Available online www.jsaer.com

***Journal of Scientific and Engineering Research*, 2018, 5(4):296-302**

Research Article

ISSN: 2394-2630
CODEN(USA): JSERBR

Comparison of Different Prediction Approaches For Precipitation Distribution by Kriging Methods

Ismail B. Gundogdu*, Mustafa Husrevoglu

Selcuk University, Geomatics Engineering Department, Konya, Turkey

Abstract It is expected that optimal prediction results will be obtained by carrying out the Kriging estimations in accordance with the assumptions of geostatistical analysis. In order to map precipitation distribution correctly, it is important to determine the appropriate estimation approach. In this study, two different applications with spatial estimation of rainfall data samples located in Central Anatolia have been performed. The first application has been carried out by ignoring the assumptions of geostatistical analysis. The second application has been carried out by providing the assumptions of geostatistical analysis. The errors of the estimation results of the two applications are determined by the cross validation method and the applications are compared. According to the cross validation findings, the first application has been produced better estimation results than the second. This is the results of outliers in dataset. Thus, by using geostatistics analysis of precipitation or other environmental variables, estimates of differences between considering and not considering assumptions have been observed and an appropriate estimation approach has been suggested with statistical results and maps.

Keywords geostatistics, kriging, precipitation, regionalized variable

1. Introduction

The most accurate estimation of the spatial distribution of precipitation leads to appropriate decisions on issues such as the economy and the prevention of disasters. Various spatial prediction methods are used for this purpose. In the past, the methods of kriging, mostly used in mining and geological disciplines, have spread to a wide variety of disciplines such as hydrology, ecology, environmental science under the name geostatistics [1, 2]. Geostatistical analysis is also referred to as spatial estimations performed by Kriging methods. Geostatistics is a set of methods that perform analysis and estimation of spatial or spatial-time related variables and provide information about the uncertainty of these estimates [3]. Geostatistical analysis is also known as 'Theory of Regionalized Variables' [4]. Many researchers have given detailed information about spatial prediction methods, regional variables, mathematical principles of geostatistical analysis and application principles [2-21]. In Turkey, several studies that investigated the spatial distribution of rainfall is available in the literature [1, 22-29] The basic assumptions of geostatistical analysis, spatial dependence, stationarity and distribution of the data set should be examined before estimates. Estimating with the kriging methods will not produce the expected optimal results if the data are spatially independent of each other [8, 14]. The assumption that the variance between random variables depends only on the distance between them is referred the intrinsic stationary assumption [19]. To provide the stationary assumption, random variables must not have a trend in time or position. Whether the data provides stationary assumption is investigated and if the data has a trend effect, it is removed or the appropriate method is selected [14]. Another important issue for geostatistical analysis is whether the random variables studied are normal (gaussian) distributions. According to [14], geostatistical analysis methods are better applied with normally distributed data. In addition, according to[18], the geostatistical analysis does not require normally distributed data, but the variogram function that forms the basis of the geostatistical analysis indicated that it cannot have appropriate value with the skewed data and outliers. In the Kriging estimates, the distance-based relationship between sample points is investigated with semi-variogram functions. In the Kriging estimates, the distance-based relationship between sample points is investigated with semi-variogram functions. Anisotropic variation exists if spatial autocorrelation or variation is not similar in various directions [6]. In this case, it is expected that more appropriate estimation results will be observed using an anisotropic semi-variogram model. In order to apply the principles of geostatistics analysis, it is necessary to search the data set and provide assumptions before estimating. However, it is directly related to the set of data used that all these assumptions and application principles can be ensured and the corresponding predicted values can be obtained. In this study, the estimations performed in accordance with the geostatistical analysis principles and the estimates realized without paying attention to the assumptions have been compared and the differences between them have been revealed together with the reasons. For this purpose, a data set which has a skewed distribution, anisotropic variation and spatial trend effect, has been used. In the material and method section, the theoretical basis of the Ordinary Kriging method and the statistical and spatial properties of the data have been given. In the results and discussion section, the estimates of the implemented applications have been evaluated and compared. In the conclusion section, it is recommended that researchers should follow a estimating way.

2. Material and Method

Data sets used in this study have been obtained from the General Directorate of Meteorology Turkey. This dataset contains long-term monthly average precipitation data measured from 87 meteorological stations and location information of the stations. Study area is between of 37°N - 41°N latitudes and 30°E-38°E longitudes. This region forms the inner part of Turkey. Different climatic conditions are observed in the north-south and east-west directions in the study area. The high precipitation values observed in the coastal areas are reduced due to the topography towards the interior parts and cause a spatial trend effect. This region is the most arid region compared to other regions of Turkey. It is important to investigate the factors that affect the Kriging estimates that the amount of precipitation varies spatially in the study area. The spatial distribution of the meteorological stations and study area are given in Figure 1.

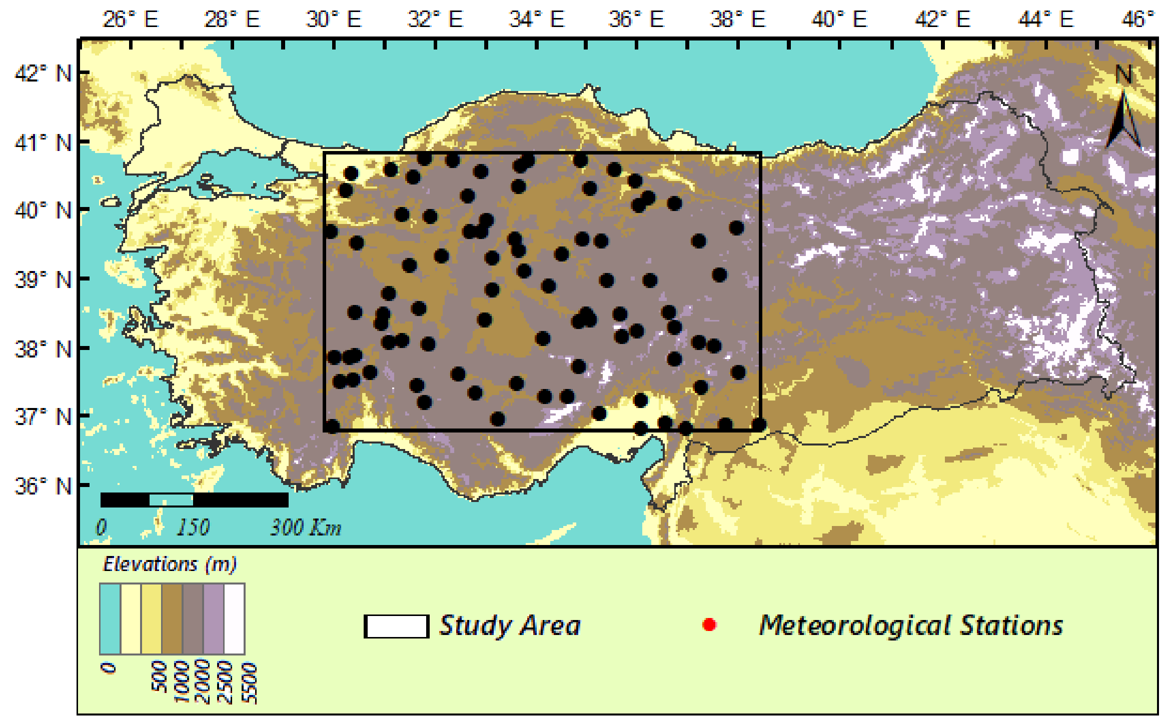

Figure 1: Spatial distribution of meteorological stations and study area

2.1. Data Analysis

By reviewing the summary statistics of the data, it can be understood whether the data are skewed or normal distribution. According to the information in Table 1, the data are skewed distribution. By applying a logarithmic transformation to the data, the distribution of data can be approximated to normal distribution.

Table 1: Summary statistics of dataset

Summary Statistics (cm)			
Number of data	87	1-st Quartile	3.333
Min	2.241	3-st Quartile	4.644
Max	10.156	Skewness	1.484
Mean	4.424	Kurtosis	4.716
Median	3.993	Standard Deviation	1.659

Table 2 summarizes the statistics of the dataset where the logarithmic transformation of the data and the normal distribution are approximated.

Table 2: Summary statistics of transformed to normal distribution dataset

Summary Statistics (cm) - After Transformation			
Number of data	87	**1-st Quartile**	1.204
Min	0.807	**3-st Quartile**	1.536
Max	2.318	**Skewness**	0.743
Mean	1.43	**Kurtosis**	3.156
Median	1.385	**Standard Deviation**	0.328

Spatial dependence has been investigated with directional experimental semi-variogram graphs. In Figure 2, the presence of different spatial relationships in different directions represents an anisotropic variation.

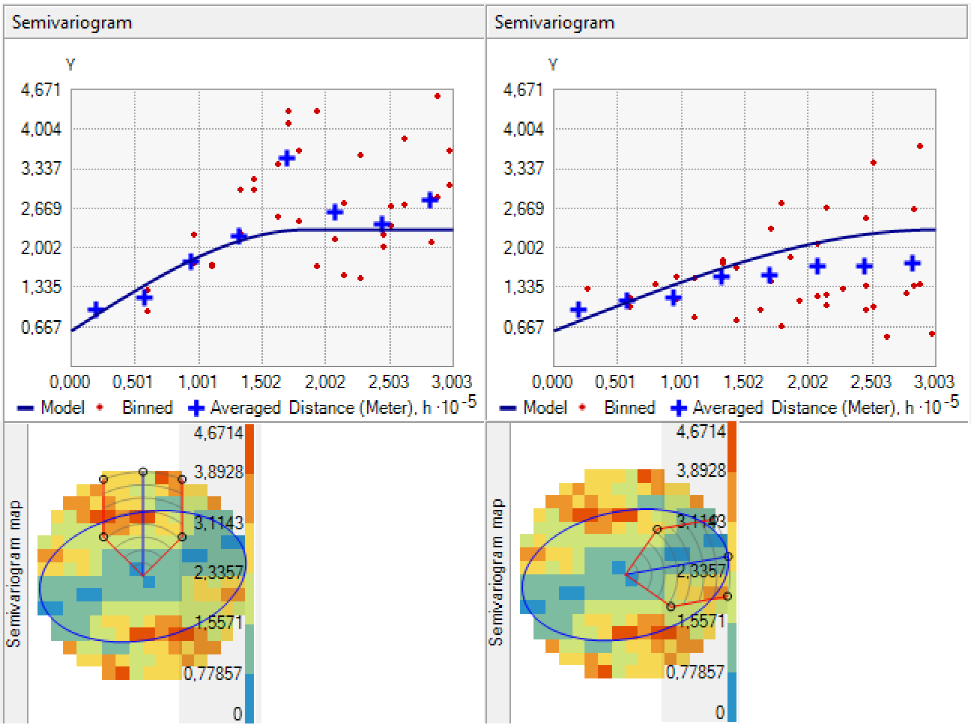

Figure 2: Investigating of spatial dependence for different direction

Whether or not there is a spatial trend effect among the data has been investigated by a trend analysis graph. According to the graph of the spatial trend analysis given in Figure 3, the precipitation values increasing from the inner part to the outer part of the study area illustrate the effect of the spatial trend. In such a case theoretically the trend effect is removed and the estimates are applied and then this effect is added to the estimates.

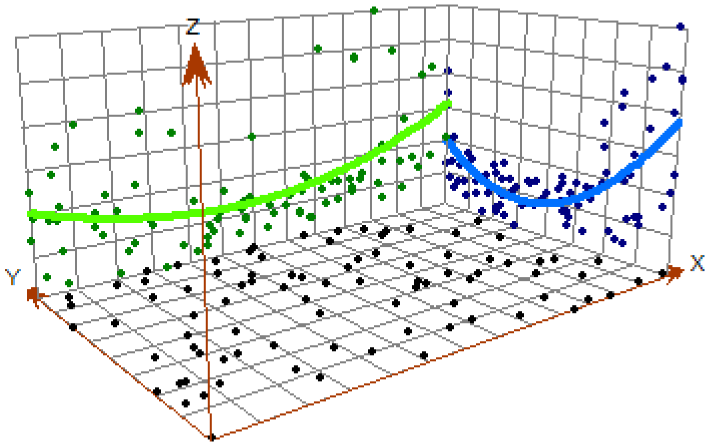

Figure 3: Trend analysis graph

2.2. Ordinary Kriging

Ordinary Kriging method has been used to map the spatial distribution of precipitation. Prior to the Kriging estimates, experimental semi-variogram models have been used to investigate spatial dependence.
The formula of OK is given in Webster and Oliver (2007) as follows;

$$\hat{Z}(x_0) = \sum_{i=1}^{N} \lambda_i z(x_i)$$

Where, x_0 is the predicted point, $\hat{Z}(x_0)$ is the predicted value of the random variable at point x_0, λ_i is the weight, N is the number of sampling points, and $z(x_i)$ is the random variable value of the x_i sample.
The semi-variogram function, which describes the spatial relationship of regional variables to each other, is defined as [30];

$$\gamma(h) = \frac{1}{2} Var\left[z(x+h) - z(x)\right]$$

and the semi-variogram function according to this equation [19];

$$\hat{\gamma}(h) = \frac{1}{2N} \sum_{i=1}^{N} \{z(x_i + h) - z(x_i)\}^2$$

Where, N is number of sample pairs compared, $z(x_i + h)$ is expected value of the regionalized variable at position $x_i + h$. Between determined by the distance h is calculated for each sample pair is semi-variogram. An experimental semi-variogram is obtained as a result of graphing the calculated semi-variogram values for the different h distances against the determined lag increments [18]. Theoretical semi-variogram models are obtained to model the spatial dependence for all distances and to apply a regional variable to the kriging prediction model [20]. In this study, Spherical semi-variogram model has been used for modeling of spatial dependency.
In this study, the differences between Kriging estimates obtained by providing geostatistical assumptions and without paying attention to these assumptions are examined. Using the isotropic semi-variogram model with Application-1, the estimation results are directly obtained. For Application-2, the trend effect has been removed, the anisotropic semi-variogram model has been used and the logarithmic transformation has been used to approximate the normal distribution. The error findings of the estimates have been obtained by the cross validation method and the two applications have been compared with each other.

3. Results and Discussion

The spatial distribution maps obtained by Application-1 and Application-2 are given in Figures 4 and 5 respectively.

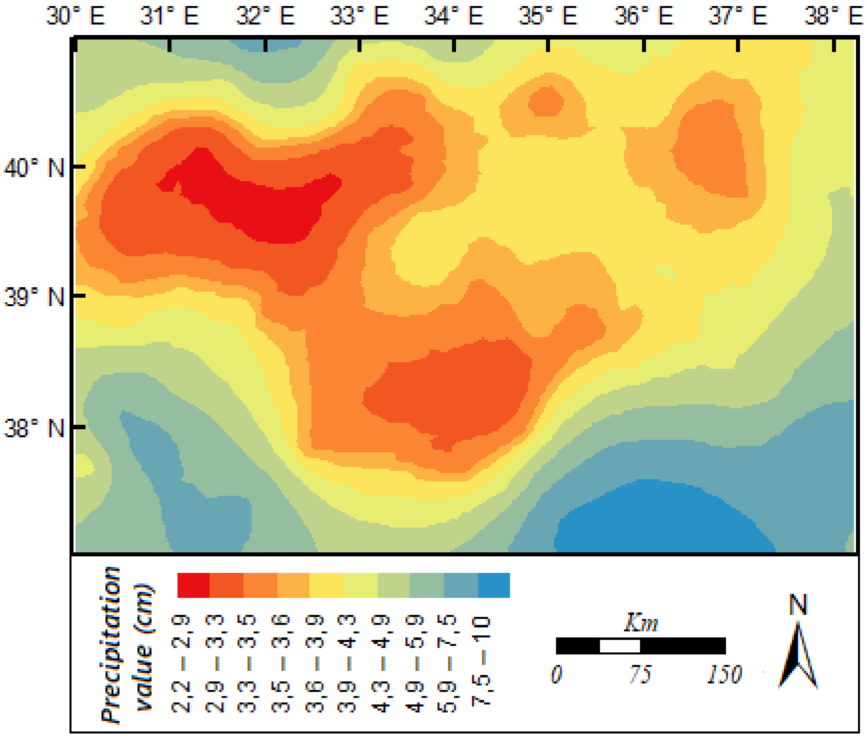

Figure 4: Map of Spatial distribution of precipitation by Application-1

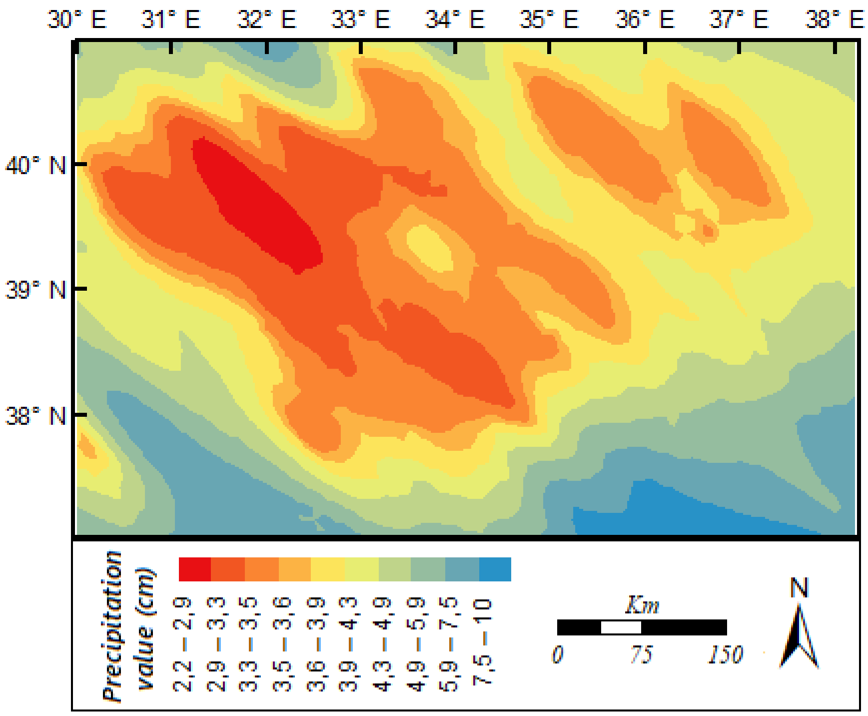

Figure 5: Map of Spatial distribution of precipitation by Application-2

Estimation errors, which are the result of evaluation of the estimation results by cross valuation method, are given in Table 3 and the applications have been compared.

Table 3: Cross validation result of applications

Cross Validation Results	Application-1	Application-2
Mean Error (cm)	-0.0220	0.0107
Root-Mean-Square Error (cm)	1.0707	1.1857
Mean Standardized Error	-0.0160	-0.0268
Root-Mean-Square Standardized Error	0.9723	1.0372
Average Standard Error (cm)	1.0950	1.0616

Mean value (ME), Mean Standardized Error (MSE) and Root Mean-Square Error (RMSE) values of the cross validation indices are low and the RMSE and Average Standard Error (ASE) value is close and (RMSSE) close to 1 indicates the appropriate estimation results[3]. These values are used to compare applications. According to the error values in Table 3, more accurate and unbiased estimation results have been obtained with Application-1. In fact, the reason for the more appropriate results with Application-1 is the outliers in the data set. However, more accurate estimation results have been expected with Application-2. It may make sense to remove the outlier values before the estimates. Reducing of limited number of data for any region with a small number of meteorological stations is an undesirable state for mapping the precipitation distribution. Actually an outlier data can be also represents current condition. For this reason, the data set has been evaluated with outliers.

4. Conclusion

Geostatistical analysis produces different estimation results with different estimation approaches. The reason of getting more appropriate results from Application-1 is it's data set contain outliers. If outliers don't remove from data set before estimation, empirical applications must be enforced.

References

[1]. Gundogdu, I.B., (2017). Usage of Multivariate Geostatistics In Interpolation Processes For Meteorological Precipitation Maps. *Theoretical and Applied Climatology*. 127(1-2): P. 81-86.

[2]. Webster, R. and M.A. Oliver, (2007). *Geostatistics For Environmental Scientists Second Edition*. John Wiley &Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England. 332.

[3]. Johnston, K., et al., (2001). *Using Arcgis Geostatistical Analyst*. Vol. 380. Esri Redlands. 306.

[4]. Kitanidis, P.K., (1997). *Introduction to Geostatistics: Applications In Hydrogeology*. Cambridge University Press.

[5]. Armstrong, M., (1998). *Basic Linear Geostatistics*. Springer Science &Business Media. 160.

[6]. Bohling, G., (2005). Introduction to Geostatistics and Variogram Analysis. *Kansas Geological Survey*: P. 1-20.

[7]. Chiles, J. and P. Delfiner, (1999). *Modelling Spatial Uncertainty*. John Wiley & Sons. 734.

[8]. Cressie, N., (1993). *Statistics For Spatial Data, Revised Edition, 1993*. Wiley-Interscience Google Scholar. 928.

[9]. Englund, E.J. and A.R. Sparks, (1988) *GEO-EAS (Geostatistical Environmental Assessment Software) User's Guide*. Battelle Columbus Labs., Washington, DC (USA).

[10]. Goovaerts, P., (1997). *Geostatistics For Natural Resources Evaluation*. Oxford University Press. 496.

[11]. Haining, R., (2014). *Spatial Data and Statistical Methods: A Chronological Overview*, In *Handbook of Regional Science*. Springer. P. 1277-1294.

[12]. Hengl, T., (2009). *A Practical Guide to Geostatistical Mapping*. Tomislav Hengl. 293.

[13]. Isaaks, E.H. and M.R. Srivastava, (1989). *Applied Geostatistics*. New York Oxford Oxford University Press. 582.

[14]. Krivoruchko, K., (2005). Introduction to Modeling Spatial Processes Using Geostatistical Analyst. *Environmental Systems Research Institute, Redlands. (Http://Www.Esri.Com/Library/Whitepapers/Pdfs/Intro-Modeling.Pdf)*. 6: P. 27.

[15]. Matheron, G., (1963). Principles of Geostatistics. *Economic Geology*. 58(8): P. 1246-1266.

[16]. Matheron, G., (1973). The Intrinsic Random Functions and Their Applications. *Advances In Applied Probability*. 5(3): P. 439-468.
[17]. Oliver, M.A., (2010). *The Variogram and Kriging*, In *Handbook of Applied Spatial Analysis*, M.M. Fischer, Editor., Springer-Verlag Berlin Heidelberg 2010. P. 319-352.
[18]. Oliver, M.A. and R. Webster, (2015). *Basic Steps In Geostatistics: The Variogram and Kriging*. Springer. 106.
[19]. Sarma, D.D., (2010). *Geostatistics With Applications In Earth Sciences*. Springer Science &Business Media. 219.
[20]. Sen, Z., (2009). *Spatial Modeling Principles In Earth Sciences*. Springer. 358.
[21]. Wackernagel, H., (2003). *Multivariate Geostatistics: An Introduction With Applications*. Springer Science &Business Media. 381.
[22]. Aydin, O., (2014) *Turkiye'de Yillik Ortalama Toplam Yagisin Kriging Yontemiyle Belirlenmesi*, In *Sosyal Bilimler Enstitusu*. Anakara Universitesi: Ankara.
[23]. Bostan, P.A., (2013) *Analysis and Modeling of Spatially and Temporally Varying Meteorological Parameter: Precipitation Over Turkey*, In *The Graduate School of Natural and Applied Sciences*. Middle East Technical University: Ankara. P. 94.
[24]. Cetin, M. and K. Tulucu, (1998). DoguAkdeniz Bolgesi'nde Aylik Yagislarin Yersel Degisimlerinin Jeoistatistik Yontemle Incelenmesi. *Turkish Journal of Engineering and Environmental Science*. 22: P. 279-288.
[25]. Citakoglu, H., et al., (2017). Mevsimsel Yagislarin Jeoistatistiksel Yontemle Modellenmesi Ve Gozlemi Olmayan Noktalarda Tahmin Edilmesi. *IMO Teknik Dergi*. 7725(7745): P. 469.
[26]. Eris, E., (2011) *Determination of Spatial Distrabution of Precipitation On Poorly Gauged Coastal Regions*, In *Fen Bilimleri Enstitusu*. Istanbul Teknik Universitesi: Istanbul.
[27]. Ilker, A., (2012) *Akdeniz Bolgesinde Yagisin Alansal Dagilimi*, In *Fen Bilimleri Enstitusu*. Suleyman Demirel Universitesi: Isparta. P. 83.
[28]. Ozcakal, E., (2014) *Ege Bolgesi Yagis Ve Sicaklik Degerlerinin Zaman Ve Mekân Boyutundaki Degisiminin Istatistiksel Analizi*, In *Fen Bilimleri Enstitusu*. Ege Universitesi: Izmir. P. 312.
[29]. Sensoy, S., et al., (2007) *Klimatolojik Uygulamalarda Arcgis Kullanimi*, In *12. ESRI Kullanicilari Grubu Toplantisi*. Meteoroloji Genel Mudurlugu: ODTU.
[30]. Matheron, G., (1967). Kriging Or Polynomial Interpolation Procedures. *CIMM Transactions*. 70: P. 240-244.