

## A geostatistical method for Texas NexRad data calibration

Bo Li<sup>1\*,†</sup>, Marian Eriksson<sup>2</sup>, Raghavan Srinivasan<sup>2</sup> and Michael Sherman<sup>3</sup>

<sup>1</sup>National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000, U.S.A.

<sup>2</sup>Spatial Sciences Laboratory, Texas A&M University, College Station, TX 77843, U.S.A.

<sup>3</sup>Department of Statistics, Texas A&M University, College Station, TX 77843, U.S.A.

### SUMMARY

Rainfall is one of the most important hydrologic model inputs and is recognized as a random process in time and space. Rain gauges generally provide good quality data, however they are usually too sparse to capture the spatial variability. Radar estimates provide a better spatial representation of rainfall patterns, but they are subject to substantial biases. Our calibration of radar estimates using gauge data takes season, rainfall type, and rainfall amount into account, and is accomplished via a combination of threshold estimation, bias reduction, regression techniques, and geostatistical procedures. We explore the varying-coefficient model to adapt to the temporal variability of rainfall. The methods are illustrated using Texas rainfall data in 2003, which includes Weather Surveillance Radar-1988 Doppler (WSR-88D) radar-reflectivity data and the corresponding rain gauge measurements. Simulation experiments are carried out to evaluate the accuracy of our methodology. Copyright © 2007 John Wiley & Sons, Ltd.

KEY WORDS: linear regression; NexRad data; threshold; variogram estimation; varying-coefficient

### 1. INTRODUCTION

Precipitation is one of the most important hydrologic model inputs and is characterized by spatial and temporal variability. Traditionally, point precipitation measurements at rain gauges have been used with hydrological models. Since rain gauges physically measure the depth of rainfall at the points of measurement, they generally provide good quality data. However, rain gauge networks are usually too sparse to capture the spatial variability of precipitation over the hydrologic system. This problem becomes more critical when simulating large river basins.

Compared to rain gauges, weather radars such as the Next generation weather Radar (NexRad), formally known as the Weather Surveillance Radar-1988 Doppler (WSR-88D) of the United States provide precipitation data with much better spatial and temporal sampling frequencies. A better representation of rainfall variability can be accomplished in a hydrologic model by using radar rainfall data. However, there are several possible sources of errors and possibilities of 'data contamination' in the radar estimates of precipitation, as the radar estimates are not the real measurement of rainfall and

\*Correspondence to: B. Li, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000, U.S.A.

†E-mail: boli@ucar.edu

thus subject to errors (Jayakrishnan *et al.*, 2004). For this reason, we seek to improve the quality of radar precipitation data over the study area using rain gauge measurements.

Radar data are formed over three stages. Originally, a radar measures the reflectivities by volume scans over a fixed polar grid with a radial resolution of  $1^\circ$  in azimuth by 1 km in range. These reflectivities are then converted into rainfall rates by the convective Z–R relationship or the Rosenfeld tropical Z–R relationship contingent on the rainfall types, and further converted to rainfall depth on a grid, called the Stage I output. Stage II processing corrects Stage I output for the individual radar using a bias adjustment factor which is calculated as the ratio of the sum of all positive rain gauge data over a specific radar umbrella to the sum of all nonzero Stage I gridded output at the same gauge locations over the same spatio-temporal window of sampling ([http://www.srh.noaa.gov/wgrfc/resources/projects/stageiii\\_paper/default.html](http://www.srh.noaa.gov/wgrfc/resources/projects/stageiii_paper/default.html)). Finally, Stage III processing mosaics the data from multi-radar for the areas under the umbrella of more than one radar.

In this paper, we propose a methodology to calibrate daily NexRad data using rain gauge data based on Texas rainfall data sets in 2003. There are three sets of daily rainfall data: one is daily NexRad estimates over a  $4 \times 4$  km grid over Texas; another one is gauge data that is of relatively good quality and contains 664 rain gauges; the last one is station data which is of high quality, but contains only 60 weather stations. The NexRad data is Stage III WSR-88D precipitation data from National Weather Service (NWS), and the gauge data is compiled at the National Climate Data Center (NCDC) of NWS. Both NexRad data and station data give daily rainfalls from 7am to 7am the next day, but not all of the gauges are collected from 7am to 7am. Only 311 out of 664 are collected at 7am and thus are appropriate for use in the real data analysis, as they are the only ones that are comparable with the station and NexRad data. Although the 60 stations are the most reliable, they are too sparse to capture the spatial features of the rainfall data, but they will be helpful to assess the quality of our predictions. The 311 gauges are dense enough to give some spatial information, so we use the 311 gauges to calibrate the NexRad data. Among the 60 stations, 14 are collocated with gauges, while the measurements at the remaining 46 station locations are not being used in our calibration, hence it is appropriate to use these 46 stations to assess the quality of our prediction.

A lot of research has been done to explore the rainfall random field and make inferences from it. For example, Barancourt *et al.* (1992) proposed a geostatistical scheme for modeling rainfall using two mutually independent random fields,  $Z(\mathbf{s}) = I(\mathbf{s})F(\mathbf{s})$ , where  $\mathbf{s}$  denotes locations.  $I(\mathbf{s})$  is a binary random field which models the intermittency of the rainfall.  $I(\mathbf{s})$  is estimated by indicator kriging, which gives a map of rain probabilities. This map is further converted into a contour of the rain areas by assigning a threshold to the probabilities. Inside the rainy areas,  $F(\mathbf{s})$  accounts for the variability of the nonzero rainfall value, and it is also estimated by kriging. Their method is useful when only one data set is available and the interest focuses on delineation of rainy areas. If the mean areal rainfall over a fixed and large domain is desired, Barancourt *et al.* (1992) does no better than spatial kriging. De Oliveira (2004) introduced an analogous model for short time rainfall based on the separate modeling of spatial occurrence of rainfall and the spatial distribution of positive rainfalls. This model is essentially the same as Barancourt *et al.* (1992) but is formulated differently.

Gel *et al.* (2004) described a geostatistical output perturbation method to deal with two sets of temperature data. Their situation is similar to ours in that we both have less accurate values on a regularly spaced grid and more accurate values at irregularly spaced locations, and we share the same goal of finding a methodology to improve the performance of less accurate values. However, our situation differs from theirs in several ways, principally because the rainfall data is a mixture of zeroes and continuous nonzero values while the temperature data are absolutely continuous. Brown *et al.* (2001) presented

an approach for real-time prediction by calibrating the radar data against the rain gauge data. In their approach, they first proposed a linear relationship between the logarithmic gauge measurements,  $G_t$ , and radar estimates,  $R_t$ ,  $\log(G_t) = A_t + B_t \log(R_t) + \epsilon_t$ , where  $\epsilon_t$  is assumed to be uncorrelated Gaussian noise with mean 0 and variance  $\sigma_\epsilon^2$ , and then investigated spatial-temporal models for the biases  $A_t$  and  $B_t$ . Driven by focusing on the calibration rather than the prediction, we can bypass the computation involved in estimating the spatio-temporal correlation structure of  $A_t$  and  $B_t$ . Besides, their specific data set allows them to ignore the zero values in the data, while our data set does not display this feature. Widespread methods in hydrology such as mean field bias adjustment based on the difference between the radar rainfall estimates and the gauge accumulations (e.g., Anagnostou and Krajewski, 1998), and range-dependent gauge adjustment techniques classifying the distance between gauges and radar into range bins and deriving the adjustment factor as a function of distance from the radar (e.g., Amitai *et al.*, 2002), may not be able to exploit the spatial information efficiently.

We develop a geostatistical method combined with threshold estimation, bias reduction, and regression techniques to calibrate the NexRad estimates using the rain gauge measurements. To account for zeroes, we include a binary random variable in the model to delineate the dry and wet areas based on the NexRad estimates. We apply regression to reduce the bias of NexRad estimates and obtain the spatially correlated errors. We model the errors using a geostatistical method, then finally make predictions with the bias reduced NexRad estimates and the estimated errors. In order to adapt to the possible temporal variability of rainfall, we employ the varying-coefficient model proposed by Hastie and Tibshirani (1993). Our proposed method preserves the spatial occurrence pattern and spatial variability within the wet areas of the true rainfall.

In Section 2, we describe the geostatistical method, including the basic statistical model, parameter estimation, prediction using the estimated model, and prediction validation to evaluate our method and assess our predictions. In Section 3, we apply our method to the NexRad estimates and the corresponding gauge measurements in Texas in 2003, and give the prediction validation results. In Section 4, we describe simulation experiments and compare the simulation results with those of the data analysis. Finally, in Section 5, we discuss the features of the proposed method and possible future work.

## 2. GEOSTATISTICAL METHOD

We now describe our geostatistical method. First we outline the statistical model underlying the rainfall. Then we describe how to estimate model parameters from data and how to use the estimated model to make predictions. Finally, we propose a validation method to assess the accuracy of our model.

### 2.1. Statistical model

Let  $X(\mathbf{s}_i, t)$  denote the gauge measurement at location  $\mathbf{s}_i$ ,  $i = 1, \dots, 311$ , and time  $t$ ;  $W(\mathbf{s}'_j, t)$  denote NexRad estimate at location  $\mathbf{s}'_j$ ,  $j = 1, \dots, 50151$ , and time  $t$  and  $t = 1, \dots, 365$  for both notations. For  $\mathbf{s}'_j$  close to  $\mathbf{s}_i$ , we expect that the value of  $W(\mathbf{s}'_j, t)$  is close to the value of  $X(\mathbf{s}_i, t)$  if they are recorded on the same day. For  $W(\mathbf{s}'_j, t)$  with  $\min_j \|\mathbf{s}'_j - \mathbf{s}_i\|$ , define  $W(\mathbf{s}'_j, t) = W(\mathbf{s}_i, t)$ . There are four cases in the data:

- (i)  $X(\mathbf{s}_i, t) = 0$ ,  $W(\mathbf{s}_i, t) = 0$ ;
- (ii)  $X(\mathbf{s}_i, t) > 0$ ,  $W(\mathbf{s}_i, t) = 0$ ;
- (iii)  $X(\mathbf{s}_i, t) = 0$ ,  $W(\mathbf{s}_i, t) > 0$ ;
- (iv)  $X(\mathbf{s}_i, t) > 0$ ,  $W(\mathbf{s}_i, t) > 0$ .

Cases (i) and (iv) are what we hope to see in the data, while cases (ii) and (iii) are not. The goal is to calibrate  $W$  using  $X$ . If we use  $W(\mathbf{s}_i, t) = 0$  as the classifier for  $X(\mathbf{s}_i, t) = 0$ , we misclassify the  $X(\mathbf{s}_i, t)$ 's in cases (ii) and (iii). To reduce the misclassification error, we introduce a new classifier defined by a threshold  $\beta_2$  and we assume:  $X(\mathbf{s}_i, t) = 0$ , if  $W(\mathbf{s}_i, t) \leq \beta_2$ ;  $X(\mathbf{s}_i, t) > 0$ , if  $W(\mathbf{s}_i, t) > \beta_2$ . We expect the introduction of  $\beta_2$  to reduce the misclassification rate. For NexRad data that are above the threshold, we apply linear regression to remove bias from the NexRad data. In order to adapt to the possible temporal variability of rainfall, we employ a varying-coefficient model (Hastie and Tibshirani, 1993) to allow the coefficients to vary with time.

Based on the above discussion, we construct the following model:

$$X(\mathbf{s}, t)^{1/k} = \{\beta_0(t) + \beta_1(t)W(\mathbf{s}, t)\}^{1/k} + \xi(W(\mathbf{s}, t), \boldsymbol{\beta}_t)\epsilon_t I\{W(\mathbf{s}, t) > \beta_2(t)\} \quad (1)$$

where  $X(\mathbf{s}, t)$  and  $W(\mathbf{s}, t)$  are random processes of rain gauge and NexRad over locations  $\mathbf{s}$  at time  $t$ . The index  $t$  also serves as an effect modifying variable (Hastie and Tibshirani, 1993).  $I(\cdot)$  denotes an indicator function and  $\beta_2(t)$  is a time-dependent threshold. The functions  $\beta_0(t)$  and  $\beta_1(t)$  are time-dependent additive and multiplicative bias terms. Let  $\boldsymbol{\beta}_t$  denote  $(\beta_0(t), \beta_1(t))^T$  and  $\xi(W(\mathbf{s}, t), \boldsymbol{\beta}_t)$  is a function to account for nonconstant variance. The spatial errors  $\epsilon_t \sim F(\mathbf{0}, \boldsymbol{\Sigma}_t)$ , where  $F$  is a multivariate distribution function. We parameterize  $\boldsymbol{\Sigma}_t$  from the Matérn class of covariance functions as in Brown *et al.* (2001). Unlike other covariance functions, the Matérn allows to estimate the smoothness of a random field from the data. We follow the parameterization recommended by Handcock and Wallis (1994) for the Matérn covariance function:

$$C(r) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)} \left(\frac{2\nu^{\frac{1}{2}}r}{\phi}\right)^{\nu} \mathcal{K}_{\nu}\left(\frac{2\nu^{\frac{1}{2}}r}{\phi}\right) \quad (2)$$

where  $\Gamma(\cdot)$  is the gamma function;  $\mathcal{K}_{\nu}$  is the modified Bessel function of the second kind of order  $\nu$ ;  $r$  is Euclidian distance between two locations, and  $\sigma^2$  is the variance, also called the sill in geostatistics;  $\nu$  is the smoothness parameter. The larger the  $\nu$  is, the smoother the random process. The parameter  $\phi$  measures how quickly the correlations of the random field decay with distance. In this parameterization, it has the attractive feature that its interpretation is largely independent of  $\nu$ .

The power  $1/k$  is used to make data well behaved, since we seek to use likelihood methods to fit model parameters. Although the logarithmic transformation is one common choice to transform rainfall data to normality, and it demonstrates good performance for the accumulated rainfall amount, it is inappropriate for point estimates due to the zeroes in the data set. Adding a small constant to the zero values allows the logarithmic transform, but different constants lead to different parameter estimates. Our power transform avoids this choice.

Model (1) is the product of a Gaussian random field with linear mean function and a binary random field. It is related to the model of Barancourt *et al.* (1992) and De Oliveira (2004). They proposed a model as a product of two independent random fields  $Z(\mathbf{s}) = I(\mathbf{s})F(\mathbf{s})$ . Their model has the flexibility to reproduce the spatial association structure of rainfall fields, because the spatial association structure of the rainfall occurrence and rainfall amounts process are governed by different association structures. One drawback of their method being applied to our data is the lack of smooth transitions between the dry and wet sub-regions which are the norm for most rainfall patterns. Moreover, their procedures are mainly developed for modeling the rainfall random field using one data set, so it is often not efficient for calibrating one data set using another data set. However our threshold, which is estimated directly

from the NexRad data, lessens this problem by following the true transition patterns from dry to wet area whether the transition is smooth or not.

## 2.2. Parameter estimation

We estimate the model parameters given by Equations (1) and (2) using data in Texas from 2003. Our methodology, however, is sufficiently general to apply to other years and other states. The sequence of parameter estimation is  $k$ ,  $\beta_2(t)$ ,  $\beta_0(t)$ ,  $\beta_1(t)$ , followed by the parameters involved in  $\Sigma_t$  of  $\text{var}(\epsilon_t)$ . If weighted least squares are needed,  $\xi(W(\mathbf{s}, t), \boldsymbol{\beta})$  will be estimated simultaneously with  $\beta_0(t)$  and  $\beta_1(t)$ .

We choose an appropriate  $k$  based on the Shapiro–Wilk test of the transformed gauge data. We estimate the time-dependent coefficients by the plane fitting procedure described by Hastie and Tibshirani (1993), noting that our varying-coefficient model contains only a single effect modifying variable  $t$ . The main idea of the plane fitting procedure is to construct a small neighborhood for a specific  $t_0$  in the  $(t, W)$ -plane, fit the model of  $X$  on  $W$  using the data only within the neighborhood of  $t_0$  to obtain point estimates of  $\hat{\beta}_2(t_0)$ ,  $\hat{\beta}_0(t_0)$ , and  $\hat{\beta}_1(t_0)$ , and repeat this for different  $t_0$ .

To choose  $\beta_2(t)$ , we pair the gauge,  $X(\mathbf{s}_i, t)$ , and the average of the nearest four NexRad neighbors,  $\bar{W}(\mathbf{s}_i, t)$ , in the neighborhood of  $t$ , and employ logistic regression to estimate the threshold through the decision boundary between  $X(\mathbf{s}_i, t) = 0$  and  $X(\mathbf{s}_i, t) > 0$  in terms of  $\bar{W}(\mathbf{s}_i, t)$  when  $\bar{W}(\mathbf{s}_i, t) > 0$  (e.g., Hastie *et al.*, 2001). The reason for averaging is to stabilize the NexRad data. Averaging over the nearest eight neighbors gives similar results. Let  $P(\bar{W}(\mathbf{s}_i, t)) = \Pr(X(\mathbf{s}_i, t) > 0 \mid \bar{W}(\mathbf{s}_i, t) > 0)$ . The logistic regression model has the form

$$\text{logit}\{P(\bar{W}(\mathbf{s}_i, t))\} = \log \left\{ \frac{P(\bar{W}(\mathbf{s}_i, t))}{1 - P(\bar{W}(\mathbf{s}_i, t))} \right\} = \alpha_0(t) + \alpha_1(t)\bar{W}(\mathbf{s}_i, t)$$

The threshold is the value of  $W$  for which the log-odds are zero, and this is the point defined by  $\{W \mid \alpha_0(t) + \alpha_1(t)W = 0\}$ , which leads directly to  $\hat{\beta}_2(t) = -\hat{\alpha}_0(t)/\hat{\alpha}_1(t)$ . We explore several methods of estimating  $\beta_0(t)$  and  $\beta_1(t)$  and settle on linear regression on the transformed value of  $W(\mathbf{s}, t)$ .

Before we determine the form of  $\xi(W(\mathbf{s}, t), \boldsymbol{\beta}_t)$ , we first assess the heteroscedasticity of the residuals. Carroll and Ruppert (1988) suggested several methods of assessing heteroscedasticity based on a fairly large number of observations and varying degrees of data density. One is to compute the Spearman rank correlation coefficients of the absolute studentized residuals with the fitted values. Another way to cope with the density effect in large data sets while at the same time obtaining information about a model for the variability is to use nonparametric regression techniques. We carry out both in our analysis. If the analysis results detect heteroscedasticity, we employ weighted least squares to estimate  $\beta_0(t)$ ,  $\beta_1(t)$ , and  $\xi(W(\mathbf{s}, t), \boldsymbol{\beta}_t)$  iteratively.

Finally, we calculate residuals from the linear regression and find the underlying spatial pattern of the residuals. To estimate the parameters in  $\Sigma_t$ , or equivalently, the parameters in the variogram of  $\epsilon_t$ , we use daily residuals since the rainfall patterns differ appreciably from day to day. However calculating residuals at only the closest NexRad location whose rainfall value is above the threshold for each gauge, gives at most 311 residuals at a very limited number of spatial lags. This gives information which is too sparse to capture the spatial pattern. Instead we assume that the nearest eight NexRad locations and the corresponding gauge have the same true rainfall, as discussed by Jayakrishnan *et al.* (2004), and calculate the residuals at the nearest eight NexRad locations for each gauge. This allows for a sufficient number of residuals at a variety of spatial lags enabling accurate variogram estimation.

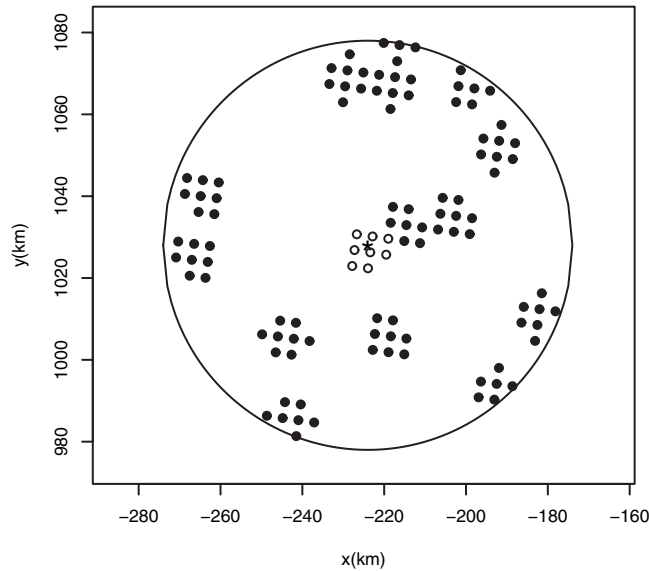


Figure 1. Illustration of the moving window

We calculate

$$\hat{\epsilon}_{ij}(t) = \frac{1}{\hat{\xi}(W_{ij}(t), \hat{\beta}_t)} \{X(s_i, t)^{1/\hat{k}} - \hat{\beta}_0(t) - \hat{\beta}_1(t)W_{ij}(t)^{1/\hat{k}}\}, \quad i = 1, \dots, 311, \quad j = 1, \dots, 8 \quad (3)$$

where  $W_{ij}(t)$  is the  $j$ -th nearest NexRad whose rainfall is above the threshold for the  $i$ -th gauge and  $\hat{\beta}_t = \{\hat{\beta}_0(t), \hat{\beta}_1(t)\}^T$ .

It is known, e.g., Stein (1999), that using empirical semivariograms for model selection can work disastrously for smooth processes. For this reason, we use likelihood-based methods. The full maximum likelihood method, however, tends to be prohibitively time consuming in this setting. We seek a simpler and faster procedure that approximates the full maximum likelihood method and works well for our implementation. Due to the well-behaved transformed data and the fact that the spatial dependence of residuals decays quickly as distance between residuals increases, it is natural to use an approximate likelihood over moving windows (Stein 1986, 1999; Vecchia, 1992). We center the moving windows on each gauge that is used to calculate the residuals in Equation (3), obtaining overlapped moving windows across the whole domain. Figure 1 shows an example of a moving window. The circle window is centered on the ‘\*’ which denotes a gauge location. The eight small circles around the ‘\*’ are the nearest eight NexRad locations for that gauge. The black dots belong to the nearest eight NexRad locations for some other gauges. We write the likelihood function using the residuals on all the dots and small circles within each window, and then pool up all the likelihood functions of all windows to obtain the joint likelihood function.

Specifically, the parameters in the variogram model are  $\theta = (\tau^2, \sigma^2, \phi, \nu)^T$  which represent the nugget and Matérn sill, range and smoothness, respectively. Let  $K$  denote the total number of gauges that are involved in Equation (3). Then,  $K$  is also the total number of moving windows. Let  $\mathbf{e}_k$  denote the residual vector in the  $k$ -th window,  $k = 1, 2, \dots, K$ ,  $\Sigma_k(\theta)$  denote the covariance matrix of  $\mathbf{e}_k$ , and let  $\det(\cdot)$  denote the determinant of a matrix.

Our joint log likelihood, up to a constant that does not depend on  $\theta$ , is as follows:

$$l(\theta) = - \sum_{k=1}^K [\log\{\det(\Sigma_k(\theta))\} + \mathbf{e}_k^T \Sigma_k(\theta)^{-1} \mathbf{e}_k]$$

We estimate the variogram parameters by maximizing  $l(\theta)$ .

### 2.3. Prediction

Given model (1) and the estimated parameters, we now address the prediction at unobserved locations on any given day. For example, let  $\mathbf{s}_0$  denote an unobserved location. To predict at  $\mathbf{s}_0$  on the  $t$ -th day, we calculate the residual vector  $\hat{\boldsymbol{\epsilon}}_t$  using the pairs of gauge,  $X(\mathbf{s}_i, t)$ , and the average of its nearest eight NexRads,  $\bar{W}(\mathbf{s}_i, t)$ ,  $\hat{\boldsymbol{\epsilon}}_{ti} = \frac{1}{\xi(\bar{W}(\mathbf{s}_i, t), \hat{\boldsymbol{\beta}}_t)} \{X(\mathbf{s}_i, t)^{1/\hat{k}} - \hat{\beta}_0(t) - \hat{\beta}_1(t) \bar{W}(\mathbf{s}_i, t)^{1/\hat{k}}\}$ ,  $i = 1, \dots, 311$ . We then kriging  $\hat{\boldsymbol{\epsilon}}_t$  to location  $\mathbf{s}_0$  via ordinary kriging (e.g., Cressie, 1993) with the estimated variogram model based on our approximate likelihood method:

$$\hat{\boldsymbol{\epsilon}}(\mathbf{s}_0, t) = \boldsymbol{\lambda}(\mathbf{s}_0, t)^T \hat{\boldsymbol{\epsilon}}_t \quad (4)$$

where  $\boldsymbol{\lambda}(\mathbf{s}_0, t)$  is the vector of kriging weights.

Finally, we plug  $\hat{\boldsymbol{\epsilon}}(\mathbf{s}_0, t)$  into the following formula:

$$\hat{X}(\mathbf{s}_0, t) = ([\hat{\beta}_0(t) + \hat{\beta}_1(t) \bar{W}(\mathbf{s}_0, t)^{1/\hat{k}} + \hat{\xi}\{\bar{W}(\mathbf{s}_0, t), \hat{\boldsymbol{\beta}}_t\} \hat{\boldsymbol{\epsilon}}(\mathbf{s}_0, t)]^{\hat{k}} - \text{bias}) I\{\bar{W}(\mathbf{s}_0, t) > \hat{\beta}_2(t)\} \quad (5)$$

where ‘bias’ is the estimated bias induced by the power transformation.

### 2.4. Validation

We predict at each of the 46 stations, and compare the prediction with the station rainfall measurements by computing several statistics that are widely used by hydrologists, for example, Jayakrishnan *et al.* (2004). Let  $Z_{il}$ ,  $i = 1, \dots, 46$ ,  $l = 1, \dots, L$ , denote the  $i$ -th station measurement on the  $l$ -th day, where  $L$  is the number of days being used in the comparison. Let  $\hat{Z}_{il}$  denote the predicted value of  $Z_{il}$ ,  $\hat{Z}_{il} = \hat{X}(s_{il})$  in Equation (5), where  $s_{il}$  is the location of  $Z_{il}$ . Define  $Z_l = \sum_{i=1}^{46} Z_{il}$ , and  $\hat{Z}_l = \sum_{i=1}^{46} \hat{Z}_{il}$ .  $Z_l$  is the daily total of the station measurements and  $\hat{Z}_l$  is the predicted value of  $Z_l$ . Let  $\bar{Z}$  denote the mean of the  $Z_l$ 's. To assess accuracy, we calculate:

1. Total Difference in precipitation,  $D = \sum_{l=1}^L (\hat{Z}_l - Z_l)$ .
2. Estimation Bias (%),  $EB = 100D / \sum_{l=1}^L Z_l$ .
3. Estimation Efficiency,  $EE = 1.0 - \frac{\sum_{l=1}^L (\hat{Z}_l - Z_l)^2}{\sum_{l=1}^L (Z_l - \bar{Z})^2}$ .
4. Sum of Square Prediction Error,  $SSPE = \sum_{l=1}^L \sum_{i=1}^{46} (\hat{Z}_{il} - Z_{il})^2$ .

It is common to use NexRad as rainfall estimates in hydrology. For example, Bedient *et al.* (2000, 2003) utilized NexRad data in hydrological model for flood prediction. For this reason, we also compute the above four statistics using the nearest NexRad value as the estimate of the rainfall. Let  $\tilde{W}_{il}$  denote the nearest NexRad for  $Z_{il}$  on the  $l$ -th day. We compute the corresponding  $\tilde{D}$ ,  $\tilde{EB}$ ,  $\tilde{EE}$ , and  $\tilde{SSPE}$  for

the nearest NexRad predictor by replacing  $\widehat{Z}$  by  $W$  in 1–4. We compare our prediction with the nearest NexRad to see what (if any) improvement our method accomplishes.

### 3. RESULTS

#### 3.1. Data

We estimate the model and assess our predictions using the rainfall data in Texas in 2003. The spatial locations of the 311 gauges and 60 weather stations are shown in Figure 2. The 50 151 daily NexRad estimates over a  $4 \text{ km} \times 4 \text{ km}$  grid are shown in Figure 3. The units of measurement are millimeters (mm) for rainfall and kilometers (km) for distance throughout.

#### 3.2. Parameter estimation

We compare several power transformations with different powers. Using the Shapiro–Wilk test statistics of gauge data for different powers, we find that  $1/3$  is the best power among  $1/k$ ,  $k = 1, 2, \dots$ , for both seasons. This choice is also strongly supported by Stidd (1953), who found the cube root of a large variety of precipitation types, no matter of daily or annual time resolution, is normally distributed. The cube root transformation also allows for a closed form bias expression (see the Appendix).

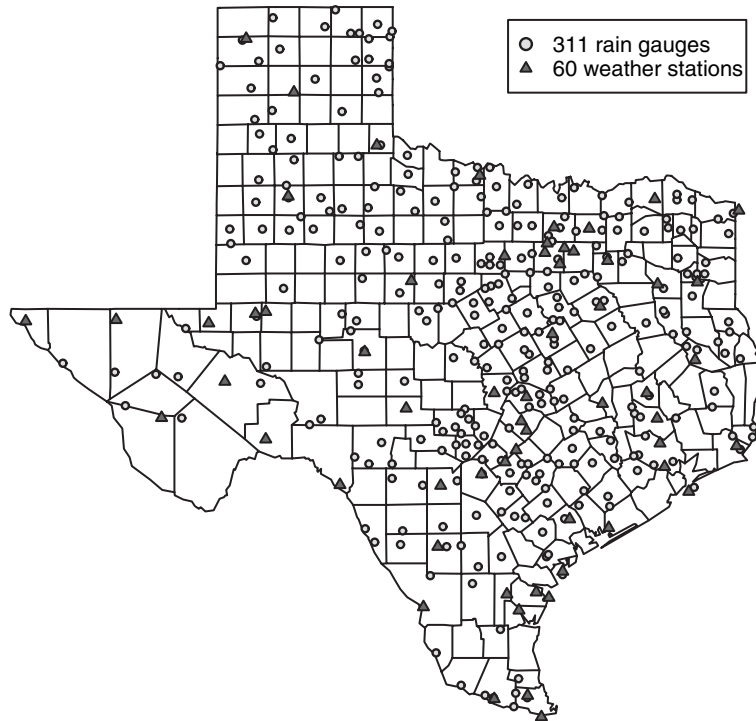


Figure 2. Distribution of rain gauges and weather stations

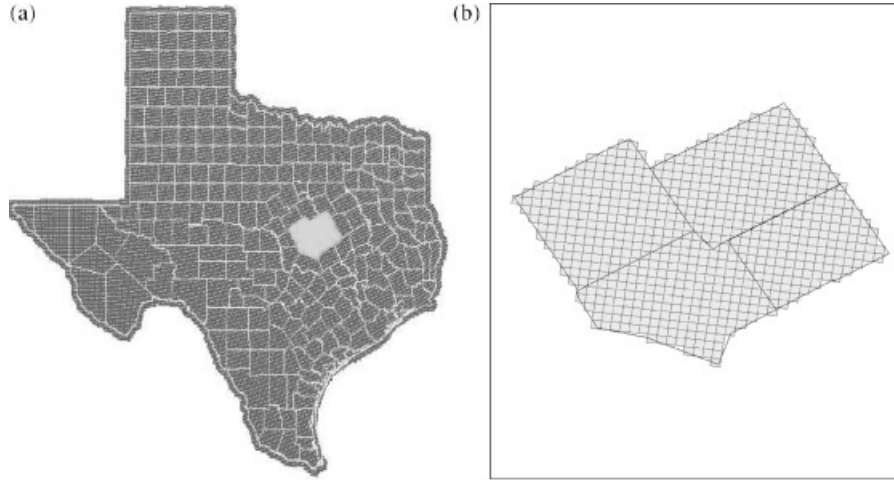


Figure 3. The NexRad grid. (a) The NexRad grid over Texas and (b) details of NexRad showing four counties highlighted in (a)

To estimate time-dependent coefficients using the plane fitting method discussed in Subsection 2.2, we choose  $t_0$  as 7pm for each day and the neighborhood of  $t_0$  as 24 h spans from 7am to 7am the next day, that is, we fit the model using daily data to obtain point coefficients estimates at discrete  $t = 1, 2, \dots, 365$ . For each day, equivalently, for each discrete value of  $t$ , we select pairs  $(X(s_i, t), \bar{W}(s_i, t))$  with  $\bar{W}(s_i, t) > 0$ ,  $i = 1, \dots, 311$ , and fit a logistic regression model to the binary outcome of  $X(s_i, t)$  in terms of rain or no rain. We calculate  $\hat{\beta}_2(t)$  for each  $t$  such that there are at least 20 pairs of  $(X(s_i, t), \bar{W}(s_i, t) > 0)$  on the  $t$ -th day. Finally we obtain 219 point estimates of  $\beta_2(t)$  together with their fitted cubic smoothing spline using two inner knots as depicted in Figure 4. With the observation that  $\hat{\beta}_2(t)$  is noisy, this picture shows only the estimates that are between  $-10$  and  $20$

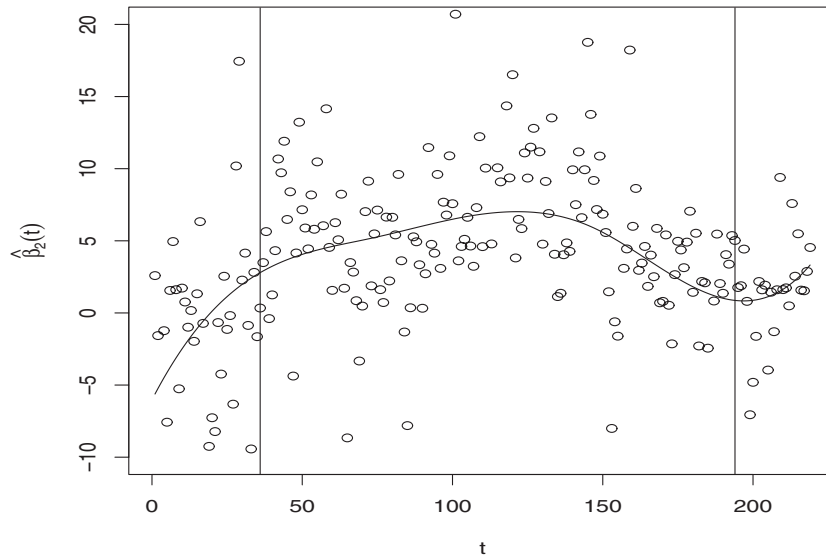


Figure 4. Daily threshold estimates

to make the main body of  $\hat{\beta}_2(t)$  clear. Using the unstable  $\hat{\beta}_2(t)$  as the corresponding threshold, we fit linear regression for each of 219  $t$ 's. We then find  $\hat{\beta}_0(t)$  and  $\hat{\beta}_1(t)$  are even noisier, which suggests to combine daily data for threshold estimation.

It is well known that the rainfall in the warm season (April–October) and the cold season (November–March) is of qualitatively different types (Anagnostou *et al.*, 1998). The warm season is dominated by convective rainfall, while the cold season is dominated by widespread stratiform rainfall. Convective rainfall is localized rainfall, like a thunderstorm. It is formed when cold currents from the north meet warm currents from the south. On the contrary, stratiform rainfall is more spread out and not focused on a few small areas. The two vertical lines shown in Figure 4 divide the whole plot into the cold season and the warm season. The cold season lies in the left area to the left line and the right area to the right line. The warm season lies in the area between the two vertical lines. It can be seen from Figure 4 that the estimates of  $\beta_2(t)$  are approximately constantly high in the warm season and approximately constantly low in the cold season. This suggests estimating  $\beta_2$  by combining the seasonal data, which corroborates the suggestion by the noise of the estimates. We get  $\hat{\beta}_2 = 0.579$  in cold season and  $\hat{\beta}_2 = 5.675$  in warm season. These two values roughly match the value of estimates displayed in Figure 4.

Let  $|A|$  denote the cardinality of the set  $A$  and  $M$  denote the total misclassification error rate:

$$M = \frac{|\{W(\mathbf{s}_i, t) \leq \hat{\beta}_2, X(\mathbf{s}_i, t) > 0\}| + |\{W(\mathbf{s}_i, t) > \hat{\beta}_2, X(\mathbf{s}_i, t) = 0\}|}{|\{W(\mathbf{s}_i, t), X(\mathbf{s}_i, t) > 0\}| + |\{W(\mathbf{s}_i, t), X(\mathbf{s}_i, t) = 0\}|}$$

where  $i = 1, \dots, 311$  and  $t = 1, \dots, 365$ . The estimated thresholds reduce the misclassification error rates from 11% to 9.7% in the cold season, and from 15.9% to 11.3% in the warm season compared to using  $\beta_2 = 0$  as the threshold.

Given the specified seasonal  $\hat{\beta}_2$ , we employ the plane fitting method again to estimate  $\beta_0(t)$  and  $\beta_1(t)$ . In order to obtain a successful linear regression, we select 194 days and fit a linear model on the pairs  $(X(\mathbf{s}_i, t)^{1/3}, \bar{W}(\mathbf{s}_i, t)^{1/3} | \bar{W}(\mathbf{s}_i, t) \geq \hat{\beta}_2)$  on each of those days. The resulting  $\hat{\beta}_0(t)$  and  $\hat{\beta}_1(t)$  together with their fitted cubic smoothing splines using two inner knots (e.g., Ruppert *et al.*, 2003) are illustrated in Figure 5. The two vertical lines in each plot of Figure 5 play the same role as in Figure 4.

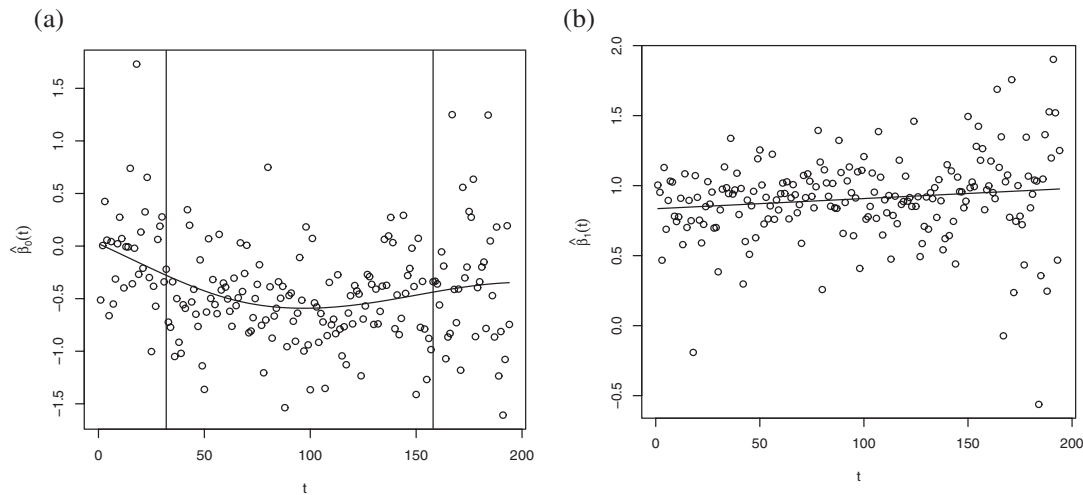


Figure 5.  $\hat{\beta}_0(t)$  and  $\hat{\beta}_1(t)$  given the specified seasonal  $\hat{\beta}_2$ . (a)  $\hat{\beta}_0(t)$  given  $\hat{\beta}_2$  and (b)  $\hat{\beta}_1(t)$  given  $\hat{\beta}_2$

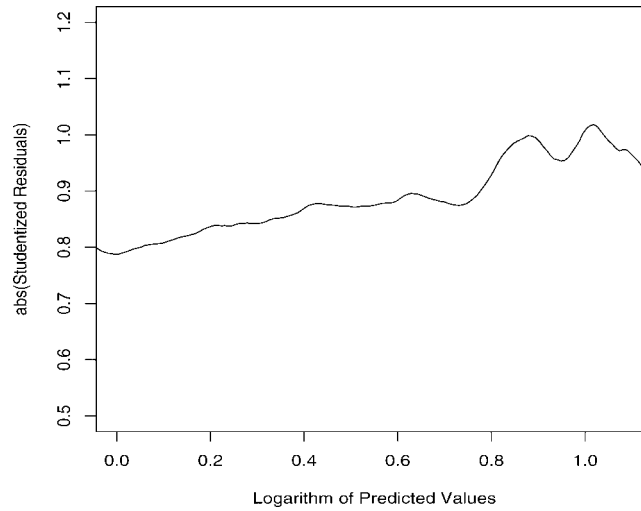


Figure 6. A plot of residuals versus predicted values

$\hat{\beta}_0(t)$  is then also approximately constant over each season, while  $\hat{\beta}_1(t)$  is approximately constant over the year. Nevertheless we estimate both  $\beta_0$  and  $\beta_1$  separately in the two seasons, as suggested by Anagnostou *et al.* (1998).

We pool up the pairs  $(X(s_i, t)^{1/3}, \bar{W}(s_i, t)^{1/3} | \bar{W}(s_i, t) \geq \hat{\beta}_2)$  within each season and fit them by a linear model, having assumed there is no rainfall for pairs with  $\bar{W}(s_i, t) < \hat{\beta}_2$ . The data behave well after the cubic root transform, which allows the use of Least Squares (LS) regression. Using the seasonal LS regression, we find  $R^2 = 0.369$  in the cold season and  $R^2 = 0.241$  in the warm season, where  $R^2$  is the coefficient of determination which measures the goodness-of-fit of linear regression. Considering a possible quadratic relationship between NexRad and gauge, we fit a linear regression with an extra quadratic term in the model. We then obtain  $R^2 = 0.369$  in the cold season and  $R^2 = 0.242$  in the warm season. Thus there is little if any benefit to introduce a quadratic term. The following are the LS parameters estimates(standard deviation):

$$\begin{aligned} \text{Cold season: } & \hat{\beta}_0 = -0.314(0.0320), \hat{\beta}_1 = 0.938(0.0166); \\ \text{Warm season: } & \hat{\beta}_0 = -0.717(0.0578), \hat{\beta}_1 = 1.029(0.0221). \end{aligned}$$

In the warm season  $\hat{\beta}_1$  is not significantly different from 1. This means there is little multiplicative bias in the warm season. These LS parameter estimates also approximately correspond to the values shown in Figure 5(a) and (b).

To assess possible heteroscedasticity, we use kernel regression with a global bandwidth to smooth the absolute studentized residuals (Carroll and Ruppert, 1988). The bandwidth is selected by trial and error. The kernel regression shows the main body of standard deviation increases mildly from 0.8 to 1.0 as shown in Figure 6, so we do not weight residuals, that is, we set  $\xi(W(s_i, t), \hat{\beta}_t) = 1$ . This is confirmed by the Spearman's rank correlation test which gives  $\hat{\rho} = 0.020$ , and a two sided  $p$ -value = 0.137.

As an example, Figure 7 shows the empirical semivariograms of residuals on 20 February together with the fitted variogram model using an approximate likelihood method as illustrated in Subsection 2.2. Although the fitted model does not match the empirical variogram very well, we still decide to use

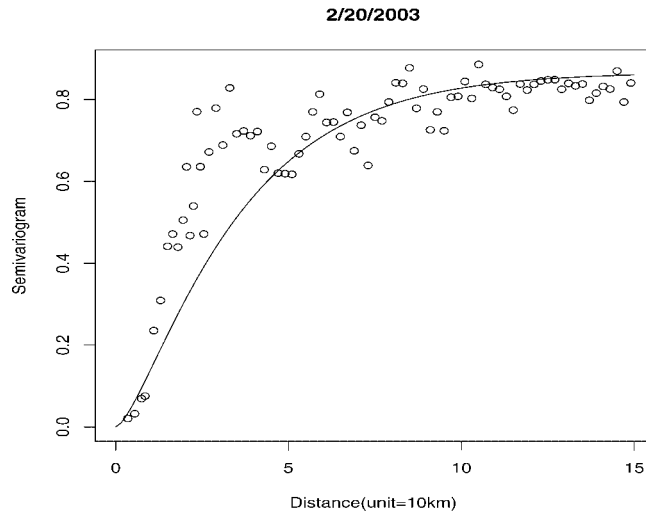


Figure 7. An example of the semivariogram from the regression residuals

this likelihood based estimate. Stein (1999) highly recommends using maximum likelihood in plug-in predictions even though it shows a worse match to the empirical variogram than moment-based estimators. Viewing several semivariogram plots indicates it is not necessary to include a nugget term in the covariance model. In estimating variogram model parameters by the approximate likelihood method over moving windows, the radius of the moving window is set to 50 km by examining the difference between the maximum likelihood estimated variogram models of different window radii and the required number of observations within each window suggested by Stein (1999, p. 172). In 2003, we selected 146 days that have the largest wet coverage to analyze the daily data. Two of the days do not have station data, and one day has too few NexRad locations whose values are above the threshold, so we remove those three days, and get the mean (standard deviation) of the sill, range, and smoothness over the remaining 143 days: 1.023 (0.285), 34.134 (13.905), 0.858 (0.192). The median of those parameters are 1.000, 31.071, 0.843.

### 3.3. Prediction and validation over rainy days

If we assume  $\widehat{\beta}_0(t)$  and  $\widehat{\beta}_1(t)$  in Equation (5) are the true parameters, when  $\xi\{W(\mathbf{s}_0, t), \widehat{\beta}_t\} = 1$  and  $\widehat{k} = 3$  (as obtained in Subsection 3.2), the bias in Equation (5) can be estimated unbiasedly by:

$$\widehat{\text{bias}} = 3\{\widehat{\beta}_0(t) + \widehat{\beta}_1(t)W(\mathbf{s}_0, t)^{1/3}\}\{\sigma^2(\mathbf{s}_0, t) - 2\boldsymbol{\lambda}(\mathbf{s}_0, t)^T \boldsymbol{\gamma}_0(\mathbf{s}_0, t)\} \quad (6)$$

where  $\boldsymbol{\lambda}$  is the vector of kriging weights as presented in Equation (4),  $\boldsymbol{\gamma}_0(\mathbf{s}_0, t)$  is the vector of variogram values between  $\mathbf{s}_0$  and observed locations, and  $\sigma^2(\mathbf{s}_0, t)$  is the kriging variance. The proof of Equation (6) is given in the Appendix.

The assumption that  $\widehat{\beta}_0(t)$  and  $\widehat{\beta}_1(t)$  are the true parameters is reasonable since we pooled the seasonal data to estimate the intercept and slope. The large amount of seasonal data justifies our assumption that  $\widehat{\beta}_0(t)$  and  $\widehat{\beta}_1(t)$  can be treated as fixed.

Table 1. Validation results over 46 weather stations in data analysis

$D$ (mm)	EB	EE	SSPE (mm <sup>2</sup> )
1344.852	7.044	0.853	332543.4
$\tilde{D}$ (mm)	$\tilde{EB}$	$\tilde{EE}$	$\tilde{SSPE}$ (mm <sup>2</sup> )
7557.220	39.584	0.601	374383.9

We predict rainfall at the 46 stations for the 143 days as discussed in Subsection 3.2, and calculate the statistics as defined in Subsection 2.4. Note that  $L = 143$  here. The results are summarized in Table 1. Comparing  $D$  to  $\tilde{D}$ , EB to  $\tilde{EB}$ , and EE to  $\tilde{EE}$ , we see our prediction has a significantly less total difference in precipitation, lower estimation bias, and higher estimation efficiency. This result indicates that our prediction does much better than the NexRad in estimating the total rainfall.

To evaluate our method in doing point estimation, we compare SSPE with  $\tilde{SSPE}$ , which shows that our prediction improves by approximately 11.2% over the nearest NexRad based prediction. 11 September is the highest rainfall day in 2003. The total rainfall amount of the 60 stations on that day is 1917.7 mm, which is about twice the next largest rainfall amount of 1092.5 mm on 20 February. Our prediction method does not work well on this ‘outlier’ day. If we remove 11 September, the results are as following:  $SSPE' = 306418.6 \text{ mm}^2$  and  $\tilde{SSPE}' = 367913 \text{ mm}^2$ . Our prediction improves by 16.7% compared to using the closest NexRad as the predictor.

For further evaluation of our prediction in terms of point estimation, we also calculate the mean (variance) of  $\hat{Z}_{il} - Z_{il}$ , which is 0.204 mm (50.390 mm<sup>2</sup>); and the mean (variance) of  $W_{il} - Z_{il}$ , which is 1.146 mm (55.463 mm<sup>2</sup>).  $\hat{Z}_{il}$ ,  $Z_{il}$ , and  $W_{il}$  are defined as in Subsection 2.4. It shows that our prediction is less biased and slightly sharper than the closest NexRad to predict the true rainfall.

Figure 8 shows the comparison of predictions on 23 April and 6 July as examples to illustrate how the predictions behave. The side by side box plot of differences between prediction and the station data indicates that the predictions using our method are better than using the nearest NexRad location in the sense of both unbiasedness and low variability.

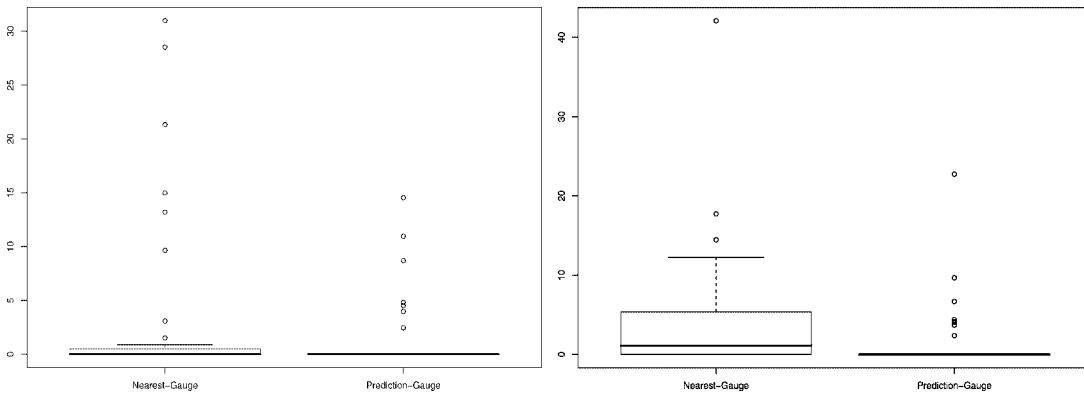


Figure 8. A Comparison between predictors on 23 April 2003 and 6 July 2003

## 4. A SIMULATION

## 4.1. Setup

To verify our procedures in the analysis, we carry out simulations to check their validity. Our goal is to make the simulation as comparable as possible to the actual data to judge the reliability of our improvement over using the closest NexRad. We simulate random fields on a  $224 \times 224$  grid with a grid interval equal to 1. Each simulated random field contains 50 176 grid points, which correspond to the 50 151 NexRad grid points. Since the actual NexRad grid interval is 4, we shrink all the distance parameters in the actual data by the approximate ratio of 1/4 to get the corresponding ones in the simulation. The set up of the simulation is as follows.

*4.1.1. Generating rainfall.* We generate NexRad data,  $W^{1/3}$ , as a random field on the  $224 \times 224$  grid with an exponential covariance structure:  $(\mu = 0.5, \sigma^2 = 2.0, \phi = 13)$ , where  $\mu$  and  $\sigma^2$  are the mean and variance of the random field, and  $\phi$  is the range parameter of the exponential covariance. The value of  $\sigma^2$  reflects the real data, and  $\phi = 13$  corresponds to  $\phi = 52$  km in the real data. The value of  $\mu$  is selected to make the ratio of positive values in the random field close to the actual rainfall data.

We set  $\beta_0$  and  $\beta_1$  to be exactly the same as estimates from the data in the data analysis.

We generate spatial errors,  $\epsilon$ , as a Gaussian random field on the  $224 \times 224$  grid with mean zero and Matérn covariance ( $\tau^2 = 0, \sigma^2 = 1, \phi = 7.7, \nu = 0.84$ ). These values correspond to the median of estimates from the 2003 NexRad data.

We generate the corresponding gauge data,  $X$ , as:

$$X^{1/3} = [\beta_0 + \beta_1 W^{1/3} + \epsilon]^+$$

where  $[z]^+ = \max(z, 0)$  denotes the clipped random field with zero as the threshold. We generate a clipped random field as in De Oliveira (2000). Specifically we generate the NexRad and gauge data over the same grid. The NexRad is the observed data, while the gauge data is considered the target data. Only a small part of the gauge data is assumed to be observed, and our goal is to predict the unobserved gauges using the observed gauges and the NexRad Data.

*4.1.2. Parameter estimation.* We randomly pick 311 grid points from the  $224 \times 224$  grid, and assume we observe the gauge data at these locations. Thus we have 311 pairs of observed  $(X, W)$  which correspond to the 311 gauges in Section 3.

Completely analogous to our procedures in Section 3 we pool the seasonal data to estimate  $\beta_2, \beta_0$ , and  $\beta_1$ . Here we generate 40 NexRad random fields and their corresponding gauge random fields, pool up all the  $40 \times 311 = 12\,440$  pairs of observed  $(X, W)$  to estimate  $\beta_2$ , and then  $\beta_0$  and  $\beta_1$ . Only 7285 of 12 440 pairs have  $W > 0$  and are actually used in the estimation, which is comparable to the 7114 pairs being used in the cold season.

*4.1.3. Prediction.* Again we randomly select 311 grid points as the observed gauge locations to parallel the 311 gauges in the data. Fix these locations in this step.

We find the nearest nine NexRad  $W_{ij}, j = 1, \dots, 9$  for each selected gauge  $X_i, i = 1, \dots, 311$ , which corresponds to the nearest eight NexRad locations in the true data. The reason that we choose the nearest nine instead of eight NexRad locations is due to the geometric property of the grid over

Table 2. Validation results over 200 simulations

EB	EE	mean (std. dev.) of SSPE
-4.178	0.949	216373.2 (8968.972)
$\widetilde{EB}$	$\widetilde{EE}$	mean(std. dev.) of $\widetilde{SSPE}$
-2.211	0.933	280582.6 (11477.850)

which we generate the NexRad and gauge data set. In the real data, we assume that the gauge and its nearest eight NexRad have the same true rainfall. Likewise in the simulation we assume the gauge and its nearest nine NexRad have the same true rainfall. We then calculate residuals at each of the nearest nine NexRad locations.

We employ the approximate likelihood method as in Subsection 3.2 to estimate the parameters in the Matérn covariance function. In the data analysis, the moving window is a circle of radius 50 km centered at the gauge. This window contains about 500 NexRad locations, so in the simulation we set the moving window as a  $23 \times 23$  grid centered on the gauge which contains 529 NexRad grid points.

Finally, we randomly pick 4000 grid points and predict the rainfall,  $\widehat{Z}_i$ ,  $i = 1, \dots, 4000$ , at those locations, and compare the prediction to the ‘true’ rainfall  $Z_i$ .

We repeat 200 times the steps in Subsections 4.1.1 to 4.1.3. We compute EB, EE, and  $\widetilde{EB}$ ,  $\widetilde{EE}$  which are defined as in Subsection 2.4. The only change we make in the definition in Subsection 2.4 is replacing 46 by 4000 and setting  $L = 200$ . For each simulation run, we define  $SSPE = \sum_{i=1}^{4000} (\widehat{Z}_i - Z_i)^2$  for our prediction and  $\widetilde{SSPE} = \sum_{i=1}^{4000} (W_i - Z_i)^2$  for the nearest NexRad which is collocated with its corresponding gauge in the simulation. We did not calculate  $D$  and  $\widetilde{D}$  in the simulation for two reasons. The value of  $D$  here is not comparable with the corresponding value in Subsection 3.3 due to a different number of locations at which the simulation makes predictions than the actual data analysis does, and the values of EB and  $\widetilde{EB}$  are sufficient to represent the relative relationship between  $D$  and  $\widetilde{D}$ .

#### 4.2. Simulation results and comparison with data analysis

The results of the simulation are summarized in Table 2. The simulated NexRad and gauge random field already agree in their total rainfall, so there is no room for improvement. However, comparison of SSPE with  $\widetilde{SSPE}$  shows a 22.9% improvement of our prediction over the NexRad. The simulation indicates greater improvement than seen in the actual data analysis. This is not surprising as we have generated the simulated rainfall according to the model. Nevertheless the improvement here is roughly comparable to that seen in the data.

We show one simulated NexRad random field with its corresponding gauge and prediction random field in Figure 9. The blank regions in this figure represent dry regions. Except the blank regions, the color scheme is shown in the plot. It is seen that the prediction follows the NexRad in discriminating wet areas from dry areas, but if we take a closer look at the light-colored areas within the wet regions, we see that the pattern of these areas in the prediction random field is closer to that in the gauge random field than is the NexRad random field. This indicates that we capture the essential part of rainfall by successfully correcting the pattern of NexRad data at the areas with large rainfall.

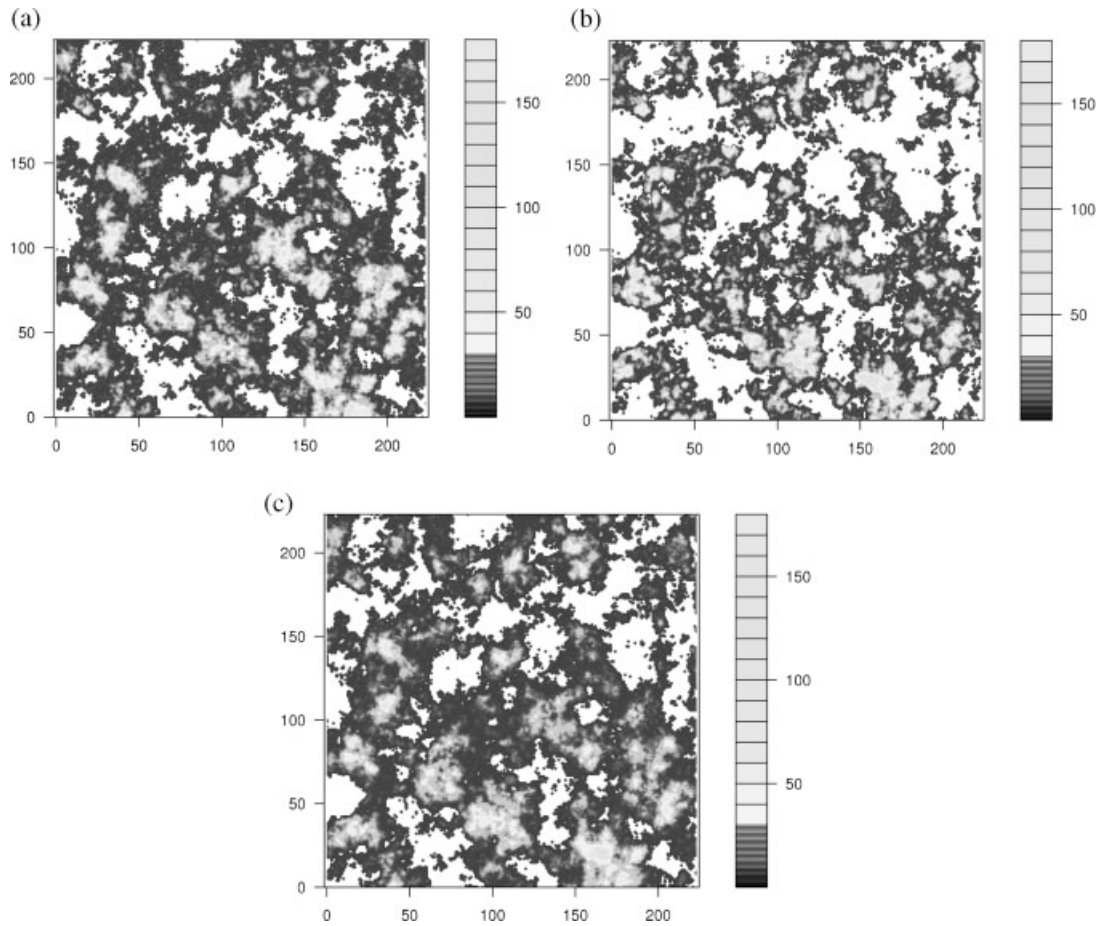


Figure 9. Simulated rainfall random fields and their prediction. (a) Simulated NexRad random field over  $224 \times 224$  grid, (b) simulated gauge random field, and (c) prediction of gauge random field

## 5. DISCUSSION

Hydrology data usually turns out to be of large size. The cumbersome body of data and consequently the massive work in manipulation increase the difficulty for data analysis. This paper proposed a simple, yet efficient method to calibrate the NexRad data using the rain gauge data. The proposed method combines spatial rainfall occurrence estimation, bias reduction by regression techniques, and geostatistical procedures. An approximate likelihood method is employed to estimate the parameters in the random process. Although the proposed method is developed based on 2003 Texas rainfall data, it provides a framework to analyze short time accumulated rainfall so that only slight modifications are needed to accommodate different data sets. For example, if the variance of the errors depends on the mean value, that is, the estimated  $\xi(W, \boldsymbol{\beta})$  in Equation (1) is some function of  $\beta_0 + \beta_1 W^{1/k}$ , we can employ weighted least squares to estimate  $\beta_0$ ,  $\beta_1$ , and  $\xi(W, \boldsymbol{\beta})$  iteratively.

The superiority of the proposed method lies in estimating total rainfall as well as point rainfall amount. The total rainfall is crucial in flood prediction, designing sewage systems, and managing other water resource decision support systems. Compared to the NexRad, our prediction of total rainfall amount has closer agreement with the true value. Further, the estimation bias, estimation efficiency, and root mean square difference are all highly improved over the NexRad itself. This method did not just provide a simple but successful solution to calibrate the NexRad rainfall data using the gauge data, it can also be used in other environmental problems with a similar structure.

Unlike the procedures in Barancourt *et al.* (1992), our method is applied to daily data and the geostatistical method is applied on the regression residuals. Thus there is no need to be concerned about nonergodicity and nonstationarity in the rainfall data.

We estimate the varying coefficients  $\beta_2(t)$ ,  $\beta_0(t)$ , and  $\beta_1(t)$  by fitting a simple linear model at the neighborhood of each  $t_0$ . However, the noise and the specific pattern of the varying coefficients motivates us to finally estimate the parameters from the seasonal data. Nevertheless, we are aware that other data sets may not exhibit such a specific pattern to allow us considering the parameters as constant over a long time period. In this situation, we need an enhancement to the plane fitting method to refine the varying coefficients estimates. The approach of Cleveland *et al.* (1991) assigns observations weights from a tricube weight function based on their Euclidian distance from  $t_0$  in addition to the restriction to neighborhood. Specifically, the closer to  $t_0$ , the bigger the weight. In that approach, varying coefficients are estimated by fitting a linear model using weighted least squares. By choosing the appropriate neighborhood size, this produces adequate smooth and stable varying coefficient estimates.

Alternatively, spatio-temporal hierarchical Bayesian modeling (Wikle *et al.*, 2001) can be an effective method to combine data from different sources. Their strategy simplifies the complex problem by formulating three primary statistical models or stages. Stage 1 models only measurement errors, stage 2 formulates the true process, and stage 3 specifies the priors.

The model in this paper does not account for the situation when the gauge is positive but the NexRad is below the threshold. This is not a big concern for the data set being used here, because our estimated thresholds for both seasons are low and consequently only approximately 4% of the data fall into this situation. In addition, most of the positive gauge measurements among these 4% are of small values which usually are not of interest in practice. However, in the future, more research is needed to address this issue.

#### ACKNOWLEDGEMENTS

The authors thank Dr. Raymond J. Carroll and Dr. Richard W. Katz for helpful discussions. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

#### REFERENCES

- Amitai E, Wolff DB, Marks DA, Silberstein DS. 2002. Radar rainfall estimation: lessons learned from the NASA/TRMM validation program. *Proceedings of European Conference on Radar Meteorology*, 255–260.
- Anagnostou EN, Krajewski WF. 1998. Calibration of the WSR-88D precipitation processing subsystem. *Weather and Forecasting* **13**: 396–406.
- Anagnostou EN, Krajewski WF, Seo D, Johnson E. 1998. Mean-field rainfall bias studies for WSR-88D. *Journal of Hydrological Engineering* **3**(3): 149–159.
- Barancourt C, Creutin JD, Rivoirard J. 1992. A method for delineating and estimating rainfall fields. *Water Resource Research* **28**: 1122–1144.
- Bedient PB, Hoblit BC, Gladwell DC, Vieux BE. 2000. NEXRAD radar for flood prediction in Houston. *Journal of Hydrological Engineering* **5**(3): 269–277.

- Bedient PB, Holder A, Benavides JA, Vieux BE. 2003. Radar-based flood warning system applied to tropical storm Allison. *Journal of Hydrological Engineering* **8**(6): 308–318.
- Brown PE, Diggle PJ, Lord ME, Young PC. 2001. Space-time calibration of radar rainfall data. *Journal of the Royal Statistical Society, Series C* **50**: 221–241.
- Carroll RJ, Ruppert D. 1988. *Transformation and Weighting in Regression*. Chapman and Hall: New York.
- Cleveland WS, Gross E, Shyu WM. 1991. Local regression models. In *Statistical Models in S*, Chambers JM, Hastie T (eds). Wadsworth and Brooks/Cole: Pacific Grove.
- Cressie N. 1993. *Statistics for Spatial Data*. Wiley: New York.
- De Oliveira V. 2000. Bayesian prediction of clipped Gaussian random fields. *Computational Statistics & Data Analysis* **34**: 299–314.
- De Oliveira V. 2004. A simple model for spatial rainfall fields. *Stochastic Environmental Research* **18**: 131–140.
- Gel Y, Raftery E, Gneiting T. 2004. Calibrated probabilistic mesoscale weather field forecasting: the geostatistical output perturbation method (with discussion). *Journal of the American Statistical Association* **99**: 575–587.
- Handcock MS, Wallis JR. 1994. An approach to statistical spatial-temporal modeling of meteorological fields (with discussion). *Journal of the American Statistical Association* **89**: 368–390.
- Hastie T, Tibshirani R. 1993. Varying-coefficient models (with discussion). *Journal of the Royal Statistical Society, Series B* **55**: 757–796.
- Hastie T, Tibshirani R, Friedman J. 2001. *The Elements of Statistical Learning*. Springer: New York.
- Jayakrishnan R, Srinivasan R, Arnold J. 2004. Comparison of raingage and WSR-88D Stage III precipitation data over the Texas-Gulf basin. *Journal of Hydrology* **292**: 135–152.
- Ruppert D, Wand MP, Carroll RJ. 2003. *Semiparametric Regression*. Cambridge University Press: Cambridge.
- Stein M. 1986. A modification of minimum norm quadratic estimation of a generalized covariance function for use with large data sets. *Mathematical Geology* **18**: 625–633.
- Stein M. 1999. *Interpolation of Spatial Data*. Springer: New York.
- Stidd CK. 1953. Cube root normal precipitation distributions. *American Geophysical Union Transactions* **34**: 31–35.
- Vecchia AV. 1992. A new method of prediction for spatial regression models with correlated errors. *Journal of the Royal Statistical Society, Series B* **54**: 813–830.
- Wikle CK, Milliff RF, Nychka D, Berliner LM. 2001. Spatiotemporal hierarchical Bayesian modeling: Tropical ocean surface winds. *Journal of the American Statistical Association* **96**: 382–397.

#### APPENDIX: DERIVATION OF TRANSFORMATION BIAS IN SUBSECTION 3.3

In order to make the derivation concise, we omit the index  $t$  in the Appendix. Suppose  $X(\mathbf{s}_0)$  is observed, when  $W(\mathbf{s}_0) > \hat{\beta}_2$ , plugging  $\xi(W, \boldsymbol{\beta}) = 1$  and  $k = 3$  into model (1), we get

$$X(\mathbf{s}_0)^{1/3} = \beta_0 + W(\mathbf{s}_0)^{1/3} \beta_1 + \epsilon(\mathbf{s}_0)$$

where  $\epsilon(\mathbf{s}_0) \sim \text{Normal}(0, \sigma^2)$ . Thus  $E\{\epsilon(\mathbf{s}_0)\} = 0$ ,  $E\{\epsilon(\mathbf{s}_0)^3\} = 0$ , and  $E\{\epsilon(\mathbf{s}_0)^2\} = \text{var}\{\epsilon(\mathbf{s}_0)\} = \sigma^2$ .

However  $X(\mathbf{s}_0)$  is not observed, so we estimate it as in Equation (5). Under the assumption that  $\beta_0$  and  $\beta_1$  are known parameters,

$$\hat{X}(\mathbf{s}_0)^{1/3} = \beta_0 + W(\mathbf{s}_0)^{1/3} \beta_1 + \hat{\epsilon}(\mathbf{s}_0)$$

where  $\hat{\epsilon}(\mathbf{s}_0) = \sum_{i=1}^n \lambda_i \epsilon(\mathbf{s}_i)$  and  $\lambda_i$  is the kriging weight consisting of the vector  $\boldsymbol{\lambda}(\mathbf{s}_0)$  in Equation (4), so  $E\{\hat{\epsilon}(\mathbf{s}_0)\} = \sum_{i=1}^n \lambda_i E\{\epsilon(\mathbf{s}_i)\} = 0$ .

Let  $f\{W(\mathbf{s}_0), \boldsymbol{\beta}\} = \beta_0 + W(\mathbf{s}_0)^{1/3} \beta_1$ .  $f\{W(\mathbf{s}_0), \boldsymbol{\beta}\}$  can be generalized to any linear or nonlinear function of  $W(\mathbf{s}_0)$  and known coefficients.

The derivation here is analogous to the one shown by Cressie (1993, p. 135) in finding the bias for a lognormal random field. When we transform the predictor back by the cubic power transformation, the

mean of the true value should be

$$\begin{aligned} E \left\{ \left( X(\mathbf{s}_0)^{1/3} \right)^3 \right\} &= E([f\{W(\mathbf{s}_0), \boldsymbol{\beta}\} + \epsilon(\mathbf{s}_0)]^3) \\ &= E([f\{W(\mathbf{s}_0), \boldsymbol{\beta}\}]^3 + 3[f\{W(\mathbf{s}_0), \boldsymbol{\beta}\}]^2\epsilon(\mathbf{s}_0) + 3f\{W(\mathbf{s}_0), \boldsymbol{\beta}\}\{\epsilon(\mathbf{s}_0)\}^2 + \{\epsilon(\mathbf{s}_0)\}^3) \\ &= E([f\{W(\mathbf{s}_0), \boldsymbol{\beta}\}]^3 + 3f\{W(\mathbf{s}_0), \boldsymbol{\beta}\}\{\epsilon(\mathbf{s}_0)\}^2) \\ &= [f\{W(\mathbf{s}_0), \boldsymbol{\beta}\}]^3 + 3f\{W(\mathbf{s}_0), \boldsymbol{\beta}\}\text{var}\{\epsilon(\mathbf{s}_0)\} \end{aligned}$$

while the mean of our predictor is

$$\begin{aligned} E \left\{ \left( \widehat{X}(\mathbf{s}_0)^{1/3} \right)^3 \right\} &= E([f\{W(\mathbf{s}_0), \boldsymbol{\beta}\} + \widehat{\epsilon}(\mathbf{s}_0)]^3) \\ &= E([f\{W(\mathbf{s}_0), \boldsymbol{\beta}\}]^3 + 3[f\{W(\mathbf{s}_0), \boldsymbol{\beta}\}]^2\widehat{\epsilon}(\mathbf{s}_0) + 3f\{W(\mathbf{s}_0), \boldsymbol{\beta}\}\{\widehat{\epsilon}(\mathbf{s}_0)\}^2 + \{\widehat{\epsilon}(\mathbf{s}_0)\}^3) \\ &= E([f\{W(\mathbf{s}_0), \boldsymbol{\beta}\}]^3 + 3f\{W(\mathbf{s}_0), \boldsymbol{\beta}\}\{\widehat{\epsilon}(\mathbf{s}_0)\}^2) \\ &= [f\{W(\mathbf{s}_0), \boldsymbol{\beta}\}]^3 + 3f\{W(\mathbf{s}_0), \boldsymbol{\beta}\}\text{var}\{\widehat{\epsilon}(\mathbf{s}_0)\} \end{aligned}$$

Thus the bias induced by the cubic transformation is

$$E \left\{ \left( \widehat{X}(\mathbf{s}_0)^{1/3} \right)^3 \right\} - E \left\{ \left( X(\mathbf{s}_0)^{1/3} \right)^3 \right\} = 3f\{W(\mathbf{s}_0), \boldsymbol{\beta}\}[\text{var}\{\widehat{\epsilon}(\mathbf{s}_0)\} - \text{var}\{\epsilon(\mathbf{s}_0)\}]$$

where

$$\begin{aligned} \text{var}\{\widehat{\epsilon}(\mathbf{s}_0)\} - \text{var}\{\epsilon(\mathbf{s}_0)\} &= \text{var}\{\widehat{\epsilon}(\mathbf{s}_0) - \epsilon(\mathbf{s}_0)\} - 2[\text{var}\{\epsilon(\mathbf{s}_0)\} - \text{cov}\{\widehat{\epsilon}(\mathbf{s}_0), \epsilon(\mathbf{s}_0)\}] \\ &= \sigma^2(\mathbf{s}_0) - 2 \left[ \text{var}\{\epsilon(\mathbf{s}_0)\} - \text{cov} \left\{ \sum_{i=1}^n \lambda_i \epsilon(\mathbf{s}_i), \epsilon(\mathbf{s}_0) \right\} \right] \\ &= \sigma^2(\mathbf{s}_0) - 2 \left[ \text{var}\{\epsilon(\mathbf{s}_0)\} - \sum_{i=1}^n \lambda_i \text{cov}\{\epsilon(\mathbf{s}_i), \epsilon(\mathbf{s}_0)\} \right] \\ &= \sigma^2(\mathbf{s}_0) - 2 \sum_{i=1}^n \lambda_i [\text{var}\{\epsilon(\mathbf{s}_0)\} - \text{cov}\{\epsilon(\mathbf{s}_i), \epsilon(\mathbf{s}_0)\}] \\ &= \sigma^2(\mathbf{s}_0) - 2 \sum_{i=1}^n \lambda_i \gamma(\mathbf{s}_i, \mathbf{s}_0) \\ &= \sigma^2(\mathbf{s}_0) - 2\boldsymbol{\lambda}^T \boldsymbol{\gamma}_0 \end{aligned}$$