. 11, pp. 6697-6718, (2011).

[25] A. Saha and S. Das, "Feature-weighted clustering with inner product induced norm based dissimilarity measures: an optimization perspective," *Machine Learning,* vol. 106, pp. 951-992, (2017).

[26] J. Yang, Y.-s. Ke, and M.-z. Wang, "An adaptive clustering segmentation algorithm based on FCM," *Turkish Journal of Electrical Engineering & Computer Sciences,* vol. 25, pp. 4533-4544, (2017).

[27] P. Novianti, D. Setyorini, and U. Rafflesia, "K-Means cluster analysis in earthquake epicenter clustering," *International Journal of Advances in Intelligent Informatics,* vol. 3, pp. 81-89, (2017).

[28] D. Olszewski, "Asymmetric K -Means Clustering of the Asymmetric Self-Organizing Map," *Neural Processing Letters,* vol. 43, pp. 231-253, (2015).

[29] J. Wu, "The Uniform Effect of K-means Clustering," *Springer Theses,* pp. 17-35, (2012).

[30] X. Zhou, J. Gu, S. Shen, H. Ma, F. Miao, H. Zhang, and H. Gong, "An Automatic K-Means Clustering Algorithm of GPS Data Combining a Novel Niche Genetic Algorithm with Noise and Density," *ISPRS International Journal of Geo-Information,* vol. 6, p. 392, (2017).

[31] M. T. H. Mark Hudson Beale, Howard B. Demuth, *Neural Network Toolbox™*. 3 Apple Hill Drive-Natick, MA 01760-2098: The MathWorks, Inc., (2015).