International Journal of Scientific and Innovative Mathematical Research (IJSIMR)
Volume 2, Issue 6, June 2014, PP 530-536
ISSN 2347-307X (Print) & ISSN 2347-3142 (Online)
www.arcjournals.org

Kriging Approach for an Epidemic Data Analysis

D. Sheeba Singh

Assistant Professor
Department of Mathematics, Noorul Islam University,
Kumaracoil, Kanyakumari Dist, Tamilnadu, India
dssheeba@gmail.com

S. Saranya

M. Phil Scholor,
Department of Mathematics,
Noorul Islam University, Kumaracoil,
KanyakumariDist, Tamilnadu, India
saranya.10590@gmail.com

M. Immaculate Mary

Professor
Department ofMathematics,
Noorul Islam University, Kumaracoil,
KanyakumariDist, Tamilnadu, India
imma_arul@yahoo.com

Abstract: *In this study, Kriging methods namely Ordinary Kriging (OK), Simple Kriging (SK), Universal Kriging (UK), Poisson Kriging (PK) and Bayesian Kriging (BK) were used to predict the number of people affected by dengue in the four Taluks of Kanyakumari District, Tamilnadu,India. The predictions by the above said Kriging methods were close to the actual number of people affected by dengue in all the areas taken for the study. The results of the study reveal that the Kriging methods can be used to predict the people affected by dengue in different locations.*

Keywords: *Ordinary Kriging, Simple Kriging, Universal Kriging , Poisson Kriging and Bayesian Kriging*

1. INTRODUCTION

Kriging is a stochastic model that provides estimates for accuracy in predictions. It is a commonly used method of interpolation for spatial data. That is, it is a method for interpolating the value of a random field at an unobserved location based on available surrounding measurements. Daniel Krige, the doyen of Geostatistics has developed empirically statistical methods to predict ore grades from spatially correlated sample data in the gold mines of South Africa. In 1960s his approach was formalized by Matheron, and the term 'Kriging' was coined in his honour. Over the past several decades, Kriging has become a fundamental spatial prediction tool in Geostatistics. The technique of Kriging is used extensively in spatial interpolation problems and has found many applications in geology as well as ecology and environmental sciences. Lina Lundberg proposed Kriging as a method for spatial prediction that gave the best unbiased linear predictor based on observations of a spatial process [1]. ZuErich et al. compared design-based and Kriging techniques for the estimation of spatial averages in the context of double sampling, as used in forest inventory, where terrestrial sample plots were combined with auxiliary information based on aerial photographs[2]. Angeles Saavedra Gonzalez adapted the Universal Kriging method of prediction to context with uncertainty and imprecise data[3]. Ebenezer Bonyah, et al. adopted Kriging approach as methodology and provided a spatial analysis of breast cancer incidence in the Ashanti region area of Ghana during the period of 2010 to 2011[4]. Srinivasan R. and Venkatesan P studied the Bayesian Kriging of tuberculosis within the wards of Chennai in India[5]. They proved the spatial analysis to be more useful for studying the spread of tuberculosis analysis and modeling of disease prediction. Allan D. Woodbury adapted the full Bayesian approach to estimate the transmissivity from hydraulic head and transmissivity measurements were developed for two-dimensional steady state groundwater flow[6].Sujit K. Sahu and et.al analyzed modeling rainfall using a Bayesian Kriged-Kalman model[7]. In this study, Kriging methods namely Ordinary Kriging (OK), Simple Kriging (SK), Universal Kriging (UK), Poisson Kriging (PK) and Bayesian Kriging (BK) were used to predict the number of people affected by the epidemic disease dengue in the four Taluks of Kanyakumari District of Tamilnadu in India.

2. ORDINARY KRIGING (OK)

In Ordinary Kriging, a probability model is used in which the bias and the error variance can be calculated. Then the weights are chosen for the nearby samples that ensure that the average error for our model is exactly zero, and the modeled error variance is minimized.

Kriging refers to making inferences on unobserved values of the random process $Z(\cdot)$ or of $S(\cdot)$ given by $Z(s)=S(s)+\varepsilon(s),\quad s\in D$ from data $Z\equiv(Z(s_1),\ldots,Z(s_n))'$ observed at known spatial locations $\{s_1,\ldots,s_n\}$. The generic predictor of $g(Z(\cdot))$ or $g(S\cdot)$ by $p(Z;g)$ should be denoted. Ordinary Kriging refers to spatial prediction under the following two assumptions.

Model Assumption

$Z(s)=\mu+\delta(s),\qquad s\in D, \mu\in R,$ and μ unknown.

Predictor Assumption

$$p(Z;B)=\sum_{i=1}^{n} w_i Z(s_i),\qquad \sum_{i=1}^{n} w_i = 1.$$

Let $\hat{Z}_0$ be the estimate of the value of Z at the point s_0. this estimate is a random variable since it is a weighted linear combination of the random variable at the sample locations.

$$\hat{Z}(s_0)=\sum_{i=1}^{n} w_i Z(s_i)$$

The estimation error is also a random variable and $Z(s_0)$ is stationary.

$$R(s_0)=\hat{Z}(s_0)-Z(s_0),\ \operatorname{var}(Z(s_0))=\sigma^2$$

$$\operatorname{var}(R(x_0))=\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j C_{ij}+\sigma^2-2\sum_{i=1}^{n} w_i C_{i0}\text{ , where } C_{ij}=Cov(Z_i,Z_j).$$

3. SIMPLE KRIGING (SK)

A general prediction model that divides a continuous spatial process into first and second order components: $Y(s)=X(s)\beta+\varepsilon(s)$ in which it is assumed that both the trend (mean) components are known and the covariance function is known. Simple Kriging is valid for this model because the residuals are zero mean, constant variance, per assumption of the model.Let us observe some stationary (WSS) random field Z (s) at some points s_j, j = 1, 2 … n. First, assume that the mean m and covariance function $C(\cdot)$ of this process are known. The case of prediction with the known mean is often called Simple Kriging. Let $\hat{Z}_0$ be the estimate of the value of Z at the point s_0. For simplicity, denote $Z_J=Z(s_j)$ and $\mathrm{C}(\mathrm{i},\mathrm{j})=\mathrm{Cov}(\mathrm{Z_i},\mathrm{Z_j})$. We may assume that the mean m = 0. The Kriging estimate and the Kriging variance of the datasets are computed as

$$\hat{Z}(s_0)=\sum_{i=1}^{n} w_i Z(s_i)$$

$$\hat{\sigma}^2(s_0)=C(0)-\begin{bmatrix} w_1\\ \vdots\\ w_n\end{bmatrix}\times\begin{bmatrix} c_{10}\\ \vdots\\ c_{n0}\end{bmatrix}$$

where n is the number of observations.

4. Universal Kriging (Uk)

The Ordinary Kriging that was discussed earlier is based on the constant mean model given by $Z(s) = \mu + \varepsilon(s)$, where $\varepsilon(\cdot)$ has mean zero and variogram $2\gamma(\cdot)$. The mean can be a function of the coordinates X,Y, in some linear, quadratic, or higher form. For example the value of Z at location s can be expressed now as

$Z(s_i) = f_0 + f_1 x_i + f_2 y_i + \varepsilon(s_i)$, linear

Or

$Z(s_i) = f_0 + f_1 X_i + f_2 Y_i + f_3 X_i^2 + f_4 X_i Y_i + f_5 Y_i^2 + \varepsilon(s_i)$, quadratic, etc.

As with Ordinary Kriging, to find the weights when a trend is present, we need to minimize the mean square error (MSE)

$$\min\left(Z(s_0) - \sum_{i=1}^{n} w_i Z(s_i) \right)^2$$

subject to the constraints $\sum_{i=1}^{n} w_i X_i = X_0,\ \sum_{i=1}^{n} w_i Y_i = Y_0,\ \sum_{i=1}^{n} w_i = 1$

The Kriging estimate and the Kriging variance are given by

$$\hat{Z}(s_0) = \sum_{i=1}^{n} w_i Z(s_i)$$

$$\hat{\sigma}^2(s_0) = C(0) - \begin{bmatrix} w_1 \\ \vdots \\ w_n \\ \mu_0 \\ \mu_1 \\ \mu_2 \end{bmatrix} \times \begin{bmatrix} c_{10} \\ \vdots \\ c_{n0} \\ 1 \\ f_1(u_0) \\ f_2(u_0) \end{bmatrix}$$

5. Poisson Kriging (Pk)

In Poisson Kriging, the diagonal of the matrix where an error variance term, $m*/n(s_i)$, is added to each of the covariance term C_{ii}. m* is the population-weighted mean of the rate dataset, and $n(s_i)$ is the population at that location. Let s_i represent the centroid coordinates for each areal supports, v_α and *z*(s) denote the (unknown) point value of the attribute *z* at location s within a study region *D*. Area-to-point spatial interpolation is to predict any point value *z*(s) using *K* areal data $\{z(v_i), i = 1, \ldots, K\}$.

$$z^*(s) = \sum_{i=1}^{K} w_i(s) z(v_i)$$

where the areal supports v_i are disjoint and the prediction locations are arbitrary. The weights $w_i(s)$ are computed to ensure the minimization of prediction mean square error under the condition of the unbiasedness of *z** (s).

The variance is calculated as

$$\sigma_{PK}^2 = C(0) - \sum_{i=1}^{K} w_i(s) C(v_i, s) - \mu(s)$$

where $C(0)=Var(Z(s))$ and $C(v_i,s)$ is the covariance between area v_i and location s and is inferred from the experimental semivariogram by using $\hat{\gamma}(h)=C(0)-C(h)$ when the variance $Var(Z(s))$ is finite.

6. BAYESIAN KRIGING(BK)

In Bayesian Kriging, there is an additional assumption that there exists prior knowledge on the value of the coefficients in the trend. This prior knowledge is specified as a prior multi Gaussian distribution for the parameters, so that expectation and variance must be specified. The basic model of a simple spatial process is

$$Z(s)=\mu(s)+\omega(s)+\varepsilon(s)$$

where $\omega(s)$ is the spatial variation and $\varepsilon(s)$ is the nonspatial variation. Here $\mu(s)=X^T\beta\equiv\mu$ constant is the expectation of Z(s), $\omega(s)=f(\sigma^2,\phi)$ is the partial sill and $\varepsilon(s)=f(\tau^2)$ is the nugget which represents the error at the sample location s. It has zero mean and variance τ^2 at the sample location.In Bayesian modeling, the conditional distribution is

$$Z\mid\theta\sim MVN\left(\mu,\sigma^2H(\phi)+\tau^2I\right)$$

with correlation matrix $H_{ij}=\rho(s_i-s_j;\phi)$

$$=e^{-\phi\|s_i-s_j\|}$$

and prior distributions for $\mu,\sigma^2,\phi\,\&\,\tau^2$ are viz. for μ - normal, τ^2,σ^2 - inverse gamma and ϕ - Gamma. $\mu,\sigma^2,\tau^2\,\&\,\phi$ are model parameters, Normal, Inverse gamma and Gamma are prior distributions.

Posterior distribution is $\pi_1(\theta\mid z)=\dfrac{f(z\mid\theta)\,\pi_0(\theta)}{\int f(z\mid\theta)\,\pi_0(\theta)\,d\theta}$

where $f(z\mid\theta)$ is the likelihood function and $\pi_0(\theta)$ is the prior distribution.

$$\pi_1(\mu,\sigma^2,\phi,\tau^2\mid z)\,\alpha\; f(z\mid\mu,\sigma^2,\phi,\tau^2)\pi_0(\mu,\sigma^2,\phi,\tau^2)$$
$$=f(z\mid\mu,\sigma^2,\phi,\tau^2)\pi_0(\mu)\pi_0(\sigma^2)\pi_0(\phi)\pi_0(\tau^2)$$

The prediction at new location is given by

$$f(z_0\mid z)=\int f(z_0,\theta\mid z)d\theta$$
$$=\int f(z_0\mid z,\theta)\pi_1(\theta\mid z)d\theta$$

The spatial correlation matrix needs to be extended to include the new locations

$$Z,Z_0\mid\theta\sim MVN\left(\mu,\sigma^2H(\phi)+\tau^2I\right)$$
$$Z\mid\theta,W\sim MVN\left(\mu+W,\tau^2I\right)$$

with $W\mid\sigma^2,\phi\sim MVN\left(0,\sigma^2H(\phi)\right)$

with prior distributions for μ and τ^2 and for the hyperparameters σ^2 and ϕ.

7. NUMERICAL STUDY

In Kanyakumari district there were four taluks viz. Agastheeswaram, Kalkulam, Thovalai, and Vilavancode. The dengue patients reported from the health centres of the above four taluks from the year 2009 to 2012 were taken for the study. The distance of the Primary Health Centres selected in every Taluk were also calculated from the common axes of reference.The number of dengue patients during 2009, 2010, 2011 and 2012 in the selected primary health centres of the four Taluks were calculated. The different types of Kriging viz. Ordinary Kriging, Simple Kriging, Universal Kriging, Poisson Kriging and Bayesian Kriging were applied to the data. The people affected by dengue in other areas were predicted by using the various Kriging techniques. The 95% prediction intervals were also computed for each taluk. For the purpose of knowing the accuracy of prediction the actual data available in the different areas were collected and are compared with the predicted value.

Predicted Number of Dengue Affected Persons using Different Kriging Methods:

Table 1. *VilavancodeTaluk*

Kriging Types	Predicted number of affected Patients in				Kriging Variance			
	2009	2010	2011	2012	2009	2010	2011	2012
OK	17	15	16	15	7.276	2.805	7.092	6.843
SK	21	19	20	19	7.851	3.023	7.641	7.373
UK	39	20	25	31	13.55	10.64	16.61	20.62
PK	17	15	16	15	7.28	3.899	7.136	6.955
BK	13	14	14	13	5.896	2.162	6.432	5.194
Actual data	12	14	13	15				

Table 2. *KalkulamTaluk*

Kriging Types	Predicted number of affected Patients in				Kriging Variance			
	2009	2010	2011	2012	2009	2010	2011	2012
OK	11	13	9	11	1.926	2.052	1.876	1.988
SK	10	13	9	11	1.93	2.057	2.06	1.985
UK	11	12	10	11	14.05	23.89	11.98	7.06
PK	11	13	9	11	0.932	2.073	1.875	2.005
BK	17	17	16	18	0.653	1.939	1.284	1.198
Actual data	18	17	15	17				

Table 3. *AgasteeswaramTaluk*

Kriging Types	Predicted number of affected Patients in				Kriging Variance			
	2009	2010	2011	2012	2009	2010	2011	2012
OK	15	12	11	12	1.074	0.954	1.621	0.528
SK	16	17	16	17	1.092	1.391	1.326	0.772
UK	15	17	14	13	29.74	1.323	29.5	1.027
PK	15	12	11	12	5.14	1.676	1.617	1.709
BK	14	13	17	17	0.789	0.638	1.163	0.276
Actual data	14	15	17	17				

Table 4. *ThovalaiTaluk*

Kriging Types	Predicted number of affected Patients in				KrigingVarience			
	2009	2010	2011	2012	2009	2010	2011	2012
OK	11	12	11	12	0.116	0.116	1.033	0.345
SK	11	11	10	11	0.117	0.117	1.047	0.349
UK	11	12	11	12	0.513	0.402	4.894	2.705
PK	11	12	11	12	0.115	0.115	1.032	0.343
BK	12	11	12	10	0.105	0.098	0.941	0.253
Actual data	13	10	12	11				

Prediction Limits:

Table 5. *VilavancodeTaluk*

Kriging	2009	2010	2011	2012
OK	(11.5,22.07)	(11.44,18.01)	(11.01,21.44)	(10.33,20.6)
SK	(15.23,26.21)	(15.27,22.08)	(14.75,25.59)	(13.63,24.27)
UK	(12.72,65.84)	(0.72,40.98)	(7.29,57.82)	(8.6,72.23)
PK	(11.52,22.09)	(10.86,18.59)	(10.93,21.39)	(10.21,20.55)
BK	(1.444,24.556)	(9.762,18.238)	(1.393,26.607)	(2.819,23.18)

Table 6. *KalkulamTaluk*

Kriging	2009	2010	2011	2012
OK	(7.99,13.44)	(10.12,15.73)	(6.76,12.13)	(8.36,13.89)
SK	(7.62,13.07)	(9.73,15.35)	(6.30,11.93)	(8.05,13.57)
UK	(3.63,18.32)	(2.17,21.33)	(3.05,16.62)	(5.89,16.30)
PK	(8.83,12.61)	(10.09,15.74)	(6.76,12.12)	(8.41,13.96)
BK	(15.72,18.279)	(13.199,20.8)	(13.483,18.517)	(15.652,20.348)

Table 7. *AgasteeswaramTaluk*

Kriging	2009	2010	2011	2012
OK	(12.61,16.67)	(10.15,13.98)	(8.26,13.25)	(10.68,13.53)
SK	(13.46,17.56)	(14.87,19.49)	(13.76,18.27)	(15.63,19.07)
UK	(11.17,17.19)	(15.2,19.71)	(10.99,17.01)	(10.69,14.66)
PK	(4.58,13.47)	(9.52,14.59)	(8.26,13.24)	(9.55,14.67)
BK	(12.454,15.55)	(11.75,14.25)	(14.72,19.28)	(16.46,17.54)

Table 8. *ThovalaiTaluk*

Kriging	2009	2010	2011	2012
OK	(10.7,12.03)	(11.13,12.47)	(9.12,13.1)	(11.02,13.32)
SK	(10.23,11.57)	(10.23,11.57)	(8.12,12.13)	(10.1,12.42)
UK	(10.08,12.89)	(10.51,12.99)	(7.13,15.8)	(8.73,15.17)
PK	(10.7,12.03)	(11.14,12.46)	(9.12,13.1)	(11.02,13.32)
BK	(11.79,12.21)	(10.81,11.19)	(10.16,13.84)	(9.5,10.49)

8. Conclusion

In this study, Kriging methods viz. Ordinary Kriging (OK), Simple Kriging (SK), Universal Kriging (UK), Poisson Kriging (PK) and Bayesian Kriging (BK) were used to predict the number of people affected by the epidemic disease dengue in the four Taluks of Kanyakumari District. From Table 1 to Table 4, the predictions by the above said Kriging methods were close to the actual number of people affected by dengue in all the areas taken for the study. The results of the study reveal that the Kriging methods can be used to predict the people affected by dengue in different locations. Among the Kriging methods applied in this study for prediction, the Bayesian Kriging is the best method as the predicted value is very close to the actual. Also from table 5 to Table 8, the prediction intervals obtained from the study have good coverage probabilities of the true values.

References

[1] Lina Lundberg, Ordinary Kriging Prediction Intervals Based on Semiparametric Bootstrap, Journal of Mathematical Statistics.22, (1999).

[2] Daniel Mentalas and ZuErich, Estimation of the Spatial Covariance in Universal Kriging: Application to Forest Inventory, Journal of Environmental and Ecological Statistics.7, (2000).

[3] Angeles Saavedra Gonzalez, Universal Kriging Method of Prediction to Context with Uncertainty and Imprecise Data, Dyna.80 (178), (2013).

[4] Ebenezer Bonyah, L. Munyakazi, N.N.N. Nsowah-Nuamah, D. Asong and Bashiru I.I. Saeed, Application of Area to Point Kriging to Breast Cancer Incidence in Ashanti Region of Ghana, International Journal of Medicine and Medical Sciences.5(2), (2013).

[5] Srinivasan R. and Venkatesan P, Bayesian Kriging of Tuberculosis in Chennai: A Small Scale Analysis, Indian Journal of Applied ResearchResearch paper of Medical Science.3(7), (2013).

[6] Allan D. Woodbury andTadeusz J. Ulrysch, A Full-Bayesian Approach to the Groundwater Inverse Problem for Steady State Flow, Water Resources Research.36(8), (2000).

[7] Sujit K. Sahu and Kanti V. Mardia,A Bayesian KrigedKalman Model for Short-Term Forecasting of Air Pollution Levels, Applied Statistics.54(1), (2005).

[8] Carlos E. Puente and Rafael L. Bras, Disjunctive Kriging, Universal Kriging, or No Kriging: Small Sample Results with Simulated Fields, Journal of Mathematical Geology. 18(3), (1986).

[9] Dale Zimmerman,ClairePavlik, Amy Ruggles, and Marc P. Armstrong, An Experimental Comparison of Ordinary and Universal Kriging and Inverse Distance Weighting, Journal of Mathematical Geology.31(4), (1999)

[10] DamijanaKastelec and Katarina Kosmelj, Spatial Interpolation of Mean Yearly Precipitation using Universal Kriging, Journal of Spatial Variability and Statistical Studies.26(2), (2002).

[11] M. Matlas, andW.Gonzalez-Manteiga, The Possible Advantages of Relaxing the Universal Kriging, Journal of Mathematical Statistics.40, (2003).

[12] Jooho Choi, Dawn An, and Junho Won, Bayesian Approach for Structural Reliability Analysis and Optimization Using the Kriging Dimension Reduction Method, ASME Journal of Mechanical Design.132(5), (2010).

[13] Leonidas Sakalauskas, Gaussian Fields to Probabilistic Modelling of Multivariate Data", Journal of Geophysical Research. 93, (2013).

[14] MarcinLigas and MarekKulczycki, Simple Spatial Prediction - Least Squares Prediction, Simple Kriging and Conditional Expectation of Normal Vector, Journal of Geodesy and Cartograph.59(2),(2010).

[15] Marvin Karson, Marie Gaudard, Ernst Linder and DebajyotiSinha, Bayesian Spatial Prediction, Journal of Environmental and Ecological Statistics.6, (1999).

AUTHORS' BIOGRAPHY

D.Sheeba Singh, is working as Assistant professor in the Department of Mathematics, Noorul Islam University, Kumaracoil, Kanyakumari District, Tamilnadu.She has published some papers in her area of interest

S. Saranya, is M. Phil Scholor in the Department of Mathematics, Noorul Islam University, Kumaracoil, Kanyakumari District, Tamilnadu.

M. Immaculate Mary,Professor in Mathematics.,presently holds the appointment as Head of the Department of Mathematics of Noorul Islam University, Kumaracoil, Kanyakumari District, Tamilnadu . She has published many papers in her area of interest