

Introducing a Novel Digital Elevation Model using an Artificial Neural Network Algorithm

Behzadi, S.^{1*} and Jalilzadeh, A.²

Abstract: Elevation is the main information of the earth, and different models are provided for a better understanding of the earth. To have a digital elevation model (DEM), the height of the area must be gathered. However, it is not always possible to conduct a comprehensive survey in the area and calculate the whole surface. The best way is surveying some points, then estimating the elevation using these points. The purpose of this paper is to use interpolation methods to estimate elevation. In this paper, the three usual methods are chosen and introduced, then their performances are compared. These methods include inverse distance weighting (IDW), the Kriging method, and artificial neural network (ANN). The results show that the ANN models the elevation better than the two other methods with root mean square error (RMSE) equals to 5.9 m.

Keywords: Digital elevation model; inverse distance weighting; kriging; artificial neural net.

Introduction

Elevation is necessary data in research involving studies of the earth's land surface [1]. This data is a fundamental input parameter for determining slope [2], agricultural applications [3], modeling for telecommunications [4], process engineering problems [5, 6]. For perceiving the earth surface, geoscientists use different methods to model ground elevation [7]. For example, digital elevation model (DEM) is one of the most used models that illustrate the earth surface. This elevation model is usually made up of equal-sized grid cells, which represent an elevation value. Since DEMs have a variety of applications, the quality of DEMs is essential [8].

The ground survey technique is one of the most used techniques to obtain elevation data. The ground survey gathers the location of the points (latitude, longitude, and elevation) with high accuracy [9]. However, it is hard work to conduct a comprehensive ground survey in areas of interest. Therefore, a requirement arises for a practical elevation estimation method. Different methods exist for this purpose such as multiple linear regression, nearest neighbor, inverse distance to power, minimum curvature, Kriging, inverse distance weighting (IDW), artificial neural network (ANN) [10, 11].

¹Assistant Professor in Surveying Engineering Department, Faculty of Civil Engineering, Shahid Rajaei Teacher Training University, Tehran, IRAN

²MSc Student in Geographic Information Systems, Faculty of Civil Engineering, Shahid Rajaei Teacher Training University, Tehran, IRAN

*Corresponding author; Email: behzadi.saeed@gmail.com

Note: Discussion is expected before November, 1st 2020, and will be published in the "Civil Engineering Dimension", volume 23, number 1, March 2021.

Received 11 November 2019; revised 21 April 2020; accepted 04 September 2020

In this article, ANN, Kriging, and IDW methods are used to model earth surface, then these methods are compared. ANN is used widely by researchers to overcome the challenges in their researches. In this paper, ANN is used to generate the elevation model of the entire study area by taking some sample points. Its performance is then evaluated and compared with Kriging and IDW methods.

Study Area

The study area is located in the central part of Iran's plateau in Yazd province. The geographical location of the study area is from 31° 53' 00"N to 31° 56' 30"N latitude and from 53° 50' 00"E to 53° 54' 30"E longitude with the dimension of about 6000x6700 m², 530 sample points are gathered for this calculation. The altitude of the study area is about 1100 meters above sea level, with a height difference of nearly 477 m. The study area with the sample points is illustrated in Figure 1.

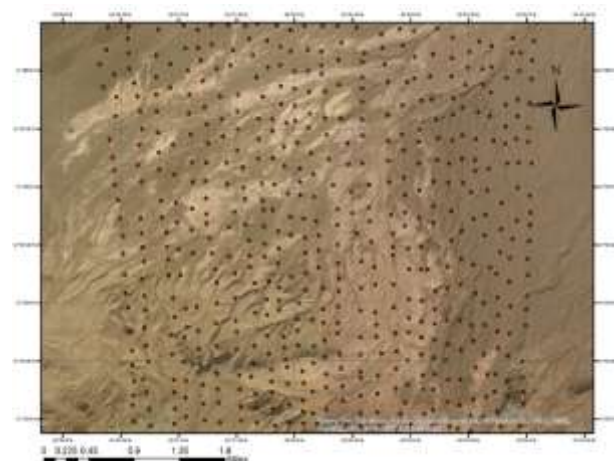


Figure 1. Study Area and Sample Points

Methods

Everything is related to the other things; however, near things are more related than far things [12]. This is the first law of Geography introduced by Waldo R. Tobler's in 1969. This law is fundamental to the analysis of geographic data. In the interpolation of elevation as geographic data, the nearest points are so important to estimate uncertain points. The techniques used for mapping surface from some sample points are called Geostatistics, which calculate values of uncertain points. A set of 530 points covering the study area is collected with known X , Y , and Z coordinates. These points are used to generate the elevation model through three different methods. At the same time, some points are also used to validate these methods. In other words, 70% of data is used to produce an elevation model, and the remaining points are used for validation and evaluating the results. These divided data sets are also applied for training and testing in the ANN method. The methodology used in this paper is illustrated in Figure 2.

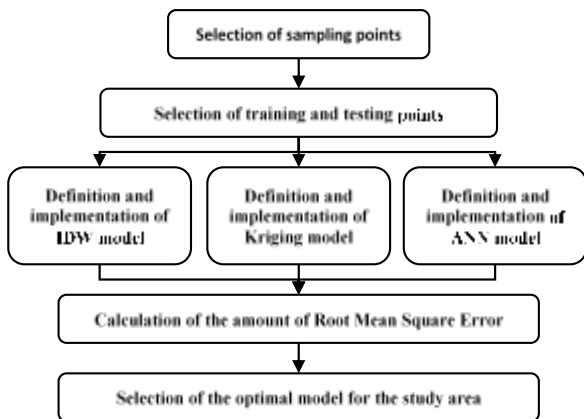


Figure 2. The Methodology

As seen in Figure 2, the methodology consists of five stages. At first, some points are gathered from the study area. These points are then divided into training and testing ones, which are going to be used in the next stage. In the third stage, three different methods are defined and implemented on gathered elevation points to generate DEM. root mean square error (RMSE) of each method is then calculated and the obtained models through each method are evaluated.

Inverse distance weighting (IDW) is the simplest interpolation method [13]. The logic behind that is approximately based on Tobler's first law, which near object to one another is more alike than those that are far away. In this method, the height of each unknown point is just obtained by measuring the average of the elevation from surrounding points with corresponding weighted values [14, 15]. The formula for this model (Equation 1).

$$x^* = \frac{\sum_{k=1}^n x_k * w_k}{\sum_{k=1}^n w_k} \quad (1)$$

$$w_k = \frac{1}{P_{kk}^j} \quad (2)$$

Where x^* is the point with an unknown value, x_k is the known point value, and w_k is the weight. The weight is calculated using Equation 2, which is the inverse distance between unknown and known points, and P is their distance.

w_k controls the effect of neighboring points on the interpolated value. If a higher power is set, it lessens the influence of far points. In this research, the common value of 2, expressed in most researches, is used for variable j in Equation 2. Finally, 12 nearest points of input data are specified for each unknown point to perform the interpolation. In Figure 3, the contours obtained through the IDW algorithm is shown.

In the next step, Kriging (Gaussian process regression as an advanced Geostatistical function) is used. This method is developed by Georges Matheron [16], The Kriging method is a powerful tool for modeling various phenomena [10] such as elevation. This method is now widely used in the spatial fields to predict values over a spatial region. The Kriging method is a computation intensive process with a local geostatic interpolation. The process starts with the exploratory statistical analysis of the data with variogram modeling to create the surface and explore the variance surface. The formula of Kriging is somewhat similar to IDW, which is based on the weight of the surrounding values, and the formula is formed as a sum of the weighted data. [17, 18].

$$x = \sum_{k=1}^n \lambda_k x'_{(s_k)} \quad (3)$$

Where, x is the prediction location, and λ_k is the weight of measured value at the k^{th} location and $x'_{(s_k)}$ is the measured value at the k^{th} location. In Figure 4, the contours obtained through the Kriging algorithm is drawn and illustrated.

The third interpolation method used in this paper is the artificial neural network (ANN). The ANN is one of the main tools in artificial intelligence and has an essential role in finding a solution in different fields of science. ANN can be trained to generate a model for the relationship between input and output data. In this paper, the neural network is trained with a feed-forward backpropagation algorithm. Feed-forward backpropagation is a simple type of ANN; the data processing moves only in the forward direction. As seen in Figure 5, the input information goes through the hidden and output layers without any loop or cycle in the network [19]. The neural network is programmed with two hidden layers, and ten neurons in each layer. Besides, 2D coordinates (X and Y) are used as training input, and the Z coordinate is also used as

training output. As shown in Figure 5, the coordinate of points (X, Y) are considered as input, and the

Elevation is set as output. The result obtained through the ANN is presented in Figure 6.

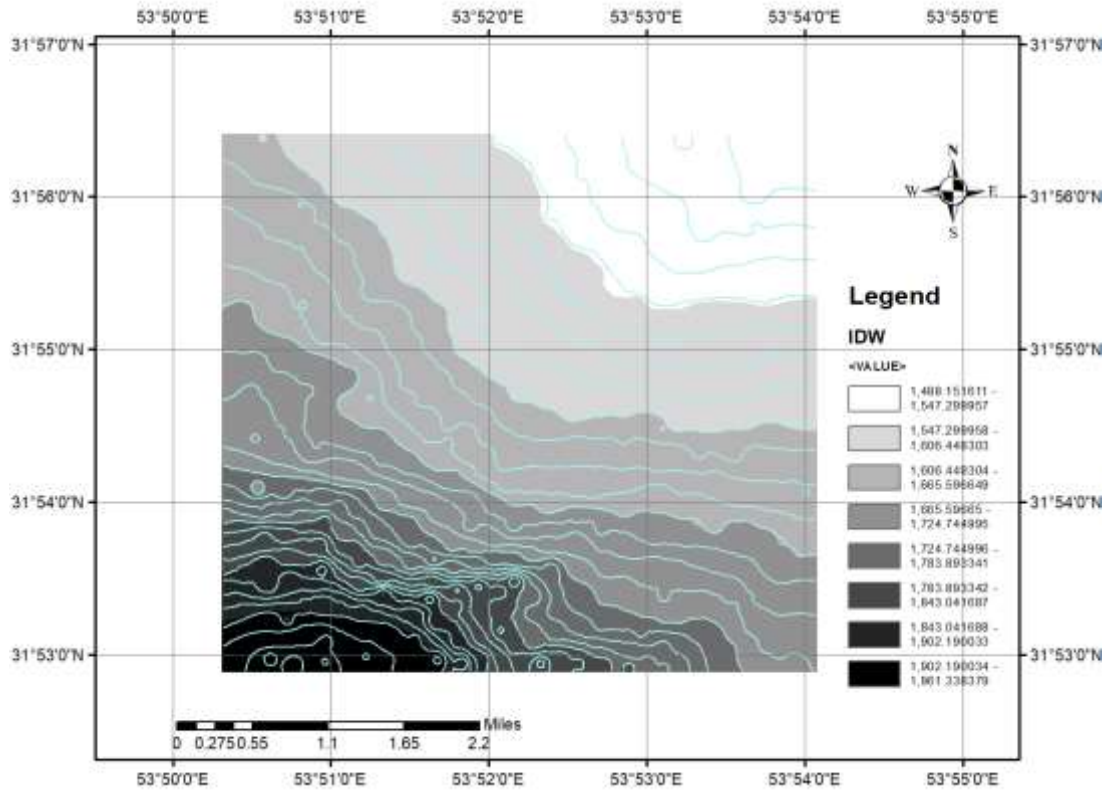


Figure 3. Contours Obtained from IDW Method

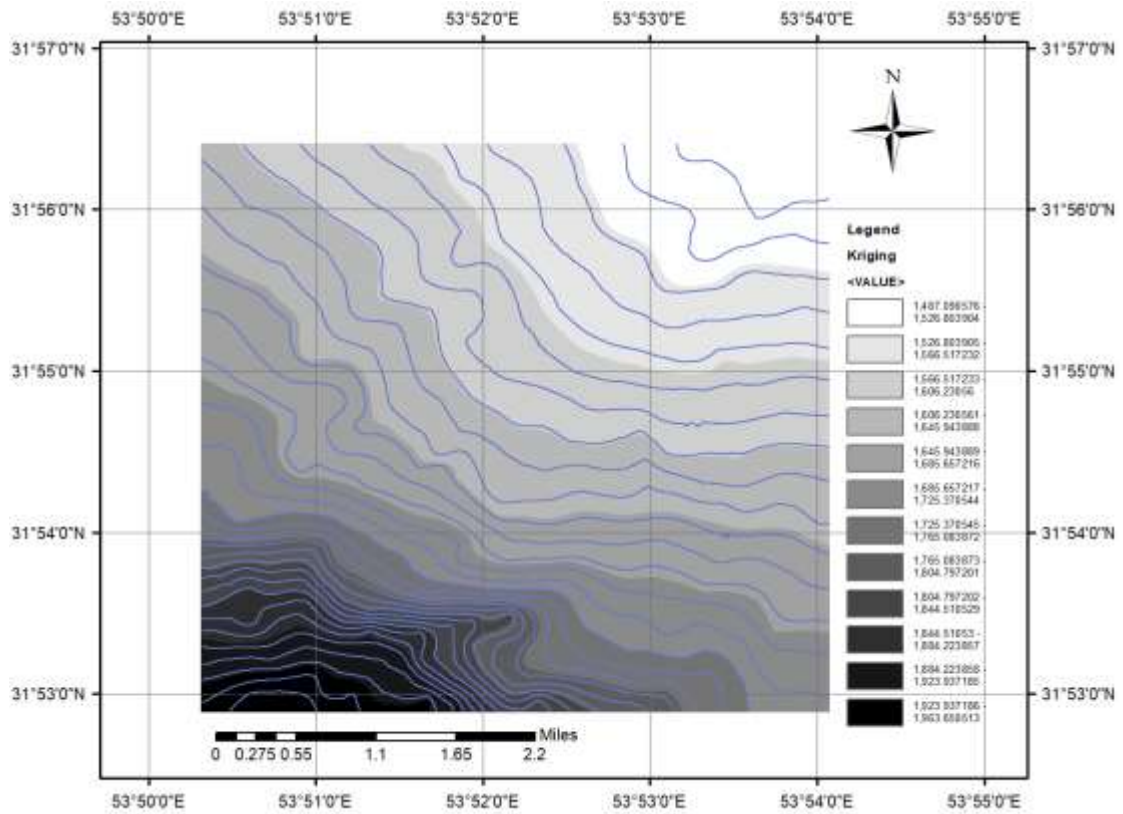


Figure 4. Contours Obtained from Kriging method

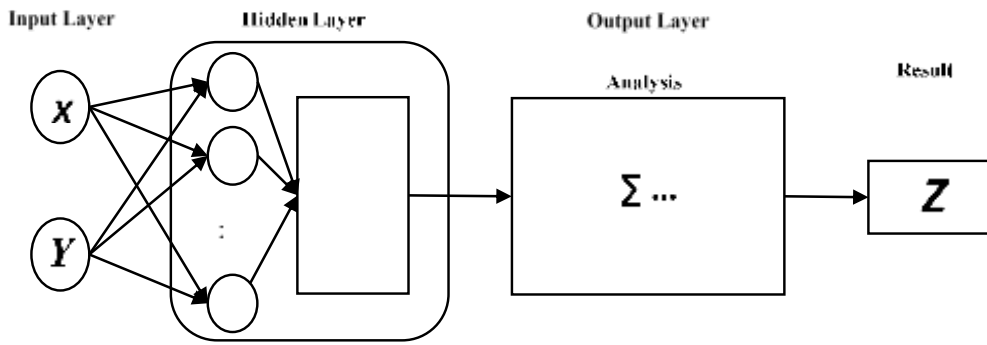


Figure 5. ANN Structure

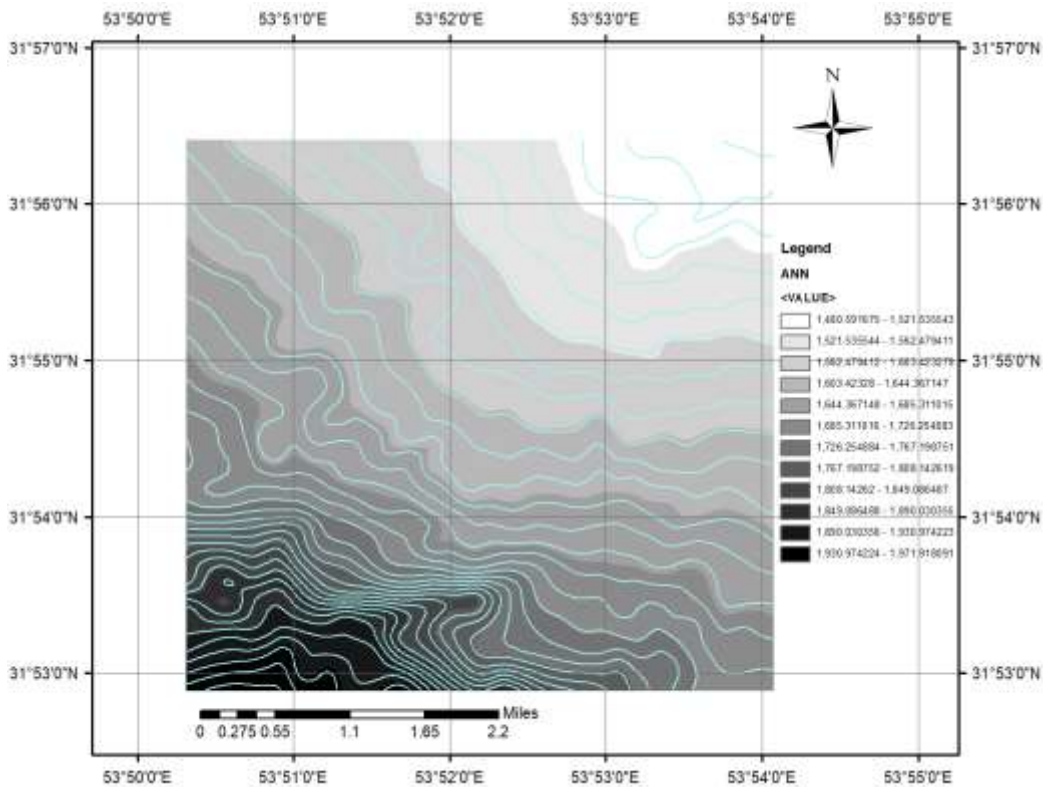


Figure 6. Contours Obtained from ANN Method

Validation

To validate each algorithm, and find out which one performs better, root mean square error (RMSE) is used as an indicator. RMSE is given as;

$$RMSE = \sqrt{\frac{1}{N-1} \sum_{k=1}^N (X - Y)^2} \quad (4)$$

Where N is the sample size and (X-Y) is the difference between the observed and obtained elevation. Finally, after implementation, the estimated values for control points are extracted and compared with real values. Table 1 shows the obtained result.

Table 1. Obtained Results for Different Algorithms.

	IDW	Kriging	ANN
RMSE	9 m	7.2 m	5.9 m

Discussion and Conclusions

Interpolation, the process of estimating the value of unmeasured points by using measured points, can be used widely in any point data set and geographic-related science such as elevation, rainfall, and so on. In this paper, three common interpolation methods were chosen and applied in the study area. As mentioned, 70 percent of the collected points were used as input data to generate DEM. Then remaining points were used for validation. So that the validation is done by comparing the elevation of test points with obtained ones through interpolation models. As seen, the obtained RMSEs in Table 1, ANN method has the best result with RMSE of 5.9 m. As expected, Kriging model has a better result than IDW with RMSE of 7.2 m. The RMSE obtained from IDW is equal to 9 m.

The differences among the results have diverse reasons. The performance of IDW was the weakest one in comparison with the other methods. The performance can be improved with changing the weight of distance and kernel function. The general type of IDW was used in this article.

Although the logic behind IDW is similar to Kriging, there are some differences between them. The main difference is that the IDW can only estimate between maximum and is not appropriate for mountainous areas, but it can perform well in terrains areas such as cliffs. On the other hand, Kriging can cause the estimated points to be lower or higher than the real amount, and the weight is based on the overall spatial arrangement of the known points. It seems this difference can have a high effect on the obtained result. ANN performed very well in interpolation. In ANN, the used algorithms and number of layers and layers can have a remarkable effect on the obtained result; however, the general setting was applied in this article. Although ANN performs are like a black box, which is not clear for decision making, it successfully deals with generating an elevation model. In this paper, the default and general values of the algorithms were used. In other words, the algorithms were compared based on their general values. Each of these algorithms depends on a series of variables. Changing these variables will improve the response of the algorithms. Finding the best value for the variables of each algorithm can be considered in future research.

References

- Hirano, A., Welch, R., and Lang, H., Mapping from ASTER Stereo Image Data: DEM Validation and Accuracy Assessment, *ISPRS Journal of Photogrammetry and Remote Sensing*, 57(5-6), 2003, pp. 356-370.
- Warren, S.D., Hohmann, M.G., Auerswald, K., and Mitasova, H., An Evaluation of Methods to determine Slope using Digital Elevation Data, *CATENA*, 58(3), 2004, pp. 215-233.
- Mokarram, M. and Hojati, M., Morphometric Analysis of Stream as One of Resources for Agricultural Lands Irrigation using High Spatial Resolution of Digital Elevation Model (DEM), *Computers and Electronics in Agriculture*, 142, 2017, pp. 190-200.
- Sawada, M., Cossette, D., Wellar, B., and Kurt, T., Analysis of the Urban/Rural Broadband Divide in Canada: Using GIS in Planning Terrestrial Wireless Deployment, *Government Information Quarterly*, 23(3), 2006, pp. 454-479.
- Behzadi, S., Mousavi, Z., and Norouzi, E., Mapping Historical Water-Supply Qanat Based on Fuzzy Method. An Application to the Isfahan Qanat (Isfahan, Iran), *International Journal of Numerical Methods in Civil Engineering*, 3(4), 2019, pp. 24-32.
- Thornton, C., Applications of DEM to Process Engineering Problems, *Engineering Computations*, 9(2), 1992, pp. 289-297.
- Toutin, T., Elevation Modelling from Satellite Visible and Infrared (VIR) Data, *International Journal of Remote Sensing*, 22(6), 2001, pp. 1097-1125.
- Erdogan, S., A Comparison of Interpolation Methods for Producing Digital Elevation Models at The Field Scale, *Earth Surface Processes and Landforms*, 34(3), 2009, pp. 366-376.
- Nelson, A., Reuter, H., and Gessler, P., DEM Production Methods and Sources, *Developments in Soil Science*, 33, 2009, pp. 65-85.
- Pavlova, A.J.O., Analysis of Elevation Interpolation Methods for Creating Digital Elevation Models, *Optoelectronics, Instrumentation and Data Processing*, 53(2), 2017, pp. 171-177.
- Jafarian, H. and Behzadi, S., Evaluation of PM_{2.5} Emissions in Tehran by Means of Remote Sensing and Regression Models, *Pollution*, 6(3), 2020, pp. 521-529.
- Tobler, W.R., A Computer Movie Simulating Urban Growth in the Detroit Region, *Economic Geography*, 46(1), 1970, pp. 234-240.
- Shepard, D., A Two-Dimensional Interpolation Function for Irregularly-Spaced Data, Proceedings of the 1968 23rd ACM National Conference, *ACM*, 1968, pp. 517-524.
- Yang, W., Zhao, Y., Wang, D., Wu, H., Lin, A., and He, L., Using Principal Components Analysis and IDW Interpolation to Determine Spatial and Temporal Changes of Surface Water Quality of Xin'anjiang River in Huangshan, China, *International Journal of Environmental Research and Public Health*, 17(8), 2020, p. 2942.
- Tanjung, M., Syahreza, S., and Rusdi, M., Comparison of Interpolation Methods Based on Geographic Information System (GIS) in the Spatial Distribution of Seawater Intrusion, *Journal Natural*, 20(2), 2020, pp. 24-30.
- Matheron, G., Principles of Geostatistics, *Economic Geology*, 58(8), 1963, pp. 1246-1266.
- Paramasivam, C. and Venkatramanan, S., An Introduction to Various Spatial Analysis Techniques, *GIS and Geostatistical Techniques for Groundwater Science*, 2019, pp. 23-30.
- Shi, G., *Data Mining and Knowledge Discovery for Geoscientists*, Elsevier, 2013.
- Yu, X., Efe, M.O., and Kaynak, O., A General Backpropagation Algorithm for Feed-forward Neural Networks Learning, *IEEE Transactions on Neural Networks*, 13(1), 2002, pp. 251-254.