

Hierarchical Bayesian space-time interpolation versus spatio-temporal BME approach

I. Hussain, J. Pilz, and G. Spoeck

Department of Statistics, University of Klagenfurt, Klagenfurt, Austria

Received: 21 October 2009 – Revised: 9 February 2010 – Accepted: 13 February 2010 – Published: 30 March 2010

Abstract. The restrictions of the analysis of natural processes which are observed at any point in space or time to a purely spatial or purely temporal domain may cause loss of information and larger prediction errors. Moreover, the arbitrary combinations of purely spatial and purely temporal models may not yield valid models for the space-time domain. For such processes the variation can be characterized by sophisticated spatio-temporal modeling. In the present study the composite spatio-temporal Bayesian maximum entropy (BME) method and transformed hierarchical Bayesian space-time interpolation are used in order to predict precipitation in Pakistan during the monsoon period. Monthly average precipitation data whose time domain is the monsoon period for the years 1974–2000 and whose spatial domain are various regions in Pakistan are considered. The prediction of space-time precipitation is applicable in many sectors of industry and economy in Pakistan especially; the agricultural sector. Mean field maps and prediction error maps for both methods are estimated and compared. In this paper it is shown that the transformed hierarchical Bayesian model is providing more accuracy and lower prediction error compared to the spatio-temporal Bayesian maximum entropy method; additionally, the transformed hierarchical Bayesian model also provides predictive distributions.

making. Monthly average precipitation data over a period of twenty-seven years (1974–2000) were collected from the meteorological department of Pakistan, providing monthly average rainfall data for fifty-one gauged sites.

Space time data are frequently analyzed through models initially developed for only spatial or temporal distributions. Kyriakidis and Journel (1999) note that the joint space-time dependence is often not fully modeled nor exploited in the estimation or forecasting at unmonitored locations. Christakos (1992) developed the idea of spatio-temporal random fields (STRF) which can take into account the composite space-time dependence and also utilize the physical knowledge of the natural processes which in turn leads to an improved estimation. This method is based on the Bayesian maximum entropy (BME) approach; BME utilizes the physical knowledge about natural processes in the form of a highly informative prior distribution whereas in absence of physical knowledge about natural process this method is similar to space-time ordinary kriging. Le et al. (1997) proposed a Bayesian hierarchical interpolation method for environmental applications which is very sensitive to non-stationarity. This approach assumes finite dimensional Gaussian distributions with the mean functions depending on an unknown parameter matrix and a vague structure for covariance. The hierarchical framework can take account of the uncertainty of the mean and covariance models and generates space-time predictions that are completely characterized by probability density functions. In the present paper a comparison between BME and hierarchical Bayesian space-time interpolation is made. For BME prediction spatial and temporal mean trends are estimated and then separable spatial and temporal co-variograms are determined. The fitted co-variograms are used for BME kriging. For the transformed hierarchical Bayesian model the generalized inverted Wishart distribution is used as prior for the covariance matrix and its hyperparameters are empirically Bayesian estimated by means of an EM-algorithm.

1 Introduction

Pakistan is located between 23° and 37° north latitude and 61° and 76° east longitude. The economy of Pakistan is highly supported from the agricultural sector; the occurrence of monsoon (June–September) rainfall being of vital importance for the said sector. The accurate prediction of precipitation in Pakistan provides useful information for decision



Correspondence to: I. Hussain
(ijaz.hussain.kherani@gmail.com)

2 Material and methods

2.1 Study area

Monthly average precipitation data of fifty-one gauged sites in Pakistan were collected from the meteorology department of Pakistan. Some gauged sites have been recording data since 1947, the year of the foundation of Pakistan, while many other gauged sites were installed later. Data from a period of twenty-seven years (1974–2000) were used. The monsoon period in Pakistan lasts from June to September. The monthly average precipitation data during the monsoon period of twenty-seven years are used in the present study.

2.2 Bayesian maximum entropy method

Most of the statistical techniques for the estimation of characteristics of space-time processes are based on hard data and do not take into account any physical knowledge (soft data) for estimation purposes. Christakos et al. (2002) proposed the Bayesian maximum entropy spatio-temporal estimation method which can take into account physical knowledge in terms of prior information which, in turn, results in improvements for the estimation of space-time processes. The space-time random field is described as $Z(s, t)$ where $(s, t) \in D \cdot T$: $D \subset R^2$ represents spatial coordinates and $T \subset R_+$ are positive real numbers describing temporal coordinates. The BME method can be described briefly as consisting of five different phases: In the first phase the estimation of space-time mean trend is performed, a smooth spatial trend is computed using an exponential spatial filter which is then applied to the average measurements of each spatial location. A smooth temporal trend for each time replication is computed using an exponential temporal filter applied to the averaged measurements at each time instant. In the second phase the space-time mean trend $m(s, t)$ is interpolated on a grid. In the third phase the residuals are computed by subtracting the space-time mean trend from the original data i.e. $R(s, t) = Z(s, t) - m(s, t)$. The residual data matrix is then used to estimate the separate spatial covariances, temporal covariances and spatio-temporal experimental co-variograms. In the fourth phase theoretical covariance models are fitted to the experimental co-variograms. In the last phase prediction is performed at unobserved locations for any time instant using the fitted co-variogram model and space-time kriging, Christakos (1992), Chiles and Delfiner (1999). Since the co-variograms are estimated based on residuals the resulting predicted data will correspond to residual values. The space-time mean trend surfaces and predicted residual surfaces are to be added to get predicted values at the original scaling.

2.3 Transformed hierarchical Bayesian model

Let $Z^{[g]}$ and $Z^{[u]}$ be the response variables of gauged and ungauged locations and correspondingly $Y^{[g]}$ and $Y^{[u]}$ be the

Box-Cox transformed response variables, respectively, i.e.

$$Y^{[g]} = \frac{(Z^{[g]\lambda} - 1)}{\lambda} : \lambda \neq 0$$

$$Y^{[u]} = \log(Z^{[u]}) : \lambda = 0 \tag{1}$$

The transformed response variables follow a Gaussian model conditional on the hyper-trend parameter β and the covariance matrix Σ . The hyper-parameter β itself follows a Gaussian model conditional on the hyper-parameter β_0 and the covariance matrix Σ . Finally, the covariance matrix Σ follows a generalized inverted Wishart distribution. Thus, the suggested model of Le et al. (1997) for the transformed response variable can be expressed as;

$$Y/\beta, \Sigma \sim N(V\beta, I_n \otimes \Sigma)$$

$$\beta/\Sigma\beta_0 \sim N(\beta_0, F^{-1} \otimes \Sigma)$$

$$\Sigma \sim GIW(\Theta, \delta) \tag{2}$$

Here $N(., .)$ denotes the multivariate normal distribution and V represents the matrix of l covariates describing the time replications i.e. $V_t = (V_{t1}, \dots, V_{tl})$ remaining constant for all sites at each time point t . Since four months (June–September) are taken into consideration for a twenty seven years period, we have $l = 4$ and $t = 27$. The β_0 is an $l \times (g + u)l$ matrix of regression hyper-parameters and F^{-1} is an $l \cdot l$ positive definite matrix specifying covariance between the rows of β . In Le et al. (1997) an empirically Bayesian EM-algorithm is suggested to estimate these hyper-parameters. With regard to the prior specification in the last line of Eq. 2, we follow Brown et al. (1994), where a generalized inverted Wishart (GIW) distribution for the covariance matrix Σ is proposed and detail this prior as follows:

$$\tau^{[u]} / \Gamma^{[u]} \sim N(\tau_{00}, H_0 \otimes \Gamma^{[u]})$$

$$\Gamma^{[u]} \sim IW(\Delta_0 \otimes \Omega, \delta_0)$$

$$\tau_j / \Gamma_j \sim N(\tau_{0j}, H_j \otimes \Gamma_j), j = 1, 2, \dots, n_g - 1$$

$$\Gamma_j \sim IW(\Delta_j \otimes \Omega, \delta_j), j = 1 \dots n_g \tag{3}$$

The abbreviation IW stands for the Inverted Wishart distribution. Here τ^u is the slope of optimal linear predictors of $Y^{[u]}$ based on $Y^{[g]}$ and the residual covariance of the optimal linear predictor $\Gamma^{[u]}$; also τ_j and Γ_j for $j = 1, 2, \dots, n_g - 1$ have similar interpretations. Suppose H is the set of hyper-parameters in Eq. 3, i.e. $H = (\Theta, \delta, F, \beta)$, where $\Theta = \{(\tau_{00}, H_0, \Delta_0), \Omega, \dots, (\tau_{0n_g-1}, H_{n_g-1}, \Delta_{n_g-1}), \Delta_{n_g}\}$ with degree of freedom parameters $\delta = (\delta_0, \delta_1, \dots, \delta_{n_g})$. The dimension of the H_j 's and Δ_j 's are $(g_{j+1} + \dots + g_{n_g})l \cdot (g_{j+1} + \dots + g_{n_g})l$ and $g_j \cdot g_j$ respectively. The Ω is a hyper scale matrix between responses and is assumed to be common across all sites.

2.3.1 Heterogeneous covariance

Sampson and Guttorp (1992) introduced a non-parametric estimation method for spatially nonstationary covariance

structures of random functions $Z_{it} = Z(x_i, t)$; here $x_i = (1, 2, \dots, N)$ are locations and $t = 1, 2, \dots, T$ are time replications. The estimation method is based on the assumption of temporal stationarity but spatial non-stationarity. Sampson and Guttorp (1992) use the spatial dispersions $d_{ij}^2 = \text{var}(Z_{it} - Z_{jt}) = s_{ii} + s_{jj} - 2 \cdot s_{ij}$ as natural metric for the spatial covariance structure model. Their method constructs a smooth mapping of the geographic space of gauged locations onto the dispersion space (D-space) and does not need any assumption of stationarity. An isotropic variogram model is fitted on the basis of observed correlations and distances in D-space. The spatial correlations of ungauged locations can be estimated in conjunction with an estimated isotropic variogram model by using a thin plate smoothing spline as a nonlinear 1-1 mapping function. The estimated nonstationary covariance model is of the form $\widehat{D}^2(x_a, x_b) = g|\widehat{f}(x_a) - \widehat{f}(x_b)| = g(|y_i - y_j|)$, here for covariance estimation y_i are referred as D-plane coordinates and for the details about function f and g , see Sampson and Guttorp (1992).

2.3.2 Predictive distributions

The predictive distribution can be specified as matrix t-distribution. The transformed random variable $Y_{n \cdot m}$ is said to have a matrix t-distribution, $Y_{n \cdot m} \sim t_{n \cdot m}(Y_{n \cdot m}, A \otimes B, \delta)$, where A is $n \cdot n$ and B is $m \cdot m$, if its density function has the form

$$f(Y) \propto \left| A^{-\frac{m}{2}} \right| \left| B^{-\frac{n}{2}} \right| \left[I_n + \delta^{-1} \left\{ A^{-1} (Y - Y^{(0)}) \right\} \left\{ (Y - Y^{(0)}) B^{-1} \right\}^t \right]^{-\frac{\delta + m + n - 1}{2}} \quad (4)$$

The normalizing constant of this density is given by $K = \left[\frac{(\delta \pi^2)^{\frac{nm}{2}} \Gamma_{n+m} \left\{ \frac{\delta + m + n - 1}{2} \right\}}{\Gamma_n \left(\frac{\delta + n - 1}{2} \right) \Gamma_{n+m} \left(\frac{\delta + m - 1}{2} \right)} \right]$, where $\Gamma_p(t)$ is the multivariate Gamma function and $Y^{(0)}$ is the mean of Y . The predictive distribution of the unobserved transformed responses conditional on the observed transformed data $Y^{[g]}$ and hyperparameter H is given by

$$\left(Y^{[u]} / Y^{[g]}, H \right) \sim t_{n_u \cdot n_{ul}} \left(\mu^{[u/g]}, (\delta_0 - n_u l + 1)^{-1} \Phi^{[u/g]} \otimes (\Delta_0 \otimes \Omega), \delta_0 - n_u l + 1 \right) \quad (5)$$

where $\mu^{[u/g]} = V \beta_0^{[u]} + \hat{\epsilon}^{[g]} \tau_{0n_u}$, $\Phi^{[u/g]} = I_{n_u} + V F^{-1} \hat{V} + \hat{\epsilon}^{[g]} H_0 \hat{\epsilon}^{[g]}$ and $\hat{\epsilon}^{[g]} = Y^{[g]} - V \beta_0^{[g]}$. The software provided by Le and Zidek (2006) offers the possibility to simulate realizations $Y_i^{[u]}, i = 1, 2, \dots, 5000$ from the predictive distribution Eq. 5. These simulated values have just to be back-transformed to the original scale by means of the inverse Box-Cox transformation. Multivariate histograms based on these back-transformed simulated values are then used to approximate the predictive distributions at the original scale.

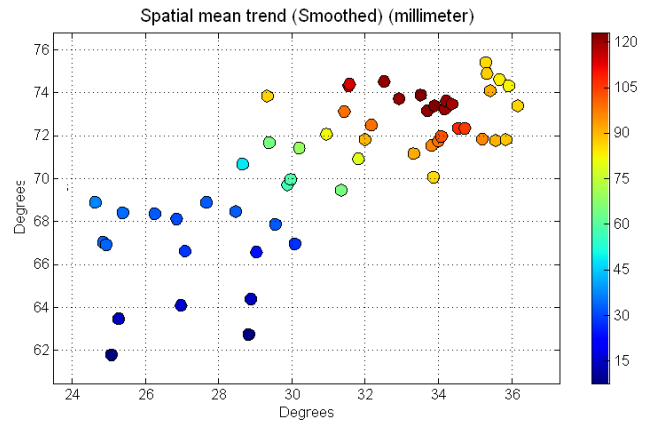


Fig. 1. The map of smoothed mean trend aggregated over time.

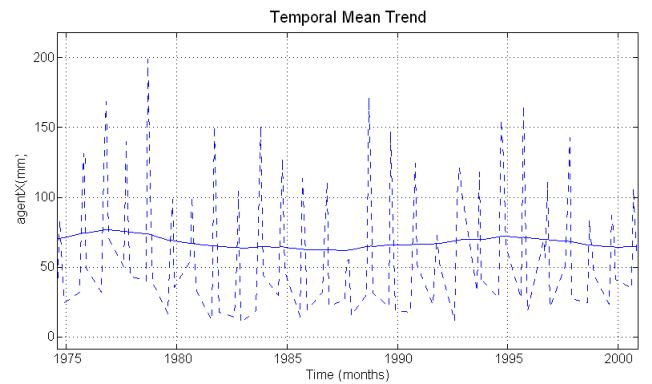


Fig. 2. The temporal mean trend aggregated over the regions.

3 Results and discussions

3.1 Estimation by the Bayesian maximum entropy method

The spatial mean trend and temporal mean trend for precipitation are estimated as described in Sect. 2.2. The smooth spatial mean trend for the whole period of study is displayed in Fig. 1. The regions between 73° – 75° longitude and 32° – 34° latitude show very heavy precipitation compared to other regions. The regions falling between 62° – 64° longitude and 24° – 29° latitude are getting very little rainfall. The temporal mean trend during the whole period of study is presented in Fig. 2. The temporal trend of precipitation is cyclical and the amount of precipitation is decreasing while time increases. In the next phase of the BME method the spatial and temporal dependence of precipitation is estimated. A nested spatial covariance model is fitted as a combination of a Gaussian and an exponential covariance function i.e. $C(r, \tau = 0) = 5000 \exp\left(\frac{-3r}{12}\right) + 5000 \exp\left(\frac{-3r^2}{1.5^2}\right)$. The nested temporal covariance model is fitted as a combination of hole-sine and an exponential covariance function

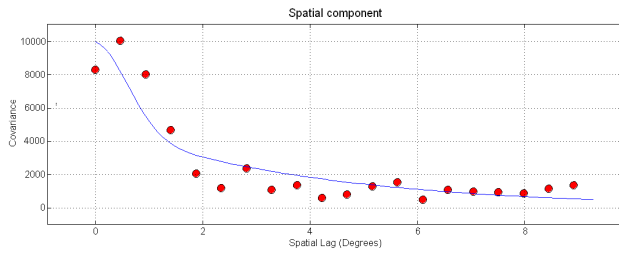


Fig. 3. The spatial experimental covariance and nested spatial covariance model.

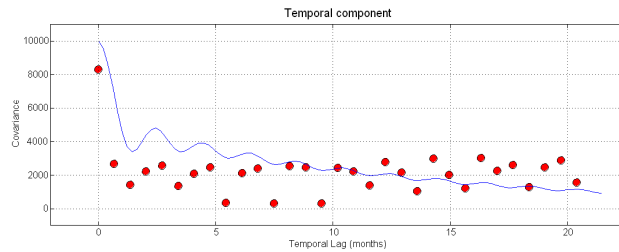


Fig. 4. The experimental temporal covariance and the nested covariance model.

i.e. $C(r = 0, \tau) = 5000 \exp\left(\frac{-3\tau}{40}\right) + 5000 \frac{\sin(\pi\tau)}{\pi\tau}$. The nested spatial covariance model and the experimental covariance are shown in Fig. 3, the dots show the experimental covariances and the line represents the fitted covariance model. The nested temporal covariance model and the experimental temporal covariance are displayed in Fig. 4. According to Sect. 2.2, in the last phase, the Bayesian maximum entropy predictions are calculated for the year 2000 using the fitted co-variograms. The prediction maps for the months of July and September in 2000 are displayed in the left two panels of Fig. 5. These prediction maps suggest that heavy rainfall occurs in the regions between 73°–75° longitude and 32°–34° latitude while the remaining regions of Pakistan have very little rainfall.

3.2 Estimation of hyper-parameters

A time series trend analysis is made and temporal correlations are removed by fitting an autoregressive model of order 2. To meet the assumption of normality for hierarchical Bayesian space time interpolation the Box-Cox transformation is used and it is observed that the data fulfill the assumption of normality for Box-Cox transformation parameter $\lambda = 0.1575$. The posterior distributions of β_0 and Σ are estimated on the basis of the EM algorithm in Le and Zidek (2006) using their software. The hyper-parameters $H = \{\Theta, \delta, F, \beta_0\}$ for gauged locations are estimated by the

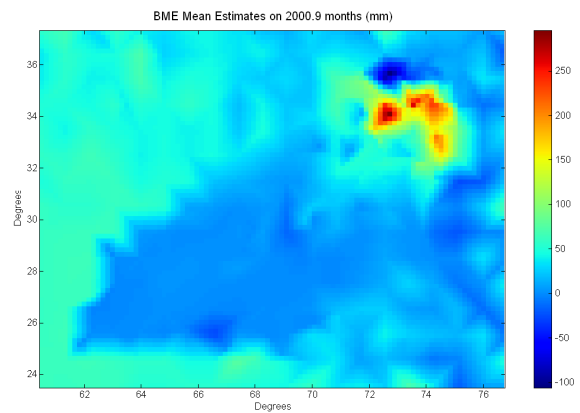
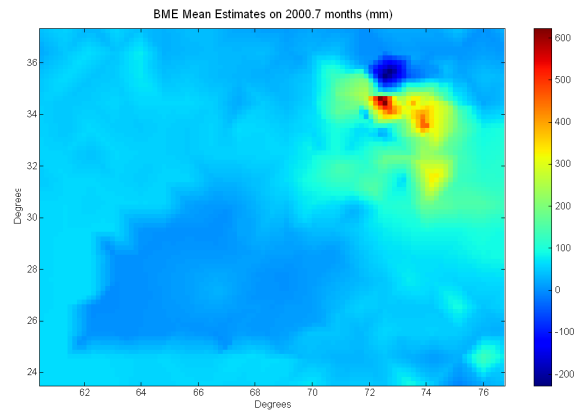


Fig. 5. The maps of average predicted rainfall BME spatio temporal method for the month of July and September of year 2000.

EM algorithm, too. The method for estimating a nonstationary spatial covariance matrix for ungauged locations is applied in three phases. In the first phase, the method utilizes the empirical correlations and fits an empirical Gaussian variogram model; $\gamma(h) = a_0 + (2 - a_0)(1 - \exp(-t_0 \cdot h))$, where $a_0 = 0.0260$ and $t_0 = 1.032$. For the fitted variogram based on empirical correlations see Fig. 6. In the second phase, a thin plate smoothing spline parameter $\lambda_0 = 20$ is used to estimate the coefficients of the thin plate spline smoothing splines. In the third phase, the correlations between the gauged sites and 196 ungauged sites are estimated. The correlations for ungauged locations are obtained by first converting ungauged geographic coordinates to D-plane coordinates and then calculating the correlations based on fitted variogram parameters and the inter-distances in D-space. A comparison of the empirical correlations of gauged sites and ungauged sites is shown in Fig. 7.

3.2.1 Predictive distributions

To predict unobserved transformed responses, 5000 realizations are simulated at a 14 by 14 grid ungauged locations by using the predictive distributions. The predicted transformed

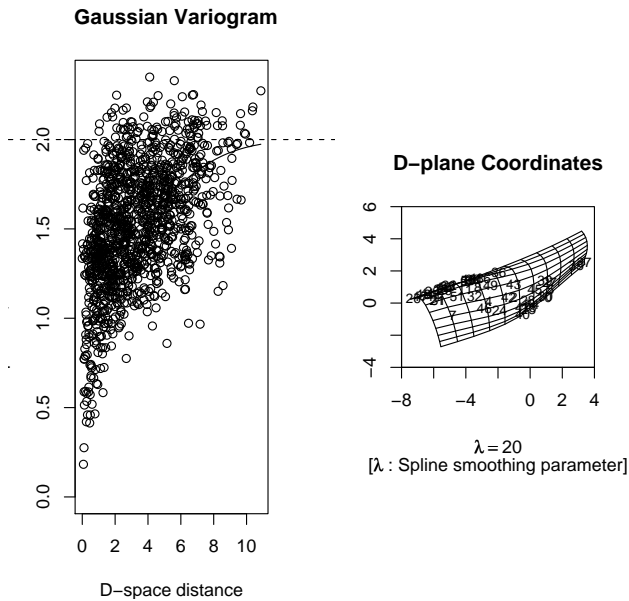


Fig. 6. The fitted Gaussian variogram and mapping from G-plane to D-plane.

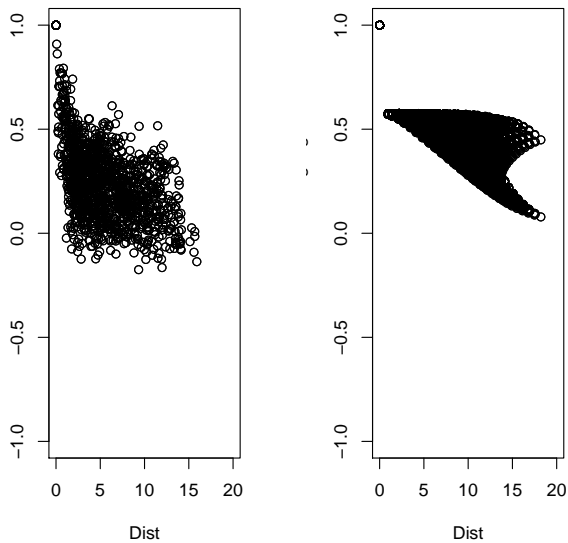


Fig. 7. Comparison of empirical correlations of gauged sites and estimated correlations by using the Sampson and Guttorp (1992) method

responses are back transformed to the original scale by using the inverse Box-Cox transformation. The predicted maps of mean values for July and September of the year 2000 are shown in Fig. 8. The South-eastern part of Pakistan had very little rainfall for the selected months of this year. During the investigated period the native regions of Islamabad (capital of Pakistan) received much rainfall. The maximum rainfalls occurred in July.

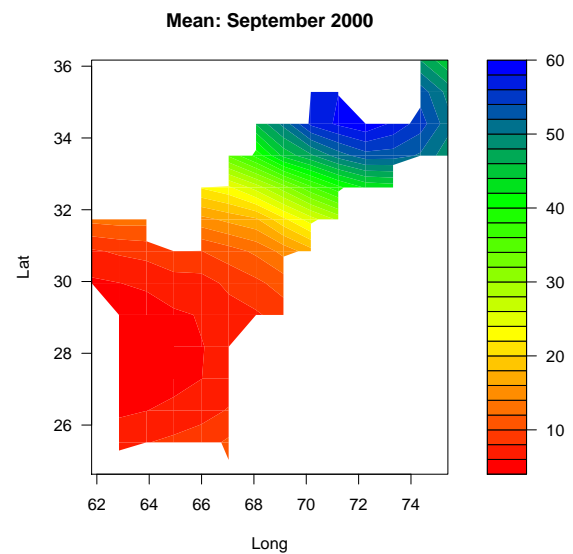
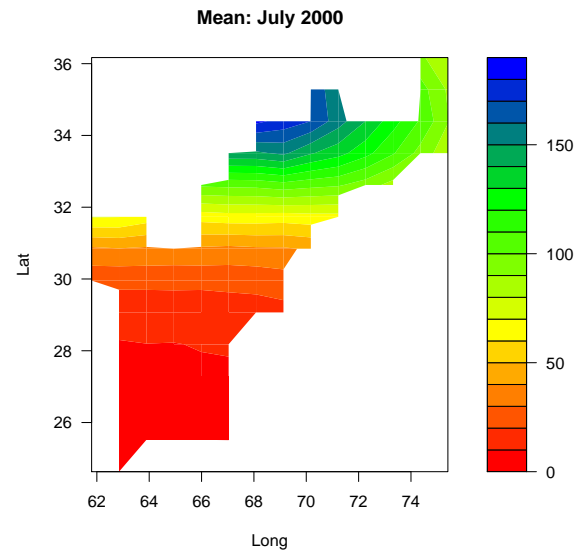


Fig. 8. The predicted maps for the months of July and September of year 2000 using the transformed hierarchical Bayesian model

4 Conclusions

The BME and the Hierarchical Bayesian space-time model indicate heavy rain for the same regions, however, the Bayesian hierarchical model is also showing variations in the amount of rainfall for regions which received lower rainfall. The BME maps are showing spatial consistency in the amount of rainfall for the remaining regions except the heavy rainfall regions i.e. 73°–75° longitude and 32°–34° latitude. Fig. 1 shows spatial variation in the amount of rainfall, excluding heavy rainfall regions i.e. (73°–75° longitude and 32°–34° latitude), the BME spatio-temporal prediction method does not take this variation into account. The

hierarchical Bayesian model is accounting for more variation compared to the BME spatio-temporal method. The Bayesian non-stationary multivariate space-time interpolation method also provides more accuracy compared to the BME spatio-temporal method. Moreover, from the transformed hierarchical Bayesian method it is possible to identify regions of Pakistan with low or high precipitation during the 27 years of investigation. From the maps produced it can be seen that heavy rainfall occurs during the months of July and low rainfall occurs in June. The neighboring regions of Islamabad, the capital of Pakistan, are getting more rainfall for the whole time period. Most of the regions with low rainfall are having low elevation and are agricultural areas and some regions are characterized as deserts. A prediction for ungauged locations can be made for any time point in the given range. Moreover, from the simulated realizations the uncertainties about droughts and heavy rainfall for any threshold values can be obtained from the predictive distribution. The cross validation for the Hierarchical Bayesian model led to 95.1% of the values of gauged locations falling within the 95% credible intervals, which suggests that this prediction method is appropriate for predicting precipitation in Pakistan during the monsoon periods of the years 1974–2000.

Acknowledgements. We would like to thank the Higher Education Commission of Pakistan for financial support of this project and the Pakistan meteorology department for providing the data. We are grateful to S. C. Michaelides for comments and suggestions to improve the paper.

Edited by: S. C. Michaelides

Reviewed by: two anonymous referees

References

- Brown, P. J., Le, N. D., and Zidek, J. V.: Inference for a covariance matrix, *Aspects of Uncertainty: A Tribute to DV Lindley*, edited by: Freeman, P. R. and Smith, A. F. M.: Chichester, Wiley, 77–92, 1994.
- Chiles, J. P. and Delfiner, P.: *Geostatistics, Modeling Spatial Uncertainty*, Wiley-Interscience, 1999.
- Christakos, G.: *Random Field Models in Earth Sciences*, Academic Press, 1992.
- Christakos, G., Bogaert, P., and Serre, M. L.: *it Temporal Gis, Advanced Functions for Field-Based Applications*, Springer Verlag, 2002.
- Kyriakidis, P. C. and Journel, A. G.: *Geostatistical SpaceTime Models: A Review*, *Mathematical geology*, 31, 651–684, 1999.
- Le, N. D., Sun, W., and Zidek, J. V.: *Bayesian Multivariate Spatial Interpolation with Data Missing by Design*, *J. Roy. Stat. Soc. B*, 59, 501–510, 1997.
- Le, N. D. and Zidek, J. V.: *Statistical Analysis of Environmental Space-Time Processes*, Springer Verlag, 272–294, 2006.
- Sampson, P. D. and Guttorp, P.: *Nonparametric Estimation of Non-stationary Spatial Covariance Structure*, *J. Am. Stat. Assoc.*, 87, 108–119, 1992.