ANALYSIS OF LAND SURFACE TEMPERTURE AND PREDICTING THE UNKNOWN VALUES USING SPATIAL INTERPOLATION METHODS

Parvathi.R
School of Computing Sciences and Engineering,
VIT University- Chennai Campus,
Chennai-640 127, Tamil Nadu, India
Parvathi.r@vit.ac.in, parvathyraman@gmail.com

Abstract:

This paper compares five spatial interpolation methods of Kriging-exponential, Krigingspherical,Kriging-Gaussian with different weighting values are unity and 1/Standard Deviation in temperature interpolation,utilizing GIS software spatial analysis functions. The research data are the minimum temperature, mean air temperatures and maximum temperature of the first ten days of January, April, July and October from 1990 to 2005 in 14 weather stations of Alaska. The result shows that Kriging-doublespherical,Matern and Gaussian interpolation methods are the highest-accuracy methods for unity weightage, Matern and Gaussian interpolation methods are highest-accuracy methods for 1/Standard Deviation weightage and Matern with unity and 1/variance weightage is best suitable interpolation methods for the spatial data

Keywords: Data mining, Spatial Interpolation, Temperature

I. INTRODUCTION

Interpolation is a method of getting new data from some known data points. Researchers have adopted lots of spatial interpolation methods in practical studies, such as of Kriging-exponential, Kriging spherical, Kriging-Gaussian.

Air temperature is a vital meteorological element, which is often used in engineering and science work. In Alaska, many weather data is from official departments and there are not many weather sites, some weather data is difficult to obtain in some areas, so people use interpolation methods to get some air temperature data [1-11]. Though in slim number studies, people compared some interpolation methods [7-11], there are rarely study on comparison of different interpolation methods when interpolating temperature in Alaska.

II.INTERPOLATION METHODS AND DATA PROCESS

Interpolation Methods

Kriging

Kringing was developed by a French mathematician Georges Matheron, and it's described as follows,

$$Z = \sum_{i=1}^{n} \lambda_i Z(x_i) \tag{1}$$

In ordinary Kriging, the weight *I* λdepends on a fitted model to the measured points, the spatial relationships among the measured values around the predicted location, and the distance to the predicted location [12]. These characters can be obtained from variogram. Kriging adopts "semi-variogram function" to describe the spatial structure of variable, which is shown in the following formula:

$$\gamma_i(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) + Z(x_i + h)]^2 \tag{2}$$

where $N(h)$ is the number of points pair, Z is the interpolated variable, h is distance between two interpolated points, xi is the starting location, and $xi + h$ is the ending location [13-16]. The graph will show the calculated variogram. The dots are the calculated value, and the blue line is the current estimate of the model. The colour represents the number of pairs for variance estimate at each lag. The colour ranges from pink (smallest no. pairs) to blue (highest no. pairs) as indicated by the legend on the bottom right hand-side.

The ordinary Kriging module includes four semivariogram models; this paper only focuses on three of them, including Exponential Model, Spherical Model and Gaussian Model.

Exponential Model (EM):

$$\gamma(h) = c_0 + c[1 - e^{\frac{-h}{r}}] , \gamma(0) = 0 \quad (3)$$

Spherical Model (SM):

$$\gamma(h) = c_0 + c[\frac{3h}{2a} - \frac{1}{2}(\frac{h}{a})^3] \quad 0 < h \leq a \quad (4)$$

Gaussian Model (GM):

$$\gamma(h) = c_0 + c[1 - e^{\frac{k^2}{r^2}}] , \gamma(0) = 0 \quad (5)$$

$c0$, c above are unknown parameters that should be determined by least squares.`

Local/Global variogram

Global variogram refers to calculating a variogram for the whole area. And kriging uses this whole area variogram for prediction. Local variogram is intended for field with high data density, where "local" variogram was calculated for each interpolation point.

SPHERICAL

if (h < A1) then

rho = 1- 1.5 h/A1 + 0.5 * (h/A1)^3

else

rho = 0

endif

gamma = C0 + C1 * (1- rho)

EXPONENTIAL

rho = exp(-h/A1)

gamma = C0+ C1 * (1 - rho)

GAUSSIAN

rho = exp(-(h/A1)^2)

gamma = C0+ C1 * (1 - rho)

LINEAR WITH SILL

if(h < A1) then

rho = 1- (h/A1)

else

rho = 0

end if

gamma = C0+ C1 * (1 – rho)

STABLE

rho = exp[-(h/A1)^{alfa}]

gamma = C0+ C1 * (1 – rho)

(0<alfa<2)

GENERALISED CAUCHY

rho = (1 + (h/A1)^2)$^{-alfa}$

gamma = C0+ C1 * (1 – rho)

(alfa>0)

MATERN

rho = $1/[2^{(SMOOTH-1)} * \Gamma(SMOOTH)] * (h/A1)^{SMOOTH} * Bess_{SMOOTH}(h/A1)$

gamma= C0 + C1*(1 –rho)

where

Γ (…) is Gamma function,

$Bess_{SMOOTH}$(…) is the modified Bessel function of the third kind of order smooth.

(0<SMOOTH<2)

Matern is a general model that is flexible and can be used to approximate function behaving as exponential (smooth = 0.5), power, or Whittle (Bessel function) model (smooth = 1).

DOUBLE_SPHERICAL

if (h < A2) then

rho1 = $1-1.5*h/A1+0.5*(h/A1)^3$

rho2 = $1-1.5*h/A2+0.5*(h/A2)^3$

if (h > A1) then

rho1=0

end if

else

rho1 = 0

rho2 = 0

end if

gamma = C0+C1*(1-rho1)+C2*(1-rho2)

DOUBLE_EXPONENTIAL

rho1 = exp(-h/A1)

rho2 = exp(-h/A2)

gamma = C0+C1*(1-rho1)+C2*(1-rho2)

For local variogram, most crop yield data can be fitted with spherical and exponential model. The recommended model for local variogram is the exponential model, Gaussian model is not recommended as it can produce unstable kriging equation.

Weight for Fitting Variogram

Variogram model is fitted to the data by using weighted nonlinear least-squares method (Jian et al., 1996), minimising:

$$R = \sum_{i=1}^{n} w_i \left[\hat{\gamma}(h_i) - \hat{\gamma}^*(h_i) \right]^2 \quad (6)$$

User can specify the type of weighting for *w*: Unity (no weighting) No. of pairs, no. of pairs calculated from semivariance $N(h)$ 1/std.dev, the standard deviation of the average of semivariance for particular lag.

No_pairs/std_dev, combination of no. of pairs & std. deviation of the semivarainace estimate.

The goodness of fit can be assessed by the SSE (sum of squared error) or AIC (Akaike Information Critereon). The lowest AIC pertains to the best model (Webster and McBratney, 1989).

AIC is defined as:

AIC = -2 ln(maximum likelihood) + 2 (number of parameters), and is estimated by:

AIC = $n \ln(R) + 2p$ where R is the sum of squares of residuals, and p is the number of parameters.

Data sets

The research data is provided by Alaska Meteorological Administration, including the Temperature from January to June (T_F_JAN), Temperature from July to December(T_F_JULY) and Mean_Temperature in 14 meteorological stations of Alaska, and longitude, latitude and altitude of the Weather sites. The sites involved are located in east longitude 75.98°E to 131.98°E degrees, north latitude 18.23°N to 53.47°N.

Data Process

Step1: The raw data is text format, so those text-format data of 14 weather station (latitude, longitude, altitude, and temperature values) are changed into shp format in ArcView.

Step 2: The temperature is always affected by the elevation, and many researchers interpolated the temperature combing with DEM data. In order to guarantee all sites are on the same level, the elevation effect is removed on all interpolation sites before interpolation, and the method is described as follows,

$$T_a = T_b + \frac{E}{100} \times 0.6 \times 10 \tag{7}$$

where *Ta* is the temperature value after removing the elevation effect, *T*b is the temperature value before removing the elevation effect, E is elevation value.

Step 3: Interpolation methods for the given dataset are carried out by VSEPER software and corresponding graphs are generated.

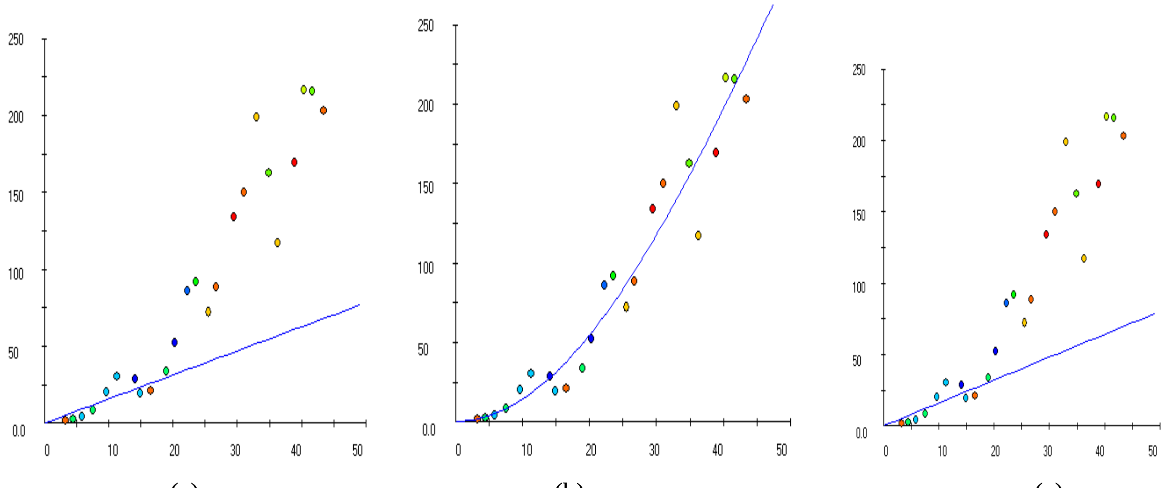
Step 4: Integrating with the elevation information, the elevation effect is finally considered on the temperature data obtained from step 4, so the final interpolation results of all interpolation sites are got. The process is described as follows,

$$T_f = T_a - \frac{E}{100} \times 0.6 \times 10 \tag{8}$$

where *Tf* is the final interpolation temperature value of points, *T*a is the value got in formula (11), E is elevation value.

III RESULTS AND ANALYSIS

Mean Absolute Error (MAE) and Root Mean Squared Interpolation Error (RMSIE) are chosen as the two indicators that evaluate the precision of the interpolation methods, the result is shown as follows (the values are in Celsius), Figure 1(a-l) shows the results of Exponential, Gaussian, Spherical, Double Exponential and Double Spherical interpolation methods with Unity and 1 / standard deviation as the weightage values for the given Alaska climate dataset respectively. The x axis to graphs represents the temperature and y axis of the graphs represents the number of days for the each slot like January to June, July to December. The data represented in the graphs is the mean temperature of the particular day



(a) (b) (c)

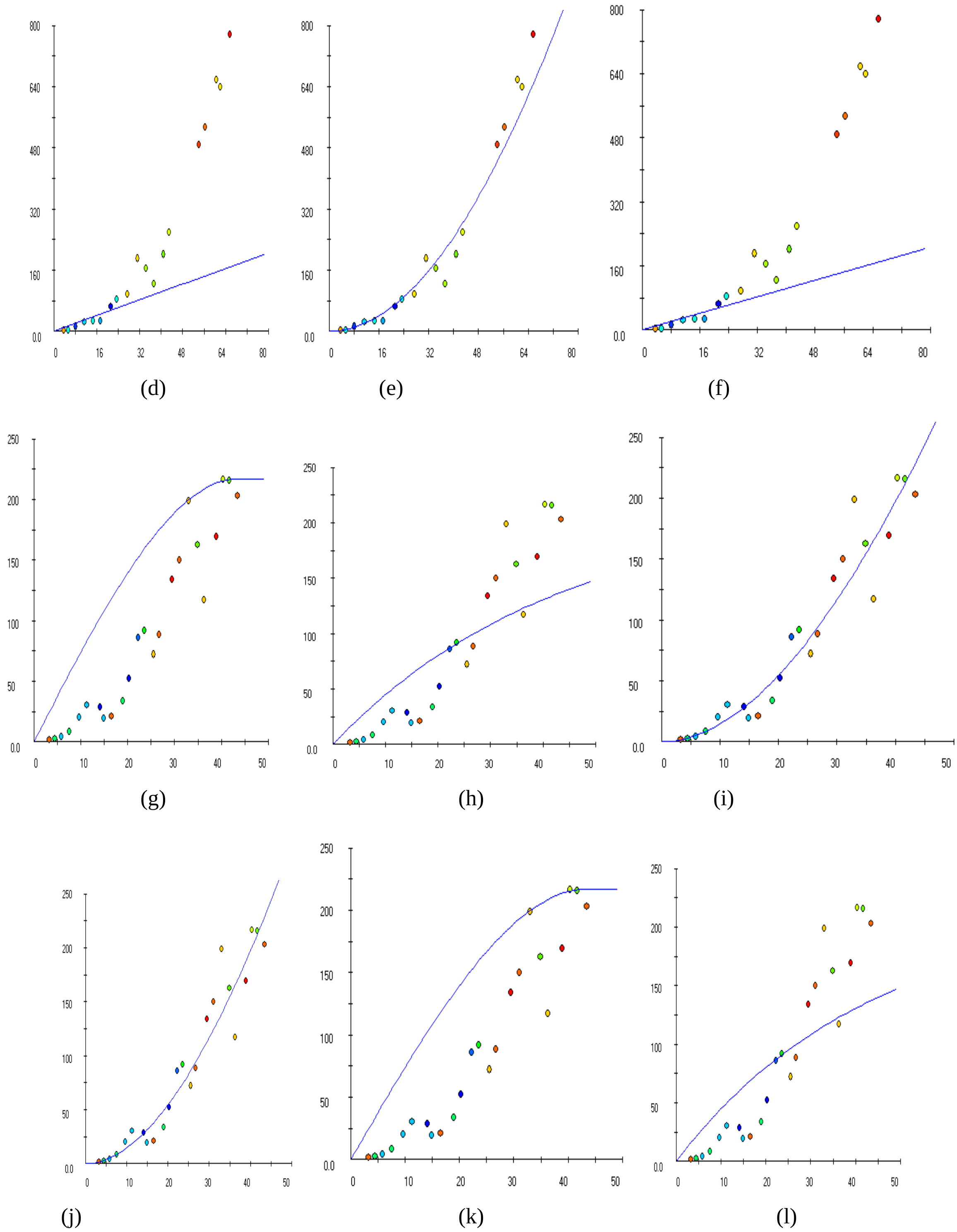


Figure 1.(a – l) Spherical, Double Exponential and Double Spherical interpolation methods with Unity and 1 / standard deviation as the weightage values for the given Alaska climate dataset

Table 1 Analysis of Spatial Interpolation Methods

Model	Weighting	Parameters				Unweighted			Weighted
		C0	C1	A1	C2/ Alpha	SSE	RMSE	AIC	SSE
Exponential	Unity	0	9891.3	6166.1		138663	76.01	290.2	124.5
Gaussian	Unity	0.2409	829.0	76.72		9967.8	20.38	227.0	18.96
Spherical	Unity	0	9985.2	9328.2		138121	75.86	290.1	124.3
Double Spherical	Unity	15.27	1.082	10000	2.942	8580.9	46.32	44.23	10.37
Double Exponential	Unity	0	0.0000100	50000	217.0	42354	42.01	265.7	866.3
Matern	Unity	0.05712	10000	214.9	1.343	10227	20.64	229.6	18.94
Exponential	1/Variance	0	9636.0	3723.9		1187969	250.0	271.8	323.6
Gaussian	1/Variance	0.3215	10000	253.5		50709	51.66	211.8	25.92
Spherical	1/Variance	0	10000	5789.3		1180502	249.3	271.6	322.4
Double Spherical	1/variance	0	0.0000100	22780	217.0	92941	62.23	284.6	2758.9
Double Exponential	1/variance	0	0.0000100	50000	217.0	42354	42.01	265.7	866.3
Matern	1/variance	0.05712	10000	214.9	1.343	10227	20.64	229.6	18.94

IV CONCLUSION

To select the optimal method in this paper, six interpolation methods – Exponential method,Gaussian method,Spherical method, Double Spherical method,Double Expontential method and Matern method with different weighting values Unity and 1/standard deviation -were compared, and then the optimal interpolation method was used to give the spatial distribution of temperature dataset. We draw the following conclusions from this study: Kriging-doublespherical, Matern and Gaussian interpolation methods are the highest-accuracy methods for unity weightage, Matern and Gaussian interpolation methods are highest-accuracy methods for 1/Standard Deviation weightage and Matern with unity and 1/variance weightage is best suitable interpolation methods for the spatial data

REFERENCES

[1] Hubert Hasenauer, Katarina Merganicova, and Richard Petritsch.“Validating daily climate interpolations over complex terrain in Austria,“ Agricultural and Forest Meteorology,vol. 119, pp. 87-107, 2003.

[2] Hong Yan, “Spline Interpolation of Spatial-temporal Climate Data for China,” Geography and Geo-Information Science,vol. 19, no. 5, pp. 27-31, 2003.

[3] Patrizia Basili, Stefania Bonafoni, Piero Ciotti, Frank Silvio Marzano, Giovanni d’Auria, and Nazzareno Pierdicca, “Retrieving Atmospheric Temperature Profiles by Microwave Radiometry Using a Priori Information on Atmospheric Spatial-Temporal Evolution, “Transactions on Geoscience and Remote Sensing, vol. 39, No. 9, pp. 1896-1904, September 2001

[4] P.A.Hancock, M.F.Hutchinson, “ Spatial interpolation of large climate data sets using bivariate thin plate smoothing splines, “ Environmental Modelling & Software, vol. 1, no.11 , pp. 1-11,2005.

[5] Shuming Fang, Jiangwei Qin, Yongfei Li, and Taibao Yang, “Method of spatial interpolation of air temperature based on GIS in Gansu Province, ” Journal of Lanzhou University (Natural Sciences), vol. 41, no. 2, pp. 6-9, 2005.

[6] Shunbao Liao and Zehui Li, “Some Practical Problems Related to Rasterization of Air Temperature,” Meterological Science and Technology, vol.32, no. 5, pp. 352-356, 2004.

[7] ANDERSON Sharolyn, “ An Evaluation of Spatial Interpolation Methods on Air Temperature in Phoenix, AZ, “ Online paper at http://www.ucgis.com from UCGIS Summer Assembly 2001 Program, July 2001.
[8] Eischeid, J.K., P.A. Pasteris, H.F. Diaz, M.S. Plantico, and N.J. Lott, “ Creating a serially complete, national daily time series of temperature and precipitation for the western United States, ” J. Appl. Meteorol., vol.39, pp. 1580-1591, 2000.
[9] Fred C. Collins and Paul V. Bolstad, “A Comparison of Spatial Interpolation Techniques in Temperature Estimation,” . Third International Conference/Workshop on Integrating GIS and Environmental Modeling, 1996
[10] Jarvis, C.H. and N. Stuart. “A comparison among strategies for interpolating maximum and minimum daily air temperatures. Part II: The interaction between number of guiding variables and the type of interpolation method,” J. Appl. Meteorol, vol. 40, pp. 1075-1084, 2001.
[11] T. Ishida and S. Kawashima, “Use of cokriging to estimate surface air temperature from elevation, ” Theoretical and Applied Climatology, vol.47, no. 3, pp. 147-157, 2004.
[12] Volker Janssen, Linlin Ge and Chris Rizos. “Troposheric delay corrections to differential INSAR results from GPS obeservations,” 6th International Symposium on Satellite Navigation Technology Including Mobile Positioning and Location Services (SatNav), pp.1-12, July 2003.
[13] Geol N S and Norman J M. “Biospheric models, measurements,and remote sensing of vegetation,” International Journal of Remote Sensing, Vol. 47, pp. 163-188, 1992.
[14] Huizan Wang, Ren Zhang, Kefeng Liu, and etc, “Improved Kriging Interpolation Based on Support Vector Machine and Its Application in Oceanic Missing Data Recovery,” 2008 International Conference on Computer Science and Software Engineering, pp.726-729, 2008.
[15] Roth M, Oke T R, and Emery W J. “Satellite derived urban heat islands from three coastal cities and the utilization of such data in urban climatology,” International Journal of Remote Sensing, Vol.10, pp. 1699-1720, 1989.
[16] Wenhui Zhao, Huili Gong, and Zhaoning Gong, “A preliminary study of urban airborne inhalable particle spatial distribution and their mechanism in Beijing using RS and GIS,” IGARSS 2008, pp. 1366- 1369, 2008.