RAINFALL ESTIMATION BY USING THIESSEN POLYGONS, INVERSE DISTANCE WEIGHTED, SPLINE, AND KRIGING METHODS: A CASE STUDY IN PONTIANAK, WEST KALIMANTAN

Iin Arianti[1], Soemarno[2], Hasyim A.W.[3], and Sulistyono R.[4]
[1]Doctoral Program of Environmental Science, Brawijaya University, Malang
[2]Department of Soil Science, Faculty of Agriculture, Brawijaya University, Malang, Indonesia
[3]Department of Regional and City Planning, Faculty of Engineering, Brawijaya University, Malang, Indonesia
[4] Postgraduate Lecturers, Brawijaya University, Malang, Indonesia
*Corresponden email***:** *iin_arianti@yahoo.com*

Abstract

There are four methods for estimating rainfall in a region with GIS; they are Thiessen polygons, Inverse Distance Weighted (IDW), Spline, and Kriging. Each has its own advantages and disadvantages. The selection of these four methods must be adjusted with the appropriation of the rainfall data. This article will show the advantages and disadvantages of the four methods for the study area of Pontianak, West Kalimantan, Indonesia.

Keywords: Rainfall, Thiessen Polygons, IDW, Spline, Kriging

1. Introduction

Rainfall in an area can be measured by using a device called Ombrometer. The data obtained from a rain gauge is a data concerning rain that occurs only in one area or point (*point rainfall*). Given that rain varies with space (space), then for large areas, one rain gauge still cannot represent the rain condition in the area. In this case, regional rain is needed from the average rainfall value of some rain gauge stations in / or around the area.

Interpolation is a method for obtaining data based on several data already known (Wikipedia, 2008). In mapping, interpolation is a process for estimating values in areas that are not sampled or measured, so that maps or distribution of values are made across regions (Gamma Design Software, 2005).

Testing methods for predicting the rainfall distribution by using applications of geographic information system can be presented as a map of the average rainfall in each month and daily rain of each month. In making rainfall distribution maps, it can be done in ArcGis by using several methods, namely Thiessen, IDW, Spline, and Kriging Methods.

This article uses monthly rainfall data of 2017 at the nearest climatology or meteorological stations in Pontianak study area, namely data from the Maritime station, Supadio meteorological station, and Mempawah meteorological station. The data is in a spreadsheet format such as the *.csv format that already has a geographical location.

2. Discussion and Result Research

2.1 Thiessen Polygons

Thiessen is a method that is determined by making polygons between stations in an area then the average rainfall height is calculated from the number of multiplications between each polygon area and the rain height is divided by the entire area. The Thiessen polygon method is usually used to detect the average rain height if the rain station is not evenly distributed.

Thiessen's polygon is also one of the local interpolation methods. This method is also called the proximal method, i.e. an attempt to weight the data points in an area. The measures taken are a number of triangles drawn by connecting control points (for example, meteorological stations) using the Delaunay triangulation technique (also used for TIN). The line is drawn perpendicularly to the side of the triangle at the midpoint. Polygon is defined by intersection of lines.

The weighted average of each rain station is determined by its area of influence according to the formed polygon (drawing axis lines on the connecting lines between two adjacent rain stations). This method is obtained by making a polygon that perpendicularly cuts the middle of the connecting line of two rain stations. Thus each measuring station R_n will lie in a certain polygon A_n. By calculating the area ratio for each station = A_n/A, where A is the wide of shelter area or the total area of the entire area sought for the high rainfall. Average rainfall is obtained by summing up on each point with an area of influence which is formed by drawing axis lines perpendicularly to the connecting line between two measuring stations.

$$Q = \frac{A_1 R_1 + A_2 R_2 + A_3 R_3 + \cdots + A_n R_n}{A_1 + A_2 + A_3 + \cdots + A_n} = \frac{\sum A_i R_i}{\sum A_i}$$

Where Q is the average rainfall; $A_1, A_2, A_3, \ldots A_n$ are areas represented by the rainfall recorded at stations 1, 2, 3, and so on; R_1, R_2, R_3, …, R_n are rainfalls at stations 1, 2, 3, and n is the station number.

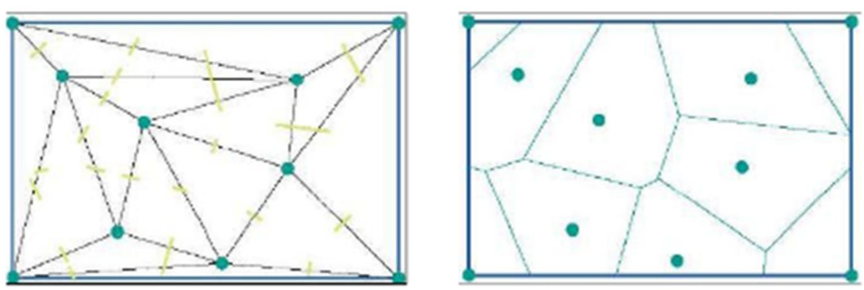

Figure1. Thiessen Polygons Method

Classification of rainfall classes is started from less than 1500, 1500 - 2000, 2000 - 2500, 2500 - 3000, 3000 - 3500 and above 3500 (according to rainfall classification in general). Because the appropriate data in Pontianak is only limited to 2560 to 3402, so it is classified into 2 classes only. As seen in figure 2, Pontianak is only classified in the 3000 - 3500 rainfall class (high).

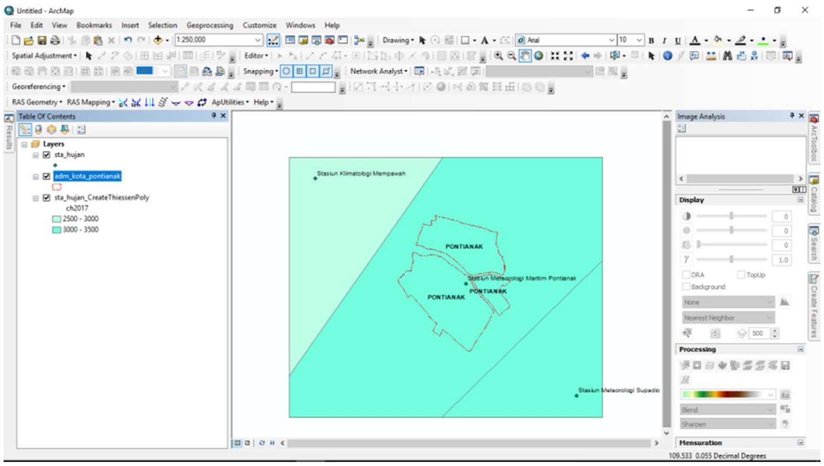

Figure 2. Results of the Classification Process based on three observing stations

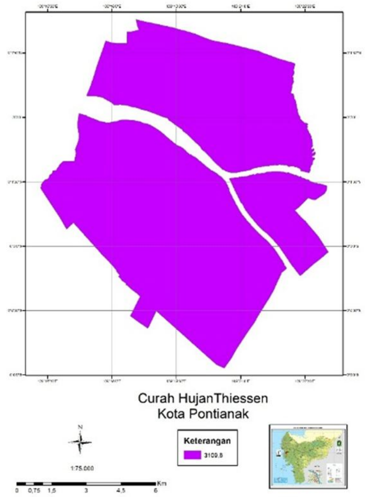

Figure 3. Map of Rainfall with Thiessen Polygon Method

2.2 IDW (*Inverse Distance Weighted*)

IDW (*Inverse Distance Weighted*) method assumes that each input point has a local effect that decreases with distance. This method emphasizes more on the closest cell to the data point than the further cell. Points at a certain radius can be used to determine the output value for each location

The IDW method is generally influenced by inverse distance obtained from mathematical equations. In this interpolation method, we can adjust the relative influence of the sample point. The power value in IDW interpolation determines the effect towards input points, where the influence will be greater at closer points so as to produce a more detailed surface. The influence will be smaller with the increasing distance where the resulting surface is less detailed and looks smoother. If the power value is increased, it means that the output value of the cell becomes more localized and has a lower average value. The decreasing power value will result in an output with a larger average because it will give influence to a wider area. If the power value is reduced, a smoother surface is resulted. The weight used for the average is a distance function between the sample point and the interpolated point (Philip & Watson, 1985).

The weighting general function is the inverse of the distance square, and this equation is used in the inverse *invers distancewighted* method which is formulated in the following formula:

$$Z^* = \sum_{i-1}^{N} w_i Z_i$$

WhereZ_i (i= 1,2,3,.....,n) is the data height value will be interpolated by a number of N points and weights (*weight*) w_i which are formulated as follows:

$$w_i = \frac{h_i^{-p}}{\sum_{j=1}^{n} h_j^{-p}}$$

p is a changeable positive value called the *power* parameter (its value usually 2) and h_j is the distance from the point distribution to the interpolation point which described as follows:

$$h_i = \sqrt{(x - x_i)^2 + (y - y_i)^2}$$

(x,y) are coordinates of the interpolation point and (x_i,y_i) are coordinates for each spread point. The function of weight variable is vary for the whole data point distribution to a value close to zero, where the distance increases with the spread point.

The determination of results in IDW method is based on the assumption that the value of z attribute (the estimated value) at the unrecorded point is a distance function and the average value of points around it. The interpolation results are depend on how strong a data point is known to affect the surrounding area. In addition, the number of points around it is used to calculate the average value, as well as the desired pixel / raster size.

The advantages of IDW interpolation method are that the interpolation characteristics can be controlled by limiting the input points used in the interpolation process. Points located far from the sample point and which are estimated to have spatial correlations can be eliminated from the calculation. The points used can be directly determined or selected based on the distance to be interpolated. The disadvantage of IDW interpolation is that it cannot estimate values above the maximum value and below the minimum value of the sample points (Pramono, 2008).

The resulted effect when IDW interpolation is applied is the peaks and valleys flattening unless the highest and lowest points are part of the sample point. Due to the estimated value as the average value as well, the surface results will not exactly pass through the sample points. Another disadvantage of this interpolation method is the *bull-eye* effect.

. Interpolation using the IDW method gives values with a wider range of areas. The interpolation value will be more similar to the sample data in adjacent location rather than data in further location. Because this method uses an average of sample data, thus the value cannot be smaller than the minimum or greater than the sample data. Therefore, top of the hill or the deepest valley cannot be displayed from the interpolation results of this model (Watson & Philip, 1985). To obtain good results, the data sample used must be dense that meet the local variations. If the sample is rather rare and unevenly distributed, the results are likely not to be as expected.

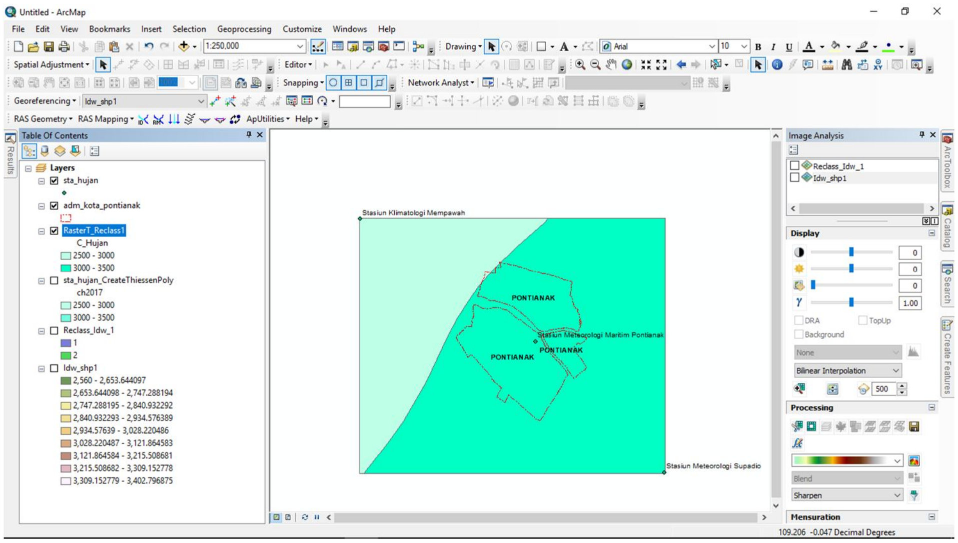

Figure 4. IDW Interpolation Results

Based on the results of IDW's interpolation, most of Pontianak area has a rainfall of 3000 - 3500 and the North West area has a rather small rainfall number of 2500 - 3000 mm per year.

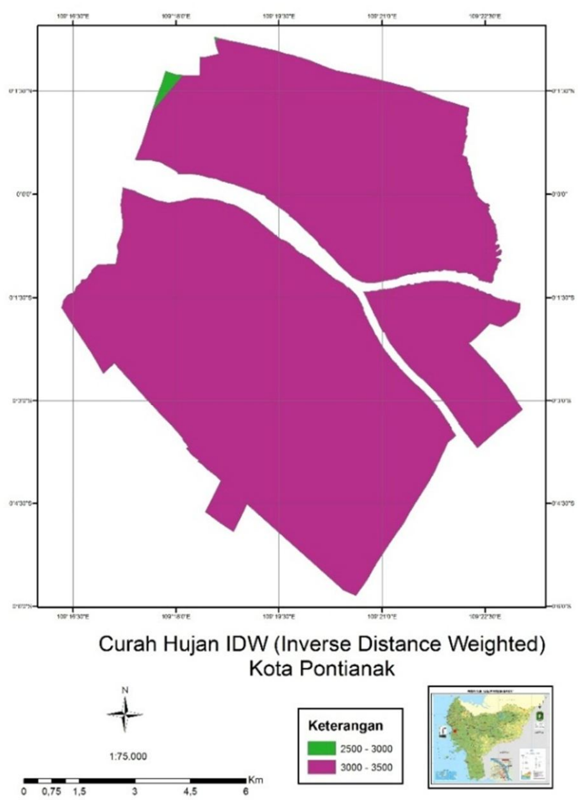

Figure 5. Rainfall map of Pontiianak city with IDW method

2.3 *Spline*

Spline is a method which estimates values by using mathematical functions that minimize the total surface curvature. The stretching effect of the spline is very useful to estimate the value below the minimum value and the value above the maximum which possibly found in the data set used. This makes the Spline interpolation method a good method for estimating low and high values that are not found on the data sample.

The advantage of Spline method is its ability to produce a good accuracy result eventhough the data used is limited. This method is very appropriate in making surfaces such as the height of earth's surface, groundwater level, and concentration of air pollution. This method is not appropriate if applied to situations where there are significant differences in values at very close distances. The drawback of this Spline method is that when adjacent sample points have significant differences in values, then the Spline method cannot work properly. This is because due to Spline method uses a calculation of slope that changes based on distance to estimate the shape of the surface.

The equation used in the Spline method is a surface interpolation formula:

$$S_{(x,y)} = \tau_{(x,y)} + \sum_{i=1}^{n} \lambda_i R_{(r_i)}$$

Where n is the number of data points; λ_i are weighting coefficients; r_i is the distance from the desired point (x,y) to the measurement point i;$\tau_{(x,y)}$and $R_{(r_i)}$are defined by

$$\tau_{(x,y)} = a_1 + a_2 x + a_3 y$$

$$R_{(r_i)} = \frac{1}{2\pi}\left\{\frac{r^{\varphi}}{4}\left[ln\left(\frac{r}{2\tau}\right) + c - 1\right] + \tau^{\varphi}\left[K_0 \frac{r}{\tau} + c + ln\left(\frac{r}{2\pi}\right)\right]\right\}$$

Where a_i is coefficient; r is thedistance between the desire location; τ & φ define the weight of the third derivative of the surface; K_0 is a modified Bessel function; and c is a constant equal to 0.5772.

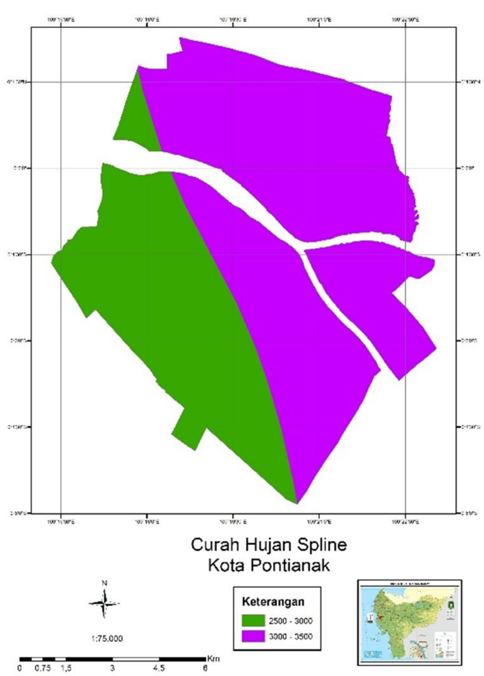

Figure 6. Rainfall map of Pontianak city with *Spline* method

2.4 Kriging

Kriging method is a stochastic estimation similar to IDW, it uses linear combinations of weights to estimate the value between sample data. This method was developed by D.L. Krige to estimate the value of mining materials. The assumptions behind this model are distance and orientation between sample data showing spatial correlation. This model gi vesa measure of error and confidence.

The determination of results in the Kriging method is based on the assumption that every point in the landscape is interconnected and has a trend. Trends (mathematical equations) are used to predict points that do not have data / information.

In the Kriging interpolation, the coverage of interpolation area is narrower. At close distances (horizontal axis), *semivariance* is small, but at larger distances, it has a higher value which indicates that variations in the z value are no longer related to the distance of sample point.

The kriged or predicted value $Z^*(x_0)$ is a linear combination of observations at Nnb neighbor's stations:

$$Z^*(x_0) = \sum_{i=1}^{N_{nb}} \lambda_i Z(x_i)$$

Where $\lambda_i$are kriging weights which are estimated as solution of the kriging system. In the case of external drift kriging using a drift variable Y, the system is:

$$\sum_{i=1}^{N_{nb}} \lambda_j \, \gamma(x_j - x_i) + \mu_1 + \mu_2 Y(x_j) = \gamma(x_j - x_0) \quad j = 1, \dots, N_{nb}$$

$$\sum_{i=1}^{N_{nb}} \lambda_i = 1$$

$$\sum_{i=1}^{N_{nb}} \lambda_i \, Y(x_i) = Y(x_0)$$

where μ_1, μ_2 are two Lagrange parameters accounting for the two constraints on the weights. So, Y has to be known at locations x_i as well as the target location x_0. The ED kriging variance is obtained by:

$$\sigma^2 = \gamma(0) - \sum_{i=1}^{N_{nb}} \lambda_i \, \gamma(x_i - x_0) - \mu_1 - \mu_2 \, Y(x_0)$$

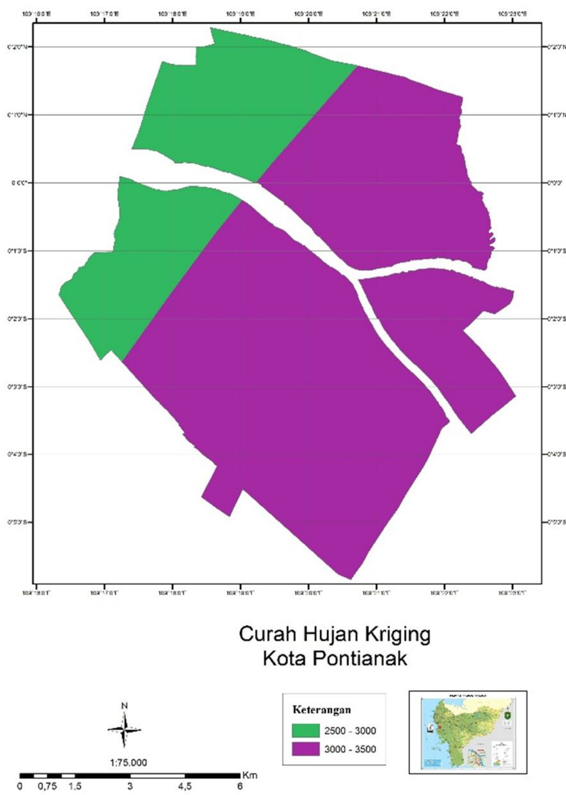

Figure 7. Rainfall map of Pontianak city by using the Kriging method

3. Conclusion

Spatial interpolation is needed in meteorology and climatology to predict rainfall intensity, by overcoming the limited data, environment, and other physiographic conditions, so that the modeling of climate element can still be done.

Thiessen polygons are probably the most common approach for modelling the spatial distribution of rainfall. The approach is based on defining the area closer to a particular gauge than any other gauges and the assumption that the best estimation of rainfall on that area is represented by the point measurement at the gauge. However, an impact resulted regarding the use of Thiessen polygons is the development of discontinous functions defining the rainfall depth over the catchment. This impact arises at the boundaries of the polygons where a discrete change in rainfall depth, or intensity, occurs (Luk. 1997).

The IDW method assumes that each input point has a local effect which decreases with distance. This method emphasizes more on the cell closest to the data point than the further cell. Points at a certain radius can be used to determine the output value for each location.

In contrast, *Spline* is a local fitting technique which is based on the interpolation between several adjacent measurement points by using some low-order polynomials. This approach avoids overfitting on the measurement points by high-order polynomials. In addition, estimation of the order of polynomials is not required. Surfaces of this type have been found to be a robust spatial interpolation for many meteorological problems (Luk, 1997).

The use of Kriging for interpolating rainfall at locations is not monitored based on the assumption of spatial homogenuity, which means that the same pattern of variation can be observed in all locations within the catchment. Furthermore, it is assumed that the spatial variability of the rainfall is a function of a structural component which is correlated to a constant trend, a spatially correlated component, and a random residual error term (Luk. 1997).

The IDW interpolation method provides more accurate interpolation results. This is because there are only three rain gauges for Pontianak city. The results of IDW method give the value close to the minimum and maximum values of the sample data.

References

The Meteorology, Climatology, and Geophysical Agency.2017: Analisis Hujan Tahunan Januari- Desember 2017. *BMKG Buletin*.

Bargaoui, Z.K., & Chebbi, A. 2009. Comparison of two kriging interpolation methods applied to spatiotemporal rainfall. Journal of Hydrology 365, pp. 56-73.

Burrough, P.A.& McDonnell R.A. 1998. *Principles Of Geographical Information System*. London: Oxford University Press Inc.

Citra Satelit Landsat7-ETM+ Terhadap Data Observasi Stasiun Cuaca/Iklim Darat (Kasus Provinsi Jawa Barat Bagian Selatan). *Agroclimatology Journal*, IPB, Bogor.

Demers & Michael, N. 2000. *Fundamentals of Geographic Information System* Second Edition. New York(US): John Wiley and Sons.

Elfaig, A.H.I., Eltom, I.M., & Salih, A. 2013. Rainfall Variability in the Sahel: A Study from Sudan (1970-2010). International Journal of Scientific and Research Publications. Vol. 3, Issue 11, Nov. pp 1-8.

ESRI. 1996. *Using the ArcView Spatial Analist*. Redlands (US): Environmental Systems Research Institute, Inc.

Gamma Design Software. 2005. Interpolation in GS+. http://www.geostatistics.com/ OverviewInterpolation.html (23 Juni 2018).

Hadi, B.S. 2013. Metode Interpolasi Spasial dalam Studi Geografi. *Geomedia Journal*. Vol 11(2): pp. 231-240.

Linglingue, M. 2017. Assessment of the variability of spatial interpolation methods using elevation and drill hole data over the Magmont Mine area, South-East Missouri. Thesis. University of Missouri-Columbia.

Lu, G.Y.& Wong, D.W. 2008. An Adaptive Inverse Distance Weighting Spatial Interpolation Technique. *Computers and Geosciences* 34: pp. 1044-1055.

Luk, K.C., & Ball, J.E. 1997. Modelling The Instantaneous Spatial Variability of Rainfall. International Congress on Modelling and Simulation. Hobart, Tasmania, pp. 71-76.

Malla, M.A., Rather, M.A., Teli, M.N., & Kuchhay, N.A. 2014. Comparison of Spatial Interpolation Techniques – A Case Study of Anantnag District J & K, India. International Journal of Modern Engineering Research. Vol. 4, Iss. 11, Nov. 2014, pp 59-67.

Naoum, S. & Tsanis, I.K. 2004. Ranking Spatial Interpolation Techniques Using A GIS-Based DSS. Global Nest The Int. J. Vol 6, No. 1, pp. 1-20.

Ningsih D.H.U. 2012. Metode Thiessen Poligon untuk Ramalan Sebaran Curah Hujan Periode Tertentu pada Wilayah yang Tidak Memiliki Data Curah Hujan. *Journal of Dynamic Information Technology*. 17(2): pp. 154-156.

Pasaribu, J.M. & Haryani, N.S.2012. Perbandingan Teknik Interpolasi DEM SRTM dengan Metode Inverse Distance Weighted (IDW), Natural Neighbor, dan Spline. *Remote Sensing Journal*, Vol. 9 No. 2.pp. 126-139.

Prahasta, E. 2015. SIG: *Tutorial ArcGis untuk bidang geodesi dan geomatika*, Informatika Bandung.

Pramono,G.H. 2008. Akurasi Metode IDW dan Kriging untuk Interpolasi Sebaran Sedimen Tersuspensi. *Forum Geografi*. Vol. 22 No. 1. pp. 145-158.

Thiessen, A.H. 1991. Precipitation Averages for Large Areas. Monthly Weather Review. Vol. 39, pp. 1082-1084.

Watson D.F & Philip G.M. 1985. A Refinement of Inverse Distance Weighted Interpolation. *Journal Geo Processing*. Vol 2.pp. 315-327.

Zimmerman, D., Pavlik, C., Ruggles, A., &Amstrong, M.P. 1999. An Experimental Comparison of Ordinary and Universal Kriging and Inverse Distance Weighting. *Mathematical Geology*, 31 (4), pp. 375-390.