To my Parents, my Family, and the memory of my Grandfather

Anandkumar Shetiya
New Mexico Institute of Mining and Technology
December, 2005

HIDDEN RANDOM FIELD MODELING OF OROGRAPHIC EFFECTS ON MOUNTAINOUS PRECIPITATION

by

Anandkumar Shetiya

Submitted in Partial Fulfillment

of the Requirements for the Degree of

Master of Science in Mathematics with Emphasis in Industrial Mathematics

New Mexico Institute of Mining and Technology

Socorro, New Mexico

December, 2005

ABSTRACT

This work uses a Markov chain Monte Carlo (MCMC) method to model the moisture flux direction random field (MFD RF) for a mountainous region. It builds on the approach originally proposed by Guan et al. (2005) in their precipitation model named ASOADeK (Auto-Searched Orographic and Atmospheric effects De-trended Kriging). The ASOADeK assumes the MFD to be constant throughout the study region. However, our model allows for non-constant MFD RF. It takes into account the influence of local orographic and atmospheric effects on spatially correlated gauge precipitation data. It combines linear regression with subsequent spatial interpolation (*kriging*). It also performs 'all in one' estimation of regression and other parameters using a Gibbs sampler. We tested the model with the precipitation data collected for the mountainous region of semi-arid northern New Mexico. The results reported include the estimated MFD random field, as well as detailed precipitation maps. Knowledge of MFD RF may offer insight into the region's precipitation patterns.

ACKNOWLEDGMENT

Many people deserve thanks for extending their help in whichever way it was possible for them. I would like to first thank Dr.Oleg Makhnin, my academic advisor. He has been very supportive and understanding and created an ideal atmosphere for me to complete this project on a good note.

Let me acknowledge the fact that as a Mathematics graduate student at New Mexico Tech, one has to complete considerable amount of coursework. This gave me enough opportunities to know the teaching styles of eminent professors like Dr.William Stone, Dr.Brian Borchers, Dr.Subhasish Mazumdar, Dr.Rakhim Aitbayev, and Dr.Oleg Makhnin.

Life is different for an international graduate student, and there are times when they need some moral support. I was fortunate enough to find people like Dr.Stone, Dr.Aitbayev, Dr.Makhnin, and Emma (Mathematics department secretary) to whom I could talk about problems in my personal life.

Thanks to my fellow graduate students - Andrey Novoseltsev, Satya Sai Vaddadhi, Qian Xia, Raphael Clancy, Christian Lucero, Don Clewett, Leny Mathew, and Amber Polizzi. Some of them also helped me with their LaTeX skills.

As to personal friends beyond work, I could ask for none better than Tushar Shitole, and Sujit Tatke. Whether it's been trekking, a regular get-together, or just tele-conferences, I've enjoyed it all! To my closely-knit Indian

community at Tech with whom I have had a great time cooking, listening music, celebrating festivals, and playing cricket. Special thanks to all my other friends whose emails and chat sessions make me smile.

And of course, with all my heart I thank my parents, my family, and my grandfather to whom this work is dedicated.

I would also like to thank Huade Guan for providing the DEM maps and converting them to retrieve the aspect and elevation data, and Dr.John Wilson for his suggestions, comments and interest in this project.

Finally, I gratefully acknowledge the initial concept of estimating hidden moisture flux direction random field using a Bayesian approach to my advisor Dr.Makhnin.

Anandkumar Shetiya

This report was typeset with LaTeX[1] by the author.

[1] LaTeX document preparation system was developed by Leslie Lamport as a special version of Donald Knuth's TeX program for computer typesetting. TeX is a trademark of the American Mathematical Society. The LaTeX macro package for the New Mexico Institute of Mining and Technology report format was adapted from Gerald Arnold's modification of the LaTeX macro package for The University of Texas at Austin by Khe-Sing The.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

This report is accepted on behalf of the faculty of the Institute by the following committee:

Dr. Oleg Makhnin, Advisor

Anandkumar Shetiya Date

CHAPTER 1

INTRODUCTION

Mountainous regions have complex topography typically characterized by continuously varying elevation. Varying elevation is one of the factors causing considerable difference in the climatic conditions throughout such regions. While measuring precipitation in such regions, it is important to allow for the local and precise climatic conditions prevailing at different elevations because spatial variability of precipitation greatly influences many different hydrologic and ecologic studies. Long-term averages of precipitation measurements done using a limited number of gauges scattered throughout the mountainous region are used to develop precipitation mapping products which can estimate precipitation at other locations within that region. The objective of this work is to test a method for developing such a product by taking into account orographic[1] factors like the terrain elevation, the terrain aspect, and the moisture flux direction (MFD).

Several different models have been developed to estimate precipitation with each one of them following some basic philosophy or combination of various philosophies. For instance, there are simple models based on Theissen polygon [9],[12], and inverse square distance [9] which do not consider spatial

[1] *Orography* is the science of mountains.

covariance structure of precipitation data nor do they consider orographic and atmospheric effects. Orographic factors like the terrain elevation and the terrain aspect cause more precipitation on the windward side of the mountain than on the lee side. The direction of gradient along the mountain topography is given by the terrain aspect. Orographic effects on precipitation measurement are well documented, see e.g., [2],[5]. Models based on kriging technique [9],[15] take into account spatial covariance structure and yield better estimates than the above simple models. But they too fail to address the relation between precipitation and orographic effects. Regression-based models provide a partial solution to this problem by considering some orographic features but ignoring spatial covariance. But there are models that consider both spatial covariance and the climatic conditions, e.g., cokriging with terrain elevation, and de-trended residual kriging [9].

PRISM (Precipitation-elevation Regression on Independent Slopes Model) [4, 16], and ASOADeK (Auto-Searched Orographic and Atmospheric effects De-trended Kriging) [9] are two precipitation mapping products developed especially for the mountainous terrains. They consider both the terrain elevation and the terrain aspect which play an important role in determining orographic effects. These products provide better precipitation estimates. The downside to using PRISM is the need to have sufficient regional climatic knowledge in order to obtain reliable estimates. In particular, to account for orographic effects, one needs to find the direction from which moisture comes, i.e., the moisture flux direction (MFD). Unlike PRISM, ASOADeK addresses this difficulty by explicitly estimating the MFD. Also, spatial resolution of PRISM

product depends on its input DEM (Digital Elevation Map) grid size [9][2].

1.1 Details of ASOADeK

ASOADeK uses a multivariate linear regression approach to produce high spatial resolution precipitation maps. Regression step is done using the following function:

$$Y_i = \beta_0 + \beta_1 E_i + \beta_2 N_i + \beta_3 Z_i + \beta_4 \cos(A_i - W) + V_i \tag{1.1}$$

where Y_i is the precipitation measured in mm at gauge i, E_i is the UTM (Universal Transverse Mercator)[3] easting coordinate [17] expressed in km, N_i is the UTM northing coordinate expressed in km, Z_i is the above sea-level terrain elevation expressed in km, A_i is the terrain aspect measured in radians, W is the MFD measured in radians, V_i is the error term (a.k.a. residual), and $(\beta_0, \beta_1, \beta_2, \beta_3, \beta_4)' = \boldsymbol{\beta}$ are regression parameters to be estimated. Y_i can be a single precipitation event or a long term average of different precipitation events. MFD was specifically introduced in the ASOADeK model for the purpose of accounting for orographic effects due to moisture direction. Note that equation (1.1) is linear in regression parameters $\boldsymbol{\beta}$, but not in W. Equation (1.1) captures the local and regional climatic and orographic effects like the effective terrain elevation, and the effective terrain aspect. However, to avoid possible complications with nonlinear estimation for W, the mathemati-

[2]DEM's used in our work are generated by the ESRI ArcMap GIS tool by Guan et al.

[3]There are 60 longitudinal projection zones and within each zone the tranverse Mercator projection is used to give the easting and northing coordinates in meters. The UTM easting and northing coordinates thus define a location within a UTM projection zone either north or south of the equator. Refer to [17] for more details about UTM coordinates.

cal model in equation (1.1) is transformed as:

$$Y_i = \beta_0 + \beta_1 E_i + \beta_2 N_i + \beta_3 Z_i + \beta_5 \cos(A_i) + \beta_6 \sin(A_i) + V_i \tag{1.2}$$

where $\beta_5 = \beta_4 \cos W$ and $\beta_6 = \beta_4 \sin W$ implicitly contain the information about the constant moisture flux direction W. This implies that ASOADeK assumes the moisture flux direction to be constant throughout the study region. Once $\boldsymbol{\beta}$ and W are estimated, equation (1.2) can then be applied to predict precipitation at the locations where no observations are made.

Assuming that E_i, N_i, Z_i, and A_i are measured precisely, and applying equation (1.2) to observations $Y_1, Y_2, \ldots, Y_n$, where n is the number of gauge stations (a.k.a. observation sites), we obtain a system of equations with n rows and $p = 6$ columns that relates the precipitation data Y_i, to the model parameters β_j, $j = 0, 1, 2, 3, 5, 6$. The above system of equations can be written using matrix notation as:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{V} \tag{1.3}$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & E_1 & N_1 & Z_1 & \cos A_1 & \sin A_1 \\ 1 & E_2 & N_2 & Z_2 & \cos A_2 & \sin A_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & E_n & N_n & Z_n & \cos A_n & \sin A_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_5 \\ \beta_6 \end{bmatrix} + \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix}$$

When there are more data points Y_i than model parameters β_j (which is usually the case with precipitation data), it is impossible to find a model $\boldsymbol{\beta}$ that satisfies every equation exactly. However, we can still find model parameters which fit the given set of data in an approximate "best fit" sense [1]. This problem of linear regression is solved using a least squares approach which minimizes the

2-norm of the residuals, $\mathbf{V} = \mathbf{Y} - \mathbf{X}\boldsymbol{\beta}$. The least squares solution is obtained using the normal equations [1] as[4]:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \tag{1.4}$$

In the next step, ASOADeK removes $\mathbf{X}\hat{\boldsymbol{\beta}}$, the mean part of the precipitation random field $\mathbf{Y}$ (a.k.a. "trend"), leaving behind stationary, mean 0 (detrended) random field $\mathbf{V}$ obtained as the residuals from regression. These residuals are further spatially interpolated by ordinary kriging to generate a residual precipitation surface. The final precipitation map is produced by adding the regression surface to the kriged residual surface.

However, there are a few shortcomings of ASOADeK model which can significantly influence the precipitation estimates. They are:

1. Precipitation estimation using ASOADeK model is a multi-step approach (linear regression, de-trending, and spatial interpolation). Multi-step approach leads to aggregation of errors.

2. Least squares solution for linear regression doesn't account for any kind of correlation that may be present between the gauge precipitation data.

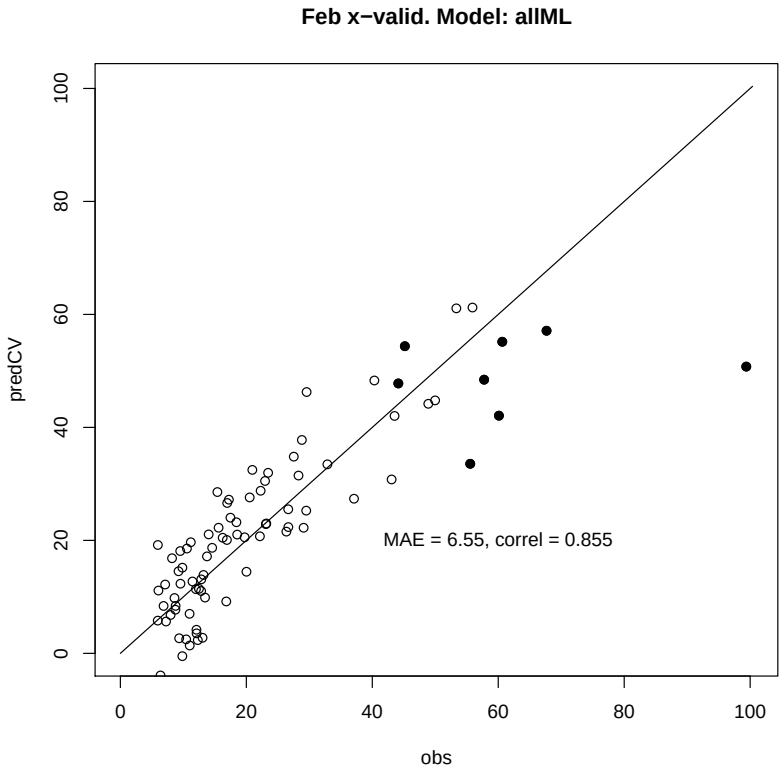
3. The moisture flux direction W, is assumed to be constant throughout the mountainous region.

4. Error magnitudes are proportional to the gauge precipitation measurements - the higher the measurement, the higher is the error and vice versa. This is shown in Figure 1.1.

[4]Hereafter, variable$'$ stands for Transpose of that variable

5. Within linear regression given by equation (1.2), negative estimates of precipitation are possible, although they make no sense physically.

Figure 1.1: Predicted values from ASOADeK model for February. Dark circles - SNOTEL sites. (Courtesy - Dr. Makhnin)



1.2 Structure Overview

This work is an attempt to refine the ASOADeK model by:

1. transforming multi-step estimation process into an 'all in one' estimation process using Gibbs sampling,

2. taking into account the spatial covariance structure of the gauge data,

3. doing away with the assumption of W being constant throughout the region, and

4. taking the natural log of the precipitation measurements to achieve the homogeneity of error variances, and prevent negative estimates.

The newly proposed model uses a Bayesian approach to estimate the moisture flux direction random field (MFD RF) $\mathbf{W}$ (a column vector of all W_i's), instead of estimating a single value of W, and using it for all the observation and prediction locations. The model assumes some prior distribution for the MFD RF that insures local uniformity of $\mathbf{W}$, therefore keeping the effective number of parameters low. The structure of this prior is inspired by Ising model from image processing [11]. The model implements a Gibbs sampler with an efficient 'Metropolis within Gibbs' step to compute the full conditional posterior density for the MFD RF. It also does 'all in one' estimation of the parameters responsible for the spatial covariance structure, and regression parameters by implementing Markov chain Monte Carlo (MCMC) algorithm through Gibbs sampling. This model is subjective since it involves choosing a prior distribution for MFD RF. It also requires us to choose the parameter regulating the smoothness of the estimated $\mathbf{W}$. The results obtained in this work are encouraging but there is a need for more analysis.

Chapter 2 begins by describing the mathematical model used in this work. It also includes the underlying geostatistical model for the continuous spatial process of precipitation, following a standard representation as given in [15]. It concludes by presenting the theory behind spatial interpolation (kriging). Chapter 3 focusses on the key algorithms like Gibbs sampling, and

'Metropolis within Gibbs' which are used to estimate the model parameters. It also describes the Bayesian framework for sampling from the full conditional posterior distributions. At the end of this chapter, we present the implementation issues of our algorithm. Chapter 4 is devoted to the results obtained for the real dataset used in this work. It also includes some simulation and cross-validation results. Chapter 5 is the concluding chapter of this work where we analyze the results, draw some important conclusions, and talk about possible future work.

CHAPTER 2

MODEL SPECIFICATION

In this chapter, we first describe the mathematical model for precipitation used in this work. Then we show that this model is equivalent to a hierarchical geostatistical model for continuous spatial processes developed in [15]. Finally, we conclude the chapter by presenting some theory behind kriging.

2.1 Regression

The multivariate linear function (linear only in regression parameters) defining the proposed model can be written as:

$$Y_i \equiv \ln(P_i) = \beta_0 + \beta_1 E_i + \beta_2 N_i + \beta_3 Z_i + \beta_4 \cos(A_i - W_i) + V_i \qquad (2.1)$$

where P_i is the precipitation measured in mm at gauge i, E_i is the UTM easting coordinate [17] expressed in km, N_i is the UTM northing coordinate expressed in km, Z_i is the above sea-level terrain elevation expressed in km, A_i is the terrain aspect measured in radians, W_i is the moisture flux direction (MFD) measured in radians, V_i is residual (possibly, spatially correlated), and $(\beta_0, \beta_1, \beta_2, \beta_3, \beta_4)' = \boldsymbol{\beta}$ are regression parameters.

The elevation effect on the gauge precipitation is determined by β_3, and the spatial gradient effect is determined by β_1 and β_2. The orographic

effects due to the effective moisture flux direction and the effective terrain aspect are jointly determined by β_4 and $\mathbf{W}$, a vector of all W_i's.

We chose the above functional form for the following reasons:

1. $\beta_1 E_i$ and $\beta_2 N_i$ can account for the "spatial trend" (different from "trend" in geostatistics) in precipitation measurement.

2. $\beta_3 Z_i$ can account for the effect of the terrain elevation Z_i on long-term average precipitation Y_i, as is evident from prior studies [9].

3. $\cos(A_i - W_i)$ is chosen in order to incorporate orographic effects; there is a positive correction ($\cos(A_i - W_i) = 1$) when A_i and W_i have the same direction, negative correction ($\cos(A_i - W_i) = -1$) when they have exactly opposite directions, and zero orographic effect when they are perpendicular ($\cos(A_i - W_i) = 0$).

The direction of the slope orientation at each gauge location is given by the terrain aspect A_i, where $A_i = 0$ corresponds to the linear elevation gradient pointing north-south, angle increasing clockwise, and $A_i = \pi$ corresponds to the linear elevation gradient pointing south-north. The moisture flux direction W_i, follows the same convention. It captures interaction with the terrain aspect and is averaged over many potential precipitation events.

We have seen that ASOADeK is developed on the assumption of a constant W_i ($W_1 = W_2 = \ldots = W_n$) over the entire region having n observation sites. If the mountainous region under study is small, and carefully selected, then this assumption may hold true, but for bigger regions, this may

no longer be true. This is the main difference between ASOADeK and the model proposed in this work. Our primary goal is to first model the moisture flux direction for the entire mountainous region irrespective of its size and then use these estimates to achieve the secondary goal of constructing precipitation maps for the entire region by the technique of spatial interpolation (a.k.a. **kriging** in geostatistics parlance).

We have the precipitation data from the gauge measurements, and also obtain the variables like easting, northing, elevation, and aspect for all the observation sites from DEM obtained using ESRI ArcMap GIS tool. If we also knew W_i for all the observation sites, then we could find a parametrized surface that approximately fits this set of data. This procedure of surface-fitting is known as "regression" [1]. By applying equation (2.1) to each observation, we obtain a system of equations with n rows and $p = 5$ columns that relates the natural log precipitation data $\ln Y_i$, to the model parameters β_j, $j = 0, 1, 2, 3, 4$. The above system of equations can be written using matrix notation as:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{V} \tag{2.2}$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & E_1 & N_1 & Z_1 & \cos(A_1 - W_1) \\ 1 & E_2 & N_2 & Z_2 & \cos(A_2 - W_2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & E_n & N_n & Z_n & \cos(A_n - W_n) \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} + \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix}$$

Since the model in equation (2.1) is linear in regression parameters $\boldsymbol{\beta}$, we have a multivariate linear regression problem to solve. For now we assume that W_i's are known, and solve the linear regression problem using "generalized least squares" [10]. Later, we will address the joint estimation of $\mathbf{W}$ and $\boldsymbol{\beta}$ in the Bayesian framework. Generalized least squares solution minimizes the 2-norm

of the residuals, $\mathbf{V} = \mathbf{Y} - \mathbf{X}\boldsymbol{\beta}$, by taking into account the covariance structure of the data. Such a solution is obtained from the normal equations [1] as:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{Y} \tag{2.3}$$

where $\boldsymbol{\Sigma}$ is the covariance matrix of $\mathbf{Y}$ with elements $\sigma_{i,j} = \text{Cov}(Y_i, Y_j)$. Cov stands for covariance. Hereafter, $\hat{\boldsymbol{\beta}}$ is denoted only by $\boldsymbol{\beta}$. These residuals are further spatially interpolated by kriging to generate a residual precipitation surface. The final precipitation map is produced by adding the regression surface to the kriged residual surface.

2.2 Geostatistical Model

The gauge precipitation samples collected over a certain region can be thought of as a realization of some spatial continuous stochastic (random) process which can be modeled as a hierarchical linear Gaussian model [15].

We follow a model-based approach in which we begin by assuming a hierarchical linear Gaussian model [15]. Consider a finite set of precipitation observation sites $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$, within a region D. In our case, each $\mathbf{u}_i$ is characterized by a unique pair of UTM easting and northing coordinates, and D is the mountainous region in semi-arid northern New Mexico. (More about this study area is given in Chapter 4.) The precipitation data vector is denoted by $\mathbf{y}(\mathbf{u}) = (y(\mathbf{u}_1), y(\mathbf{u}_2), \ldots, y(\mathbf{u}_n))'$, measurements of a random vector $\mathbf{Y}(\mathbf{u}) = (Y(\mathbf{u}_1), Y(\mathbf{u}_2), \ldots, Y(\mathbf{u}_n))'$. Consider any one variable $Y(\mathbf{u}_i)$ from the random vector $\mathbf{Y}(\mathbf{u})$. The model splits $Y(\mathbf{u}_i)$ into a mean (or trend) part, the stationary signal $S(\mathbf{u}_i)$, and noise $\varepsilon(\mathbf{u}_i)$

This hierarchical model has the following features [15]:

1. *Covariates:* The "mean part" of the model is given by the term $\mathbf{X}(\mathbf{u}_i)'\boldsymbol{\beta}$. $\mathbf{X}(\mathbf{u}_i)$ represents a vector of spatially referenced non-random variables at observation site $\mathbf{u}_i$. In our case, these non-random variables are easting E_i, northing N_i, the terrain elevation Z_i, and the term $\cos(A_i - W_i)$. $\boldsymbol{\beta}$ is a vector of regression parameters.

2. *Underlying spatial process:* $S(\mathbf{u}_i)$ is a stationary Gaussian random process with zero mean, variance σ^2, and variogram function $\boldsymbol{\Phi}(\mathbf{h}, \boldsymbol{\varphi})$, where $\boldsymbol{\varphi}$ is a vector of correlation function parameters, and $\mathbf{h}$ is the vector distance between any two observation sites, independent of $\mathbf{u}_i$. The variogram function is usually assumed to have a parametric form. In our case, we assume it as isotropic exponential because it is one of the simplest forms which reflects the geostatistical principle that the correlation is highest between the observation sites that are close by.

$$\boldsymbol{\Phi}(\mathbf{h}, \sigma^2, \varphi, \tau_R^2) = \sigma^2 \left[1 - \exp\left(-\frac{|\mathbf{h}|}{\varphi}\right) + \tau_R^2\right] \tag{2.4}$$

where φ is the "range", and τ_R^2 is the "relative nugget" in geostatistics parlance. τ_R^2 can be written as:

$$\tau_R^2 = \frac{\tau^2}{\sigma^2} \tag{2.5}$$

where τ^2 is the "nugget" and σ^2 is the "partial sill" in geostatistics parlance.

3. *Conditional independence:* Variables $Y(\mathbf{u}_i), i = 1, \ldots, n$, are assumed to be normally distributed and conditionally independent given the signal, i.e.,

$$Y(\mathbf{u}_i)|S \sim N(\mathbf{X}(\mathbf{u}_i)'\boldsymbol{\beta} + S(\mathbf{u}_i), \tau^2) \tag{2.6}$$

$Y(\mathbf{u}_i)$'s are related through a "semivariogram (or simply variogram) function as:

$$\mathbf{\Phi}(\mathbf{h}) = \frac{1}{2} \text{ Var } (Y(\mathbf{u}_i) - Y(\mathbf{u}_i + \mathbf{h})) \tag{2.7}$$

where Var stands for variance.

The model given in equation (2.6) corresponds to a spatial linear mixed model and it can be specified in a hierarchical scheme for the random vector $\mathbf{Y}(\mathbf{u})$ as follows [15]:

$$\text{Level 1:} \quad \mathbf{Y}(\mathbf{u}) = \mathbf{X}(\mathbf{u})\boldsymbol{\beta} + \mathbf{S}(\mathbf{u}) + \boldsymbol{\varepsilon}(\mathbf{u}) \tag{2.8}$$

where $\boldsymbol{\varepsilon}(\mathbf{u}) \sim N(0, \tau^2\mathbf{I})$, and $\mathbf{I}$ is $n \times n$ identity matrix.

$$\text{Level 2:} \quad \mathbf{S}(\mathbf{u}) \sim N(0, \sigma^2\mathbf{R}_y(\varphi)) \tag{2.9}$$

where $\mathbf{R}_y(\varphi)$ is the $n \times n$ correlation matrix (hereafter denoted only by $\mathbf{R}_y$) with elements

$$r_{i,j} = \text{Corr}(Y_i, Y_j) = \exp\left(-\frac{\sqrt{(E_i - E_j)^2 + (N_i - N_j)^2}}{\varphi}\right) \tag{2.10}$$

where Corr stands for correlation.

$$\text{Level 3:} \quad (\boldsymbol{\beta}, \sigma^2, \boldsymbol{\varphi}, \tau^2) \sim p(\cdot) \tag{2.11}$$

where $p(\cdot)$ is a *prior* distribution. The model parameters for the above hierarchical scheme can be described as follows:

1. $\mathbf{Y}(\mathbf{u})$ is a random vector with components $Y(\mathbf{u}_1), Y(\mathbf{u}_2), \ldots, Y(\mathbf{u}_n)$, related to the precipitation measurements at the observation sites.

2. $\mathbf{X}(\mathbf{u})\boldsymbol{\beta} = \boldsymbol{\mu}(\mathbf{u})$ is the mean of $\mathbf{Y}(\mathbf{u})$ (hereafter denoted only by $\mathbf{Y}$). In geostatistical parlance, this is known as the trend. $\mathbf{X}(\mathbf{u})$ (hereafter denoted only as $\mathbf{X}$) is a matrix of fixed covariates measured at observation sites $\mathbf{u}_i$. $\boldsymbol{\beta}$ is a vector of regression parameters.

3. $\mathbf{S}(\mathbf{u})$ has zero mean and the covariance matrix $\sigma^2\mathbf{R}_y$. The elements of $\mathbf{R}_y$ are determined by the variogram function expressed in equation (2.4).

4. $\boldsymbol{\varepsilon}(\mathbf{u})$ denotes the Gaussian white noise vector at the observation sites with zero mean and covariance matrix $\tau^2\mathbf{I}$.

5. Level 3 specifies the prior distribution for the model parameters.

In our case, precipitation is the final result which can be determined by the summation of components of the above linear model and exponentiating. Note that only the variable $\mathbf{Y}$ is observable. Trend, signal, and variogram parameters all have to be estimated.

2.3 Kriging

The measured precipitation data can be used to estimate precipitation at some other locations in the study region. This estimation is done using kriging, a least squares linear prediction procedure, which, under certain stationarity assumptions, requires at least the knowledge of the covariance parameters (a.k.a. variogram parameters in geostatistics parlance) and the functional form for the mean of the underlying stochastic process. More often the covariance parameters are not known and hence spatial interpolation should be done using different approaches [15]. We chose a Bayesian approach primarily because we

can incorporate parameter uncertainties when deriving all the posterior distributions for these parameters and in the kriging prediction. It also allows us to exploit the modular structure of a Gibbs sampler (explained later in Chapter 3), thus incorporating the estimation of $\mathbf{W}$, as well.

Suppose we want to predict natural log precipitation at some unobserved location $\mathbf{u}_0$. Let's denote this unknown value as Y_0. Given the estimates of parameters σ^2, φ, and τ_R^2, we compute the correlation matrix $\mathbf{R}_y$, and the covariance matrix $\mathbf{C} = \mathbf{R}_y + \tau_R^2\mathbf{I}$ (so that $\text{Cov}(\mathbf{Y}) = \sigma^2\mathbf{C}$). We then predict "signal" S_0 at location $\mathbf{u}_0$ as:

$$\hat{S}_0 = \mathbf{r}'\mathbf{C}^{-1}\mathbf{V} \tag{2.12}$$

where $\mathbf{V} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$ is residual vector from regression, and $\mathbf{r}$ is a vector with elements $r_{0,i}$ given by the following expression:

$$r_{0,i} = \exp\left(-\frac{\sqrt{(E_0 - E_i)^2 + (N_0 - N_i)^2}}{\varphi}\right) \tag{2.13}$$

where E_0 and N_0 are the UTM easting and northing coordinates [17] of location $\mathbf{u}_0$. To predict Y_0, we will add the signal from equation (2.12) to regression surface, i.e.,

$$\hat{Y}_0 = \mathbf{X}_0'\hat{\boldsymbol{\beta}} + \hat{S}_0 \tag{2.14}$$

where $\mathbf{X}_0$ is the column vector of covariates at location $\mathbf{u}_0$. The entire procedure is known as "kriging with external trend" [6]. This method's prediction usually differs from the optimal Bayesian estimate $E[Y_0|\mathbf{Y}]$ [3], especially it underestimates the prediction variance. However, kriging is still very popular because of its high efficiency. We will employ this method to obtain precipitation maps,

and for cross-validation. Computation of $E[Y_0|\mathbf{Y}]$ can be incorporated into the Gibbs sampler.

CHAPTER 3

HIDDEN RANDOM FIELD MODELING

This chapter describes the entire methodology behind the proposed algorithm to estimate the hidden moisture flux direction random field (MFD RF), the model parameters, and further predict precipitation at unobserved locations. In the first section, we present a brief overview of Markov chain Monte Carlo (MCMC) methods, followed by Gibbs Sampling and the Metropolis algorithm used to construct a Markov chain having a stationary distribution. In the next couple of sections, we discuss the actual procedure of estimating MFD RF and the model parameters. Last section is devoted to the details of the sampling algorithm.

3.1 Background on Markov Chain Monte Carlo

Markov chain Monte Carlo (MCMC) is a method of doing Monte Carlo integration using Markov chains. It is used in Bayesian inference to draw a sample from the posterior distribution of model parameters given the data. When such a sample is drawn from a suitably constructed Markov chain which is run for a long time, we get Markov chain Monte Carlo integration [8]. Let's first try to understand the Bayesian formulation to obtain the posterior distribution.

Under Bayesian framework, all parameters in a statistical model are

considered as random quantities. Let $\mathbf{Y}$ denote the data, and θ denote model parameters. Bayesian inference requires setting up a joint probability distribution $p(\mathbf{Y}, \theta)$ over all random quantities. A full probability model can be written as:

$$p(\mathbf{Y}, \theta) = p(\mathbf{Y}|\theta) \cdot p(\theta) \tag{3.1}$$

where $p(\theta)$ is a prior distribution, and $p(\mathbf{Y}|\theta)$ is the likelihood. Using Bayes theorem, we can determine the distribution of θ conditional on $\mathbf{Y}$ as:

$$p(\theta|\mathbf{Y}) = \frac{p(\theta) \cdot p(\mathbf{Y}|\theta)}{\int p(\theta) \cdot p(\mathbf{Y}|\theta) d\theta} \tag{3.2}$$

Left hand side of equation (3.2) is called the posterior distribution of θ, and is the main objective of Bayesian inference. Any desirable feature of the posterior distribution can be expressed in terms of posterior expectations of functions of θ [8]. This posterior expectation can be written as:

$$E[f(\theta)|\mathbf{Y}] = \int f(\theta) \cdot p(\theta) \cdot p(\mathbf{Y}|\theta) d\theta \tag{3.3}$$

The integrations in equation (3.3) can be evaluated using MCMC.

Restating the above problem in more general terms, let X be a vector of random model parameters, with posterior distribution $\pi(\cdot) = p(\cdot|\mathbf{Y})$. So, the task is to evaluate the expectation

$$E[f(X)] = \frac{\int f(x)\pi(x)dx}{\int \pi(x)dx} \tag{3.4}$$

for some function of interest $f(\cdot)$.

3.1.1 Monte Carlo Integration

Monte Carlo integration evaluates $E[f(X)]$ by drawing samples X_t, $t = 1, \ldots, m$ from the posterior $\pi(\cdot)$ and then approximating

$$E[f(X)] \approx \frac{1}{m} \sum_{t=1}^{m} f(X_t)$$

So, a population mean of $f(X)$ is estimated by a sample mean [8]. Law of large numbers ensures that by increasing the sample size m, we can increase the accuracy of our approximation. One way of generating the sample set X_t is through a Markov chain having $\pi(\cdot)$ as its stationary distribution.

3.1.2 Markov Chains

A stochastic model has the Markov property if what happens at time $t + 1$ depends only on the state at time t, and not on previous history. That is, given X_t, the next state X_{t+1} is sampled from a distribution $p(X_{t+1}|X_t)$ which depends only on the current state of the chain, X_t. This sequence is called a Markov chain. The chain eventually converges to a unique stationary distribution, which is independent of t or initial state X_0. Thus, as t increases, the sampled points X_t resemble dependent samples from the stationary distribution. We can now estimate the expectation $E[f(X)]$ from the Markov chain output as:

$$\overline{f} = \frac{1}{m-k} \sum_{t=k+1}^{m} f(X_t) \tag{3.5}$$

Here, k denotes "burn-in" (a.k.a. pre-convergence time), the number of iterations needed for the Markov chain output to resemble stationary distribution. In some of our computations, we neglected the effect due to burn-in because

the output converged quickly. The expression in equation (3.5) is known as ergodic average and convergence to the desired expectation is ensured by the ergodic theorem [8].

3.1.3 Gibbs Sampling

Gibbs sampling, a phrase coined by Geman and Geman (1984) [8], is an iterative Monte Carlo method which produces Markov chains based on so-called full or complete conditional distributions. Gibbs sampling is useful when the multivariate conditional distributions are not in the closed form, which is more often the case for real-world problems, or in scenarios where it is impractical to sample from a large number of univariate conditional distributions [3]. For a M-dimensional problem with M random variables $(U_1, U_2, \ldots, U_M)$, a univariate substitution approach would require sampling from a chain of $M(M-1)$ univariate conditional distributions. A pre-requisite for using Gibbs sampling is to know all the full conditional distributions $p_i(U_i|U_{j\neq i}), i = 1, 2, \ldots, M$, either completely or partially (when sampled using the Metropolis algorithm which is discussed in Section 3.1.4). Under mild conditions (Besag, 1974) [3], these full conditional distributions uniquely determine the full joint distribution $p(U_1, U_2, \ldots, U_M)$, and marginal distributions $p(U_i), i = 1, 2, \ldots, M$. The steps involved in Gibbs sampling algorithm are as follows:

1. Starting with $\left(U_1^{(0)}, U_2^{(0)}, \ldots, U_M^{(0)}\right)$, complete one Gibbs iteration as fol-

lows:

$$\begin{aligned}
\text{Draw } U_1^{(1)} &\sim p_1\left(U_1|U_2^{(0)},\ldots,U_M^{(0)}\right)\\
\text{Draw } U_2^{(1)} &\sim p_2\left(U_2|U_1^{(1)},U_3^{(0)},\ldots,U_M^{(0)}\right)\\
\text{Draw } U_3^{(1)} &\sim p_3\left(U_3|U_1^{(1)},U_2^{(1)},U_4^{(0)},\ldots,U_M^{(0)}\right)\\
&\vdots\\
\text{Draw } U_M^{(1)} &\sim p_M\left(U_M|U_1^{(1)},\ldots,U_{M-1}^{(1)}\right)
\end{aligned}$$

2. Repeat the first step t times to obtain $\left(U_1^{(t)},U_2^{(t)},\ldots,U_M^{(t)}\right)$ such that it resembles the true stationary distribution of the Markov chain.

Hence, we can say that Gibbs sampling sequentially updates each estimated parameter from its full conditional until satisfactory convergence obeying the following theorem [3] is achieved.

Theorem: For the Gibbs sampling algorithm outlined above,

1. $\left(U_1^{(t)},\ldots,U_M^{(t)}\right)\rightarrow(U_1,\ldots,U_M)\sim p(U_1,\ldots,U_M)$ as $t\rightarrow\infty$.
2. The convergence in part (1) is exponential in t using the L_1 norm.

From a practical viewpoint, it is reasonable to say that MCMC algorithm converges after time T, when the output resembles the true stationary distribution of the Markov chain for all $t > T$ [3].

3.1.4 The Metropolis Algorithm

The Metropolis algorithm is an efficient way of constructing a Markov chain such that its stationary distribution coincides with the posterior distribution $\pi(\cdot)$. It is a special case of the "Metropolis-Hastings" (or "Hastings-Metropolis") algorithm due to Hastings (1970), and first proposed by Metropolis et al. (1953) [8].

Consider a *proposal* distribution $q(\cdot|X_t)$. We can sample a proposal value Y at time t from the proposal distribution. If the sampled value is accepted, the next state becomes $X_{t+1} = Y$. If not, the Markov chain does not move, i.e., $X_{t+1} = X_t$. The proposal distribution of the form $q(X|Y) = q(Y|X)$ is said to be symmetric for all X and Y, and is the one defining the Metropolis algorithm [8]. The probability of acceptance of Y is given as:

$$\alpha(X, Y) = \min \left(1, \frac{\pi(Y)}{\pi(X)}\right) \tag{3.6}$$

Thus, we can easily generate a Markov chain using the Metropolis algorithm as follows:

1. Draw a sample (proposal value) Y from $q(\cdot|X_t)$, where X_t is the current state of the Markov chain.

2. Compute the joint posterior densities at both the proposal value Y and the current value X_t.

3. Sample a uniform random variable U from $(0, 1)$.

4. If $U < \alpha(X, Y)$, then accept the proposal value, i.e., set the next state of Markov chain $X_{t+1} = Y$.

5. If $U > \alpha(X, Y)$, then reject the proposal value, i.e., set the next state of Markov chain $X_{t+1} = X_t$.

One attractive thing about using the Metropolis algorithm to generate a Markov chain is that we don't have to know the joint posterior distributions exactly, which is more often the case in Bayesian inference. Instead, as seen from equation (3.6), we only need to know them up to a constant.

3.2 Estimation of Moisture Flux Direction Random Field

The value of hidden moisture flux direction random field (MFD RF) W_i, is fitted at each of the observation sites i, $1 \leq i \leq n$. Thus, it is different for different sites, unlike the ASOADeK model which assumes a uniform MFD (recall from Chapter 1). Also, the values of W_i are taken modulo 2π, which renders a certain unique character to the model, compared to other random field models. For example, Gaussian priors for W_i are not applicable. On the other hand, issues of tail behavior are less important here, since the values of W_i belong to a compact set.

Using Bayesian approach and trying to mimic results from image processing [11], we specify a prior distribution for the values of W_i as:

$$p(W_1, \ldots, W_n) \propto \exp\left[\gamma \sum_{k \sim l} \omega_{kl} \cos(W_k - W_l)\right] \tag{3.7}$$

where the sum is taken over the pairs of neighboring sites k and l (for example, those that are less than a critical distance apart) and weights ω_{kl} reflect the distance between k^{th} and l^{th} gauge locations. The weights are chosen to be highest for gauge locations that are closest. For example, we can use weighting

by inverse square distance, or some other decreasing function of distance. We chose the critical distance $d_0 = 60$ km in our work. We chose the weights using the following relation:

$$\omega_{kl} = \frac{(d_0/2)^2}{d_{kl}^2} \tag{3.8}$$

where

$$d_{kl} = \sqrt{(E_k - E_l)^2 + (N_k - N_l)^2}$$

is the Euclidean distance between k^{th} and l^{th} gauge locations. As a result, uniformity is rewarded, but the differences between points are allowable, especially for the points that are far apart.

The choice of prior as given in equation (3.7) is analogous to the prior used in image processing models based on Ising model [11], with $\gamma \geq 0$ playing the role of the phase constant. High values of γ should lead to a more uniform random field, whereas $\gamma = 0$ will result in the orientation of W_i either parallel or opposite to the aspect A_i. Thus, γ can be considered as a "smoothing parameter.

In order to obtain the complete likelihood for $\mathbf{W}$, we should also consider $\mathbf{W}$ conditional on other model parameters (refer to equation (3.16) in Section 3.3). The complete likelihood can then be written as[1]:

$$\begin{aligned} & p(\mathbf{W}|\mathbf{Y}, \gamma_*, \boldsymbol{\beta}, \sigma^2, \varphi, \tau_R^2) \\ & \quad \propto \quad p(\mathbf{W}) \cdot p(\mathbf{Y}|\boldsymbol{\beta}, \sigma^2, \varphi, \tau_R^2, \mathbf{W}, \gamma_*) \\ & \quad \propto \quad \exp\left[\gamma_* \sum_{k \sim l} \omega_{kl} \cos(W_k - W_l) - \frac{1}{2}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})' \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})\right] \end{aligned} \tag{3.9}$$

[1]Hereafter, the subscript '$*$' indicates that the indexed parameter is assumed to be known.

where $\mathbf{\Sigma} = \sigma^2\mathbf{C}$ is the covariance matrix of $\mathbf{Y}$. Using equation (3.9), we can obtain the full conditional posterior densities[2], $p_i(W_i|W_1, \ldots, W_{i-1}, W_{i+1}, \ldots, W_n, \mathbf{Y}, \gamma_*, \boldsymbol{\beta}, \sigma^2, \varphi, \tau_R^2)$. These will be known only up to a constant, but we don't have to know the constant in order to use the Metropolis algorithm as discussed in Section 3.1.4. This part of the procedure is known as 'Metropolis within Gibbs' with single-site updating. That is, we loop over all the sites, each time sampling from the full conditional, for the implementation of a Gibbs sampler. The variogram parameters φ and τ_R^2 are also estimated using MCMC algorithm through 'Metropolis within Gibbs' sampling. Variogram parameter σ^2 and regression parameters $\boldsymbol{\beta}$, are also estimated using MCMC algorithm through Gibbs sampling without Metropolis step, using the approach described in Section 2.1, and it's equivalent to generalized least squares solution given in equation (2.3).

3.3 Estimation of Model Parameters

Geostatistical Bayesian framework developed by Ribeiro and Diggle [15] is used in this work. The joint prior distribution for variogram and regression parameters (a.k.a. the model parameters) can be factorized and written as:

$$p(\boldsymbol{\beta}, \sigma^2, \varphi, \tau_R^2) = p(\varphi, \tau_R^2) \cdot p(\boldsymbol{\beta}, \sigma^2|\varphi, \tau_R^2) \tag{3.10}$$

The uncertainty in the model parameters is considered when deriving all the posterior distributions. The joint posterior distribution for the model

[2]We use the terms 'distribution' and 'density' interchangeably.

parameters can be factorized and written as:

$$p(\boldsymbol{\beta}, \sigma^2, \varphi, \tau_R^2 | \mathbf{Y}, \gamma_*, \mathbf{W}) = p(\boldsymbol{\beta}, \sigma^2 | \mathbf{Y}, \gamma_*, \mathbf{W}, \varphi, \tau_R^2) \cdot p(\varphi, \tau_R^2 | \mathbf{Y}, \gamma_*, \mathbf{W}) \quad (3.11)$$

Therefore,

$$p(\varphi, \tau_R^2 | \mathbf{Y}, \gamma_*, \mathbf{W}) = \frac{p(\boldsymbol{\beta}, \sigma^2, \varphi, \tau_R^2 | \mathbf{Y}, \gamma_*, \mathbf{W})}{p(\boldsymbol{\beta}, \sigma^2 | \mathbf{Y}, \gamma_*, \mathbf{W}, \varphi, \tau_R^2)} \quad (3.12)$$

The joint posterior distribution in the numerator of equation (3.12) can also be expressed in terms of the prior distribution and likelihood as:

$$p(\boldsymbol{\beta}, \sigma^2, \varphi, \tau_R^2 | \mathbf{Y}, \gamma_*, \mathbf{W}) \propto p(\boldsymbol{\beta}, \sigma^2, \varphi, \tau_R^2) \cdot p(\mathbf{Y}, \gamma_*, \mathbf{W} | \boldsymbol{\beta}, \sigma^2, \varphi, \tau_R^2) \quad (3.13)$$

The posterior distribution in the denominator of equation (3.12) can be factorized and written as:

$$p(\boldsymbol{\beta}, \sigma^2 | \mathbf{Y}, \gamma_*, \mathbf{W}, \varphi, \tau_R^2) = p(\boldsymbol{\beta} | \mathbf{Y}, \gamma_*, \sigma^2, \mathbf{W}, \varphi, \tau_R^2) \cdot p(\sigma^2 | \mathbf{Y}, \gamma_*, \mathbf{W}, \varphi, \tau_R^2) \quad (3.14)$$

Using equations (3.13) and (3.14), the joint posterior distribution given in equation (3.12) can now be expressed as:

$$p(\varphi, \tau_R^2 | \mathbf{Y}, \gamma_*, \mathbf{W}) \propto \frac{p(\boldsymbol{\beta}, \sigma^2, \varphi, \tau_R^2) \cdot p(\mathbf{Y}, \gamma_*, \mathbf{W} | \boldsymbol{\beta}, \sigma^2, \varphi, \tau_R^2)}{p(\boldsymbol{\beta} | \mathbf{Y}, \gamma_*, \sigma^2, \mathbf{W}, \varphi, \tau_R^2) \cdot p(\sigma^2 | \mathbf{Y}, \gamma_*, \mathbf{W}, \varphi, \tau_R^2)} \quad (3.15)$$

The distributions in the numerator of equation (3.15) are given by the joint prior distribution from equation (3.10) and the likelihood function:

$$L(\boldsymbol{\beta}, \sigma^2, \varphi, \tau_R^2 | \mathbf{Y}, \gamma_*, \mathbf{W}) \propto (\sigma^2)^{-\frac{n}{2}} |\mathbf{C}|^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{C}^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \right] \quad (3.16)$$

where

$$\mathbf{C} = \mathbf{R}_y + \tau_R^2 \mathbf{I} \quad (3.17)$$

Recall from Chapter 2 that $\mathbf{R}_y$ is the correlation matrix, τ_R^2 is the relative nugget and $\mathbf{I}$ is $n \times n$ identity matrix. By using the covariance matrix $\mathbf{C}$, we account for the fact that the gauge precipitation data are spatially correlated.

The conditional posterior distributions in the denominator of equation (3.15) have the following forms [15]:

$$(\boldsymbol{\beta}|\mathbf{Y}, \gamma_*, \sigma^2, \mathbf{W}, \varphi, \tau_R^2) \sim N(\hat{\boldsymbol{\beta}}, \sigma^2 \mathbf{V}_{\hat{\boldsymbol{\beta}}}) \tag{3.18}$$

$$(\sigma^2|\mathbf{Y}, \gamma_*, \mathbf{W}, \varphi, \tau_R^2) \sim \chi^2_{ScI}(n-p, S^2) \tag{3.19}$$

where

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{C}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{C}^{-1}\mathbf{Y} \tag{3.20}$$

$$\mathbf{V}_{\hat{\boldsymbol{\beta}}} = (\mathbf{X}'\mathbf{C}^{-1}\mathbf{X})^{-1} \tag{3.21}$$

$$S^2 = \frac{1}{n-p}(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})\mathbf{C}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \tag{3.22}$$

Again recall from Chapter 2 that n is the number of observation sites, and p is the number of regression parameters (elements of $\boldsymbol{\beta}$). Also, the expression for fitted parameters $\hat{\boldsymbol{\beta}}$ given in equation (3.20) is equivalent to the generalized least squares expression given in equation (2.3) where $\boldsymbol{\Sigma} = \sigma^2\mathbf{C}$. $N(\hat{\boldsymbol{\beta}}, \sigma^2\mathbf{V}_{\hat{\boldsymbol{\beta}}})$ is a p-dimensional Multivariate Normal (MVN) distribution with mean $\hat{\boldsymbol{\beta}}$ and covariance matrix $\sigma^2\mathbf{V}_{\hat{\boldsymbol{\beta}}}$. $\chi^2_{ScI}(n-p, S^2)$ is a Scaled-Inverse-χ^2 distribution with $(n-p)$ degrees of freedom.

The corresponding multivariate Normal and Scaled-Inverse-χ^2 probability density functions for $\boldsymbol{\beta}$ and σ^2 can be written as:

$$p(\boldsymbol{\beta}) = \frac{1}{(2\pi)^{\frac{p}{2}}|\sigma^2\mathbf{V}_{\hat{\boldsymbol{\beta}}}|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})'(\sigma^2\mathbf{V}_{\hat{\boldsymbol{\beta}}})^{-1}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})\right] \tag{3.23}$$

$$p(\sigma^2) = \frac{\left(\frac{n-p}{2}\right)^{\left(\frac{n-p}{2}\right)}}{\Gamma(\frac{n-p}{2})}(S^2)^{\left(\frac{n-p}{2}\right)}(\sigma^2)^{-\left(\frac{n-p}{2}+1\right)}\exp\left[-\frac{(n-p)S^2}{2\sigma^2}\right] \tag{3.24}$$

The conditional posterior distributions for $\boldsymbol{\beta}$ and σ^2 as given in equations (3.18) and (3.19) are obtained using an improper prior:

$$p(\boldsymbol{\beta}, \sigma^2|\varphi, \tau_R^2) = \frac{1}{\sigma^2} \tag{3.25}$$

This is a commonly adopted prior distribution for $(\boldsymbol{\beta}, \sigma^2)$ encountered in Bayesian inference for Gaussian linear models [15]. Some of its properties are:

- It is improper because it doesn't integrate to one.
- It corresponds to zero 'prior observations', reflecting our ignorance about σ^2 and $\boldsymbol{\beta}$.
- It also corresponds to the Jeffrey's prior [15].
- It also coincides with the generalized least squares approach [10].

Using the conditional posterior distributions from equations (3.18) and (3.19), the joint posterior distribution for $(\boldsymbol{\beta}, \sigma^2)$ can be written as a Normal-Scaled-Inverse-χ^2, i.e., a product of Normal and Scaled-Inverse-χ^2 densities [15]:

$$(\boldsymbol{\beta}, \sigma^2|\mathbf{Y}, \gamma_*, \mathbf{W}, \varphi, \tau_R^2) \sim N(\hat{\boldsymbol{\beta}}, \mathbf{V}_{\hat{\boldsymbol{\beta}}}) \cdot \chi^2_{ScI}(n-p, S^2) \tag{3.26}$$

The joint posterior distribution for (φ, τ_R^2) as given in equation (3.15) can be greatly simplified using equations (3.10), (3.16), (3.23), (3.24), and (3.25), and written as:

$$p(\varphi, \tau_R^2|\mathbf{Y}, \gamma_*, \mathbf{W}) \propto p(\varphi, \tau_R^2) \cdot |\mathbf{V}_{\hat{\boldsymbol{\beta}}}|^{\frac{1}{2}}|\mathbf{C}|^{-\frac{1}{2}}(S^2)^{\left(-\frac{n-p}{2}\right)} \tag{3.27}$$

The joint prior distribution for (φ, τ_R^2) can be factorized and written as:

$$p(\varphi, \tau_R^2) = p(\varphi) \cdot p(\tau_R^2) \tag{3.28}$$

We chose *uniform* priors for φ and τ_R^2 to obtain the joint posterior distribution for (φ, τ_R^2) using equation (3.27).

$$p(\varphi) = \text{ Uniform}(\varphi_{min}, \varphi_{max}) \tag{3.29}$$

$$p(\tau_R^2) = \text{ Uniform}(\tau_{R_{min}}^2, \tau_{R_{max}}^2) \tag{3.30}$$

$\varphi_{min}, \varphi_{max}, \tau_{R_{min}}^2$, and $\tau_{R_{max}}^2$ are sensibly chosen and they may be different for different months.

Equation (3.27) doesn't define a standard probability distribution [15]. Hence, we adopt the method of inference by simulation by taking samples from the above joint posterior distribution. Originally in [15], the authors propose a *grid* approach to sample (φ, τ_R^2), however, we found that using the Metropolis algorithm to jointly sample (φ, τ_R^2) is computationally much more efficient. In the grid approach, the distribution of (φ, τ_R^2) is discretized on a two-dimensional grid. Then we sample from this discrete distribution in order to perform a Gibbs step for (φ, τ_R^2). Thus, a single Gibbs step becomes computationally intensive. In order to obtain reliable MCMC results assuring convergence and low sampling error, we need to iterate for many such Gibbs steps (say, on the order of $100,000$). Besides, gridding causes discretization errors. We have to evaluate equation (3.27) - the costliest in our computation - just twice (using old and proposed values) per cycle using the Metropolis approach.

3.4 The Sampling Algorithm

The sampling algorithm described above can be detailed as follows:

1. Draw a $\mathbf{W}$ sample using $p_i(W_i|W_1,\ldots,W_{i-1},W_{i+1},\ldots,W_n,\mathbf{Y},\gamma_*,\boldsymbol{\beta},\sigma^2,\varphi,\tau_R^2)$, with i going from 1 through n. This step is done using random-walk Metropolis algorithm which is described below in section on "Metropolis Sampling of $\mathbf{W}$". Initially we assume some suitable values for $\varphi,\tau_R^2,\sigma^2$, $\boldsymbol{\beta}$, and $\mathbf{W}$ but for each subsequent iteration, we use $\varphi,\tau_R^2,\sigma^2,\boldsymbol{\beta}$, and $\mathbf{W}$ values computed from the previous iteration.

2. Draw a sample from $p(\boldsymbol{\beta},\sigma^2,\varphi,\tau_R^2|\mathbf{Y},\gamma_*,\mathbf{W})$ to estimate variogram and regression parameters. This involves the following substeps:

   (a) Draw a (φ,τ_R^2) pair using $p(\varphi,\tau_R^2|\mathbf{Y},\gamma_*,\mathbf{W})$ as given in equation (3.27). We use $\mathbf{W}$ previously computed in the same iteration. This step also involves a Metropolis substep for selecting (φ,τ_R^2) as explained below in Section on "Metropolis Sampling of (φ,τ_R^2) Pair".

   (b) Draw a σ^2 sample using $(\sigma^2|\mathbf{Y},\gamma_*,\mathbf{W},\varphi,\tau_R^2)$ as given in equation (3.19). We use φ and τ_R^2 values previously computed in the same iteration.

   (c) Draw a $\boldsymbol{\beta}$ sample using $(\boldsymbol{\beta}|\mathbf{Y},\gamma_*,\sigma^2,\mathbf{W},\varphi,\tau_R^2)$ as given in equation (3.18). We use φ,τ_R^2, and σ^2 values previously computed in the same iteration.

   $\mathbf{W}$ used in this step is always the one computed in the same iteration using step 1.

3. Apply steps 1 and 2 repeatedly for sufficiently large number of iterations to ensure convergence and high enough sample size sufficient for the Monte Carlo inference.

Metropolis Sampling of W

1. Generate proposal W_i^* as:

$$W_i^* = W_{i,old} + \text{Uniform}(-h, h) \tag{3.31}$$

2. Compute d_i^*, the relative log-likelihood of W_i^*, and d_i, the relative log-likelihood of $W_{i,old}$ using equation (3.9).

3. If $d_i^* > d_i$, then accept proposal, i.e., set $W_{i,new} = W_i^*$.

4. If $d_i^* < d_i$, then accept or reject proposal, W_i^*, with probability depending on d_i^* and d_i. The probability expression chosen is [11]:

$$\alpha(W_i^*, W_{i,old}) = \min(1, \exp(d_i^* - d_i)) \tag{3.32}$$

5. Sample a uniform random variable U from $(0, 1)$.

6. If $\alpha(W_i^*, W_{i,old}) > U$, then accept proposal, i.e., set $W_{i,new} = W_i^*$, otherwise reject proposal, i.e., set $W_{i,new} = W_{i,old}$.

The maximum random-walk step-size h is chosen adaptively in order to maintain the overall acceptance rate between $20 - 30\%$ which is considered most efficient in Metropolis algorithm [3]. Choice of h is slightly sensitive to the smoothing parameter γ but highly sensitive to the data (signal); the higher the *signal-to-noise* ratio is, the smaller h is, and vice versa. Here, signal-to-noise ratio is defined as β_4/σ.

Metropolis Sampling of (φ, τ_R^2) Pair

1. Generate proposals φ_i^* and $\tau_{R_i}^{2*}$ as:

$$\varphi_i^* = \text{Uniform}(\varphi_{min}, \varphi_{max}) \tag{3.33}$$

$$\tau_{R_i}^{2*} = \text{Uniform}(\tau_{R_{min}}^2, \tau_{R_{max}}^2) \tag{3.34}$$

2. Compute a relative likelihood of $(\varphi_i^*, \tau_{R_i}^{2*})$ pair as l_i^*, and a relative likelihood of $(\varphi_{i,old}, \tau_{R_{i,old}}^2)$ pair as l_i, by taking natural log of equation (3.27).

3. If $l_i^* > l_i$, then accept proposals, i.e., set $\varphi_{i,new} = \varphi_i^*$, and $\tau_{R_{i,new}}^2 = \tau_{R_i}^{2*}$.

4. If $l_i^* < l_i$, then accept or reject proposals $(\varphi_i^*, \tau_{R_i}^{2*})$, with probability depending on l_i^* and l_i. The probability expression chosen is [11]:

$$\alpha((\varphi_i^*, \tau_{R_i}^{2*}), (\varphi_{i,old}, \tau_{R_{i,old}}^2)) = \min(1, \exp(l_i^* - l_i)) \tag{3.35}$$

5. Sample a uniform random variable U from $(0, 1)$.

6. If $\alpha((\varphi_i^*, \tau_{R_i}^{2*}), (\varphi_{i_{old}}, \tau_{R_{i,old}}^2)) > U$, then accept proposals, i.e., set$(\varphi_{i,new}$, $\tau_{R_{i,new}}^2) = (\varphi_i^*, \tau_{R_i}^{2*})$, otherwise reject proposals, i.e., set $(\varphi_{i,new}, \tau_{R_{i,new}}^2) = (\varphi_{i,old}, \tau_{R_{i,old}}^2)$.

Implementation Issues

The actual implementation of the sampling algorithm is done in a slightly different way. Instead of updating W_i, $1 \leq i \leq n$, for a given i^{th} observation site, using the W_j's, $1 \leq j < i$, computed for the remaining sites in the same iteration, we update it using values of W_j from the previous iteration. This simplifies computation to some extent without adversely affecting the

results. Also, there is an "aliasing" effect of sign of β_4 with **W** direction. That is,

$$\beta_4 \cos(A_i - W_i) = -\beta_4 \cos(A_i - (W_i - \pi)) \tag{3.36}$$

From equation (3.36) we can say that switching the sign of β_4 is equivalent to switching the moisture flux direction to the opposite. In this work, we avoid this aliasing effect by restricting β_4 to be positive.

CHAPTER 4

RESULTS

In this chapter, we present the results obtained using simulated as well as actual precipitation data. Simulated data were used to validate the approach at each stage as summarized in the sampling algorithm given in Chapter 3, in order to test the robustness of the MCMC approach, and in "cross-validation" process. Analysis was done separately for each month. We present the following key results which serve as representatives for their categories - one complete set of MCMC outputs for the month of August, set of outputs showing the influence of the smoothing parameter γ, results showing common features in MFD RF for different months keeping the same γ, and cross-validation results which would help us choose an optimal value of the smoothing parameter γ for a particular month.

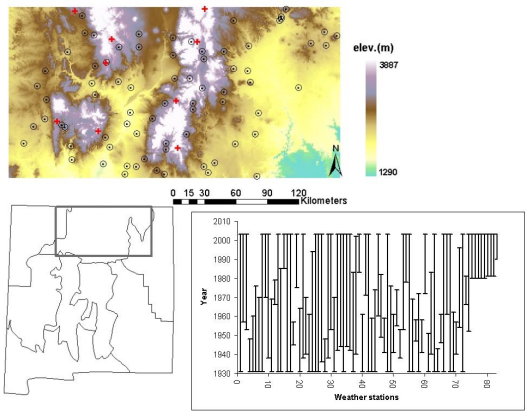
4.1 Study Area

Our model is used to estimate the moisture flux direction and subsequently obtain monthly precipitation maps for the mountainous region in semi-arid northern New Mexico.

There are three NCDC (National Climate Data Center) climate divisions in northern New Mexico [9]. Figure 4.1 shows the DEM and the weather stations for climate division 2. Division 2 consists of mountainous region (San-

gre de Cristo mountains) with inter-mountain valleys. Division 5 covers the central valley along the Rio Grande rift, and Division 6 covers the central highlands. Division 2, due to its mountainous terrain, is an ideal candidate for our study. The precipitation data from Division 2 were originally picked for ASOADeK model [9]; we use them in this study, partially to be able to compare our results with ASOADeK results. In division 2, the mountain elevation ranges from 1290m to 3887m as per the 1km-resolution DEM map shown in Figure 4.1, which was resampled from a 60m-resolution DEM [9]. The precipitation is measured using 83 gauge stations, i.e., $n = 83$. 74 of these stations are the NCDC stations which have at least 10-year data available in the period from 1931 to 2005. The remaining 9 are SNOTEL (SNOwpack TELemetry) stations. They are fairly new with data available from 1980 to 2005. The duration of the available data (NCDC and SNOTEL) is also shown in Figure 4.1.

Figure 4.1: Study Area in Northern New Mexico - A Division 2 DEM with gauge stations ('+' indicate the SNOTEL gauges), and the period of available data



4.2 MCMC Results

4.2.1 One Complete MCMC Output

We tested our model using observed precipitation data for different months and obtained the model parameters as shown in Tables 4.1 and 4.2. As our final estimates, we take mean values of $\boldsymbol{\beta}$ and median values of $\sigma^2, \varphi, \tau_R^2$ from their respective Markov chains with burn-in $\approx 2,000$. The medians are chosen because of highly skewed nature of those distributions. We ran the sampling algorithm for $40,000$ iterations for approximate CPU time of 16 mins on AMD Athlon(tm) XP 2800+ processor to obtain the entire set of estimates for a single month. We chose the initial values as $\boldsymbol{\beta} = (0, 0, 0, 0, 1)'$, $\sigma^2 = 0.01$, $\varphi = 50$, and $\tau_R^2 = 0.5$. We used $\gamma = 2$, the one obtained from cross-validation results as explained in Section on cross-validation results. Figures 4.2 through

Table 4.1: Regression parameter estimates for 3 different months

Month	β_0	β_1	β_2	β_3	β_4
February	-0.1394	-0.0027	0.0011	1.2961	0.2398
May	2.481	0.0037	0.0002	0.4715	0.1017
August	3.2108	0.0014	-0.0005	0.4232	0.1259

Table 4.2: Variogram parameter estimates and acceptance rates for 3 different months

Month	σ^2	φ	τ_R^2	**W** Accp. Rate	(φ, τ_R^2) Accp. Rate
February	0.0707	245.708	1.2072	23.92	44.14
May	0.0395	162.293	0.4540	22.18	13.04
August	0.0751	139.384	0.048	23.095	5.2

4.6 show one complete set of MCMC results for the month of August.

Figure 4.2: One complete MCMC output for August - Markov chain values and histograms of regression parameters β_0 and β_3.

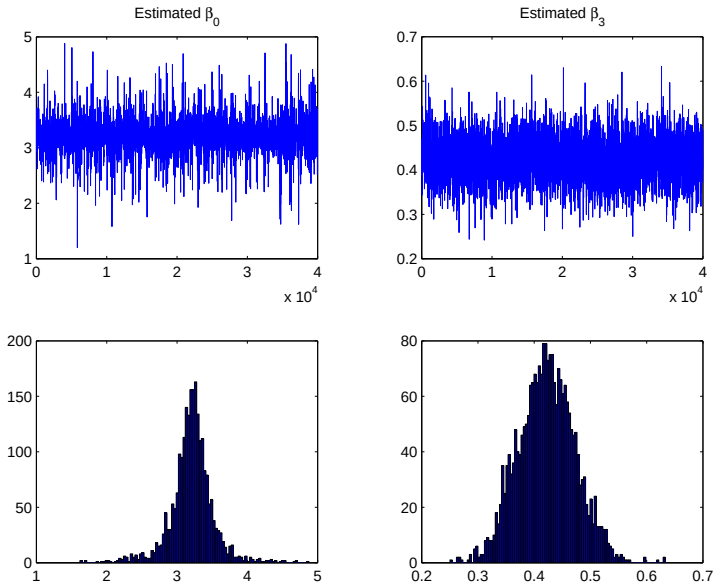


Figure 4.3: One complete MCMC output for August - Markov chain values and histograms of regression parameters β_1 and β_2.

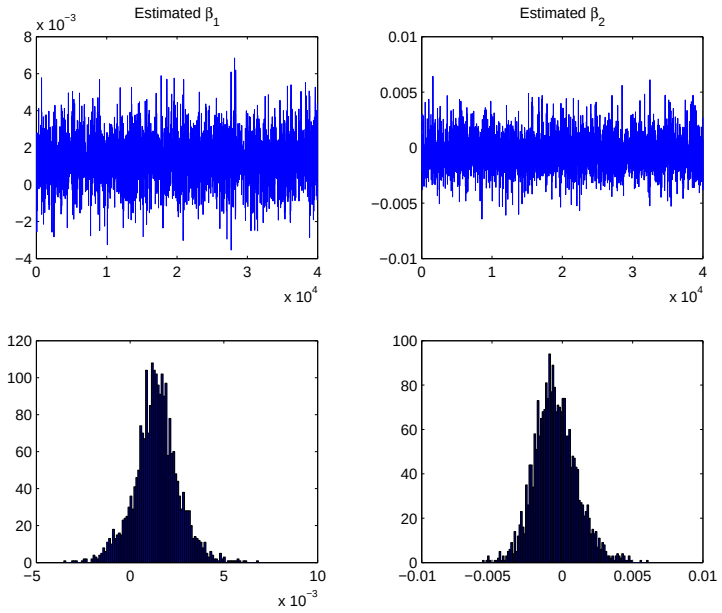


Figure 4.4: One complete MCMC output for August - Markov chain values, histograms, and Auto-correlation functions of regression parameter β_4 and variogram parameter σ^2.

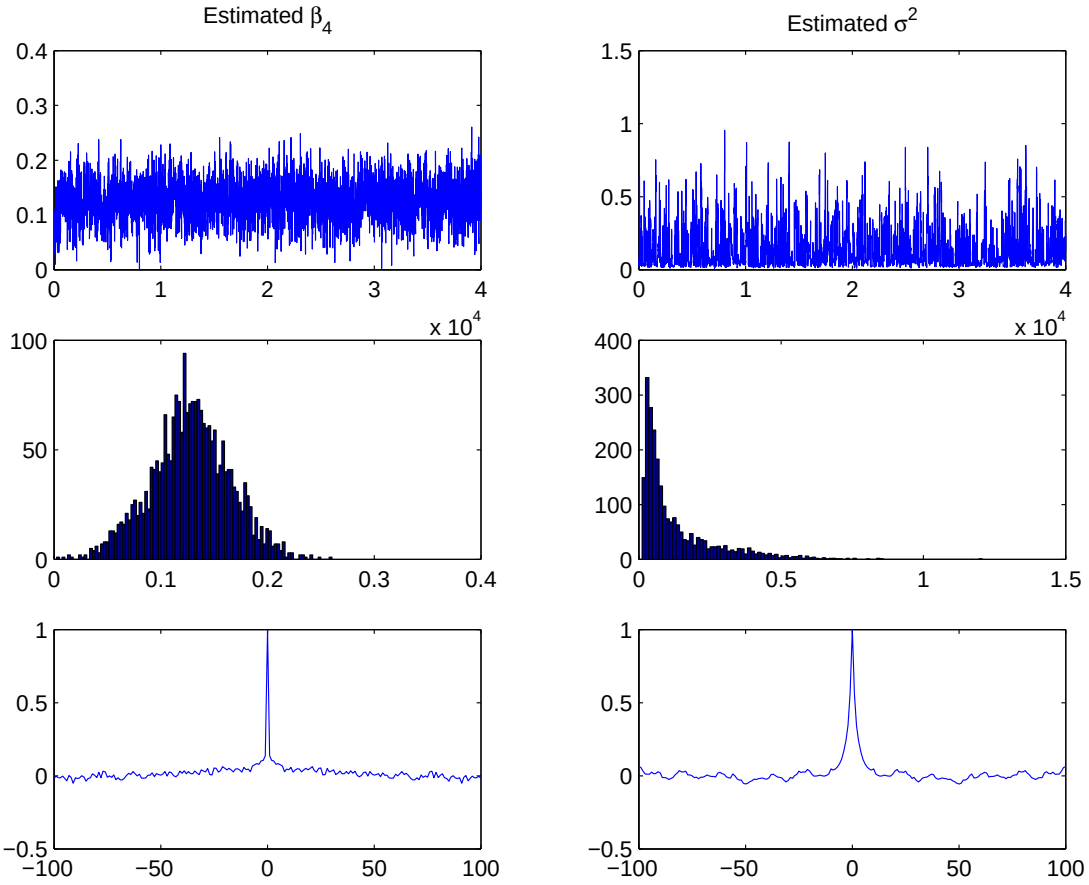


Figure 4.5: One complete MCMC output for August - Markov chain values and histograms of variogram parameters φ and τ_R^2.

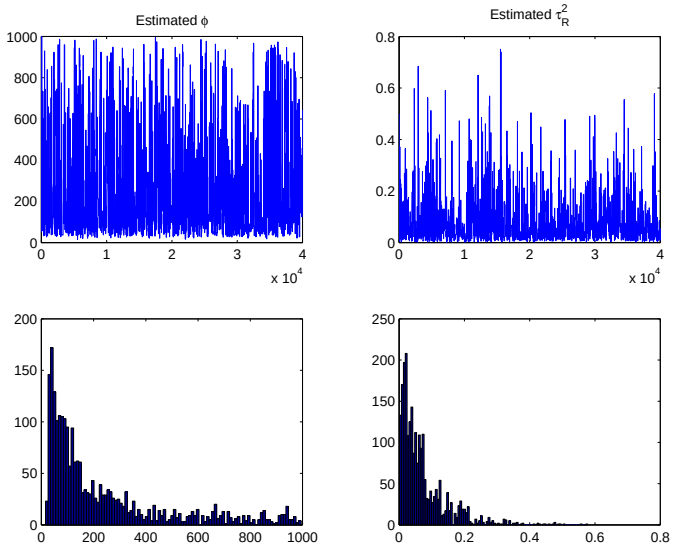
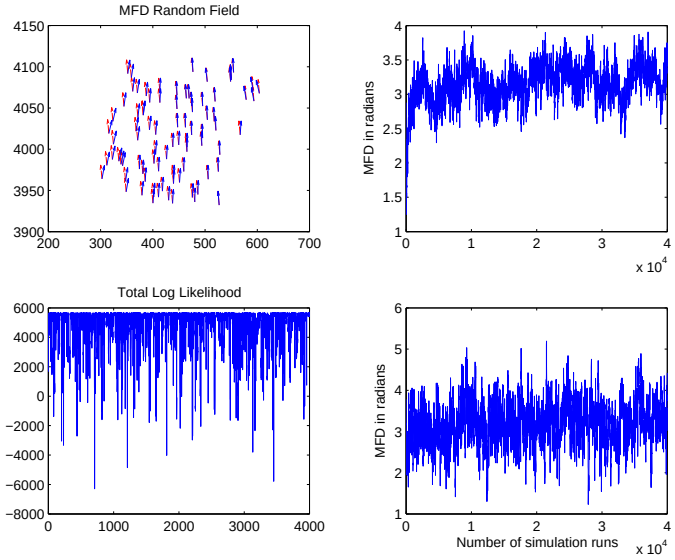


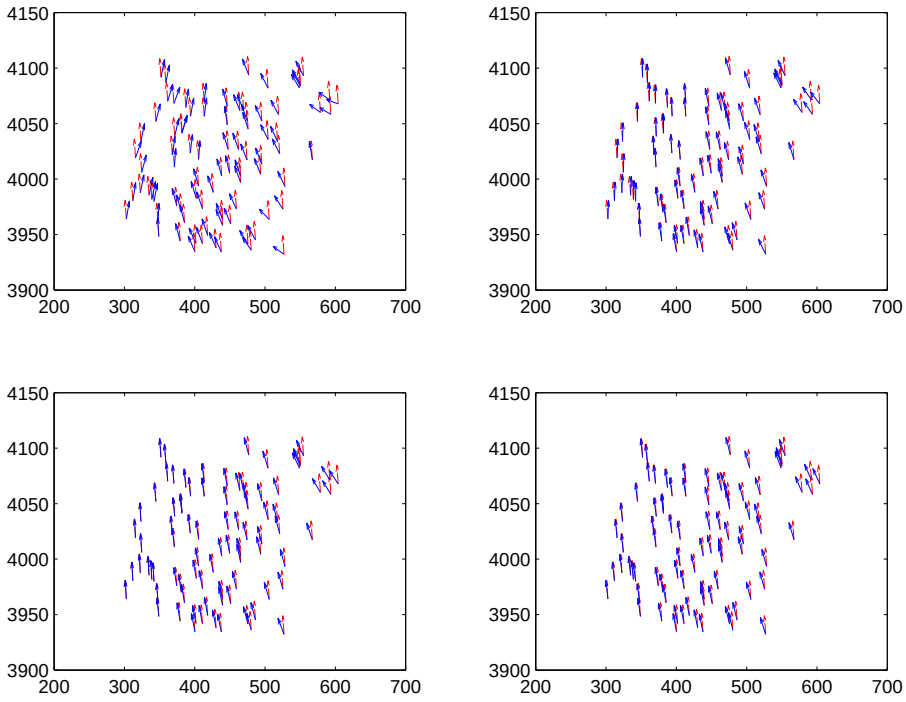
Figure 4.6: One complete MCMC output for August - Moisture flux direction random fields (dashed line - constant ASOADeK MFD, solid line - estimated MFD RF), log-likelihood, and Markov chain values for 2 observation sites.



4.2.2 Influence of γ

We also investigated the effect of four different γ's ($\gamma = 0.5, 2, 4, 8$) on the moisture flux direction random field using the precipitation data for the month of May. Once again, we ran the sampler for $40,000$ iterations with burn-in $\approx 2,000$. The plots are shown in Figure 4.7. One can observe the increasing smoothness of the field as parameter γ increases.

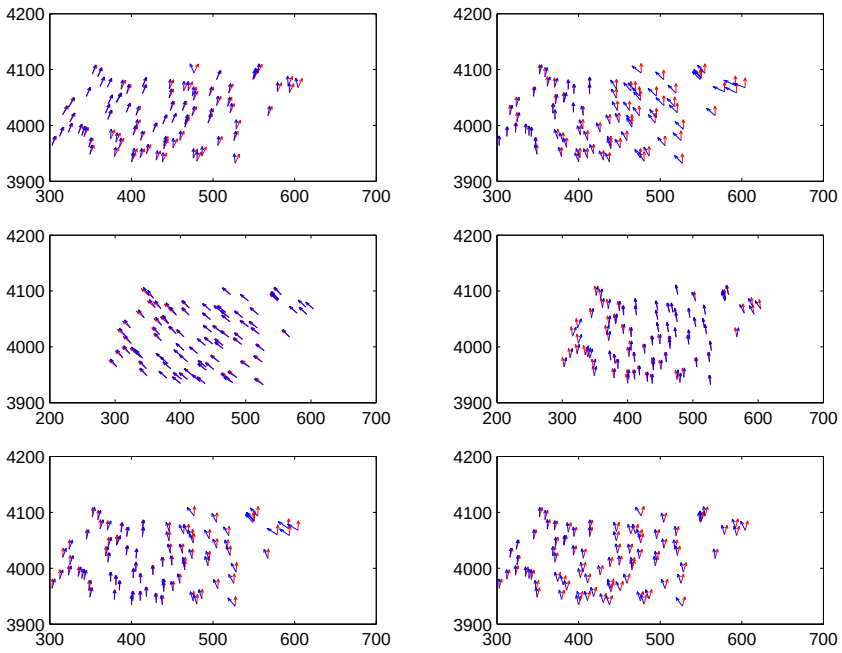
Figure 4.7: Influence of γ's - MFD RF's for May for $\gamma = 0.5, 2, 4$, and 8 (dashed line - constant ASOADeK MFD RF, solid line - estimated MFD RF).



4.2.3 MFD RF Comparison

In order to investigate the effect due to moisture flux direction random field **W** on orographic precipitation, we chose one particular value of the smoothing parameter γ, and obtained results for the random field estimates for six different months. Figure 4.8 shows the different MFD RF's. One can see that for all the six months, MFD RF is an average of two prevailing moisture directions, moisture from Pacific coast on the west, and moisture from the Gulf of Mexico on the south-west. We ran the sampler for $40,000$ iterations with burn-in $\approx 2,000$.

Figure 4.8: MFD RF Comparison - MFD RF's for February, April, June, August, October, and December obtained using $\gamma = 1$ (dashed line - constant ASOADeK MFD RF, solid line - estimated MFD RF).



4.3 Cross-validation

In order to validate our results, we used the technique of cross-validation. In cross-validation, we leave out some data points and try to estimate the values at those 'missing' points. The higher the match between the actual and the estimated values at those points is, the better the model is. We did cross-validation experiments by considering $\sim 90\%$ of the data and estimating the remaining $\sim 10\%$ values. We used simulated data, as well as real data for cross-validation. The main purpose for doing cross-validation was to 'calibrate' the smoothing parameter γ, i.e., to find out which γ produces precipitation estimates with minimum *mean absolute error* (MAE) and minimum *mean squared error* (MSE).

4.3.1 Cross-validation Using Simulation Data

We assigned the following values to the model parameters treating them as 'true' values to run the simulation - $\mathbf{W}_{true}$ is a random field with two moisture flux directions $3\pi/4$ and $5\pi/4$, $\boldsymbol{\beta}_{true} = (0, 0, 0, 0, 1)'$, $\sigma^2_{true} = 0.01$, $\varphi_{true} = 50$, and $\tau^2_{R_{true}} = 0.5$.We used the same covariates as the real data. We randomly selected $\sim 10\%$ observation sites to predict the'simulated' precipitation values for, using the remaining simulated data. Keeping the same set of sites, we ran the sampler for $\gamma = 0.5, 1, 3$, and 8 for $40,000$ Gibbs steps without burn-in. We repeated this entire procedure for 32 different combinations of '$\sim 10\%$ discarded sites' to generate large enough sample size $N = 144$ (16 combinations were obtained using kriging whereas the remaining 16 were obtained using the full Bayesian predictive distribution approach.). We calculated the t-statistics (t is Student's t distribution) and p-values (by considering $\gamma = 8$

as reference) for 3 pairwise differences namely, $\gamma = 0.5$ and 8, $\gamma = 2$ and 8, and $\gamma = 4$ and 8. From the statistical analysis, we found that all these γ's gave better estimates than $\gamma = 8$. We also compared the MAE's and MSE's for other two pairs, $\gamma = 0.5$ and $\gamma = 1$, and $\gamma = 3$ and $\gamma = 1$, and found that $\gamma = 1$ gave the best estimates. These results are listed in Table 4.3. They show improvement over ASOADeK's constant MFD (since larger γ corresponds to the case of constant MFD). The negative values for mean indicate that those γ's are better than the reference gamma, with mean for $\gamma = 1$ being the most negative. MAE for each of the 32 sets of 'discarded' sites are shown in Figure 4.9. The thick lines show the average MAEs obtained using kriging and the full Bayesian approach. The actual versus predicted values are shown in figure 4.10. Figures 4.11 and 4.12 show the MCMC outputs for cross-validation using simulated data using the full Bayesian approach.

Table 4.3: t-statistics and p-values using simulated data

γ pairs	N	Mean	Std. Dev.	t-stats	p-value
$0.5 - 8$	144	-0.1003	0.2117	-5.6854	$6.525e-09$
$1 - 8$	144	-0.1332	0.1888	-8.4661	$1.269e-17$
$3 - 8$	144	-0.0853	0.1175	-8.7115	$1.500e-18$
$0.5 - 1$	144	0.0329	0.1359	2.9051	$1.836e-03$
$3 - 1$	144	0.0479	0.1252	4.5911	$2.205e-06$

4.3.2 Cross-validation Using Northern New Mexico Data

We randomly selected $\sim 10\%$ observation sites to predict the 'true' precipitation values for May, using the remaining actual May data. Keeping the same set of sites, we ran the sampler for $\gamma = 0.5, 1, 3$, and 8 for $40,000$

Gibbs steps without burn-in. We repeated this entire procedure for 36 different combinations of '$\sim$ 10% discarded sites' to generate a sample size $N = 324$. We calculated the t-statistics and p-values (by considering $\gamma = 8$ as reference) for 3 pairwise differences namely, $\gamma = 0.5$ and 8, $\gamma = 2$ and 8, and $\gamma = 4$ and 8. From the statistical analysis, we found that all these γ's gave almost similar estimates as $\gamma = 8$. These results are listed in Table 4.4. The positive values for mean indicate that those γ's are worse than the reference gamma, with mean for $\gamma = 0.5$ being the most positive. MAE for each of the 36 sets of 'discarded' sites are shown in Figure 4.13. The thick line shows the average MAEs obtained using the full Bayesian approach. The actual versus predicted values are shown in figure **??**. Figures 4.15 and 4.16 show the MCMC outputs for cross-validation using actual data using the full Bayesian approach.

The output appears to be not very sensitive to the omission of a small fraction of the observation sites.

Table 4.4: t-statistics and p-values using May data

γ pairs	N	Mean	Std. Dev.	t-stats	p-value
$0.5 - 8$	324	0.1024	1.2198	1.5111	$6.539e-02$
$1 - 8$	324	0.0435	1.0932	0.7162	$2.369e-01$
$3 - 8$	324	0.0692	0.9882	1.2605	$1.038e-01$

4.4 Monthly Precipitation Map

Finally, with the MCMC estimates obtained for various model parameters, we constructed the precipitation map for the study region by predicting values using kriging technique. Figure 4.18 shows the precipitation map for

the month of February obtained using our model. The model parameters were estimated using $\gamma = 0.5$. Figure **??** shows the precipitation map for the same month obainted using ASOADeK model. The white patch in the lower right corner indicates negative estimates.

Figure 4.9: Cross-validation results using 32 different sets of 'discarded' sites for simulated data.

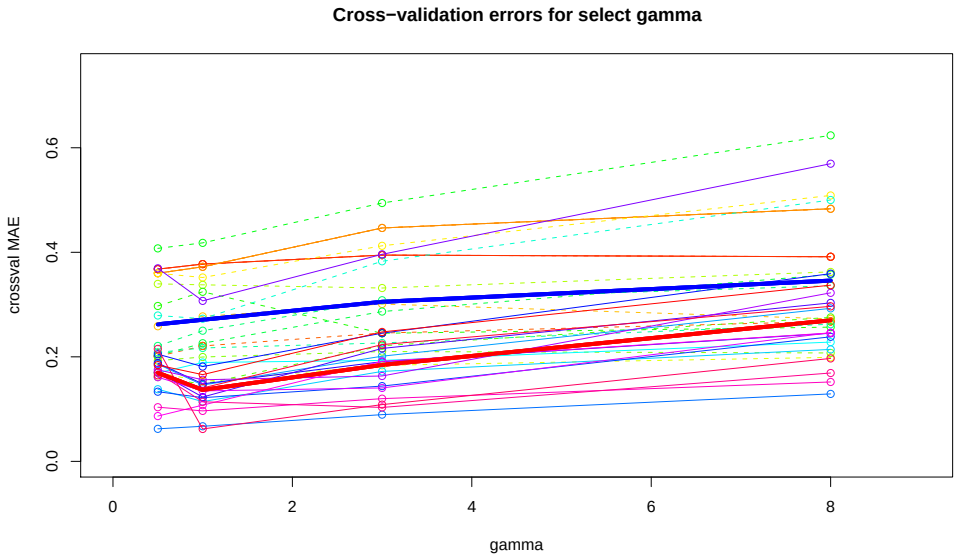


Figure 4.10: Cross-validation results showing actual versus predicted values for simulated data for $\gamma = 0.5, 1, 3, 8$.

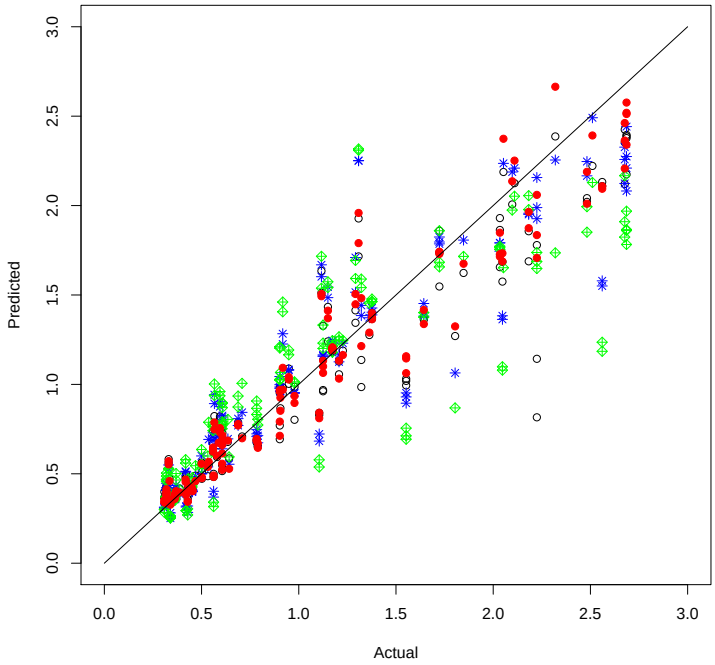


Figure 4.11: Cross-validation results using simulated data - Markov chain values, histograms, and Auto-correlation functions of regression parameter β_4 and variogram parameter σ^2.

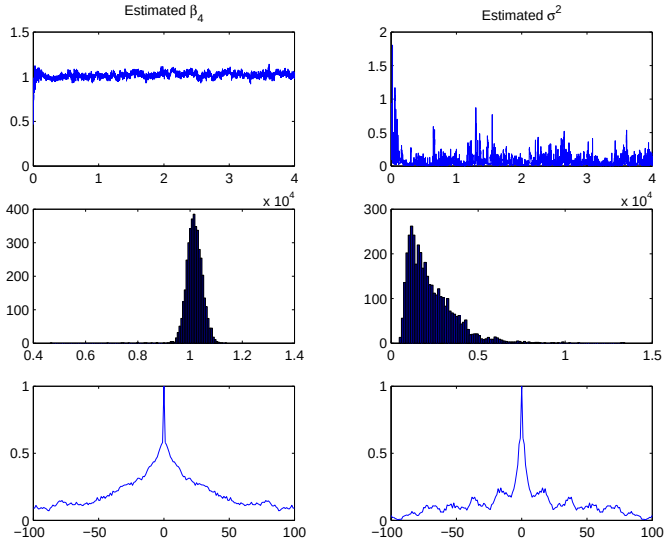


Figure 4.12: Cross-validation results using simulated data - Moisture flux direction random fields (dashed line - true MFD RF, solid line - estimated MFD RF), log-likelihood, and MCMC outputs for 2 observation sites.

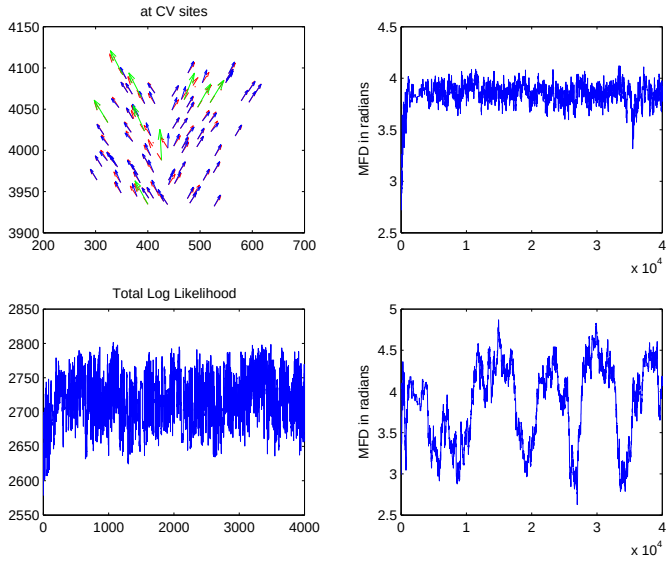


Figure 4.13: Cross-validation results using 36 different sets of 'discarded' sites for May data.

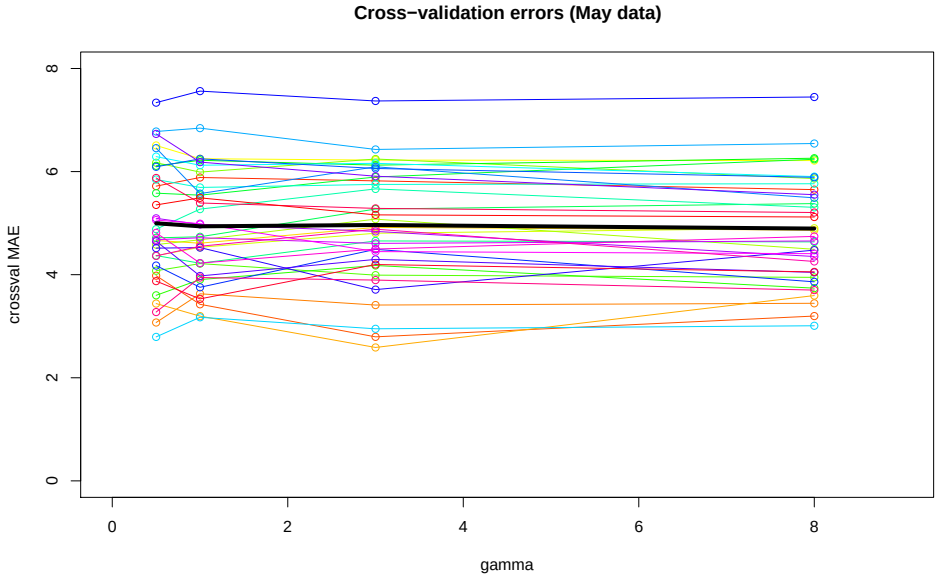


Figure 4.14: Cross-validation results showing actual versus predicted values for May for $\gamma = 0.5, 1, 3, 8$.

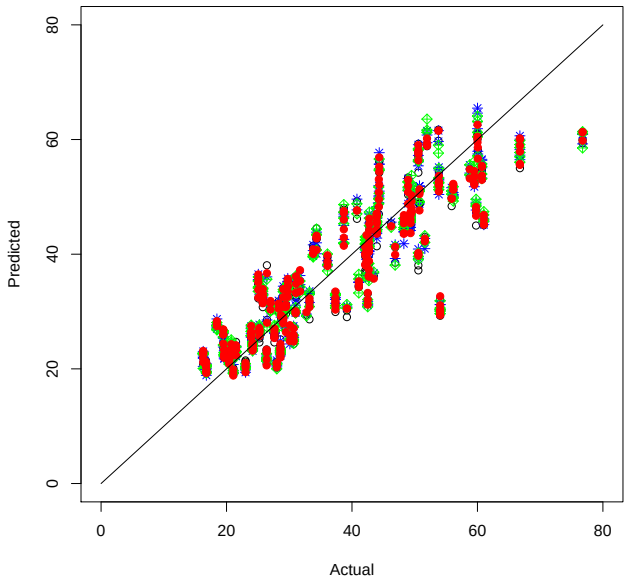


Figure 4.15: Cross-validation results using real data for May - Markov chain values, histograms, and Auto-correlation functions of regression parameter β_4 and variogram parameter σ^2.

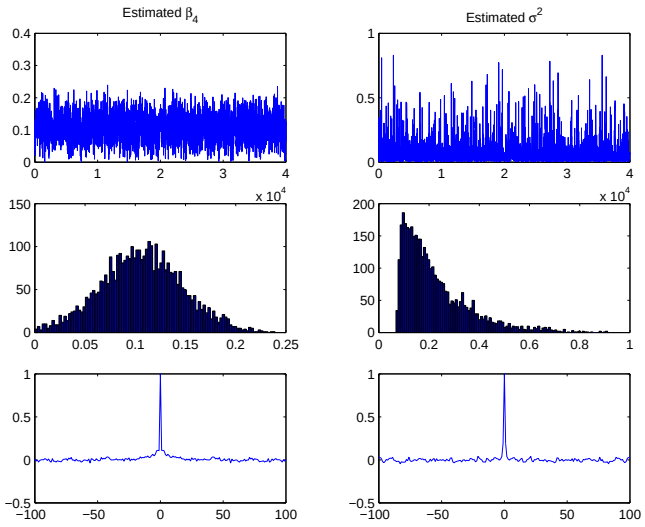


Figure 4.16: Cross-validation results using real data for May - Moisture flux direction random fields (dashed line - constant ASOADeK MFD RF, solid line - estimated MFD RF), log-likelihood, and MCMC outputs for 2 observation sites.

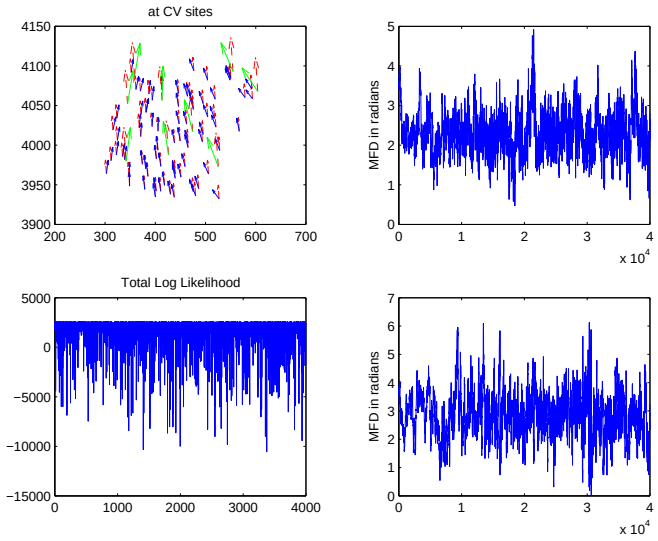


Figure 4.17: Precipitation map for February obtained using kriging and $\gamma = 0.5$

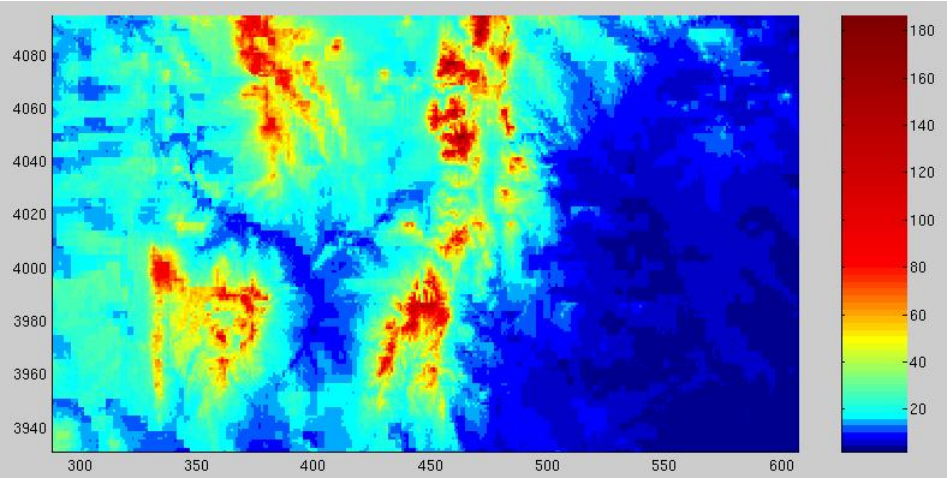
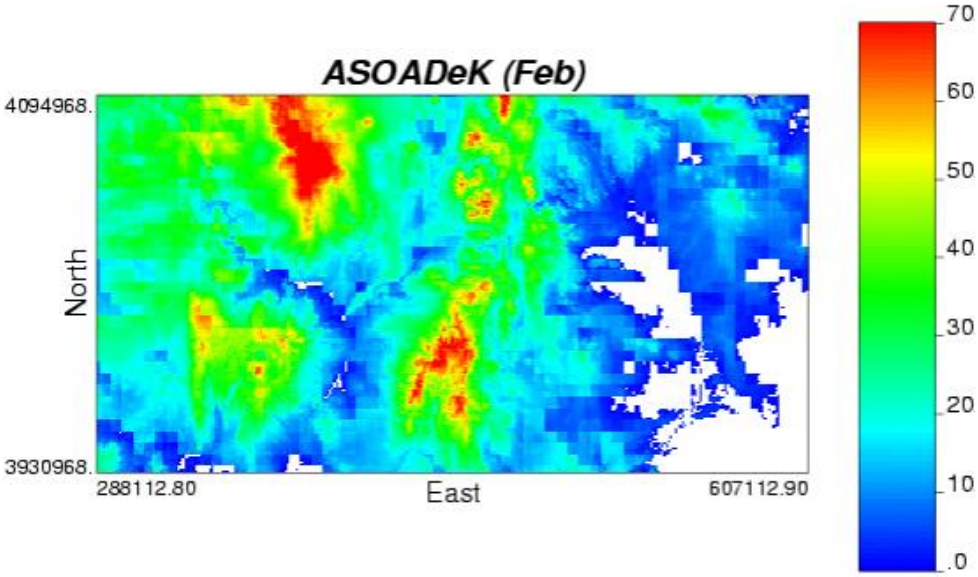


Figure 4.18: Precipitation map for February obtained using ASOADeK (Courtesy - Huade Guan).



CHAPTER 5

DISCUSSION

This chapter gives an analytical report of the results obtained using MCMC through 'Metropolis within Gibbs' approach, the core framework of the proposed model. We also present some future work that can possibly implement this MCMC through 'Metropolis within Gibbs' approach to develop more robust models.

5.1 Conclusions

What we have done:

We have implemented an extension of ASOADeK method that deals with estimation of MFD random field, an orographic factor influencing precipitation. We developed and tested a method for estimation of such MFD based on point precipitation data. Values of γ in the range 0.5 to 2 seem to give reasonable MFD random field estimates and kriging maps for the studied region. The 'optimal' γ depends on the region, the observation site locations, signal to noise ratio defined as β_4/σ, and on the nature of the 'true' MFD.

Why is that useful:

It is useful because it gives us insight into the prevailing weather patterns through MFDs. The plots for MFD random field for different months

follow consistent pattern for the Northern New Mexico mountainous region, with moisture coming from Pacific coast on the west and moisture coming from the Gulf of Mexico in south-west direction.

The model has potentially better predictive ability because of the Bayesian framework. It can also be thought of as being 'automatic' in the sense that no extensive calibration is needed except for may be γ. Previous methods like PRISM require extensive calibration. Also, this model can be applied to larger mountainous regions unlike ASOADeK. Hence, we can summarize the conclusions as:

1. Hidden random field (MFD RF) method offers an alternative to existing precipitation mapping products.

2. Bayesian nature of the method allows for 'all in one' estimation and great flexibility.

5.2 Future Work

We think that more effort should be spent on the validation of the method for larger regions (say, Switzerland, Colorado, etc.) as well as for 'single-event' data. We can also make the method more 'automatic', i.e., do something about choosing 'optimal' γ and making it less dependent on other factors. This seems to be very challenging but promising. We can choose an anisotropic variogram function and see how the model behaves. We can incorporate radar and satellite data and test the model. We can also develop a 'spatio-temporal' extension of this method provided we have enough data samples.

APPENDIX A

Some Probability Distributions

A.1 p-dimensional Multivariate Normal Distribution

If the random variables $X_1, X_2, \ldots, X_p$ have a *multivariate normal* distribution, then the joint probability density function is

$$f(\mathbf{x}) = (2\pi)^{-\frac{p}{2}} |\mathbf{C}|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\mathbf{C}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right] \tag{A.1}$$

where $\boldsymbol{\mu} = [\mu_1, \mu_2, \ldots, \mu_p]'$ is a vector of expected values along each of the coordinate directions of $X_1, X_2, \ldots, X_p$, and $\mathbf{C}$ is the *covariance matrix* which contains the covariances between the random variables

$$C_{i,j} = \mathrm{Cov}(X_i, X_j) \tag{A.2}$$

A.2 Scaled-Inverse χ^2 Distribution

If the random variable X has a *scaled-inverse* χ^2 distribution, then the probability density function is

$$f(x) = \frac{\frac{\nu}{2}^{\frac{\nu}{2}}}{\Gamma_{(\frac{\nu}{2})}} (S^2)^{\frac{\nu}{2}} x^{-(\frac{\nu}{2}+1)} \exp\left[-\frac{\nu S^2}{2x}\right] \tag{A.3}$$

where ν is the degrees of freedom, and νS^2 is the scaling parameter.

APPENDIX B

Matlab Codes

The various routines written in Matlab (Version - 6.5.0.180913a (R13)) are available in electronic format at

`euler.nmt.edu/~olegm/MFD`

.

REFERENCES

[1] Aster, R. C., Borchers, B., and Thurber, C. H. *Parameter Estimation and Inverse Problems*. 1st edition. Elsevier Academic Press, San Diego, 2005.

[2] Barros, A. P., Kim, G., Williams, E., and Nesbitt, S. W. *Probing orographic controls in the Himalayas during the monsoon using satellite imagery.* Natural Hazards and Eart System Sciences, Volume 4, Pages 1-23, 2004.

[3] Carlin, B. P., and Louis, T. A. *Bayes and Empirical Bayes Methods for Data Analysis.* 1st CRC Press reprint. Chapman & Hall/CRC, Boca Raton, 1998.

[4] Daly, C., Neilson, R. P., and Phillips, D. L. *A statistical-topographic model for mapping climatological precipitation over mountain terrain.* Journal of Applied Metereology, Volume 33, Pages 140-158, 1994.

[5] Dettinger, M., Redmond, K., and Cayan, D. *Winter Orographic Precipitation Ratios in the Sierra-Nevada - Large-Scale Atmospheric Circulations and Hydrologic Consequences.* Journal of Hydrometeorology, Volume 5, Pages 1102-1116, 2003.

[6] Deutsch, C. V., and Journel, A. G. *GSLIB Geostatistical Software Library and User's Guide*. 2nd edition. Oxford University Press, New York, 1998.

[7] Genton, M. G., and Furrer, R. *Analysis of Rainfall Data by Simple Good Sense: Is Spatial Statistics Worth the Trouble?* Journal of Geographic Information and Decision Analysis, Volume 2, Number 2, Pages 12-17, 1998.

[8] Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. *Markov Chain Monte Carlo in Practice.* 1st edition. Chapman & Hall, Great Britain, 1996.

[9] Guan, H., Wilson, J., and Makhnin, O. V. *Geostatistical Mapping of Mountain Precipitation Incorporating Auto-Searched Effects of Terrain and Climatic Characteristics*. Journal of Hydrometeorology, to appear, 2005.

[10] Hocking, R. R. *Methods and Applications of Linear Models.* John Wiley & Sons, Inc., New York, 1996.

[11] Hurn, M. A., Husby, O. K., and Rue, H. *A Tutorial on Image Analysis* in Spatial Statistics and Computational Methods. Edited by Moller, J. Pages 87-117, Springer, 2003.

[12] Hwang, Y., Clark, M., Rajagopalan, B., Gangopadhyay, S., and Hay, L. E. *Inter-comparison of spatial estimation schemes for precipitation and temperature.* Submitted to Water Resources Research, July 2004.

[13] Isaaks, E. H., and Srivastava, M. R. *Applied Geostatistics.* Oxford University Press, New York, 1989.

[14] Kim, S., Shephard, N., and Chib, S. *Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models.* Review of Economic Studies, 1998.

[15] Ribeiro, P. J. Jr., and Diggle, P. J. *Bayesian inference in Gaussian model-based geostatistics.* Technical Report ST-99-08, Lancaster University, 1999.

[16] Spatial Climate Analysis Service. *Parameter-elevation Regression on Independent Slopes Model (PRISM).* Spatial Climate Analysis Service, Oregon State University, 2003.

[17] DMAP: UTM Grid Zones of the World compiled by Alan Morton. `http://www.dmap.co.uk/utmworld.htm` accessed Nov 19, 2005.