Iraqi Journal of Physics, 2013 *Vol.11, No.21, PP. 59-66*

Comparison of three interpolation methods for the average monthly temperature in the south of Iraqi zone

Nawal K.Ghazal[1], Ebtesam Fadhel[2], Khalid Abdu-Alkareem[2], Aqeel Zuher[2]

[1]Department of Physics, College of Science, University of Baghdad

[2]Department of Astronomy & Space, College of Science, University of Baghdad

E-mail: dr.nawal@yahoo.com

Abstract

This study focuses on evaluating the suitability of three interpolation methods in terms of their accuracy at climate data for some provinces of south of Iraq. Two data sets of maximum and minimum temperature in February 2008 from nine meteorological stations located in the south of Iraq using three interpolation methods. ArcGIS is used to produce the spatially distributed temperature data by using IDW, ordinary kriging, and spline. Four statistical methods are applied to analyze the results obtained from three interpolation methods. These methods are RMSE, RMSE as a percentage of the mean, Model efficiency (*E*) and Bias, which showed that the ordinary krigingis the best for this data from other methods by the results that have been obtained .

Key words
Spline, Interpolation, Temperature, ArcGIS, Climate.

Article info
Received: Oct. 2012
Accepted: Apr. 2013
Published: Sep. 2013

مقارنة طرق الاستيفاء الثلاثة لدرجة حرارة المعدل الشهري في نطاق جنوب العراق

نوال خلف غزال[1]، ابتسام فاضل خنجر[2]، خالد عبد الكريم هادي[2]، عقيل زهيرعزيز[2]

[1]قسم الفيزياء، كلية العلوم، جامعة بغداد

[2]قسم الفلك والفضاء، كلية العلوم، جامعة بغداد

الخلاصة

في هذه الدراسة تم التركيزعلى تقييم مدى ملائمة طرق الاستيفاء الثلاثة من حيث دقتها على البيانات المناخية لبعض محافظات جنوب العراق. استخدمت مجموعتين من البيانات هي درجات الحرارة العظمى والصغرى في شهر شباط 2008، أخذت من تسعة محطات للإرصاد الجوي الواقعة جنوب العراق باستخدام ثلاثة من طرق الاستيفاء، تم استخدام برنامج نظم المعلومات الجغرافية لبيان التوزيع المكاني لبيانات درجات الحرارة باستخدام طرق الـ IDW,ordinary kriging ,and spline. تم تطبيق أربعة أساليب إحصائية لتحليل النتائج المستحصل عليها من طرق الاستيفاء الثلاثة. هذه الأساليب هي RMSE، RMSE كنسبة مئوية متوسط و نموذج الكفاءة (E) والتحيز والذي ظهر أنه الأفضل لهذه البيانات من الطرق ألأخرى من خلال النتائج التي تم الحصول عليها.

Introduction

Climate plays a significant role in flora and fauna distributions; it is usually a key to understand the interdependence between environmental and biological factors and is widely used in developing ecological zones and biodiversity assessments. Given the determinant effect of climate on ecosystems, estimates of the spatial distribution of climatic variables are required more than ever for sustainable management of natural resources. Determining spatial climate conditions,

however, is not easy, because long-term average weather observations come from sparse, discrete and irregular distributed meteorological stations [1].
These discrete data have to be extended spatially to reflect the continuously and gradually changed climate pattern. In particular, climate dependence on topography must be taken into account when developing reliable climate estimates [2].
More attention has been given to the application of interpolation techniques to climate analysis in recent years. Several interpolation approaches are available in geographical information systems (GISs) to meet the general requirements of interpolation. For climate interpolation, Spline and Kriging methods are preferable, as they into account the climate dependence on topography by using a trivariate function of latitude and longitude as two independent variables and elevation as a covariate [3].
Spatial distributed estimates of meteorological data becoming increasingly important as input to spatially explicit landscape, regional, and global models. Estimates of meteorological values such as temperature, precipitation, and evapotranspiration on rate are required for a number of landscape scale models, including those of regeneration, growth, and mortality of forest ecosystems [4].
Statistical evaluation is to understand statistical properties associated with the data set. These properties may include distribution, location, spread, and shape of the data set. There are many tools available in the univariate description statistics which can be used to describe these properties. For example, the frequency table and corresponding histogram can be used to describe how often observed values fall within certain intervals or classes. Probability plot can be used to determine how close the distribution of the data set is to a Gaussian or normal distribution [5].

Study area and Data used

The study area is located in the south of Iraq which includes a number of climatology of stations located in some of the southern provinces. The data used in this study, are Administrative map of Iraq as paper, with scale (1:100,000) and software used such as ERDAS, ArcGIS.

Input Data

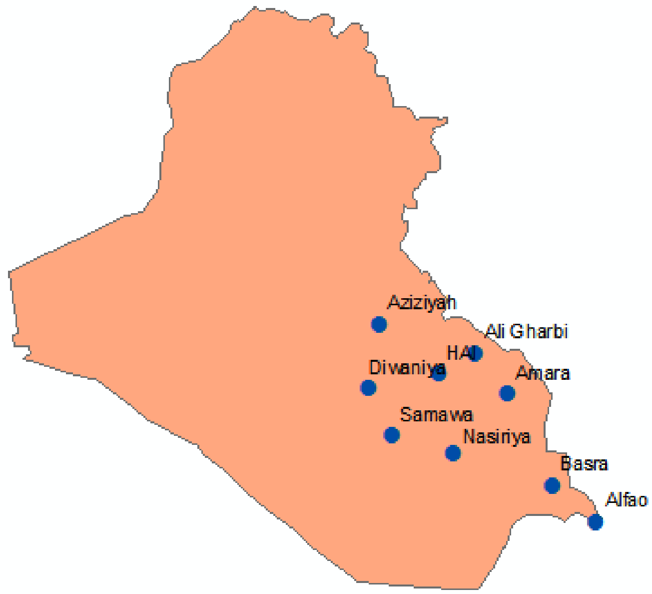

Fig.1: The location of climatological stations.

Table1: The climate input data for nine climatological station at February -2008.

ID	Provinces name	Temperature-Maximum	Temperature-Minimum
1	Aziziyah	19.3	6.8
2	Hai	20.6	8.3
3	Diwaniya	19.8	7.0
4	Amara	20.7	7.6
5	Ali-algharbi	20.5	7.6
6	Samawa	20.1	7.1
7	Nasiria	20.7	7.8
8	Basra	20.8	8.3
9	Alfao	19.8	8.9

Methodology

Accurate climatological data are collected at meteorological stations. Which are discrete point locations in space. Values at any other point must be derived from neighboring stations or from relationships with other variables. The method of spreading discrete measurements over a continuous surface represented by a regularly spaced grid is called spatial interpolation. Methods of mapping climate from point data fall into two categories: human expertise and statistical. The first group is based on human experience and knowledge and involves manual preparation of climate maps often related to topographic analysis. Various statistical methods have been developed to predict the spatial distribution of climate variables .Commonly used interpolation methods for meteorological applications include Invers-distance weighting, Spline, and Kriging. Various statistical methods have been developed to predict the spatial distribution of climate variables [6].

1-Inverse Distance Weights (IDW)

This a deterministic interpolation technique that creates surface from measured points, based on either the extent of similarity. The inverse distance weight technique provides an interpolation using s linearly combination of the temperature locations. To predict a value for any unmeasured location, IDW will use the measured values surrounding the prediction location. Those measured values closest to the prediction location will have more influence on the predicted value than those farther away. Thus, IDW assumes that each measured point has a local influence that diminishes with distance [7].

2- Spline

The Spline method can be thought of as fitting a rubber-sheeted surface through the known points using a mathematical function. In ArcGIS, the spline interpolation is a Radial Basis Function (RBF). These functions allow analysts to decide between smooth curves or tight straight edges between measured points. Advantages of splining functions are that they can generate sufficiently accurate surfaces from only a few sampled points and they retain small features. A disadvantage is that they may have different minimum and maximum values than the data set and the functions are sensitive to outliers due to the inclusion of the original data values at the sample points. This is true for all exact interpolations, which are commonly used in GIS, but can present more serious problems for Spline since it operates best for gently varying surfaces,(i.e. those having a low variance),[8].

3- Ordinary kriging

Ordinary kriging is the basic form of kriging .The prediction by ordinary kriging is a linear combination of the measured values. The spatial correlation between the data, as described by variogram, determines the weights. As the mean is unknown, fewer assumptions are made. The method assumes intrinsic stationarity, unfortunately meteorological variables are often not stationary. In some case this problem can be eliminated by using different sizes and shapes of the search nieghbourhood. Ordinary kriging is frequently applied in meteorology, often as part of residual kriging or indicator kriging[9].

Qualitative Measures of Estimation Accuracy

In this study, four statistical are used to characterize the performance of interpolation methods. They are, RMSE, RMSE as a percentage of mean observed temperature, bias and model efficiency. These statistics are described in equations (1), (2),(3),and (4).

1-Root mean square error

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum(Oi - Ei)^2} \qquad (1)$$

where:

N = no. data
O_i = Observed temperature at i
E_i = Estimated temperature at i
O_{mean} = Mean observed temperature for (i=1,2…N)

2- **Model Efficiency (*E*)**

$$E = E = 1 - (\sum_{i=1}^{N}(Oi - Ei)^2 \,/\, \sum_{i=1}^{N}(O - Omean)^2 \quad (2)$$

3- **Root mean square error as percentage of the mean:**

$$(RMSE / Omean)*100\ \% \quad (3)$$

4- Bias

$$Bias = Emean\ / Omean \quad (4)$$

Standard error of interpolation surfaces are often expressed as percentage of the mean, due to the nature of the distribution of temperature (Hutchinson 2004). The RMSE of estimated temperature is based on the number of data and it is useful to quote the RMSE as a percentage of the average temperature.

Exact interpolation methods are indicated by RMSE and (RMSE ⁄ Omean) values of zero. Accordingly, the most accurate interpolation methods are indicated by RMSE and (RMSE ⁄ Omean) values closest to zero.It should be noted that for interpolation of temperature data, estimation error is unlikely to be zero)(.Hutchinson(2004) states that standard error of fitted surfaces should be approximately 10% for monthly mean temperature data when adequate networks are available.

Model efficiency (E) equals 1 when observations and estimations are in perfect agreement. Model efficiency can be less than zero when the model estimations are worse than using the average of observed temperature as an estimator. For high model efficiency, the mean square error (MSE) of the temperature estimations will need to be small relative to stander deviation of the observed temperature.

Interpolation results that indicate no bias have a bias statistic value of 1.That is, the average temperature estimate is equal to the average observed temperature. Bias values greater than one indicate that the estimated temperature is generally over estimated, while values less than one suggest that the interpolation method resulted in underestimation [10].

Discuses and results

1- Discuses results for mean maximum temperature

Table 2 shows the average error statistics for each of the three interpolation methods based on the climate data available. For these samples, when applying IDW, ordinary kriging, and Spline are found to perform comparably to each other. The Spline method has the highest RMSE. These three methods are producing the same results for the model efficiency. Ordinary kriging and IDW provide more accurate estimation than Spline method. More frequent occurrence of extreme error is observed with Spline interpolation. The root mean square error gave good results close to zero.

Figs.2-7, illustrated continuous surfaces for spatial interpolation methods to produce the estimation values in areas where there are no measurements.

Figs. 8-10 are comparisons of estimated and observed mean maximum temperature for nine climate stations lying south of Iraq. These figures represent the IDW, Spline, and ordinary kriging methods respectively which producing the behavior of estimated mean maximum temperature with the mean maximum observer temperature. The relation between estimation and observed mean max temperatres, the weakest correlation occurred (R^2=0.119, R^2=0.118) for these climate data.

2-Discuses results for mean minimum temperature

These IDW, Ordinary kriging, and Spline interpolation methods are used for mean minimum temperature to produce the best one between them. As illustrated in Table 3, the Spline method has the highest RMSE. These three methods are producing the highest accurate results for RMSE, the same results for model efficiency and the best results for the bias method. Ordinary kriging and IDW provide more accurate estimation than Spline method. More frequent occurrence of extreme error is observed with Spline interpolation. Also for root mean square error gave good results close to zero such as the results for max temperature.

The results of the bias for these three methods have the value of one, which producing that the average minimum temperature estimate is equal to average minimum temperature observed.

Figs. 11-13, these figures illustrated the demeanor of the three interpolation methods for climate data (minimum temperature) at 9 stations. The relation between estimation and observed mean mini- temperatures, also the weakest correlation occurred (R^2=0.19, R^2=0.205) for these climate data.

3- Figures of three interpolation methods

Figure (2): The IDW method for T-max

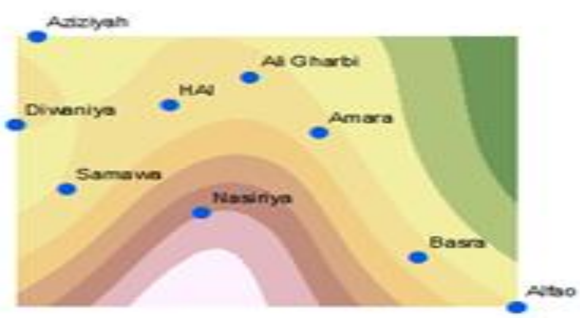

Figure (3): The Spline method for T-max

Figure (4): The Kriging method for T-max

Figure (5): The IDW method for T-mini

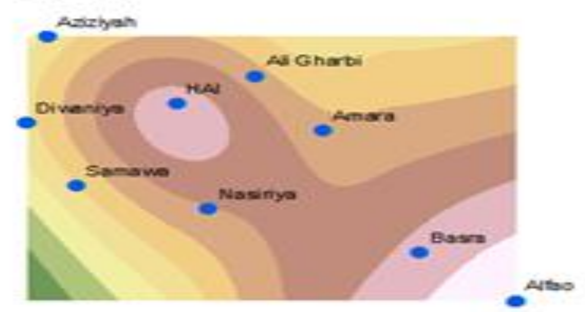

Figure (6): The Spline method for T-mini

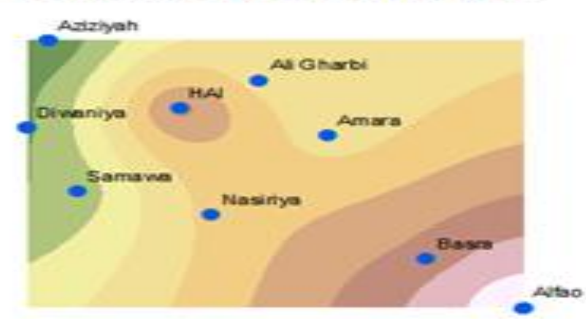

Figure (7): The Kriging method for T-mini

Table 2: Error statistics for 3 methods using climate data (maximum temperatures)

Method	RMSE	RMSE %	Bias	E
IDW	4.6	0.01%	1.17	0.9
Spline	4.7	0.01%	1.08	0.9
Ordinary Kriging	4.2	0.01%	1.17	0.9

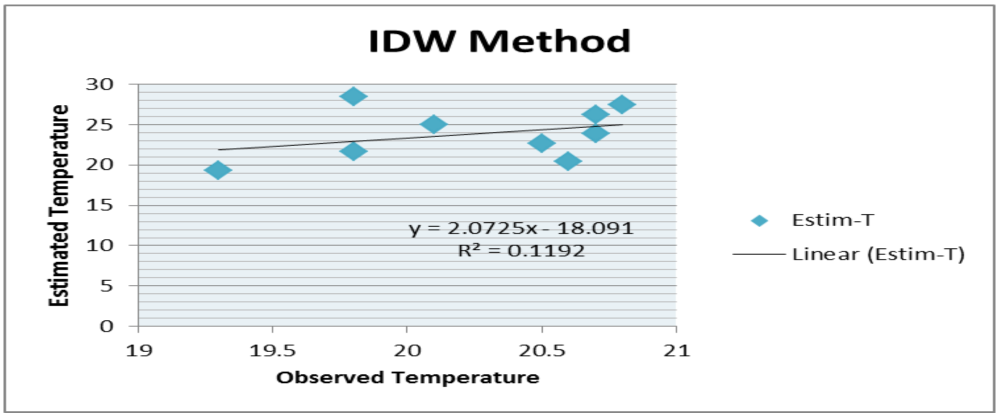

Fig.8: The IDW, comparison of estimated and observed mean Tmax.

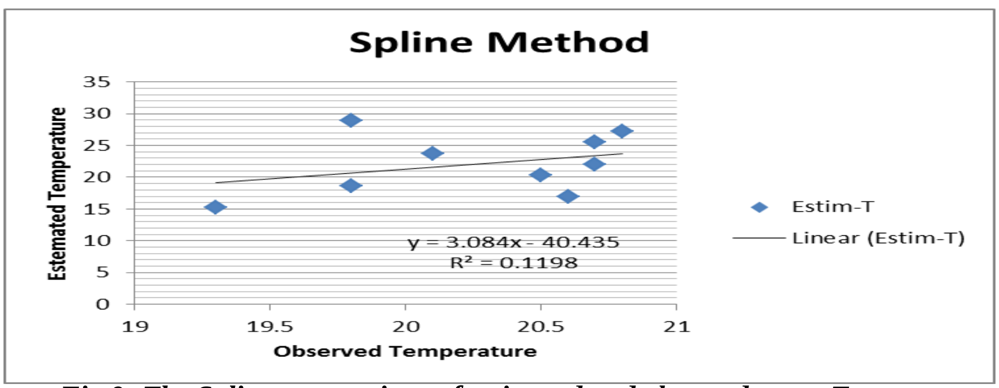

Fig.9: The Spline, comparison of estimated and observed mean Tmax.

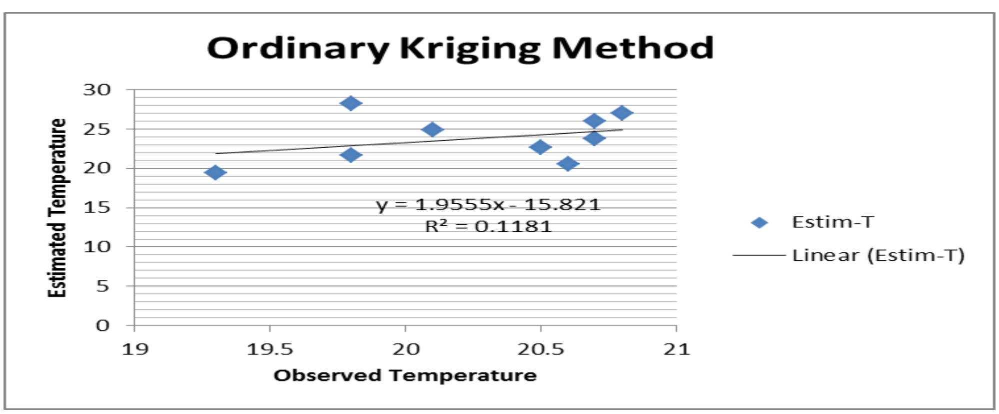

Fig.10: Ordinary kriging, comparison of estimated and observed mean Tmax.

5-Results for mean minimum temperature at February -2008

Table 3: Error statistics for 3 techniques using climate data (minimum temperatures)

Method	RMSE	RMSE %	Bias	E
IDW	0.655	0.01%	1	0.9
Spline	0.693	0.01%	0.98	0.9
Ordinary Kriging	0.656	0.01%	1	0.9

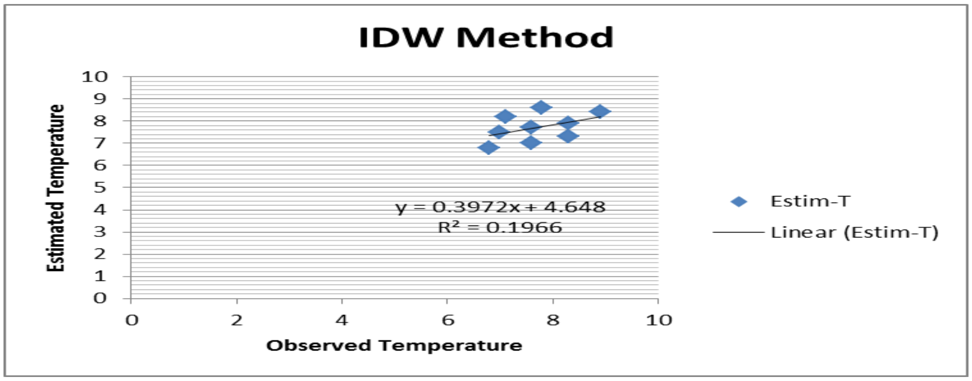

Fig. 11: The IDW, comparison of estimated and observed mean Tmini.

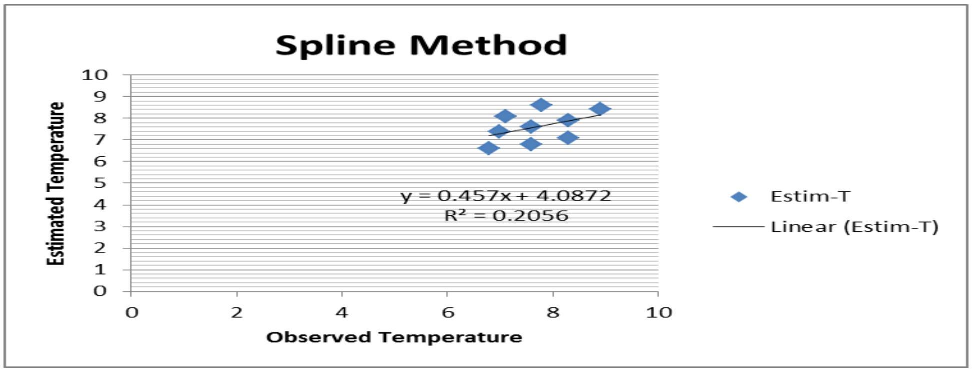

Fig.12: The Spline, comparison of estimated and observed mean Tmini.

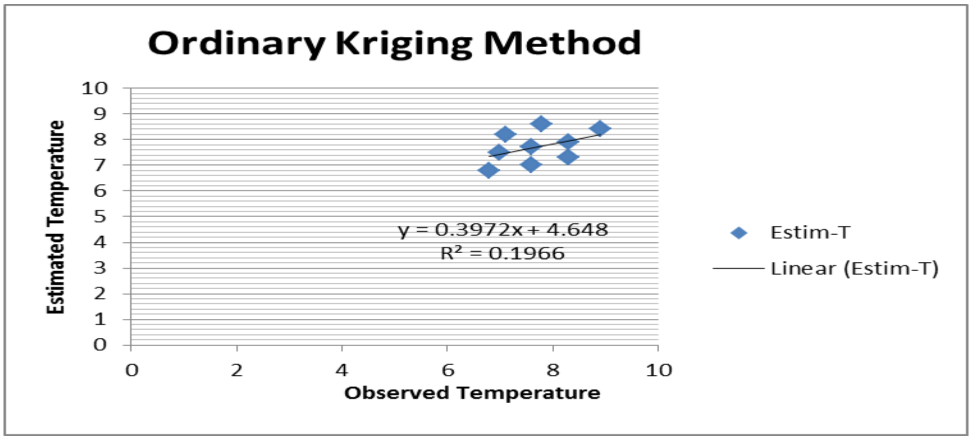

Fig.13: Ordinary kriging, comparison of estimated and observed mean Tmini.

Conclusion

High estimation error for all three interpolation methods are obtaining for maximum temperatures by using four statistics, this mean that high error implies the data became exactly interpolated and these methods are unstable and needed additional data. This should be expected that only 9 data points are used over so large an area.

Statistical results of minimum temperature by using kriging methods are the best (which has the lower estimation error) followed by IDW and then Spline.

In general, the points of samples must be well distributed on the study area and when the number of samples is greater the results will be better.

References

[1] B.Li, "Spatial Interpolation of Weather Variables Using Artificial Neural Networks", the University of Wuhan, China, 2002.

[2] J. Claireh, and S. Neil, "A Comparison among Strategies for Interpolating Maximum and Minimum Daily Air emperatures. Part II: The Interaction between Number of Guiding Variables and the Type of Interpolation method", Department of Geography, University of Edinburgh, Edinburgh, Scotland, 2001.

[3] B. William, Journal of Spatial Hydrology, Fall, 5 (2011) 2.

[4] M. G. Tewolde, T. A. Beza, and A. C. Costa, Comparison of Different Interpolation Techniques to Map Temperature in the Southern Region of Eritrea, Instituto Superior de Estatistica e Gestao de Informacao, Unevirsidade Nova de Lisboa, 2010.

[5] I. Jiri Cajtham, Spatial Modeling of Climate, Master, Study rogramme: Geodesy and Cartography, Branch of Study: Geodesy and Cartography, Faculty of Civil Engineering, Department of Mapping of Climate, Ph.D, Thesis, Czech of Technical University, 2010.

[6] L. Jin, and A. D. Heap, A Review of Spatial Interpolation Methods for Environmental Scientists, Geoscience Australia, GPO Box 378, 2008.

[7] T. Zaria, A Comparison of Thiessen-Polygon, Kriging, and Spline Models of UV Exposure, GIS Research Laboratory, Department of Geography, College of Letters, Arts and Sciences, University of Southern California, 2008.

[8] A. Sharolyn, An evaluapatial of spatial interpolation methods on air temperature in phoenix, AZ, Department of Geography, Arizona State University, Tempe, AZ 85287-0104, 2009.

[9] R. Sluiter, "Interpolation methods for climate data Literature review", KNMI,R&D Information and Observation Technology, De Built, 19 November, Intern rapport; IR 2009-04, 2008 .

[10] Y. Knight, B.Yu, G. Jenkins, and K.Morris, Comparing rainfall interpolation techniques for small subtropical urban catchments, School of Environmental Engineering, Australia 2010.