Spring 1-1-2013

Statistical Methods for Blending Satellite and Ground Observations to Improve High-Resolution Precipitation Estimates

Andrew Paul Verdin
University of Colorado at Boulder, andrew.verdin@gmail.com

Statistical Methods for Blending Satellite and Ground Observations to Improve High-Resolution Precipitation Estimates

by

Andrew P. Verdin

B.A. Statistics, University of Minnesota, Twin Cities, 2008

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Master of Science

Department of Civil, Environmental and Architectural Engineering

2013

This thesis entitled:
Statistical Methods for Blending Satellite and Ground Observations to Improve High-Resolution Precipitation Estimates
written by Andrew P. Verdin
has been approved for the Department of Civil, Environmental and Architectural Engineering

Balaji Rajagopalan

Chris Funk

Edith Zagona

Date

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Verdin, Andrew P. (M.S., Civil Engineering)

Statistical Methods for Blending Satellite and Ground Observations to Improve High-Resolution Precipitation Estimates

Thesis directed by Prof. Balaji Rajagopalan and Prof. Chris Funk

Drought and flood management practices require accurate estimates of precipitation in space and time. However, data is sparse in regions with complicated terrain, often in valleys, and of poor quality. Consequently, extreme precipitation events are poorly represented. Satellite-derived rainfall data is an attractive alternative in such regions and is being widely used, though it too fails in representing extreme events due to its dependency on retrieval algorithms and the indirect relationship between satellite infrared observations and precipitation intensities. Thus, it seems appropriate to blend satellite-derived rainfall data of extensive spatial coverage with rain gauge data in order to provide a more robust estimate of precipitation.

To this end, in this research we offer four techniques to blend rain gauge data and the Climate Hazards group InfraRed Precipitation (CHIRP) satellite-derived precipitation estimate for Central America and Colombia. In the first two methods, the gauge data is assigned to the closest CHIRP grid point, where the error is defined as $r(s) = Y_{obs}(s) - Y_{sat}(s)$. The spatial structure of $r(s)$ is then modeled using physiographic information (easting, northing, and elevation) by two methods (i) a traditional Co-kriging approach which utilizes a variogram that is calculated in Euclidean space and (ii) a nonparametric method based on local polynomial functional estimation. The models are used to estimate r at all grid points, which is then added to the CHIRP, thus creating an improved satellite estimate. We demonstrate these methods by applying them to pentadal and monthly total precipitation fields during 2009. The models predictive abilities and their ability to capture extremes are investigated. These blending methods significantly improve upon the satellite-derived estimates and are also competitive in their ability to capture extreme precipitation.

The above methods assume satellite-derived precipitation to be unbiased estimates of gauge

precipitation, which is far from being the case. Thus the third method, Bayesian Hierarchical Modeling (BHM), is offered. In this approach, first, the gauge observations are modeled as a function of satellite-derived estimates and other variables such as elevation (the satellite estimate coefficient is in effect a bias correction factor). The residual from this first hierarchical model is then subjected to a spatial kriging model. The posterior distributions of all the model parameters are estimated simultaneously in Markov Chain Monte Carlo framework – consequently, the posterior distributions and uncertainties of the blended precipitation estimates are attained. This approach provides a robust treatment of the uncertainties and the hierarchy enables incorporating all relevant covariates.

While the BHM provides a robust confidence interval of the bias correction factor for CHIRP, it is reasonable to assume this bias is not uniform over the domain. Therefore a fourth method is proposed, wherein a GLM is fit to the time series at each point ($Y_{obs}(s,t) = \beta(s)Y_{sat}(s,t) + \epsilon(s,t)$), and the satellite coefficients are interpolated using a Co-kriging model similar to the first two methods. This provides a unique bias correction factor for every time frame (pentad, month), and therefore may be applied in near-real-time. To obtain the error field (i.e. residuals $\epsilon(s,t)$) for a specific time frame t, the residuals corresponding to the appropriate time frame are extracted from the GLMs and interpolated, again using a physiographic Co-kriging model.

These blended products provide more accurate and representative initial conditions for hydrologic modeling applications that are crucial for modeling and mitigating impacts from climate hazards such as floods and landslides which are of major concern in this region.

Dedication

For the Verdin family, the Ott family, and (most of all) for my lady Carsen.

Acknowledgements

A special thank you goes out to all who have helped me through the years:

NOAA for the grant that funded this research; Chris Funk, Balaji Rajagopalan, and Edith Zagona for the guidance and support; William Kleiber for the Bayesian expertise; Pete Peterson, Diego Pedreros, Greg Husak, and the rest of the Climate Hazards Group for sharing data and tolerating my trivial questions.

This work utilized the Janus supercomputer, which is supported by the National Science Foundation (award number CNS-0821794) and the University of Colorado Boulder. The Janus supercomputer is a joint effort of the University of Colorado Boulder, the University of Colorado Denver and the National Center for Atmospheric Research.

Contents

Chapter

1 Introduction 1

2 Kriging and Local Polynomial Methods for Blending 11

2.1 Introduction 11

2.2 Study Region & Data 12

2.3 Proposed Blending Methods 15

2.3.1 Kriging and Co-kriging 16

2.3.2 K-Nearest Neighbor Local Polynomial 18

2.4 Results 21

2.4.1 Monthly Totals 21

2.4.2 Performance on Extreme Events 32

2.4.3 Application to Pentad Rainfall Estimation 34

2.5 Summary and Conclusions 41

3 A Bayesian Hierarchical Kriging Approach for Blending 44

3.1 Introduction 44

3.2 Study Region & Data 45

3.3 Proposed Blending Method 48

3.3.1 Bayesian Hierarchical Model – Bayesian Kriging 48

3.4 Results 51

3.4.1 Posterior Distributions of Model Parameters 51
3.4.2 Model Fitting 53
3.4.3 Model Validation 54
3.4.4 Model Predictions from Blending 57
3.4.5 Model Performance on Extreme Events 61
3.4.6 Application to Pentad Rainfall Estimation 63
3.5 Summary and Conclusions 70

4 A Hierarchical Modeling Approach for Time Series Blending 72
4.1 Introduction 72
4.2 Study Region & Data 73
4.3 Proposed Blending Methods 76
4.4 Results 77
4.4.1 Monthly Totals 78
4.4.2 Performance on Extreme Events 89
4.4.3 Application to Pentad Rainfall Estimation 90
4.5 Summary and Conclusions 98

5 Summary & Conclusions 100

Bibliography 104

Tables

Table

2.1 α, p, and best subset of covariates for LP model. 22

2.2 Summary statistics reflecting model-fitting performance. 24

2.3 Summary statistics reflecting drop-one cross-validated model performance. 26

2.4 Summary statistics of blending performance on extreme events for July 2009. 33

2.5 Same as Table 2.2 but for the fourth pentad of August 2009. 35

2.6 Same as Table 2.3 but for the fourth pentad of August 2009. 36

2.7 Same as Table 2.4 but for the fourth pentad of August 2009. 40

3.1 Summary statistics of blending performance on extreme events for July 2009. 62

3.2 Same as Table 3.1 but for the fourth pentad of August 2009. 69

4.1 Summary statistics reflecting model-fitting performance. 79

4.2 Summary statistics reflecting drop-one cross-validated model performance. 82

4.3 Summary statistics of blending performance on extreme events for July 2009. 90

4.4 Same as Table 4.1 but for the fourth pentad of August 2009. 92

4.5 Same as Table 4.2 but for the fourth pentad of August 2009. 92

4.6 Same as Table 4.3 but for the fourth pentad of August 2009. 97

Figures

Figure

1.1 Study region geography, rain gauge measurement locations, and country names. . . . 4

1.2 Inter-tropical Convergence Zone within the study region. 5

2.1 Study region geography, rain gauge measurement locations, and country names. . . . 13

2.2 Empirical (black) and theoretical (red) variograms for January, April, July, and October 2009. 22

2.3 Original CHIRP estimates (left) and blended estimates from Co-kriging (middle) and LP (right) models for January 2009. 23

2.4 Same as Figure 2.3 but for July 2009 24

2.5 Original CHIRP estimates (left) and cross-validated estimates from Co-kriging (middle) and LP (right) models for January 2009. 25

2.6 Same as Figure 2.5 but for July 2009. 25

2.7 Prediction skill measures from drop-25% cross-validation for July 2009. Red dots correspond to estimates from original CHIRP. 27

2.8 Same as 2.7 but from drop-50% cross-validation. 27

2.9 a) Original CHIRP estimates for January 2009, b-c) Estimated satellite error field from Co-kriging and LP models, d-e) Blended precipitation estimates from Co-kriging and LP models. 29

2.10 Estimates of standard error of the satellite error field for January 2009 from Co-kriging (left) and LP (right) models. 30

2.11 Same as Figure 2.9 but for July 2009. 31
2.12 Same as Figure 2.10 but for July 2009. 32
2.13 Locations of extreme precipitation events for July 2009 (black dots) and elevation shown in color. 33
2.14 Original CHIRP product (left) and blended estimates from cross-validated Co-kriging (middle) and LP (right) models for extreme events in July 2009. 34
2.15 Same as Figure 2.3 but for the fourth pentad of August 2009. 35
2.16 Same as Figure 2.5 but for the fourth pentad of August 2009. 36
2.17 Same as Figure 2.7 but for the fourth pentad of August 2009. 37
2.18 Same as Figure 2.8 but for the fourth pentad of August 2009. 37
2.19 Same as Figure 2.9 but for the fourth pentad of August 2009. 38
2.20 Same as Figure 2.10 but for the fourth pentad of August 2009. 39
2.21 Same as Figure 2.13 but for the fourth pentad of August 2009. 40
2.22 Same as Figure 2.14 but for the fourth pentad of August 2009. 41

3.1 Study region geography, rain gauge measurement locations, and country names. . . . 46
3.2 Posterior distributions for January 2009; a-b) regression coefficients, c-d) kriging parameters. 52
3.3 Same as Figure 3.2 but for July 2009. 52
3.4 Median Bayesian Model empirical (black) and theoretical (red) variograms for Jan, Apr, Jul, and Oct 2009. 54
3.5 Drop-one cross-validation performance for January and July 2009. 55
3.6 Drop-25% validation for July 2009, a) RMSE, b) MAE, c)% Bias (original CHIRP shown in red). 56
3.7 Same as figure 3.6 for from drop-50% validation. 56
3.8 a-b) Original and Bayesian median blended CHIRP product for January 2009; station values on same color scale, c) change from original CHIRP due to blending. 58

3.9 a-b) 5^{th} and 95^{th} percentile of Bayesian predictive samples for January 2009. 59
3.10 Same as Figure 3.8 but for July 2009. 60
3.11 Same as Figure 3.9 but for July 2009. 61
3.12 Locations of extreme precipitation events for July 2009 (black dots) and elevation shown in color. 62
3.13 Blending performance on extreme events for July 2009. 62
3.14 Same as Figure 3.2 but for the fourth pentad of August, 2009. 64
3.15 Same as Figure 3.4 but for the fourth pentad of August 2009. 64
3.16 Same as 3.5 but for the fourth pentad of August 2009. 65
3.17 Same as 3.6 but for the fourth pentad of August 2009. 66
3.18 Same as figure 3.7 but for the fourth pentad of August 2009. 66
3.19 Same as Figure 3.8 but for the fourth pentad of August 2009. 67
3.20 Same as Figure 3.9 but for the fourth pentad of August 2009. 68
3.21 Same as Figure 3.12 but for the fourth pentad of August 2009. 69
3.22 Same as 3.13 but for the fourth pentad of August 2009. 69

4.1 Study region geography, rain gauge measurement locations, and country names. . . . 74
4.2 Empirical (black) and theoretical (red) variograms for January β (a) and ϵ (b), and for July β (c) and ϵ (d). 79
4.3 Original CHIRP estimates (left) and blended estimates for January 2009. 80
4.4 Same as Figure 4.3 but for July 2009. 80
4.5 Original CHIRP estimates (left) and cross-validated estimates for January 2009. . . 81
4.6 Same as Figure 4.5 but for July 2009. 82
4.7 Prediction skill measures from drop-25% cross-validation for July 2009. Red dots correspond to estimates from original CHIRP. 83
4.8 Same as 4.7 but from drop-50% cross-validation. 84
4.9 Estimates of β (left) and ϵ (right) for January 2009. 85

4.10 Standard error associated with estimation of β (left) and ϵ (right) for January 2009. 85

4.11 a) Original CHIRP estimates, b) blended estimates, and c) change in CHIRP for January 2009. a-b) Rain gauge measurements on same scale for comparison. 86

4.12 Same as Figure 4.9 but for July 2009. 87

4.13 Same as Figure 4.10 but for July 2009. 88

4.14 Same as Figure 4.11 but for July 2009. 88

4.15 Location of extreme precipitation events for July 2009 (black dots) and elevation shown in color. 89

4.16 Original CHIRP product (left) and blended estimates from cross-validated output for extreme events in July 2009. 90

4.17 Same as Figure 4.2 but for the fourth pentad of August 2009. 91

4.18 Same as Figure 4.3 but for the fourth pentad of August 2009. 92

4.19 Same as Figure 4.5 but for the fourth pentad of August 2009. 93

4.20 Same as Figure 4.7 but for the fourth pentad of August 2009. 94

4.21 Same as Figure 4.8 but for the fourth pentad of August 2009. 94

4.22 Same as Figure 4.9 but for the fourth pentad of August 2009. 95

4.23 Same as Figure 4.10 but for the fourth pentad of August 2009. 95

4.24 Same as Figure 4.11 but for the fourth pentad of August 2009. 96

4.25 Same as Figure 4.15 but for the fourth pentad of August 2009. 97

4.26 Same as Figure 4.16 but for the fourth pentad of August 2009. 97

Chapter 1

Introduction

Drought and flood management practices require accurate estimates of precipitation in space and time. However, in regions with complicated terrain, data is often sparse, clustered in valleys, and of poor temporal consistency. Consequently, extreme precipitation events tend to be poorly represented. Satellite-derived rainfall data is an attractive alternative in such regions and is being widely used, though it too suffers from problems such as underestimation of extreme events, due to its dependency on retrieval algorithms and the indirect relationship between satellite infrared radiation observations and precipitation intensities. Thus, it seems appropriate to blend satellite-derived rainfall products of extensive spatial coverage with in situ rain gauge data in order to produce a more robust estimate of precipitation.

Motivation for this research stems from the work of the Climate Hazards Group (CHG) at the University of California, Santa Barbara. With support from the United States Agency for International Development (USAID), the United States Geological Survey (USGS), National Oceanic and Atmospheric Administration (NOAA), and National Aeronautics and Space Administration (NASA), the CHG develops early warning science to support the Famine Early Warning Systems Network (FEWS NET). At its inception, FEWS NET was used to identify vulnerable populations within Sub-Saharan Africa and the hazards that threaten their livelihoods. FEWS NET has since widened its scope to include Afghanistan, Haiti, and the Central American region. Their primary mission involves monitoring hydroclimatic hazards related to food security, droughts and floods within FEWS NET countries. Accurate monitoring requires good spatiotemporal estimates

of precipitation, thus aiding in the identification of water-stressed regions.

The efforts of this research are also aligned with the interests many global governmental bodies. For example, an intergovernmental technical body known as CRRH (Comité Regional de Recursos Hidráulicos or the Regional Water Resources Committee), which consists of members from every Central American government, has long been at the forefront of the regional water resources movement. In fact, many of the projects spearheaded by CRRH are developed with the cooperation of numerous agencies of the United States government. Their primary objective is a project involving the generation and public availability of climate information at a regional level, in turn assisting in the improvement of regional environmental management.

The research presented here is based on the assumption that these methods are portable to any region around the globe, whether Central America, Sub-Saharan Africa, or any water-insecure region. Both persistent flooding and prolonged drought may define a region as water-insecure; these disasters are destructive not only to the infrastructure of a community, but also the well-being of the people therein. For example, both Central and South America are greatly dependent on seasonal agriculture production, but because of the tropical and oceanic location, drought is a rare risk for these nations. Conversely, their location and geography are prime contributors to extreme and intense precipitation patterns. Great changes in elevation - for example from sea level to nearly 5,000 meters in Colombia - coupled with the convergence of trade winds create the perfect environment for intense and extremely variable local convective rainfall events. Annual hurricanes and persistent monsoonal rainfall can cause agricultural flooding and great socioeconomic distress. In fact, records show that a single disastrous event can destroy up to 50 years of infrastructure and progress. Access to more representative gridded high-resolution spatial time series of precipitation will better assist scientists in identifying regions vulnerable to persistent flooding or crop failure risk.

Historically, gridded time series products have been created by interpolating available rain gauge measurements over space, for every available snapshot in time, using physiographic predictors such as latitude, longitude, elevation, etc. While this method can be sufficient with adequate and

reliable data, many nations do not have access to spatially- nor temporally-consistent data series. Still, interpolation is often applied with insufficient data, thus leading to extrapolation error and edge effects. Over the years there have been efforts by numerous agencies to produce better gridded precipitation products. These products are similar in nature, in that their estimates are weighted averages of interpolated gauge measurements, numerical model output, and satellite-derived estimates. The most notable include the Tropical Rainfall Measuring Mission (TRMM), Global Precipitation Climatology Project (GPCP), and NOAA CPC Morphing Technique (CMORPH). The TRMM is available as daily-to-monthly products, beginning in 1997 until present, and has spatial resolution of 0.05°. The GPCP is a monthly product with 2.5° spatial resolution and temporal extent from 1979 to present. The CMORPH is a half-hourly product with 0.07277° spatial resolution and a temporal extent of December 3, 2002 to present. These products are widely used for climate analysis, climate change studies, and hydrologic model verification (Curtis & Adler 2000; Trenberth & Caron 2000; Janowiak et al. 1998; Dai et al. 2001).

Satellite-derived precipitation estimates are becoming more prevalent in recent years as relationships between hydroclimatic processes and rainfall are becoming better understood. Infrared (IR) geostationary satellites provide frequent coverage – at 3-hourly intervals or finer – and report brightness of cloud-tops, which can be related to the temperature of the cloud-tops. These satellite products show considerable skill, especially over the tropical region, where cloud-top temperature and precipitation correlate rather well. Microwave-based (MW) satellite products measure the scattering of raindrops and ice particles rather than cloud tops. This method is less effective over warm land surfaces and is of poor spatial resolution, yet has been used to estimate rainfall for periods dating back to 1979. It has been found that this method provides relatively accurate instantaneous estimates of precipitation, but suffers from poor temporal sampling (Adler et al. 1993, 1994). Due to the tropical location of the study region, this research will be using IR data in its satellite-derived precipitation estimates, which will allow for better spatial and temporal accuracy.

The satellite estimate used in this research is referred to as the CHIRP (Climate Hazards group InfraRed Precipitation). CHIRP is produced by the CHG using the Climate Hazards group

Precipitation Climatology (CHPCLIM) and current IR percent-anomalies; it has nearly global coverage and is of fine resolution (0.05°x0.05°). Both pentadal and monthly estimates are available for grid-cells over all landmass within 50°N and 50°S. The temporal range of the CHIRP product is bounded only by the availability of IR satellite data - 1981 to present.

Study Region

Spanning from 2°N to 18°N and from 93°W to 67°W, the study region consists of eight countries – Belize, Guatemala, El Salvador, Honduras, Guatemala, Costa Rica, Panama, and Colombia – as well as small portions of Mexico and Venezuela (see Figure 1.1). The study region for this research was selected due to its great geographic variability and its tropical location, which are main contributors to the extreme magnitude and variability of precipitation events therein. This region is consistently battered by hurricanes and monsoonal rainfall, making it a good test bed for the methods proposed in this study.

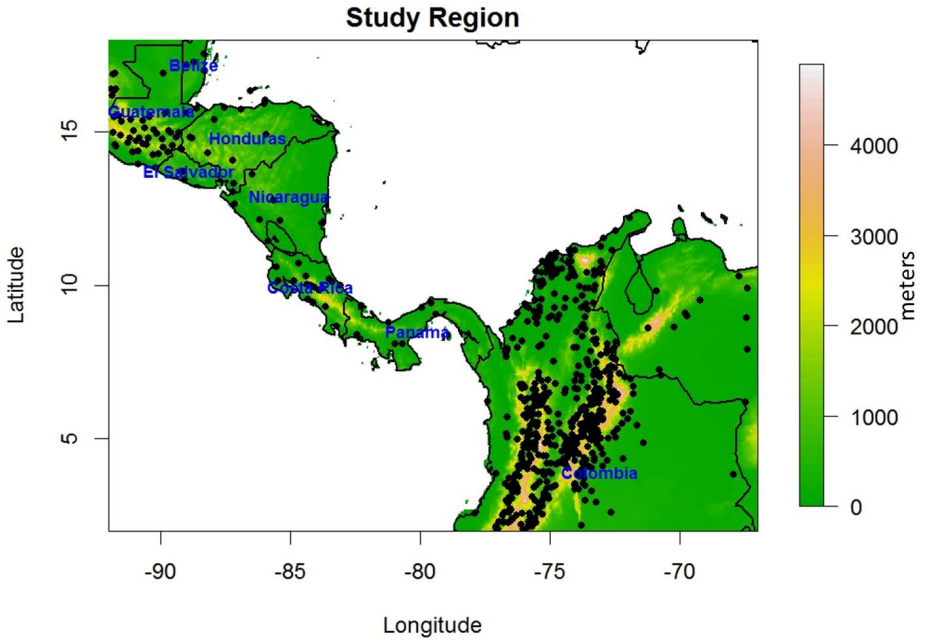

Figure 1.1: Study region geography, rain gauge measurement locations, and country names.

The dry season in the study region extends from December through early May, while the effects of both Pacific and Atlantic tropical depressions are felt during the prolonged wet season that extends from mid-May through November. A major contributor to the prolonged wet season is the inter-tropical convergence zone (ITCZ) – an asymmetric band of convection which encircles the globe. The position of the ITCZ is non-stationary, due to the seasonal shift of the trade winds. In the northern hemisphere's summer months, the ITCZ fully encompasses the study region, which results in many extreme precipitation events. In the northern hemisphere's winter months, or the study region's dry season, the ITCZ converges just to its south, resulting in fewer extreme precipitation events (see Figure 1.2).

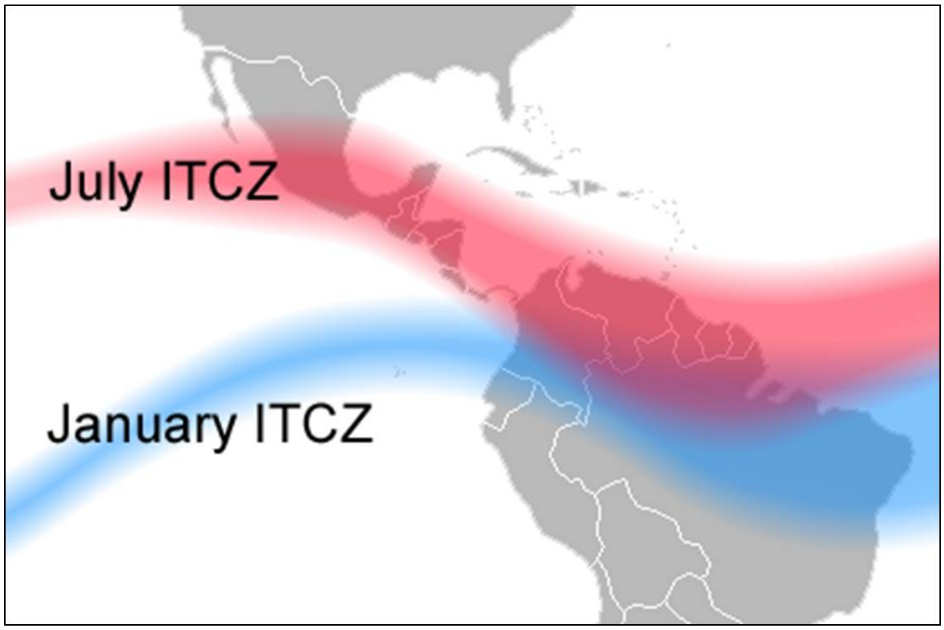

Figure 1.2: Inter-tropical Convergence Zone within the study region.

Another cause for intense spatial variability of rainfall within the study region is its complex geography. The elevation of the study region ranges from sea level to nearly 5000 meters (as can be seen in Figure 1.1) with numerous mountain ranges spanning Central America and Colombia. Costa Rica and Guatemala have steep elevation gradients along the coasts, making them vulnerable to extreme weather spawning from both Pacific and Atlantic tropical depressions. Coastal mountain

ranges cause an abrupt rising of warm, wet air. As this tropical air rises, it cools rather quickly, thus releasing the moisture from the air as precipitation. This is known as orographic precipitation, and is a major cause for the variability in extreme precipitation events in the region.

The spatial and temporal variability of rainfall in this region coupled with weaker socio-economic conditions make these nations extra susceptible to droughts and floods. History has shown that a single extreme precipitation event – whether hurricane, tropical storm, or prolonged precipitation – may result in flooding, landslides, and potentially billions of dollars in damages, thus setting the national infrastructure back many years. Conversely, a prolonged dry spell may be equally devastating to a growing nation, causing widespread crop failures. Thus, a comprehensive understanding of rainfall variability is of critical importance for resources planning and management.

Data

Two precipitation datasets are utilized in this research – gauge measurements and satellite-derived estimates. The gauge measurements (shown in Figure 1.1) were collected from partner sources of the UCSB Climate Hazards Group; they are available as pentadal sums ranging from the start of January 1981 to the end of December 2012. Due to its direct relationship with the natural process, rain gauge measurements are generally the most representative measure of extreme precipitation events. Setbacks to using gauge measurements include spatially inconsistent datasets – in that these are point measurements, reflecting to a degree the true process at the measurement location only. That being said, the assumption that gauge measurements represent the absolute truth of the unknown natural process is problematic, as these values also come with the inherent uncertainty of measurement error. Another setback with these data is they are point values that tend to be clustered in highly-populated and lower elevation areas, as well as sparsely scattered throughout the domain.

In light of these shortcomings, satellite-derived precipitation products offer an attractive alternate approach to precipitation estimation. Satellite products provide high-resolution spatial and

temporal coverage as well as a bird's-eye view of the extent of rainfall events. One major setback is these types of products have a tendency of underestimating the magnitude of precipitation, which is especially evident during a region's wet season due to an overall increase in magnitude and variability of these events. One cause for this problem is the resolution of the satellite estimate. Even at fine spatial discretization (i.e. 0.05°x0.05°, or roughly 5 km at the equator), these products estimate areal averages of rainfall over relatively large areas in comparison to point-source gauge measurements. The effects of areal averaging are emphasized in regions of complex terrain due to great variability in local convection. Also, satellite-derived estimates utilize mathematical algorithms in approximating the magnitude and occurrence of precipitation events. These algorithms include inherent bias and error due to the indirect relationship between the natural process and the interpretation of satellite data fields. For instance, satellite-derived precipitation estimates based on infrared radiation data – the kind used in this research – are based on the temperature of cloud-tops. It is common practice to specify a threshold of cloud-top temperature (i.e. 235 K) and assume colder cloud-tops imply the occurrence of precipitation, while warmer cloud-tops imply non-occurrence. A major cause for error in this assumption is the presence of cold-top thick cirrus clouds, which are typically not precipitating. It is widely accepted that satellite estimates of rainfall are better at capturing the spatial structure of precipitation, but fail in estimating magnitudes, especially of extreme events – this is because the satellite estimates are derived estimates of precipitation (Xie & Arkin 1997).

The satellite estimate used in this research is the CHIRP, a pentadal (five-day sum) product with a temporal range of Jan 1981 through near real-time. The most current product is available two days after the end of the most recent pentad. There are six pentads per month, regardless of how many days are in a given month. For example, the sixth pentad of October will consistently be the sum of the last six days, October 26-31. The sixth pentad of February, however, will be the sum of either three or four days, depending on leap year status. Pentads may also be summed to estimate monthly totals of precipitation. The algorithm in producing the CHIRP estimate is as

follows:

$$CHIRP_{i,j,k} = \%IR_{i,j,k} \cdot CHPCLIM_k \tag{1.1}$$

In the above equation, $CHIRP_{i,j,k}$ represents the satellite-derived precipitation estimate for pentad k of month j in year i; $\%IR_{i,j,k}$ represents the percent anomaly infrared radiation satellite estimate for pentad k of month j in year i; and $CHPCLIM_k$ represents the climatology for pentad k. Previous analyses have shown that this product systematically underestimates the magnitude of extreme precipitation events at monthly and pentadal scales, and overestimates the non-occurrence of rainfall at the pentadal scale.

Methods

This study is motivated by the need for a more accurate gridded time series of precipitation data to further the efforts of scientists aimed at forecasting disastrous precipitation events in regions vulnerable to hydroclimatic hazards. To this end, a suite of methods for blending satellite-derived precipitation datasets with rain gauge measurements is proposed.

In the second chapter of this research it is hypothesized that the error structure of the CHIRP product is spatially coherent, thus the following blending procedure for merging rain gauge measurements and satellite datasets is proposed. Based on their individual coordinates (latitude, longitude), the gauge data are assigned to the nearest grid cell of the CHIRP product, where the error is defined as $r(s) = Y_{obs}(s) - Y_{sat}(s)$. The spatial structure of $r(s)$ is then modeled using physiographic information (easting, northing, and elevation) via (i) a traditional Co-kriging approach which utilizes a variogram that is calculated in Euclidean space and (ii) a nonparametric method based on local polynomial functional estimation. The models are used to estimate $\hat{r}$ at all grid points thus producing a spatial map of satellite error which is added to the original CHIRP product, effectively creating an improved satellite estimate. Robust methods of model validation are carried out, and the results are analyzed.

The third chapter of this research offers a unique application of Bayesian Hierarchical Mod-

eling. A hierarchical model is defined, where the true rainfall process at a point is a realization of a normally distributed function with mean μ and standard deviation s^2 – such as $Y_{obs}(s) \sim N(\mu(s), s^2(s))$. This mean value is defined as $\mu(s) = \beta_1 Y_{sat}(s) + \beta_2 elev(s) + \epsilon(s)$. This traditional linear regression framework allows for bias correction via the β parameters. The error term $\epsilon(s)$ is a zero-mean Gaussian process model with parameters σ^2 (marginal variance) and ϕ (effective range). The nugget effect τ^2 is set to zero to force the model to be an exact estimator at observations included in the model structure. Prior distributions are defined for the β parameters, as well as the kriging parameters σ^2 (marginal variance) and ϕ (effective range). A traditional linear model is fit 15,000 times in a Markov Chain Monte Carlo (MCMC) framework, thus producing 15,000 values of all parameters (i.e. posterior distributions). Due to a necessary burn-in period, the first 5,000 values are thrown out. The burn-in period ensures a robust posterior distribution for all parameters, which allows for insight into not only the bias of the original CHIRP product but also the uncertainty that comes with the model coefficients. It is then possible to implement the Gaussian process (i.e. Bayesian Kriging) model in predictive mode – estimating the error residual $\hat{\epsilon}$ 10,000 times at each location – which is systematically added to the bias-corrected CHIRP product. Therefore 10,000 maps of the estimated rainfall process are produced. Representative statistics may then be calculated at each point – mean, median, standard deviation, and confidence interval (i.e. 5^{th} and 95^{th} percentiles). Similar methods of model validation are carried out, and the results are analyzed.

The fourth chapter of this research was spurred from the realization that, while the CHIRP product is indeed biased, it is unrealistic to assume the inherent bias in this product is uniform over any given study region. As the data in this research has a time frame of 1981-2012, it follows to utilize the entire time series at each location to shed light on the underlying spatial bias of the CHIRP product. At each point with at least 2/3 complete time series (22/32 years), a linear model is fit with the form $Y_{obs}(s,t) = \beta(s) Y_{sat}(s,t) + \epsilon(s,t)$. The value of $\beta(s)$ is extracted and saved for use in a spatial Gaussian process model. If at any location the value of $\beta(s)$ is less than 0.5 or greater than 2.0, this value is thrown out and not used in the interpolation process, thus

eliminating the chance of including a very small or large coefficient from a poorly-fit model – doing so ensures the mean function remains coherent and structurally sound. The author also assumes that the CHIRP should not be adjusted to less than half nor more than double its original value. The values of $\beta(s)$ are interpolated using a Co-kriging model similar to that of the first chapter, such that the final product is of the same resolution as the original CHIRP, in effect creating a map for bias-correcting the product. Because this method is assumed to have eliminated bias within the CHIRP product, the residuals for any given year are independent and identically-distributed. The theoretical variogram will then be a much better fit to the data, thus ensuring a more robust Co-kriging model. For any given year the residuals are extracted from the linear model output and are interpolated (again via Co-kriging). Summing this residual map with the bias-corrected CHIRP yields the final blended CHIRP product. Similar methods of model validation are carried out on the residual model, and the results are analyzed.

The methods described above are assumed portable, in the sense they can be applied to any reasonably-sized region around the globe. Applying any of the above blending methods to the entire CHIRP history will result in a more representative gridded time series of precipitation for the study region, which ultimately provides a more stable platform for decision makers, hydrologic modelers, scientists, and disaster relief agencies.

Chapter 2

Kriging and Local Polynomial Methods for Blending

2.1 Introduction

This chapter presents an analysis of two unique approaches to blending high-resolution satellite estimates with in situ rain gauge measurements. To begin, the gauge measurements are assigned to the nearest grid cell of the satellite estimate, where the difference between the measurements and the estimate is considered the satellite error r. Assuming a coherent spatial structure of the satellite error, it follows that the spatial model used for estimation will fully capture the spatial variability of the satellite error field. Therefore it follows to estimate the satellite error at the same resolution as the satellite estimate and add the estimated satellite error with the original satellite estimate to obtain a blended estimate of rainfall. Two spatial models are used for estimation – Co-kriging and K-Nearest Neighbor Local Polynomial. The methodology presented here is unique from previous approaches to blending satellite estimates with gauge measurements, which have relied on weighted averages of interpolated gauge measurements, numerical model output, and satellite-derived estimates (Huffman et al. 1995, 2007, Adler et al. 2003, Joyce et al. 2004).

Co-kriging is a natural extension to Ordinary Kriging (OK), such that covariates are no longer limited to location (Journel & Huijbregts 1978). Tropical precipitation is highly correlated with elevation, thus it makes sense to include elevation, along with location, as covariates for this application. OK is very popular in the geostatistical community, implementing both covariance and optimization methods to minimize the bias and variance of the estimated values, giving it status as a "Best Linear Unbiased Estimator" (Journel & Huijbregts 1978, Isaaks & Srivastava 1989, Cressie

1990). Kriging and Co-kriging have long been used in a variety of hydrologic and climatological applications (Hudson & Wackernagel 1994, Biau et al. 1999, Hartkamp et al. 1999, Holdaway 1996, Desbarats et al. 2002, Dingman et al. 1988, Garen et al. 1994, Apaydin et al. 2004).

K-Nearest Neighbor Local Polynomial (LP) is a nonparametric regression method that is gaining popularity due to its capability to better represent local nonlinearities in a process when compared to global estimation methods. For estimation at any given point it considers only a fraction of the available data which produces estimates with less uncertainty and a smoother spatial structure. Local polynomial functional estimation has been widely used in a variety of applications such as streamflow and salinity modeling (Prairie et al. 2005, 2006), streamflow forecasting (Grantz et al. 2005, Regonda et al. 2006a, Bracken et al. 2009), spatial estimates of climate data (Rajagopalan & Lall 1998, Apaydin et al. 2004), water quality modeling (Towler et al. 2009, 2010), and many other applications.

The proposed approaches for blending satellite and ground observations of rainfall are unique applications of the Co-kriging and LP methodologies. The theory behind Co-kriging has been around for decades and is an acceptable form of interpolation in the scientific community. LP has been gaining popularity in the geostatistical community, and has been implemented for spatial estimation of climate data, thus is also an acceptable form of interpolation. The study region and data are described in the next section, followed by an in depth description of the Co-kriging and LP frameworks, and finally the results section.

2.2 Study Region & Data

The details of the study region and data sets are described in greater detail in Chapter 1, an abstract of which is provided below.

2.2.1 Study Region

The study region covers the Central American region (see Figure 2.1). The dry season in the region extends from December through early May and the effects of both Pacific and Atlantic

tropical depressions are felt during the wet season that extends from mid-May through November. A major contributor to the prolonged wet season is the inter-tropical convergence zone (ITCZ) – an asymmetric band of convection which encircles the globe. The position of the ITCZ is non-stationary due to the seasonal shift of the trade winds.

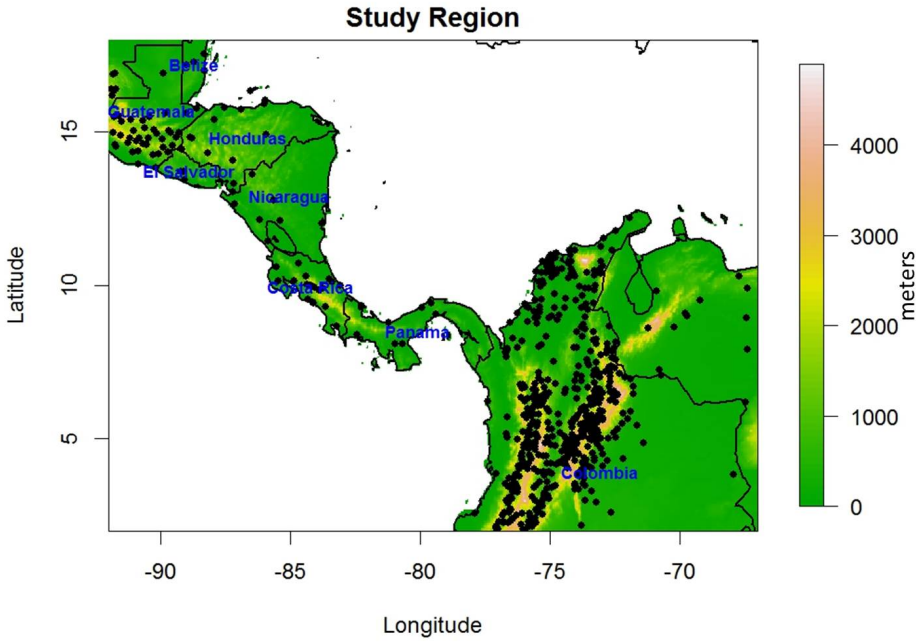

Figure 2.1: Study region geography, rain gauge measurement locations, and country names.

Another cause for intense spatial variability of rainfall within the study region is its complex geography, where the elevation ranges from sea level to nearly 5000 meters (as can be seen in Figure 2.1) with numerous mountain ranges spanning Central America and Colombia. Costa Rica and Guatemala have steep elevation gradients along the coasts, making them vulnerable to extreme weather spawning from both Pacific and Atlantic tropical depressions. Coastal mountain ranges cause an abrupt rising of warm, wet air. As this tropical air rises, it cools rather quickly, thus releasing the moisture from the air as precipitation, a major cause for the variability in extreme precipitation events in the region.

Given the spatial and temporal variability of rainfall in this region coupled with vulnerable

socio-economic conditions, droughts and floods can have a debilitating impact on these societies. It has been observed that a single extreme wet precipitation event – whether hurricane, tropical storm, or prolonged precipitation – may result in flooding, landslides, and potentially millions of dollars in damages, thus setting the national infrastructure back many years. Conversely, a prolonged dry spell may be equally devastating to a growing nation, causing widespread crop failures. Thus, a comprehensive understanding of rainfall variability is of critical importance for resources planning and management.

2.2.2 Data

Two precipitation datasets are utilized in this research – gauge measurements and satellite-derived estimates. Rain gauge measurements are point values that tend to be clustered in populous and lower elevation areas and sparsely scattered throughout the domain; the satellite-derived estimates are high-resolution – 0.05°x0.05° (approx. 5 km) – areal averages for five day periods (i.e. pentads). The satellite-derived precipitation estimate used in this research is the CHIRP (Climate Hazards group InfraRed Precipitation) covering the first pentad (five-day sum) of January, 1981 through the last pentad of December, 2012. There are six pentads per month, regardless of how many days are in a given month. For example, the sixth pentad of October will consistently be the sum of the last six days, October 26-31. The sixth pentad of February, however, will be the sum of either three or four days, depending on leap year status. Pentads may also be summed to obtain monthly totals of precipitation. It is widely accepted that satellite estimates of rainfall are better at capturing the spatial structure of precipitation, but fail in estimating magnitudes, especially of extreme events – this is because the satellite estimates are derived estimates of precipitation (Xie & Arkin 1997).

The CHIRP estimate is produced using the CHPCLIM (Climate Hazards group Precipitation Climatology), a completely unbounded global long term pentadal climatology of precipitation of the same resolution (0.05°x0.05°), as well as the infrared radiation (IR) satellite product. The CHPCLIM for the appropriate pentad is multiplied by the most recent IR percent-anomaly to

produce the CHIRP estimate, shown below:

$$CHIRP_{i,j,k} = \%IR_{i,j,k} \cdot CHPCLIM_k \tag{2.1}$$

In the above equation, $CHIRP_{i,j,k}$ represents the satellite-derived precipitation estimate for pentad k of month j in year i; $\%IR_{i,j,k}$ represents the infrared radiation satellite estimate percent-anomaly for pentad k of month j in year i; and $CHPCLIM_k$ represents the precipitation climatology for pentad k.

The station data (shown in Figure 2.1) was collected from partner sources of the UCSB Climate Hazards Group; quality control and pentadal-aggregation techniques were implemented by the CHG to produce time series data of the same temporal extent as the available CHIRP products (the first pentad of January, 1981 through the last pentad of December, 2012).

2.3 Proposed Blending Methods

Before describing the methods, the following variables are defined: $Y_{sat}(s)$ and $Y_{obs}(s)$ are the satellite estimates and observed precipitation, respectively, at locations $s = 1, ..., N_s$, where N_s is the number of stations with non-missing data for a given pentad or month. We define $r(s)$ as the satellite error, $r(s) = Y_{obs}(s) - Y_{sat}(s)$.

As mentioned earlier, satellite estimates tend to under represent extremes, which have a spatial structure in that extremes tend to be related to location and elevation. This relationship suggests that the satellite error too should inherit this spatial structure. The basis for the two blending methods used in this research is the assumption of a coherent spatial structure of satellite estimation error.

The blending methods involve developing a model of dependent variable $\hat{r}(s)$ as a function of the covariates – easting, northing, and elevation. Thus, at any desired location the satellite error is estimated using its covariates as input into the model $\hat{r}(s)$, which is then added to the original satellite estimate at that location $Y_{sat}(s)$ to obtain the estimate of true precipitation, $\hat{Y}(s) = Y_{sat}(s) + \hat{r}(s)$.

What follows are generalized descriptions of the two models used in this research – Co-kriging and K-Nearest Neighbor Local Polynomial (LP).

2.3.1 Kriging and Co-kriging

Kriging is a traditional spatial interpolation method, due to its long-term status as a "Best Linear Unbiased Estimator" (Journel & Huijbregts 1978, Isaaks & Srivastava 1989, Cressie 1990), and its popularity within the geostatistical community. Kriging is based on the concept that an estimate at an unknown point is calculated by determining the best linear combination of the available data. The weights for this combination are obtained from a 'variogram' which defines the spatial variability of the dependent variable. Under the assumption of stationarity, the observed variogram $\gamma(h)$ of a given process $r(s)$ at location s is defined as follows:

$$\gamma(h) = \frac{1}{2}\mathrm{Var}[r(s_i) - r(s_j)] \tag{2.2}$$

where $h = (s_i - s_j)$ is the Euclidean distance between observation locations s_i and s_j. Note that the value of h has units of meters because latitude and longitude are converted to easting and northing; also elevation is included as a covariate, making the value a three-dimensional Euclidean distance. This is computed for all possible pairs of observations and the plot of h versus $\gamma(h)$ is the observed variogram. To this, a theoretical variogram is fitted which has the form

$$\tilde{\gamma}(h) = \begin{cases} C_0 + C_1(1 - \exp\Big(-\frac{3|h|}{a}\Big), & |h| \geq 0 \\ 0, & |h| = 0 \end{cases} \tag{2.3}$$

where C_0 is the nugget effect, or the result of dissimilarity of sample values separated by small distances; a is the effective range, or the distance at which the variogram reaches plateau; $C_0 + C_1$ is the sill, or the maximum variogram value. The above theoretical variogram follows an exponential distribution, and is therefore called an 'exponential variogram'. There are variations to the above such as 'spherical' and 'Normal' variograms (Journel 1989), which together form a limited set of available theoretical variogram models.

Within the Kriging framework, the estimate of the dependent variable at any desired location s is obtained as a weighted combination of all the observations as follows:

$$r(s) = \sum_{i=1}^{N} \lambda_i r(s_i) \tag{2.4}$$

The non-negative weights λ_i are obtained as a solution to the constrained optimization problem:

$$\text{Minimize } L = \Big(r(s_i) - r(s_j)\Big)^2 \tag{2.5}$$

with the constraint that the non-negative weights, λ_i, must sum to 1 such as

$$\sum_{i=1}^{N} \lambda_i = 1 \tag{2.6}$$

Minimizing L (partially differentiating with respect to λ_i) results in a set of N linear equations shown as:

$$\sum_{i=1}^{N} \lambda_i \tilde{\gamma}(s_i - s_j) + \mu = \tilde{\gamma}(s_i - s_j) \text{ for } j = 1...N \tag{2.7}$$

The $\tilde{\gamma}(s_i - s_j)$ in the above equations refer to the values from the fitted theoretical exponential variogram. In order to quantify the uncertainty of the Kriging estimate, the standard error is obtained by:

$$s_{r(s)} = \sqrt{\sum_{i=1}^{N} \frac{(\hat{r}(s) - r(s))_i^2}{N}} \tag{2.8}$$

The implementation can be summarized as follows:

(1) Compute covariance of all observation points

(2) Plot covariance vs. distance between points to produce empirical variogram

(3) Produce theoretical variogram based on empirical variogram parameters (nugget, sill, range)

(4) Obtain weights via constrained optimization problem

(5) Estimate at any point s_0 using values from theoretical variogram and optimal weights

The Kriging estimates are minimum variance unbiased estimates (Journel 1989). When there is no nugget effect (i.e., $C_0 = 0$, or homogeneity in sampled values in close range), Kriging

is an 'exact estimator' at the observed locations. The theoretical variogram fitted to the observed variogram is crucial in the Kriging method, as the non-negative weights, λ_i are obtained from the fitted variogram. If the theoretical variogram is not a good fit to the data, typically in cases of non-stationarity, then the Kriging estimates tend to be poor and they do not capture the spatial structure faithfully – this is one of the major drawbacks of Kriging. However, the variogram options available are reasonable in most applications for capturing the spatial variability structure – and thus it is widely used in a variety of hydrologic and climatological applications (Hudson & Wackernagel 1994, Biau et al. 1999, Hartkamp et al. 1999, Holdaway 1996, Desbarats et al. 2002, Dingman et al. 1998, Garen et al. 1994). If other covariates are included in the variogram estimation, specifically in computing h, then it is known as Co-kriging. In this research we propose to include elevation in addition to easting and northing into the set of covariates, as precipitation is modulated by elevation (i.e. higher elevation regions tend to get more rainfall). Co-kriging is a natural extension of Kriging in the sense it allows for multiple covariates (Journel & Huijbregts 1978).

2.3.2 K-Nearest Neighbor Local Polynomial

In the context of this research, a regression model may be represented as

$$\hat{r} = f(s, \text{elev}) + \epsilon \tag{2.9}$$

where f represents a global function that is fit to a set of predictor variables s=(easting, northing) and elevation. $\hat{r}$ is the dependent variable of interest, and ϵ is the independent, normally-distributed (with mean 0 and unknown variance σ^2) estimation error.

Linear regression is the common approach, wherein f is linear with respect to the dependent variables, in this application it would be as follows:

$$\hat{r} = \beta_0 + \beta_1 \text{easting} + \beta_2 \text{northing} + \beta_3 \text{elevation} + \epsilon \tag{2.10}$$

These coefficients are estimated by minimizing the mean squared errors (Helsel & Hirsch 1995).

Linear regression has theory that is well developed and is widely used (Helsel & Hirsch 1995, Rao & Toutenburg 1999).

The linear regression, especially fitting a single ('global') model, has several drawbacks including (i) assumption of a linear relationship between the independent and predictor variables over the entire range when local nonlinearities may be present, (ii) the assumption of normality of input data and errors, (iii) model parameters greatly influenced by outliers, and (iv) higher order models require large amounts of data to fit (Rajagopalan et al. 2005, Lall 1995).

Local Polynomial Regression

To alleviate these drawbacks, 'local' or nonparametric regression methods offer an attractive alternative. In these methods, the estimate of the function at a point is based (or influenced) only by data points within its neighborhood. This neighborhood is generally much smaller than the extent of the domain, thus providing the capability to better represent local nonlinearities in the data. Many local estimation procedures have been theorized and tested to date including splines, kernel-based estimation (Bowman & Azzalini 1997), K-Nearest Neighbor polynomials (Rajagopalan & Lall 1999), and locally-weighted polynomials (Loader 1999). Local polynomials tend to be easier to understand and implement, besides offering a robust and flexible approach in estimating a variety of functions, which is implemented using the Locfit package (Loader 2001) in the public domain statistical software R. Local polynomial functional estimation has been widely used in a variety of applications such as streamflow and salinity modeling (Prairie et al. 2005, 2006), streamflow forecasting (Grantz et al. 2005, Regonda et al. 2006a, Bracken et al. 2009), spatial estimates of precipitation (Rajagopalan & Lall 1998), water quality modeling (Towler et al. 2009, 2010), and many other applications.

As mentioned above, the local polynomials are based on fitting a local polynomial of order p at a point of interested based on a neighborhood of size $K = \alpha N$ where N is the number of observations and α is a fraction of observations that vary with (0,1]. At a point of interest, K-nearest neighbors are identified; they are assigned weights based on their distances from the point

and a weighted least squares is used to fit the local polynomial or local regression. Note that a local polynomial model with $\alpha = 1$, $p = 1$, and equal weights for all neighbors will reduce to the traditional linear regression, thus making it a general framework for function estimation.

A common practice in determining the values of α and p is Generalized Cross-Validation (GCV). Craven & Wahba (1979) showed that GCV is generally unbiased in its estimate of predictive risk, provided the degrees of freedom in the model are sufficiently large. While GCV is used to select the model parameters α and p, it has also been used to select the best subset of covariates (Regonda et al. 2006a). The GCV objective function is calculated as follows:

$$GCV(\alpha, p) = \frac{\sum_{i=1}^{N} \frac{e_i^2}{N}}{(1 - \frac{q}{N})^2} \tag{2.11}$$

where q refers to the number of model parameters in the current model, N is the total number of data points (observations), and e_i is the residual for the i^{th} data point. In this research, the minimum value of α is dependent on the degree of polynomial p as well as the value of q, as shown below. The maximum value of α is 1, which would include all N data points in the model.

$$\text{Min}(\alpha) = \begin{cases} 2 \cdot \frac{q \cdot p + 1}{N}, & q \geq 3 \\ 3 \cdot \frac{q \cdot p + 1}{N}, & \text{otherwise} \end{cases} \tag{2.12}$$

Shown below is the conditional bisquare kernel

$$W(h) = \begin{cases} (1 - |h|^2)^2, & |h| < 1 \\ 0, & \text{otherwise} \end{cases} \tag{2.13}$$

where $W(h)$ is the weight at K nearest neighbor locations, with h representing the normalized Euclidean distance between the point of interest and each of the K nearest neighbors. The GCV is computed for all candidate models (covariates, α and p) and the combination with the minimum value is selected as the best model.

Confidence intervals of the estimate are based on estimates of standard error from the local regressions which are consistent with linear regression theory, testing the significance of the local

polynomial model with respect to alternate models such as linear regression are also done in the usual analysis of variance (ANOVA) approach – these are described in detail in Loader (1999).

2.4 Results

Both Co-kriging and K-Nearest Neighbor Local Polynomial (hereafter 'LP') models are applied to monthly total precipitation as well as a representative wet pentad for 2009. This research focuses on 2009 due to newly available station data for Colombia provided by partner sources of the Climate Hazards Group at UC-Santa Barbara. Model fitting and drop-one cross-validated estimates of R^2 and root mean square error (RMSE) measures are computed for each month. Further analysis such as drop-25% and -50% cross-validation and performance on extreme precipitation events are shown for representative dry season (January) and wet season (July) months. Analysis will focus on overall performance but, more importantly, the effect that blending has on the CHIRP's representativeness of extreme precipitation events. For consistency, an extreme event is defined at a location where rainfall exceeds the 85^{th} percentile of all events on record. To illustrate the utility of these blending methods at a shorter time scale, a representative wet pentad for 2009 is also analyzed. One of the wettest pentads of 2009 occurred in the midst of the region's wet season, during the fourth pentad of August, with total rainfall of nearly 400 mm.

Two packages (and functions) in the statistical software program R provide a straight-forward and user-friendly framework for implementing the Co-kriging and LP models, respectively: fields (Krig) and locfit (locfit).

2.4.1 Monthly Totals

Model Fitting

The first step in Co-kriging is the fitting of variograms, the empirical and model variograms for four select months of 2009 are shown in Figure 2.2. Note the distance h is in covariate space, which includes elevation as well as easting and northing, and therefore is a normalized Euclidean

distance with units of meters. Also, the variograms in Figure 2.2 have truncated range – and thus show only part of the covariate space – because covariance is negligibly variable at distances greater than the initial sill. The modeled variograms are very good at describing the spatial variability of

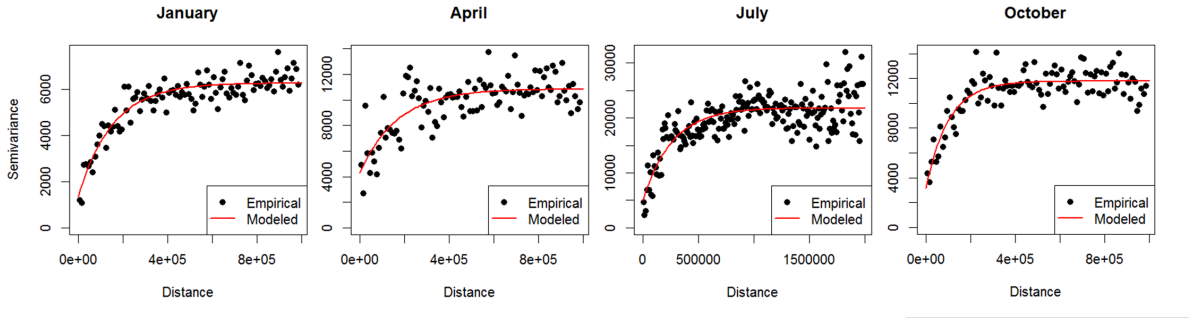

Figure 2.2: Empirical (black) and theoretical (red) variograms for January, April, July, and October 2009.

the satellite error for a variety of months. For the LP model, the best model for each month as selected by GCV is shown in Table 2.1 below.

	α	p	x	y	z
Jan	0.040	1	•	•	
Feb	0.040	1	•	•	
Mar	0.040	1	•	•	
Apr	0.042	1	•	•	
May	0.040	1	•	•	
Jun	0.047	1	•	•	
Jul	0.047	1	•	•	
Aug	0.040	1	•	•	
Sep	0.040	1	•	•	
Oct	0.042	0	•	•	
Nov	0.041	0	•	•	
Dec	0.041	1	•	•	

Table 2.1: α, p, and best subset of covariates for LP model.

Table 2.1 summarizes the values of α, or the fraction of neighbors to include in the local model, p, or the order of the local polynomial, and the best subset of covariates as selected by GCV. The fraction of neighbors is rather small and the order is generally linear (October and

November yield order 0, or weighted moving average), which implies there is great predictive skill in estimating the unknown process locally. Also, x and y (easting and northing) are consistently selected as the best subset of covariates, while z (elevation) plays no role in the predictive skill of the LP model.

The satellite error estimates are obtained from the two models at each observed location and the satellite derived precipitation estimates are added to the error estimates to obtain the blended estimates of the true precipitation. Scatterplots of observed rainfall and the above blended estimates of the true rainfall for January and July are shown in Figures 2.3 and 2.4 – for comparison the scatterplot with the raw CHIRP is also shown. It can be seen that both blending methods significantly improve over the raw CHIRP, as expected Kriging estimates are close to exact at the observed points. Table 2.2 summarizes R^2 and RMSE for all months in 2009.

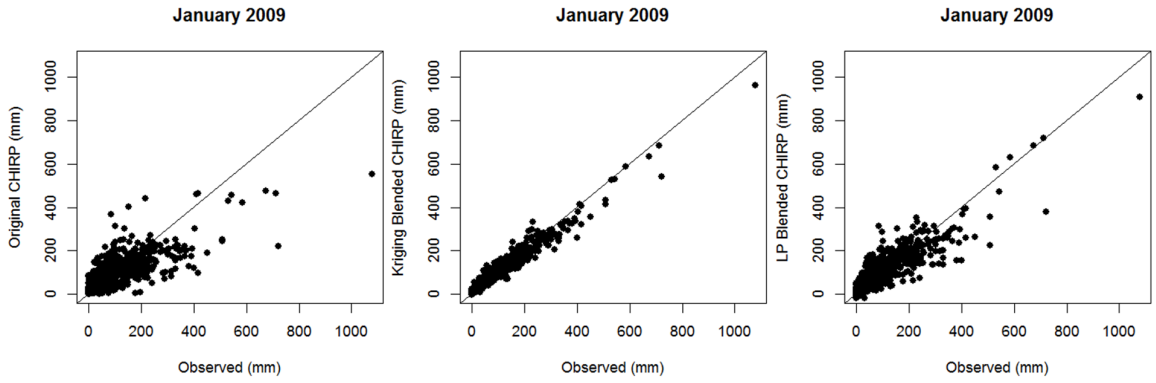

Figure 2.3: Original CHIRP estimates (left) and blended estimates from Co-kriging (middle) and LP (right) models for January 2009.

As can be seen from Table 2.2, the original CHIRP estimate has poor representativeness of the gauge measurements, with R^2 values ranging from 0.23 to 0.63; its RMSE values are quite large, with a range from 67 to 200 mm. The Co-kriging model tends to have a very low RMSE (ranging from 9.5 to 53.7 mm) and high R^2 (0.83 to 1.0), but some months show lower R^2 and higher RMSE which may be attributed to a large nugget effect, or less spatial structure in the modeled process. The LP model produces R^2 values with a range of 0.66 to 0.84; its RMSE values

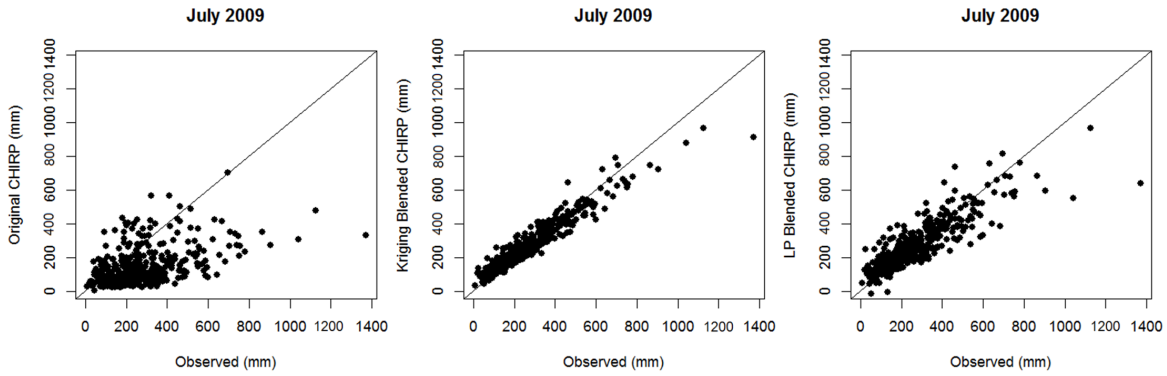

Figure 2.4: Same as Figure 2.3 but for July 2009

	CHIRP R^2	Cokrig R^2	LP R^2	CHIRP RMSE	Cokrig RMSE	LP RMSE
Jan	0.58	0.95	0.78	76.3	27.1	54.4
Feb	0.56	0.98	0.75	77.8	18.6	55.9
Mar	0.55	0.96	0.76	99.4	28.7	65.0
Apr	0.58	0.90	0.81	91.5	46.1	62.1
May	0.49	0.83	0.75	102.3	53.7	66.3
Jun	0.62	1.0	0.84	99.1	9.5	63.5
Jul	0.23	0.92	0.74	200.1	52.1	88.3
Aug	0.63	1.0	0.84	97.8	15.7	63.4
Sep	0.49	0.93	0.75	80.7	30.1	55.6
Oct	0.44	0.93	0.66	107.0	38.5	78.9
Nov	0.51	0.98	0.71	87.9	17.0	65.7
Dec	0.53	0.97	0.77	67.6	16.1	43.3

Table 2.2: Summary statistics reflecting model-fitting performance.

have a range of 44 to 90 mm, which is quite impressive. It follows that both models greatly improve the representativeness of the CHIRP product for all months in 2009.

Model Validation

To test the predictive capability of the models, cross-validations are performed. First, a drop-one cross-validation is performed, in which an observation and its covariates are dropped, the models are fitted on the rest of the data, and the dropped point is predicted using its covariates.

This process is then repeated for every observation. Figures 2.5 and 2.6 show the scatterplots of observed and cross-validated estimates for January and July, respectively, for the two methods; the scatterplots of raw CHIRP (same as in Figures 2.3 and 2.4 above) are also shown for comparison. Both the methods perform quite well in a cross-validated mode. Note that in July (the wet period) both the models perform much better than raw CHIRP. These finding are corroborated in Table 2.3 showing the R^2 and RMSE.

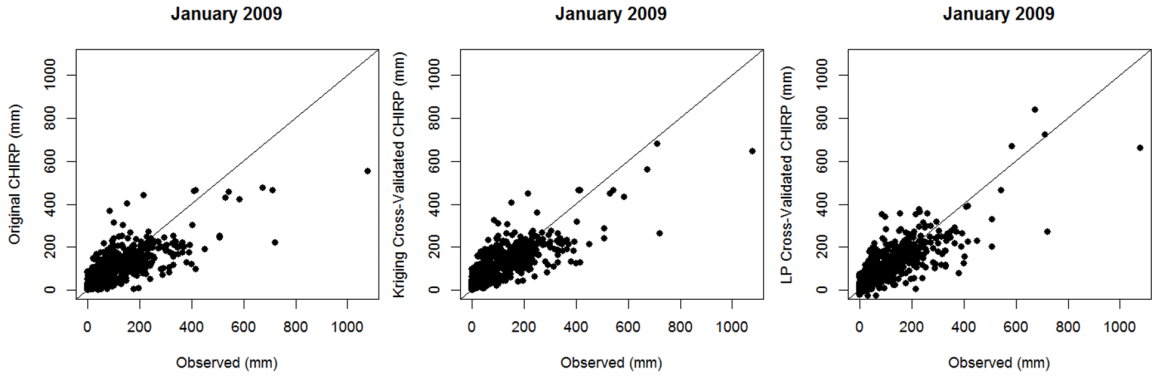

Figure 2.5: Original CHIRP estimates (left) and cross-validated estimates from Co-kriging (middle) and LP (right) models for January 2009.

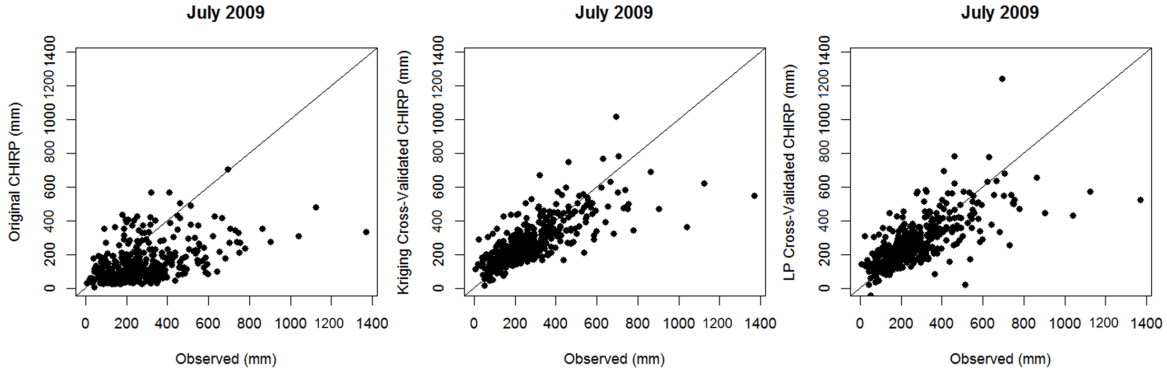

Figure 2.6: Same as Figure 2.5 but for July 2009.

The months where LP models show modest improvement in these measures are those where the original CHIRP product has good performance. Even in predictive mode, which genuinely

	CHIRP R^2	Cokrig R^2	LP R^2	CHIRP RMSE	Cokrig RMSE	LP RMSE
Jan	0.58	0.62	0.60	76.3	71.5	75.6
Feb	0.56	0.63	0.63	77.8	68.4	66.8
Mar	0.55	0.65	0.60	99.4	77.9	85.0
Apr	0.58	0.72	0.69	91.5	90.1	81.7
May	0.49	0.54	0.62	102.3	79.0	89.4
Jun	0.62	0.75	0.68	99.1	113.7	126.6
Jul	0.23	0.57	0.48	200.1	81.9	94.0
Aug	0.63	0.73	0.65	97.8	66.7	72.5
Sep	0.49	0.64	0.58	80.7	66.7	72.5
Oct	0.44	0.50	0.47	107.0	95.6	96.7
Nov	0.51	0.55	0.55	87.9	81.9	82.3
Dec	0.53	0.57	0.60	67.6	58.9	57.6

Table 2.3: Summary statistics reflecting drop-one cross-validated model performance.

stresses the models, both blending methods produce better results than the original CHIRP product.

In the second cross-validation procedure, 25% of the observations are dropped at random and are predicted with a model fitted to the remaining data – this is repeated 500 times. Three measures are computed: RMSE, MAE, and % Bias for each repetition and are displayed as boxplots in Figure 2.7 for July 2009. The original CHIRP has relatively poor performance, showing higher values for all these measures. Note that the median values of the measures calculated from drop-25% cross-validation are comparable to drop-one cross-validation. Drop-50% cross-validation was also performed to further stress the models and the results are shown in Figure 2.8.

For both models there is considerable decrease in RMSE and MAE error statistics for both drop-25% and -50% validation scenarios. Similarly, both models reduce the % Bias to nearly zero compared to the high bias in the CHIRP estimates. The validation performance of these blending methods for other months is similar, with varying magnitudes of improvement (figures not shown).

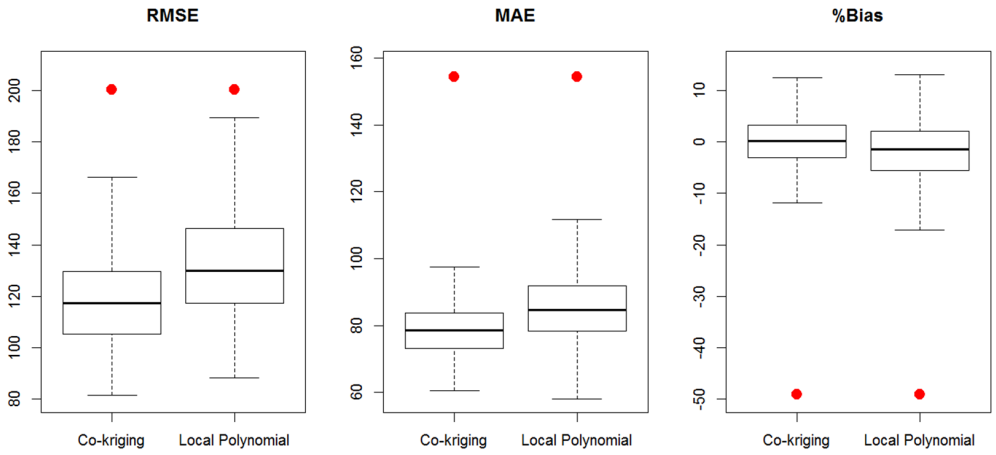

Figure 2.7: Prediction skill measures from drop-25% cross-validation for July 2009. Red dots correspond to estimates from original CHIRP.

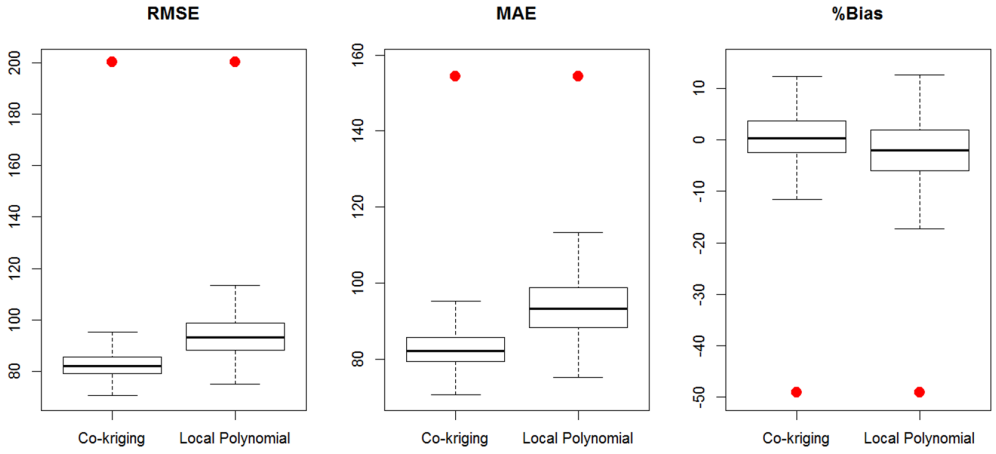

Figure 2.8: Same as 2.7 but from drop-50% cross-validation.

Spatial Estimation

The models were applied to obtain blended estimates of precipitation on the satellite grid – spatial maps of the estimates and predicted errors are shown along with the observed fields.

January 2009

Figure 2.9a shows the original CHIRP product for January 2009, with rain gauge measurements included on the same scale for comparison; Figures 2.9b and 2.9c show the estimates of satellite error using Co-kriging and LP models, respectively, with point values of satellite error on the same scale for comparison; Figures 2.9d and 2.9e show the blended CHIRP product from both models, with rain gauge measurements included on the same scale for comparison.

It can be seen in Figure 2.9a, the original CHIRP product for January 2009 does a relatively good job in capturing the spatial pattern and magnitude of precipitation events, due largely to this being a dry month. Problem areas include gauge measurements on the Caribbean coast of Honduras – where the CHIRP product both over- and underestimates the magnitude of precipitation – and along the Pacific coast of Colombia – where the CHIRP product underestimates the magnitude of precipitation events. With respect to the estimates of satellite error, the Co-kriging and LP models (Figures 2.9b and c) both indicate that the CHIRP product tends to overestimate along the Caribbean coast of Central America while underestimating along the Pacific coast of Colombia. Estimates of satellite error from the Co-kriging model match the magnitudes of observed satellite error very well at pixels surrounding observations, but in data sparse regions will reduce its estimate to the expected value of the process, which is very close to zero. The LP model also captures the magnitudes of observed satellite error very well at pixels surrounding observations, but extrapolates large estimates of error along the coast of Caribbean Nicaragua where data is sparse, resulting in estimates of satellite error that are greater than indicated by available data. Still, both the Co-kriging and LP models produce a smooth spatial map of satellite error, which is consistent with the hypothesis of a coherent spatial structure of the modeled process.

The blended estimates of precipitation from the Co-kriging model (Figure 2.9d) produce a

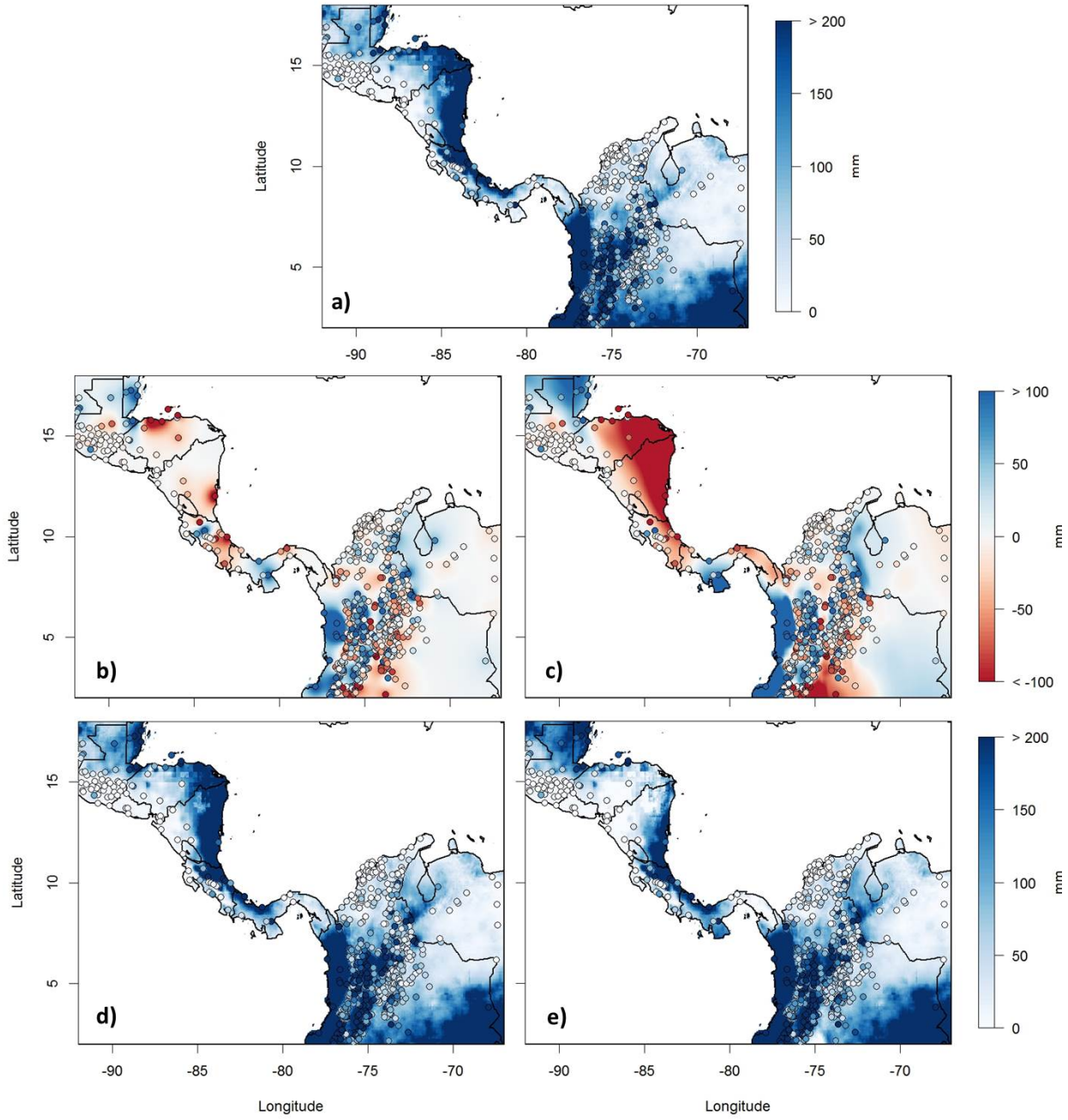

Figure 2.9: a) Original CHIRP estimates for January 2009, b-c) Estimated satellite error field from Co-kriging and LP models, d-e) Blended precipitation estimates from Co-kriging and LP models.

spatial pattern similar to the original CHIRP, due to its tendency to estimate the expected value in regions of sparse data. In the coastal Caribbean region, where data is sparse, the LP model

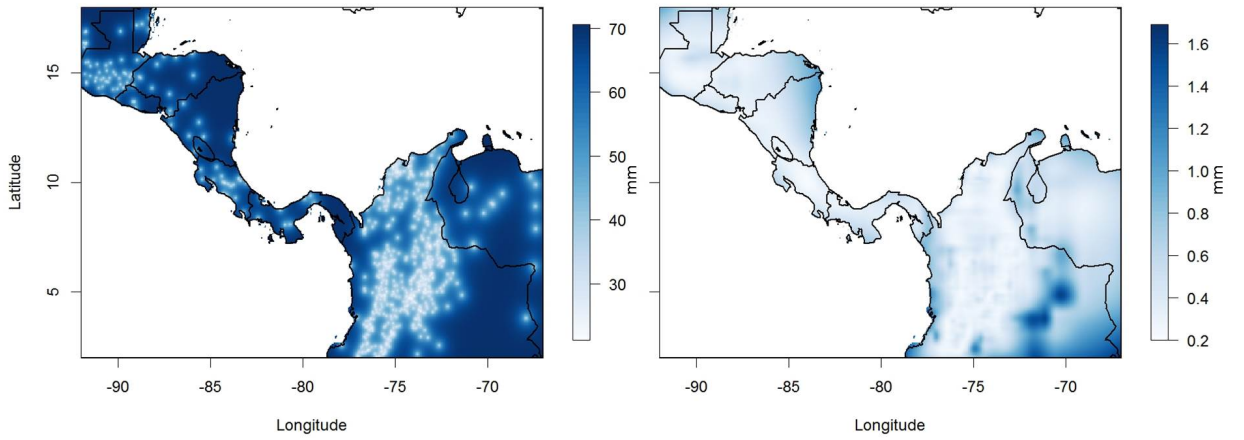

Figure 2.10: Estimates of standard error of the satellite error field for January 2009 from Co-kriging (left) and LP (right) models.

(Figure 2.9e) extrapolates the negative error to values greater than the observed error, resulting in large changes in the original CHIRP; the Co-kriging model estimates the satellite error to be close to the expected value of the process, resulting in little to no change from the original CHIRP. Note the difference between Figure 2.9d and e along the coastal Caribbean region; Co-kriging maintains the band of high rainfall consistent with coastal observations, while the LP model greatly reduces estimates in this coastal region thus interrupting the band of high rainfall. Both the Co-kriging and LP models reduced or increased rainfall to better match the observations in Colombia and Venezuela.

Standard errors of both the models are shown in Figure 2.10. The standard errors of predictions from the Co-kriging model are much lower in regions surrounding rain gauge observations. The LP model has much less uncertainty in its prediction process because it is based on local functional estimation which tends to have smaller variance (Loader 1999).

July 2009

Spatial maps for July 2009 are shown in Figure 2.11. The original CHIRP product for July 2009 (Figure 2.11a) does a relatively good job in capturing the spatial pattern, but fails in

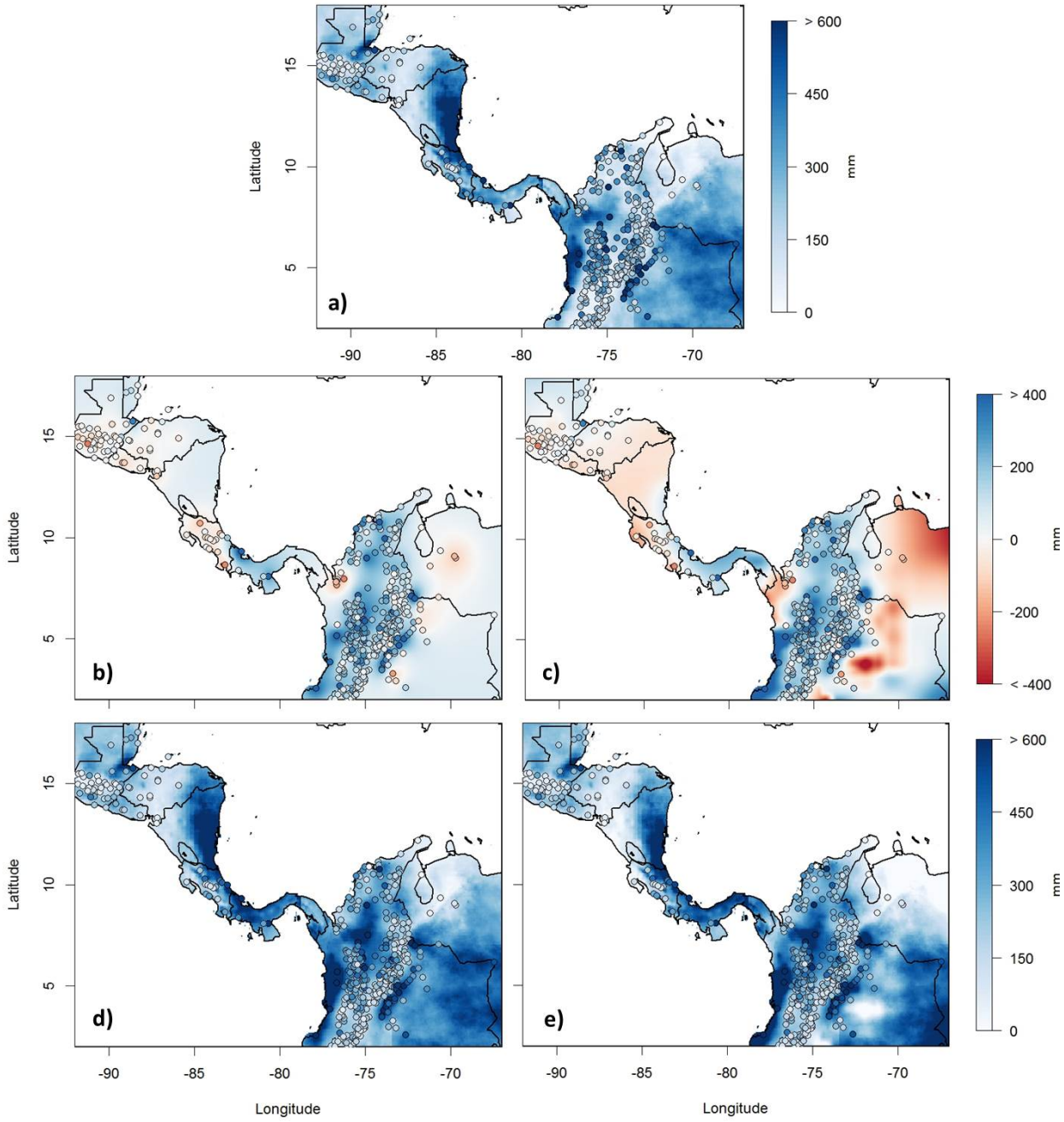

Figure 2.11: Same as Figure 2.9 but for July 2009.

representing the magnitude of precipitation events, which is due largely to this being a wet month as well as the inherent underestimation of satellite products during wet events. Problem areas include the Caribbean coast of Panama and the Pacific coast of Colombia – where the CHIRP

product underestimates the magnitude of precipitation events. Figure 2.11b and 2.11c show the interpolated satellite error fields for the Co-kriging and LP models, respectively. The Co-kriging estimates are consistent with many of the satellite error observations, yet at points furthest away from observations the estimate of error reduces to nearly zero (the expected value of the process. The LP model also estimates values consistent with the satellite error observations, but tends to extrapolate these large error values in data sparse regions, such as Caribbean Venezuela. Figure 2.11d and e shows the blended CHIRP product from Co-kriging and LP, respectively – it can be noted performance of the two models are comparable.

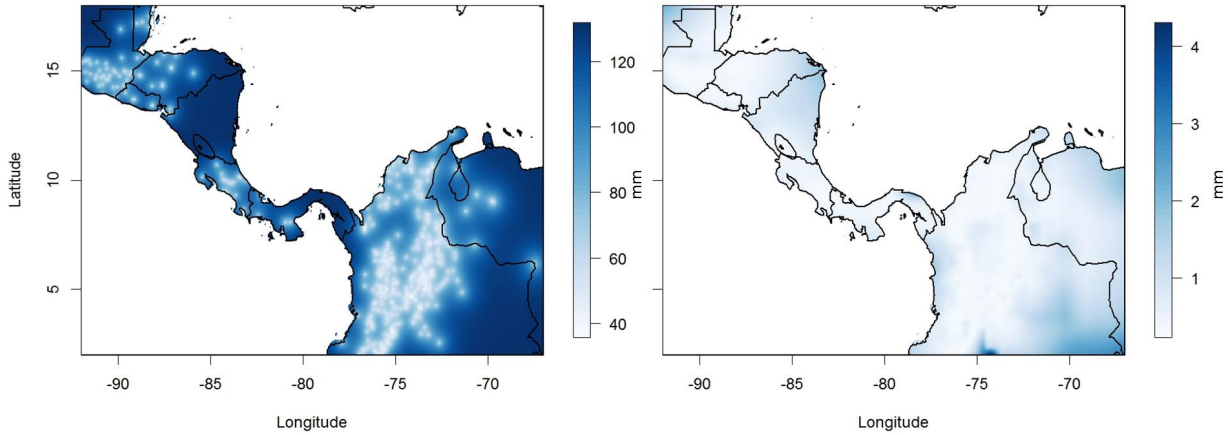

Figure 2.12: Same as Figure 2.10 but for July 2009.

Figure 2.12 shows the standard error of predictions for the Co-kriging and LP models, which is consistent with those for January 2009 (Figure 2.10).

2.4.2 Performance on Extreme Events

Good estimation of extreme events is important for hydroclimatic hazard mitigation, especially during wet months. The hazards associated with extreme events consist of floods, landslides, and agricultural overland flow, all of which are destructive to a nation's infrastructure. We define an extreme event at a location whose monthly rainfall exceeds the 85^{th} percentile of all rain gauge measurements for that month. The 85^{th} percentile of all gauged rainfall in July is 410

mm/month. The locations of these events and the majority of extreme precipitation events occur in the mountainous and coastal regions (Figure 2.13). Figure 2.14a shows the performance of the original CHIRP and 2.14b and c show the performance of the blended estimates from Co-kriging and LP models, respectively, for extreme events; Table 2.4 summarizes how blending improves performance of the satellite estimate with respect to extreme events. These statistics are computed from drop-one cross-validation output.

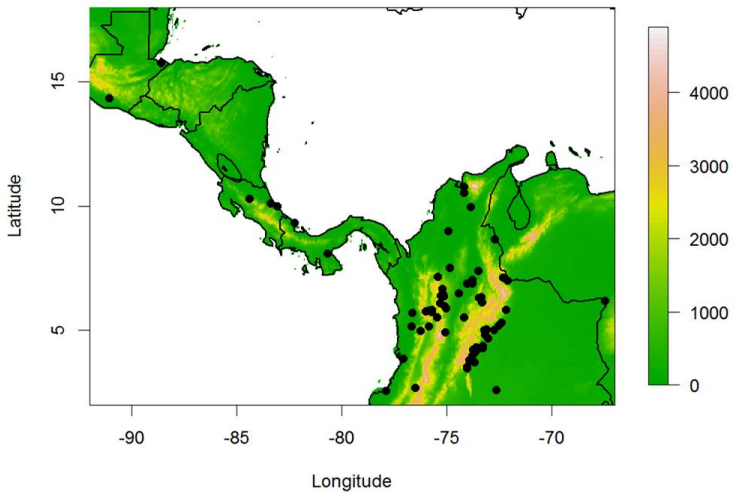

Figure 2.13: Locations of extreme precipitation events for July 2009 (black dots) and elevation shown in color.

	RMSE	MAE	% Bias
Original CHIRP	386 mm	344 mm	−58.8%
Co-kriging	227 mm	167 mm	−22.7%
Local Polynomial	257 mm	191 mm	−24.5%

Table 2.4: Summary statistics of blending performance on extreme events for July 2009.

As can be seen in Figure 2.14 and Table 2.4, both blending methods proposed in this chapter have a great impact on the CHIRP's performance with respect to extreme precipitation events. Validation statistics from Table 2.4 indicate that both blending methods reduce the error and inherent bias of the CHIRP satellite estimate. It should be noted that models for all months in

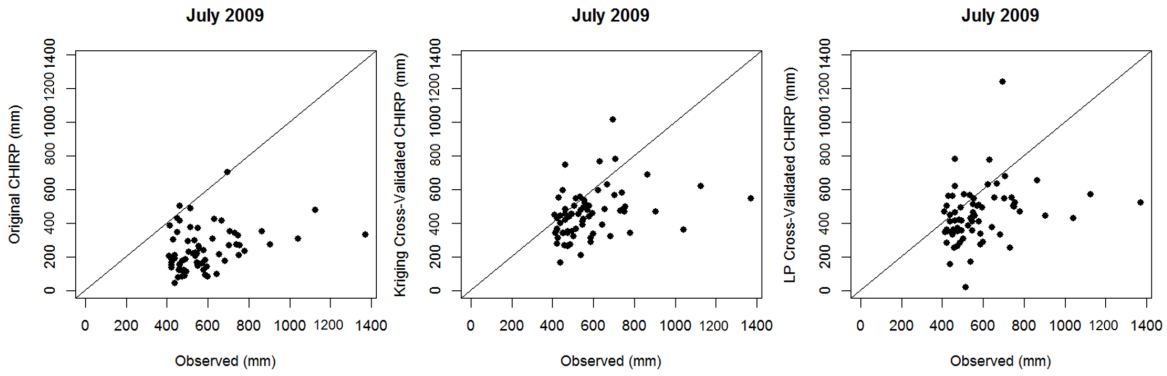

Figure 2.14: Original CHIRP product (left) and blended estimates from cross-validated Co-kriging (middle) and LP (right) models for extreme events in July 2009.

2009 – the test period – show similar skill with respect to extreme events, thus it is worthwhile to implement either of these blending methods on all available data to produce a more representative gridded time series of monthly precipitation.

2.4.3 Application to Pentad Rainfall Estimation

In this section, we investigate the utility of the two blending methods to shorter temporal scales, such as pentad rainfall. This is of interest for near-term blended precipitation estimates for natural hazard mitigation strategies. Blended estimates of precipitation on pentad time scales during the wet season can provide insight into the wetter regions of river basins, can be used to drive hydrologic models for modeling soil moisture and consequently, the potential for flooding and landslides.

The fourth pentad of August 2009 will be used in this application, as it is one of the wettest pentads of 2009, yielding a maximum observation of 385 mm where the CHIRP reports only 130 mm. The two blending methods as applied to the monthly precipitation described in previous sections are applied for this pentad data. Figure 2.15 shows scatterplots of the original CHIRP, Co-kriging, and LP models, which shows that both models fit the observations well and improve upon the CHIRP estimates. Due to the noisy spatial structure of rainfall at shorter time scales,

the model fits are less representative than those at monthly time scales.

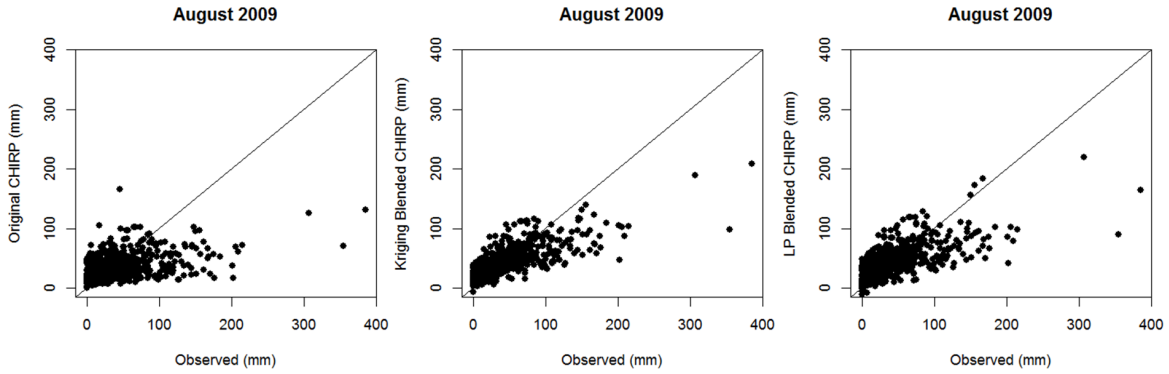

Figure 2.15: Same as Figure 2.3 but for the fourth pentad of August 2009.

As seen earlier, the CHIRP product overestimates regions of no rainfall while underestimating extreme precipitation. For this pentad, the maximum observed rainfall is 385 mm, though the CHIRP product reports only 130 mm of rain in the grid cell of this extreme event. Table 2.5 reports the relative performance of the original CHIRP, as well as the two blending methods.

	RMSE	MAE	% Bias
Original CHIRP	40 mm	26 mm	−21.4%
Co-kriging	31 mm	20 mm	0.0%
Local Polynomial	32 mm	21 mm	−0.8%

Table 2.5: Same as Table 2.2 but for the fourth pentad of August 2009.

Drop-one cross-validation analyses consistent with those of the monthly totals section were carried out on this representative wet pentad. These predictions are shown as scatterplots in Figure 2.16, and summary statistics are reported in Table 2.6.

Under drop-one cross-validation stresses, the error statistics of the two blending methods are comparable to that of the original CHIRP product. That being said, these blending methods eliminate the large inherent bias of the original CHIRP product even in predictive mode, which is encouraging of the model's predictive skill. Both drop-25% and drop-50% cross-validation tech-

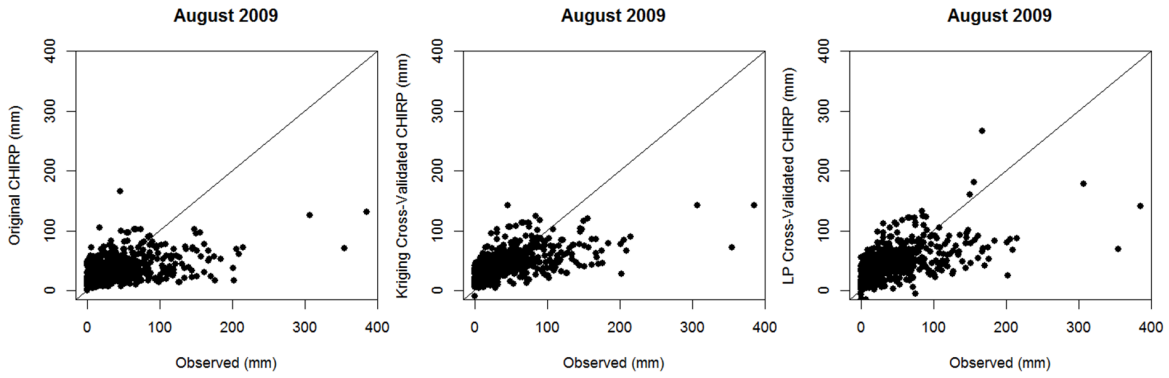

Figure 2.16: Same as Figure 2.5 but for the fourth pentad of August 2009.

	RMSE	MAE	% Bias
Original CHIRP	40 mm	26 mm	−21.4%
Co-kriging	37 mm	24 mm	0.1%
Local Polynomial	36 mm	24 mm	−0.7%

Table 2.6: Same as Table 2.3 but for the fourth pentad of August 2009.

niques were also applied to the fourth pentad of August 2009, again consistent with the monthly totals section, and statistics such as RMSE, MAE, and % Bias are shown in boxplot form in Figures 2.17 and 2.18. For completeness, the appropriate statistics for the original CHIRP product are shown as red points.

Figures 2.17 and 2.18 show that both the Co-kriging and LP models perform quite well under the predictive stresses of these validation measures, consistently reducing both error and bias with respect to the original CHIRP product.

The two blending methods were implemented in predictive mode to estimate the pentadal satellite error at the resolution of the CHIRP product, as in the monthly totals section. Spatial maps of the original CHIRP, model estimates of the satellite error, and the blended estimates from Co-kriging and LP models, respectively, are shown in Figure 2.19.

Problem areas include the Caribbean coast of Belize and eastern Guatemala, where the CHIRP product both over- and underestimates rain gauge measurements (Figure 2.19a). Similarly,

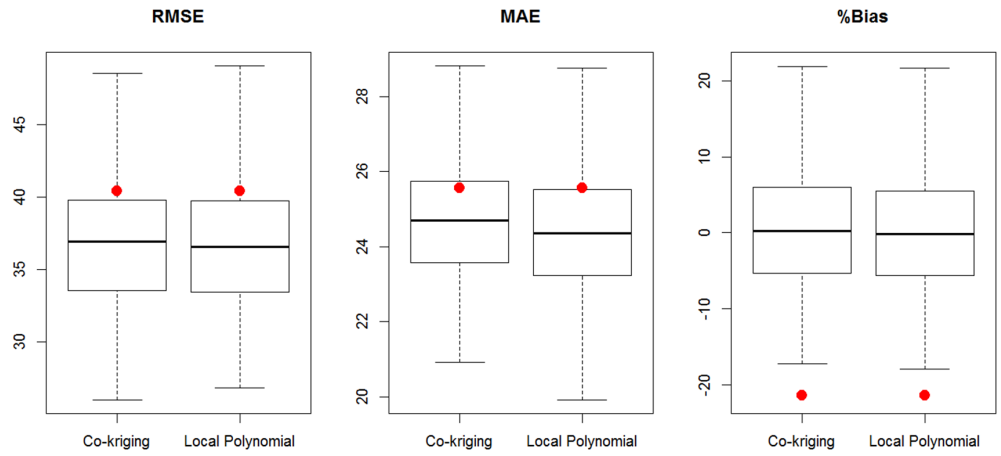

Figure 2.17: Same as Figure 2.7 but for the fourth pentad of August 2009.

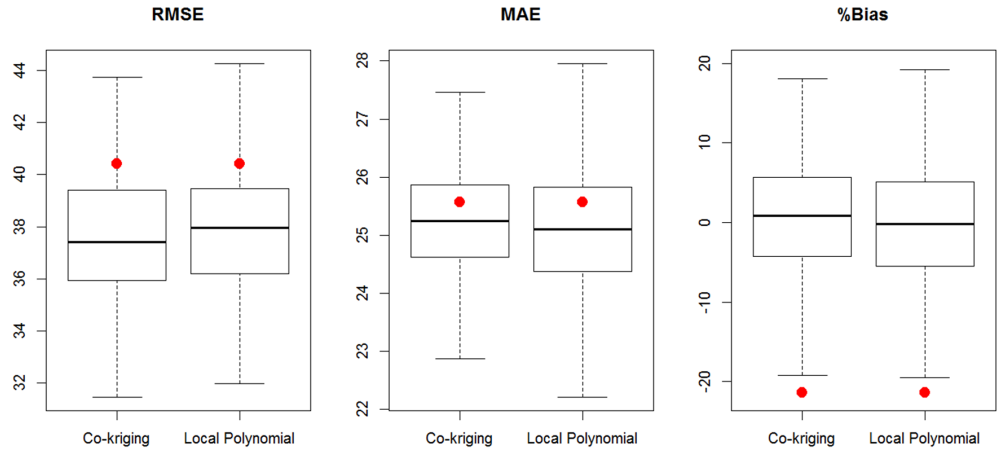

Figure 2.18: Same as Figure 2.8 but for the fourth pentad of August 2009.

Colombia is a problem area where the CHIRP tends to underestimate the magnitude of precipitation events on the Pacific coast and overestimates the magnitude of inland rainfall.

The Co-kriging model estimates relatively constant values of satellite error for the Caribbean coast of Belize, even though the observations are dissimilar, but captures the spatial structure of satellite error for Colombia (Figure 2.19b). The LP model produces a smooth spatial map of

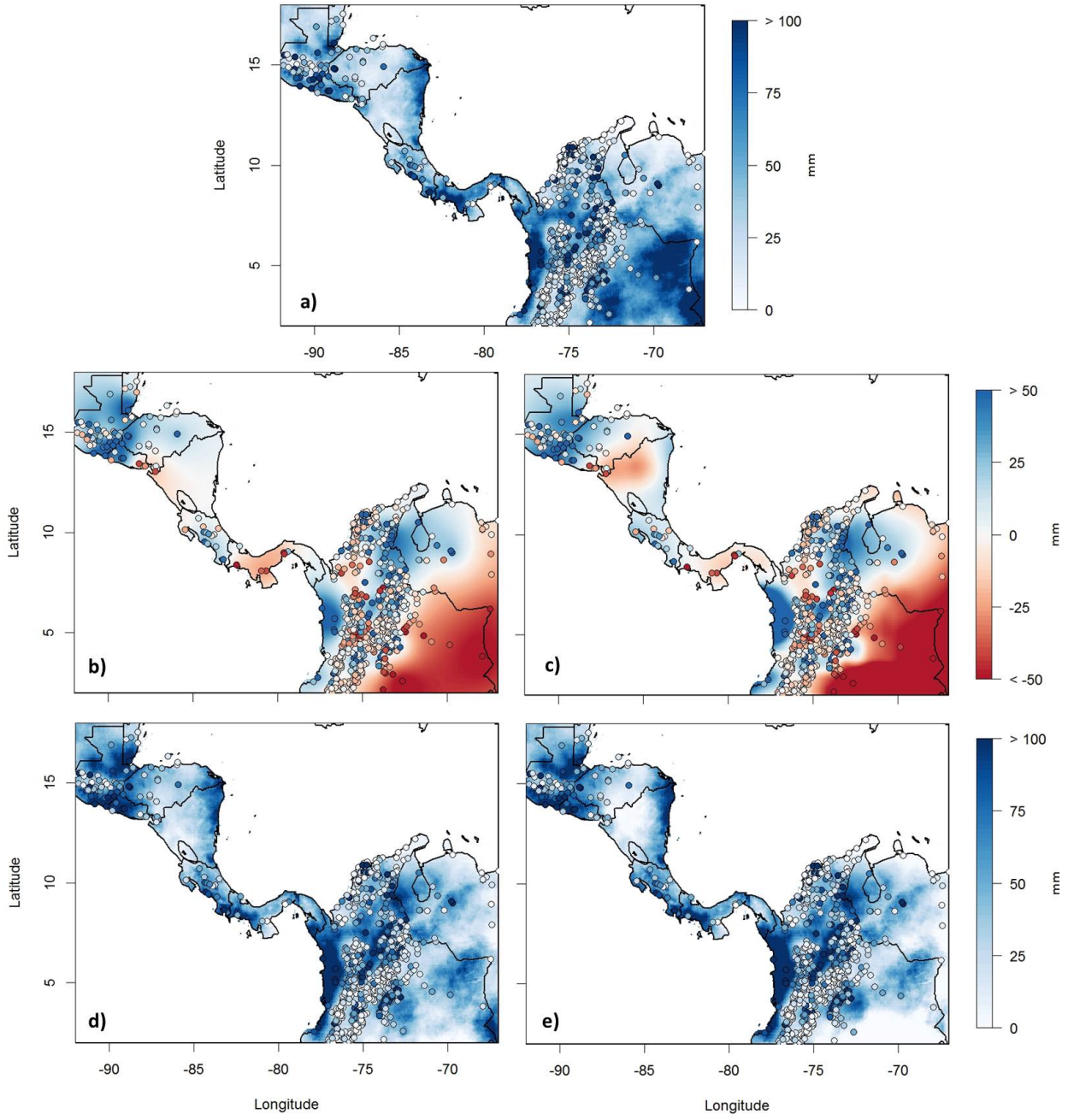

Figure 2.19: Same as Figure 2.9 but for the fourth pentad of August 2009.

estimated satellite error, consistent with the point values – this model captures both positive (Belize, Colombia) and negative (Colombia) satellite errors and better represents the sign and magnitude of predicted error along the Pacific coast of Colombia (Figure 2.19c). That being said, there seems

to be extrapolation error coupled with edge effects in the southeastern corner of the study region.

Figure 2.19d and 2.19e show the blended CHIRP products from the Co-kriging and LP models, respectively. The Co-kriging model produces a blended product similar in spatial distribution to the original CHIRP product, due to predictive issues discussed previously. The LP model tends to reduce or increase precipitation consistent with the problem regions described previously. However, the extrapolation and edge effects seen in mainland Colombia (Figure 2.19c) discussed in the previous paragraph have caused the blended product to underestimate where it previously overestimated.

Figure 2.20 shows the standard error of predictions for the Co-kriging and LP models. The spatial distribution of standard errors for the Co-kriging model is smallest at points included in the model-fitting process, which causes the sharp discontinuities present below. The LP model has largest standard errors along the edges of the domain, which is very likely due to extrapolation and edge effects.

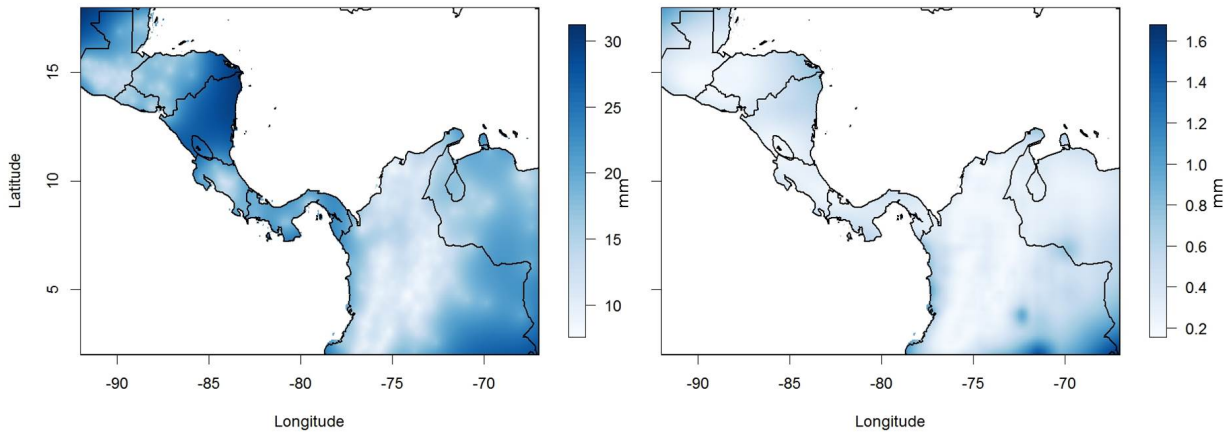

Figure 2.20: Same as Figure 2.10 but for the fourth pentad of August 2009.

To test the ability of these two blending methods in estimating extremes, as before we define an extreme event at a location where rainfall exceeds the 85^{th} percentile of all rain gauge measurements for the pentad – in this case 78 mm. The locations of these events and the majority of extreme precipitation events occur in the mountainous and coastal regions (Figure 2.21).

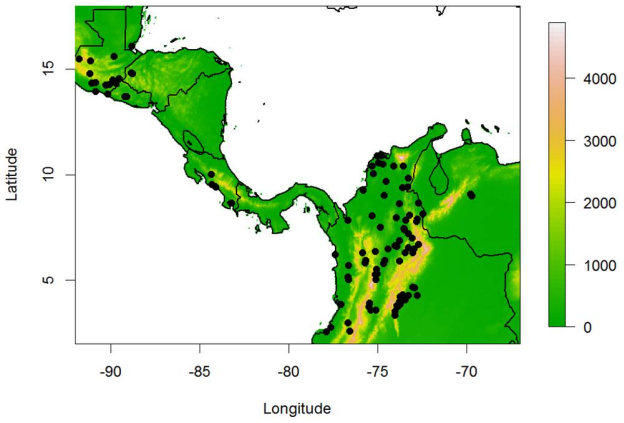

Figure 2.21: Same as Figure 2.13 but for the fourth pentad of August 2009.

Figure 2.22a shows the performance of the original CHIRP and 2.22b and 2.22c show the performance of the Co-kriging and LP models, respectively, with respect to extreme events. Table 2.7 summarizes how the two blending methods improve the performance of the satellite estimate with respect to extreme events. These statistics are computed from drop-one cross-validation output.

	RMSE	MAE	% Bias
Original CHIRP	91 mm	77 mm	−60.9%
Co-kriging	78 mm	63 mm	−48.2%
Local Polynomial	77 mm	61 mm	−44.1%

Table 2.7: Same as Table 2.4 but for the fourth pentad of August 2009.

As can be seen in Figure 2.22 and Table 2.7, the two blending methods proposed in this chapter are portable from monthly to pentadal scales. Both have positive impacts on the CHIRP's performance with respect to extreme events – validation statistics show that blending reduces the error and inherent bias of the CHIRP satellite estimate. In light of these results, it is encouraged to implement this blending technique for all available data in order to produce a more representative gridded time series of precipitation.

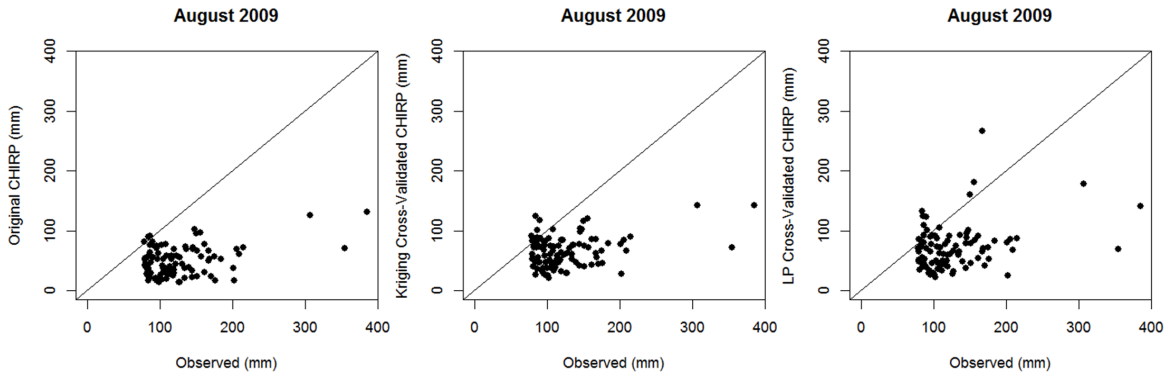

Figure 2.22: Same as Figure 2.14 but for the fourth pentad of August 2009.

2.5 Summary and Conclusions

Satellite-derived estimates of precipitation are becoming widely available and in near-real-time, which is a significant asset especially for data sparse regions. While these estimates provide a window into the spatial distribution and extent of precipitation, they consistently underestimate rain gauge measurements, thus performing poorly in the cases of extreme precipitation events. Rain gauge measurements capture the magnitude of precipitation events very well, but are sparse in their spatial coverage, more so in mountainous regions. Thus, there is a clear need to blend information from these two sources to provide robust estimates of the underlying precipitation field and its uncertainty. To this end, we offered two blending methods – Co-kriging and Local Polynomial – and demonstrated their capabilities by applying them to precipitation data from the mountainous region of Central America.

In these methods, the difference between gauge measurements and the nearest satellite grid cell is calculated; the two methods are applied to this 'satellite error' process. Thus, at any desired location the estimate of satellite error is estimated from the models and the satellite derived precipitation estimate is summed with the estimate to obtain the blended precipitation estimate.

We found that both the methods substantially improve upon the satellite-derived estimates, especially during the wet season when the satellite-derived values consistently underestimate ex-

treme precipitation events. The models were also shown to perform very well in drop-25% and drop-50% cross-validation mode – the error statistics and inherent bias are generally much lower when compared to the satellite-derived estimates. The Co-kriging model, however, tends to predict the expected value of the modeled process at prediction locations (i.e. on the satellite grid), thus the blended estimate has spatial structure very similar to the satellite-derived estimates. The LP model captures the spatial variability very well – the local functional estimation procedure enables this performance. Furthermore, the blended estimates from the LP model incorporate the precipitation magnitudes of rain gauge observations in a smooth function. However, the LP model is much more sensitive to extrapolation and edge effects than the Co-kriging model, an ideal example of which was illustrated in this chapter.

These two methods of spatially modeling the satellite errors and subsequently summing with the original satellite-derived estimates offer unique alternatives to the blending problem. With these tools it is possible to retrospectively improve satellite-derived estimates of precipitation, thus providing a robust and representative gridded time series of precipitation over the period record, which is very useful for natural hazard mitigation and management.

Since the satellite errors are spatially modeled, both satellite and gauge precipitation datasets are required for blending. The main assumption of these two proposed methods is that the CHIRP products are unbiased estimators of the true underlying precipitation field and that only the satellite errors need to be spatially modeled. It is shown in this paper that CHIRP estimates are biased, which suggests that this bias needs to be incorporated into the blending process. Thus, a hierarchical approach seems more appropriate – where the observed precipitation is first modeled as a function of CHIRP and other covariates and the residual from this first level is spatially modeled. This addresses both of the issues described above. Hierarchical modeling can be performed in a Bayesian framework enabling to provide the posterior distribution of the parameters and that of the precipitation process at any desired location – consequently, the entire uncertainty of the parameters and estimates. Another extension to this approach includes using the entire time series of precipitation at each gauged location – in this, the first level of hierarchy mentioned above is

fitted at each location and the parameters and residuals of this are spatially modeled separately. This can enable short-term forecasting that will be of great use to decision makers and resource managers.

Chapter 3

A Bayesian Hierarchical Kriging Approach for Blending

3.1 Introduction

The Bayesian Hierarchical Model (BHM) is a robust approach to data analysis, which allows for the structured treatment of random variables that together describe a spatio-temporal process. Application of this process is not limited to any one discipline. Historically it has been used to assess the uncertainty associated with ecological, economic, biological, and marketing research (Borsuk et al. 2001, Wikle 2003, Ver Hoef & Frost 2003, Allenby & Rossi 1999, Abe 2009). In the context of this research, BHMs have been implemented for the modeling and quantification of precipitation extremes (Cooley et al. 2007, Sang & Gelfand 2008, Schliep et al. 2009, Cooley & Sain 2010), prediction (Berliner et al. 2000), error analyses in hydrologic modeling (Kuczera et al. 2006, Liu & Gupta 2007, Thyer et al. 2009), general circulation models (Tebaldi & Sansó 2009, Furrer et al. 2007) and model calibration techniques (Raftery et al. 2005, Kennedy & O'Hagan 2001). Although the Bayesian hierarchical approach is gaining prominence in a variety of hydrologic and climatological applications, its use in blending satellite and gauge measures as demonstrated here is unique.

Bayesian Kriging is a novel approach to kriging due to its treatment of the covariance structure as unknown. Instead of using traditional methods of parameter estimation such as least squares and maximum likelihood, a set of prior distributions for potential parameter values is defined and routinely updated, based on observed data, via Markov Chain Monte Carlo (MCMC) simulation, in turn producing a posterior distribution of parameter values. Over the years this method has

been applied for a variety of purposes, namely to attain not only a spatial structure but also the uncertainty of geophysical and ecological processes, environmental contamination fields, topography, and subsurface processes and characteristics, among others (Omre 1987, Handcock & Stein 1993, Cui, Stein, & Myers 1995, Chen & Hubbard 2001, Aelion et al. 2009, Le & Zidek 1992, Sahu & Mardia 2005).

The proposed Bayesian Kriging (BK) approach for blending satellite and ground observations of rainfall is a new and unique application of the BK methodology. The theory behind BK has been around for decades and is an acceptable form of interpolation in the scientific community. The hierarchical framework enables the quantification of all levels of uncertainty in model parameters and response variable estimation. The study region and data are described in the next section, followed by an in depth description of the BHM and BK framework, and finally the results section.

3.2 Study Region & Data

The details of the study region and data sets are described in greater detail in Chapter 1, an abstract of which is provided below.

3.2.1 Study Region

The study region covers the Central American region (see Figure 3.1). The dry season in the region extends from December through early May and the effects of both Pacific and Atlantic tropical depressions are felt during the wet season that extends from mid-May through November. A major contributor to the prolonged wet season is the inter-tropical convergence zone (ITCZ) – an asymmetric band of convection which encircles the globe. The position of the ITCZ is non-stationary due to the seasonal shift of the trade winds.

Another cause for intense spatial variability of rainfall within the study region is its complex geography, where the elevation ranges from sea level to nearly 5000 meters (as can be seen in Figure 3.1) with numerous mountain ranges spanning Central America and Colombia. Costa Rica and Guatemala have steep elevation gradients along the coasts, making them vulnerable to extreme

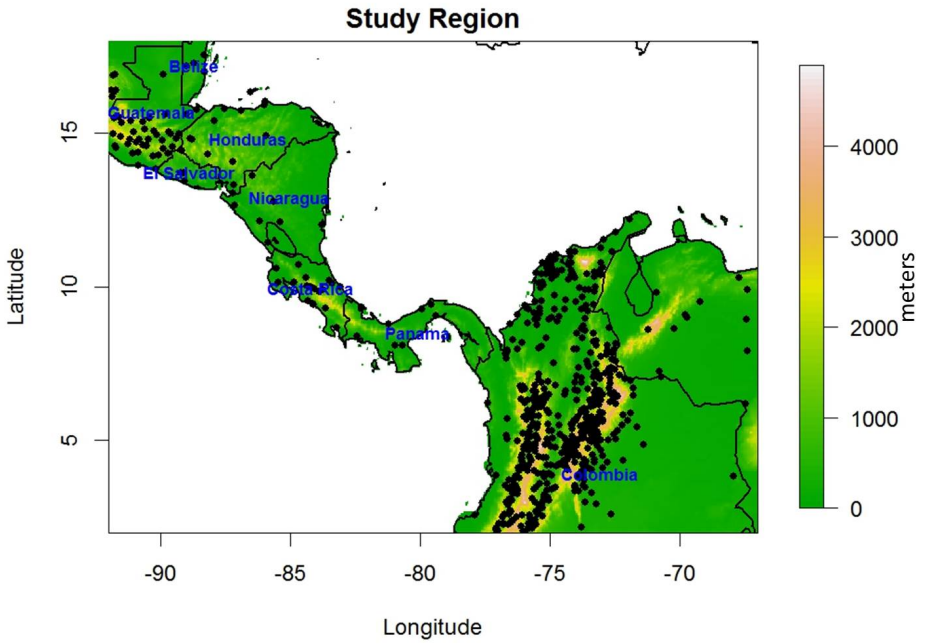

Figure 3.1: Study region geography, rain gauge measurement locations, and country names.

weather spawning from both Pacific and Atlantic tropical depressions. Coastal mountain ranges cause an abrupt rising of warm, wet air. As this tropical air rises, it cools rather quickly, thus releasing the moisture from the air as precipitation, a major cause for the variability in extreme precipitation events in the region.

Given the spatial and temporal variability of rainfall in this region coupled with vulnerable socio-economic conditions, droughts and floods can have a debilitating impact on these societies. It has been observed that a single extreme wet precipitation event – whether hurricane, tropical storm, or prolonged precipitation – may result in flooding, landslides, and potentially millions of dollars in damages, thus setting the national infrastructure back many years. Conversely, a prolonged dry spell may be equally devastating to a growing nation, causing widespread crop failures. Thus, a comprehensive understanding of rainfall variability is of critical importance for resources planning and management.

3.2.2 Data

Two precipitation datasets are utilized in this research – gauge measurements and satellite-derived estimates. Rain gauge measurements are point values that tend to be clustered in populous and lower elevation areas and sparsely scattered throughout the domain; the satellite-derived estimates are high-resolution – 0.05°x0.05° (approx. 5 km) – areal averages for five day periods (i.e. pentads). The satellite-derived precipitation estimate used in this research is the CHIRP (Climate Hazards group InfraRed Precipitation) covering the first pentad (five-day sum) of January, 1981 through the last pentad of December, 2012. There are six pentads per month, regardless of how many days are in a given month. For example, the sixth pentad of October will consistently be the sum of the last six days, October 26-31. The sixth pentad of February, however, will be the sum of either three or four days, depending on leap year status. Pentads may also be summed to obtain monthly totals of precipitation. It is widely accepted that satellite estimates of rainfall are better at capturing the spatial structure of precipitation, but fail in estimating magnitudes, especially of extreme events – this is because the satellite estimates are derived estimates of precipitation (Xie & Arkin 1997).

The CHIRP estimate is produced using the CHPCLIM (Climate Hazards group Precipitation Climatology), a completely unbounded global long term pentadal climatology of precipitation of the same resolution (0.05°x0.05°), as well as the infrared radiation (IR) satellite product. The CHPCLIM for the appropriate pentad is multiplied by the most recent IR percent-anomaly to produce the CHIRP estimate, shown below:

$$CHIRP_{i,j,k} = \%IR_{i,j,k} \cdot CHPCLIM_k \tag{3.1}$$

In the above equation, $CHIRP_{i,j,k}$ represents the satellite-derived precipitation estimate for pentad k of month j in year i; $\%IR_{i,j,k}$ represents the infrared radiation satellite estimate percent-anomaly for pentad k of month j in year i; and $CHPCLIM_k$ represents the precipitation climatology for pentad k.

The station data (shown in Figure 3.1) was collected from partner sources of the UCSB

Climate Hazards Group; quality control and pentadal-aggregation techniques were implemented by the CHG to produce time series data of the same temporal extent as the available CHIRP products (the first pentad of January, 1981 through the last pentad of December, 2012).

3.3 Proposed Blending Method

Previous efforts of merging these two datasets were based on the assumption that the CHIRP satellite product is an unbiased estimator of the station observations. Results from Chapter 2 show that this is not the case in that the CHIRP product has substantial bias in estimating observed rainfall – especially for higher rainfall amounts. To address this non-negligible bias, a Bayesian Hierarchical Model (BHM) approach is proposed.

3.3.1 Bayesian Hierarchical Model – Bayesian Kriging

Before describing the blending method the following variables are defined: $Y_{sat}(s)$ and $Y_{obs}(s)$ are the satellite estimates and observed precipitation, respectively, at locations $s = 1, ..., N_s$, where N_s is the number of stations with non-missing data for a given pentad or month. We define $X_1(s)$ and $X_2(s)$ as the covariates that are related to the observed precipitation, where $X_1(s) = Y_{sat}(s)$ and $X_2(s)$ represents the elevation at location s; any number of other covariates may be incorporated. In the hierarchical approach, the observations are assumed to be a realization of a distribution, the mean estimates of the distribution are modeled as a linear function of covariates, and lastly, the residuals from this are modeled as a spatial process. Such a hierarchical approach has been proposed for other applications (Omre 1987, Handcock & Stein 1993, Cui & Myers 1995, Chen & Hubbard 2001, Aelion et al. 2009, Le & Zidek 1992, Sahu & Mardia 2005). The BHM is defined as follows:

$$Y_{obs} \sim N(\mu, s^2)$$

$$\mu = \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

$$x = \text{ easting, } y = \text{ northing (kilometers from Prime Meridian; Equator)}$$

$$\beta_1, \beta_2 \sim flat$$

$$\sigma^2 \sim Gam^{-1}(100, 100)$$

$$\phi \sim Unif(a, b)$$

$$\tau^2 = 0 \tag{3.2}$$

The spatial model for ϵ is a zero-mean Gaussian process model with exponential variogram function with parameters σ^2 and ϕ; σ^2 is the marginal variance of the Gaussian process and ϕ is the effective range (or the distance at which the covariance reaches plateau). The nugget effect (τ^2) parameter is set to zero due to its negligible value for all months – a zero nugget also forces the kriging model to be an exact estimator of points included in the fit. We then define θ as a vector of all parameters excluding the nugget effect,

$$\theta = \{\beta_1, \beta_2, \sigma^2, \phi\} \tag{3.3}$$

The parameters are estimated in a Bayesian framework, wherein we seek the posterior distribution of θ – i.e., the conditional distribution of θ given the data, $p(\theta|Y_{obs})$, which can be expressed using Bayes' rule as follows:

$$p(\theta|Y_{obs}) = \frac{p(\theta, Y_{obs})}{p(Y_{obs})} = \frac{p(Y_{obs}|\theta)p(\theta)}{p(Y_{obs})} \propto p(Y_{obs}|\theta)p(\theta) \tag{3.4}$$

where $p(Y_{obs}|\theta)$ is the likelihood function for the observed data given the parameters and $p(\theta)$ is the prior joint distribution of the parameters. Assuming independence of the parameters, which is not unreasonable, the prior joint distribution of parameters is a product of the prior distribution of each parameter, and the posterior density simplifies to

$$p(Y_{obs}|\theta) \propto p(Y_{obs}|\theta) \prod_i p(\theta_i) \tag{3.5}$$

Because we have no a priori information regarding values of the β parameters, flat (uninformative) prior distributions are used (Finley et al. 2009). The prior distribution of the marginal variability (σ^2) is the process is assumed to follow an Inverse-Gamma distribution with shape and scale parameters defined as $\alpha = \beta = 100$, respectively (Finley et al. 2009). The prior distribution of the effective range (ϕ) of the variogram is uniform within a range. The range is obtained as follows: a linear model is fit to the data of the form $Y_{obs} = \beta_1 Y_{sat} + \beta_2 elev + \epsilon$, and an empirical exponential variogram, γ, is fitted to the residuals ϵ, of the form

$$\gamma(h) = \frac{1}{2} Var[\epsilon(s_i) - \epsilon(s_j)] \tag{3.6}$$

where $h = ||s_i - s_j||$ is the Euclidean distance between observation locations s_i and s_j. The effective range of this empirical variogram (ϕ) is used as a midpoint for the uniform prior distribution defined above, with upper and lower bounds defined as $+75\%$ and -75% of the midpoint value, such that $a = 0.25\phi$ and $b = 1.75\phi$. The range parameter theoretically has a finite range and this approach provides a wide yet realistic region in which the posterior distribution will converge.

The posterior distributions of the parameters are obtained using Markov Chain Monte Carlo (MCMC) simulations (Gallagher et al. 2009), the standard method for parameter estimation in a Bayesian estimation framework. The MCMC routine samples parameter values from their respective prior distributions, the likelihood function corresponding to this parameter set is calculated, a new set of parameter values is sampled and they are either accepted or rejected if the new scores of the likelihood function are an improvement from the current scores of the likelihood function with a certain probability depending on the proposal distribution. To obtain robust posterior distributions of these parameters, 15,000 MCMC simulations are carried out, thus yielding 15,000 values of θ. However, due to a necessary 'burn-in' time, the first 5,000 values are thrown out (Finley et al. 2009). The mean or median values of the simulations can be obtained if point estimates are desired.

For each accepted set of parameter values produced by the MCMC routine, an exponential variogram is fitted to the residuals of the traditional linear model with current parameter estimates.

This exponential variogram is used to estimate values of the residuals in an ordinary kriging framework with location parameters (easting, northing) as covariates. Because the nugget effect is set to zero, the ordinary kriging model is an exact estimator (i.e. interpolator) of these residuals, and the estimated values of the residuals match the known residuals perfectly. The estimated residuals are then summed with the mean estimate, thus reproducing the observed rainfall at each location.

In a predictive mode 10,000 estimates of the residual term ϵ and the mean estimate, and consequently the posterior distribution of Y_{obs} are obtained, thus producing a blended CHIRP distribution. With the posterior predictive PDF of the estimated true rainfall process at each grid point, it is possible to create spatial maps of the mean, median, and robust confidence intervals (i.e. 5^{th} and 95^{th} percentiles) of predicted rainfall.

BHMs have been gaining popularity in recent years, largely due to the exponential growth in computational power of affordable computers. This increase in computational power coupled with open source software packages such as R has brought this prolific theory to the foreground of statistical analysis and applications. In this research the 'spBayes' library in R was used, as it is capable of fitting multivariate spatio-temporal models using MCMC simulation.

3.4 Results

Analysis of the results is broken into the following categories: posterior distributions of the parameters, median estimates from the posteriors – i.e. fitting, prediction verification, estimated maps, and performance on extreme events. Analysis was performed on several months but we focus on January and July, 2009 – representative dry and wet months, respectively.

3.4.1 Posterior Distributions of Model Parameters

The posterior distributions of the four model parameters for January and July are shown in Figures 3.2 and 3.3, respectively – the distribution characteristics offer insight into their importance.

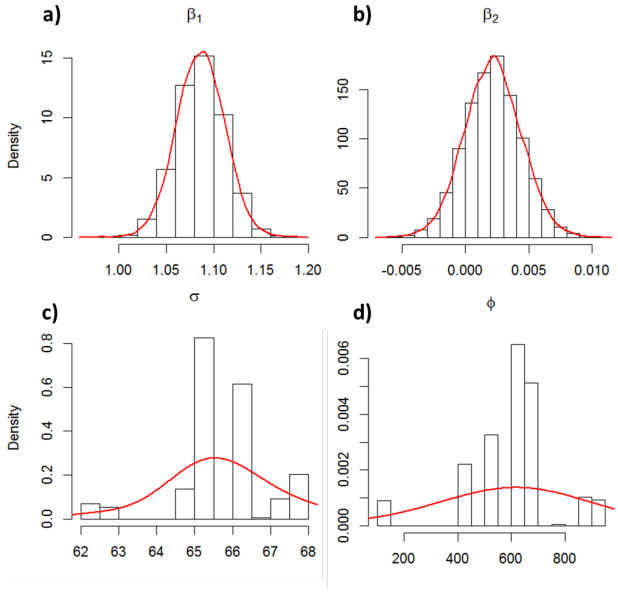

Figure 3.2: Posterior distributions for January 2009; a-b) regression coefficients, c-d) kriging parameters.

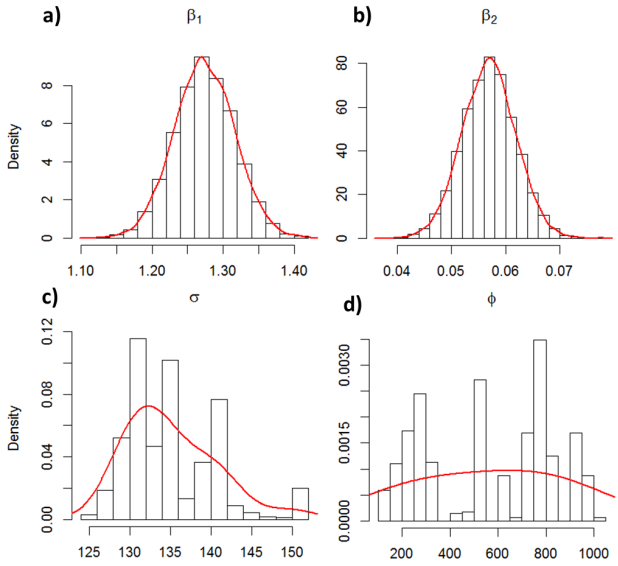

Figure 3.3: Same as Figure 3.2 but for July 2009.

For January the distribution of the CHIRP coefficient (β_1) is centered on 1.09, which implies consistent underestimation of rainfall events from the CHIRP – also, for July the distribution is

centered on 1.27, indicating substantial underestimation during this wet month. With regards to elevation, the posterior distribution of β_2 in January is centered on 0.002, but spans 0, while in July the distribution does not encompass 0 – indicating that in January the role of topography is minimal, which is to be expected given it is a dry period, but for July, during the wet period, topography plays a significant role. In January the small value of β_2 can still adjust the rainfall estimate by up to 10mm, while in July elevation has a much greater effect. The posterior distribution of the effective range ϕ for both the months has a relatively normal shape, implying a temporally-consistent spatial structure of the error term ϵ – the mean, median, and mode values of ϕ are very close to each other (approx. 630 km for January and 505 km for July), and centered on the mean value of the initial estimate that was the basis for the prior distribution. The marginal standard deviation σ (simply the square-root of the marginal variance σ^2) also has a quasi-normal shape, and the range of the distribution is relatively small (62-68 mm for January and 125-150 mm for July – slightly wider for the wet month), implying that the uncertainty in the spatial structure is consistent throughout all Bayesian simulations.

It is apparent the CHIRP product has variable performance depending on the season; during the dry season the CHIRP is close to being an unbiased estimator of the observed, but in the wet season it is substantially biased. The true rainfall process is also more dependent on elevation during the wet season – at higher elevations there tends to be more rainfall.

3.4.2 Model Fitting

The median values of the posterior distributions of the parameters were used to obtain estimates of precipitation. To illustrate the performance, Figure 3.4 shows the empirical and theoretical variograms of the residual process ϵ for four representative months in 2009 (January, April, July, and October). The theoretical variogram matches the empirical variogram very closely, which indicates a good fit.

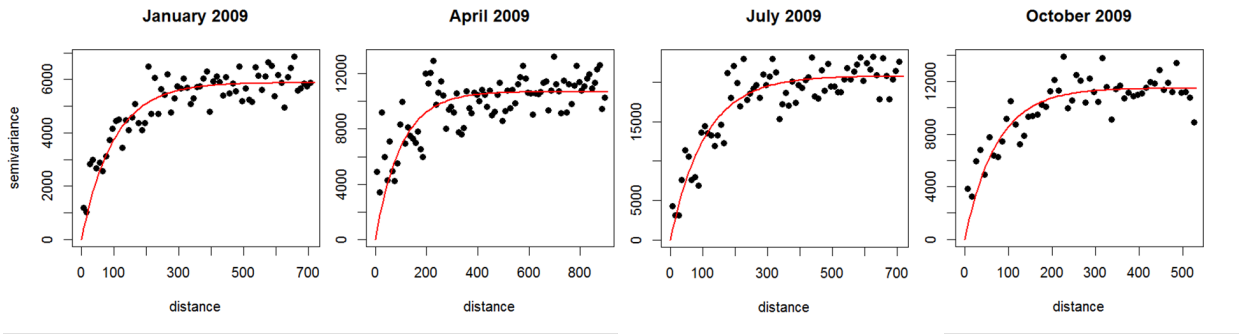

Figure 3.4: Median Bayesian Model empirical (black) and theoretical (red) variograms for Jan, Apr, Jul, and Oct 2009.

3.4.3 Model Validation

In order to test the predictive capability of the model, drop-one cross-validation – wherein one observation and its covariates are dropped from the hierarchical model fitting process and these covariates are then used to predict the dropped observation – was carried out for the test period. Because the observation and its covariates were not included in the model, its estimate may be treated as a prediction. This process is repeated until all observations have been dropped and estimated, thus producing a vector of predictions which may be compared to the known observations. Scatterplots are shown and summary statistics are reported for this validation measure. Another more intense validation method was carried out by randomly dropping 25% of the observations, fitting the hierarchical model on the remaining data, and predicting at the dropped points. Three skill measures, root mean square error (RMSE), mean absolute error (MAE), and percent bias (% Bias) were computed. To obtain the range of variability in the skill of this model, this was repeated 100 times and the skill measures computed for each repetition are shown as boxplots. For comparison, the corresponding measures of the original CHIRP products are shown as red dots. To further stress the predictive capability of the model, the aforementioned validation technique was repeated with 50% of the observations being dropped. As mentioned earlier, the median estimate of the Bayesian Hierarchical Model is an exact estimator at the observation points, so the

non-cross-validated skill measures are not shown. Figure 3.5 shows the drop-one cross-validation results for January and July 2009 as scatterplots; Figures 3.6 and 3.7 show the boxplots of skill measures from the drop-25% and drop-50% cross-validation procedures, respectively, for July 2009. Note the validation statistics are similar for all months in 2009.

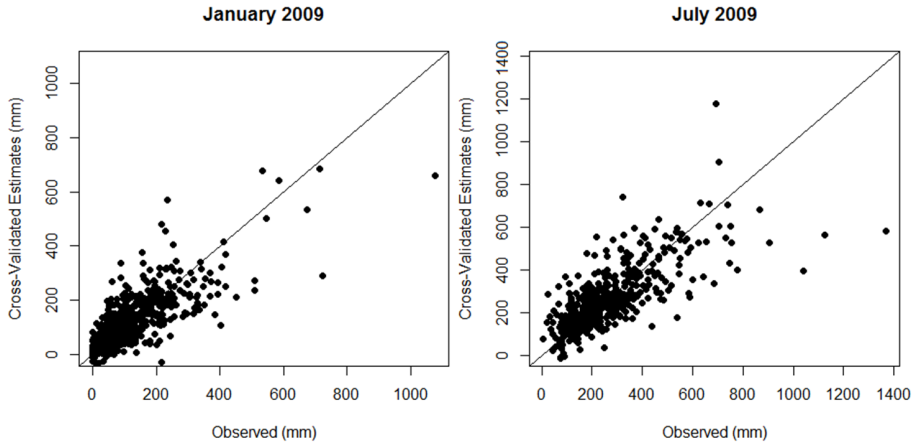

Figure 3.5: Drop-one cross-validation performance for January and July 2009.

For January 2009 or any typical dry month (Figure 3.5a), where the original CHIRP product is already a good estimator of precipitation totals, drop-one cross-validation produces results that completely eliminate bias and slightly reduce the error of observations. RMSE for drop-one cross-validation is 70 mm, while that of the original CHIRP product is 76 mm. Similarly, MAE is 45 mm for drop-one cross-validation, and the original CHIRP product has MAE of 50 mm. The resulting bias from drop-one cross-validation, however, is only 0.3% while the original CHIRP product has inherent bias of -11.1%. Conversely, July 2009 or any typical wet month (Figure 3.5b), where the original CHIRP product fails in estimating precipitation totals (especially of extremes), drop-one cross-validation produces results that reduce both bias and error that are inherent in the satellite estimate. RMSE for drop-one cross-validation is 118 mm, while that of the original CHIRP product is 200 mm. Similarly, MAE is 79 mm for drop-one cross-validation while it is 154 mm for the original CHIRP product. The resulting bias from drop-one cross-validation is 0.0%, diminished greatly from

−49%, the bias of the original CHIRP product.

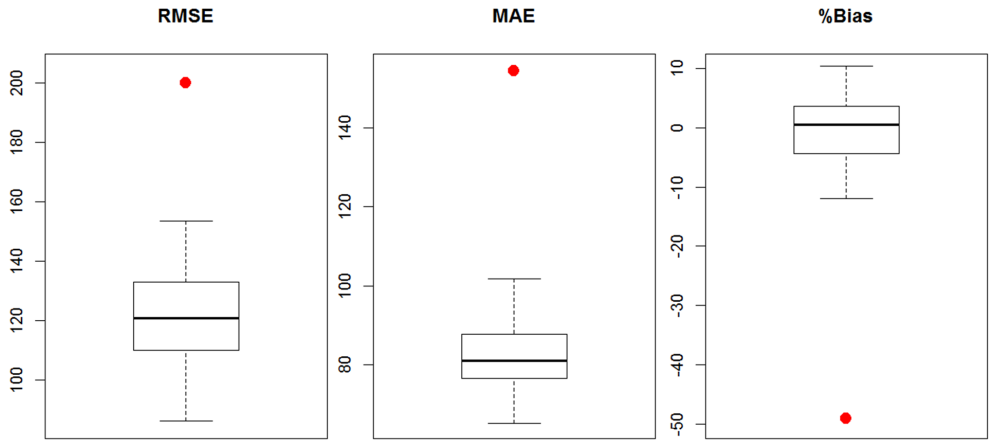

Figure 3.6: Drop-25% validation for July 2009, a) RMSE, b) MAE, c)% Bias (original CHIRP shown in red).

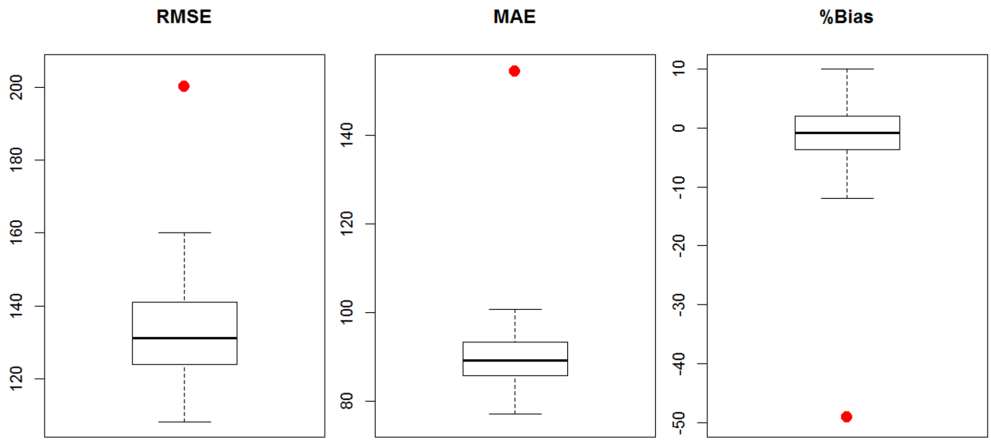

Figure 3.7: Same as figure 3.6 for from drop-50% validation.

Both intensive validation scenarios show RMSE and MAE values that are much lower compared to the original CHIRP, significantly reducing both bias and error, thus emphasizing the great

improvement achieved from blending satellite and gauge observations of precipitation. Both RMSE and MAE were reduced by at least 30%, which is very impressive, but the most impressive result is the reduction in bias from almost -50% with the original CHIRP to nearly zero for both validation scenarios. While no bias is to be expected when estimating the points included in the model-fitting process, the lack of bias in predictive mode is encouraging that this model is a robust estimator of the true rainfall process at unknown locations. Similar results were obtained for other months (figures not shown).

3.4.4 Model Predictions from Blending

The fitted model was used to make blended predictions on the satellite grid. As mentioned above, 15,000 estimates of the posterior distribution of the parameters are used to obtain 15,000 estimates of the posterior predictive distribution of rainfall at each grid point with the first 5,000 discarded as 'burn-in.' Therefore the median is computed at each grid location and shown as a spatial map – similarly, a robust confidence interval (5^{th} and 95^{th} percentile) is computed and shown – for January and July 2009. For ease of visualization, the extreme values (defined as exceeding the 85^{th} percentile of all gauge measurements for a given month) are plotted as the same deep blue color.

January 2009

Figure 3.8a shows the original CHIRP, and Figure 3.8b shows the blended estimate produced by the median values of the posterior distribution of the model parameters; the location of observations are shown as circles and their precipitation values are on the same color scale. Overall, the median blended estimates and the original CHIRP are similar for January – the greatest change occurring along the Caribbean coast of Honduras and Nicaragua, where the CHIRP overestimates the magnitude of precipitation by at least 100 mm. It appears that the area of increased rainfall is nearly equal to the area of decreased rainfall, which is consistent with the posterior distributions of the β parameters. Figure 3.8c shows the change in CHIRP due to blending (i.e. Difference =

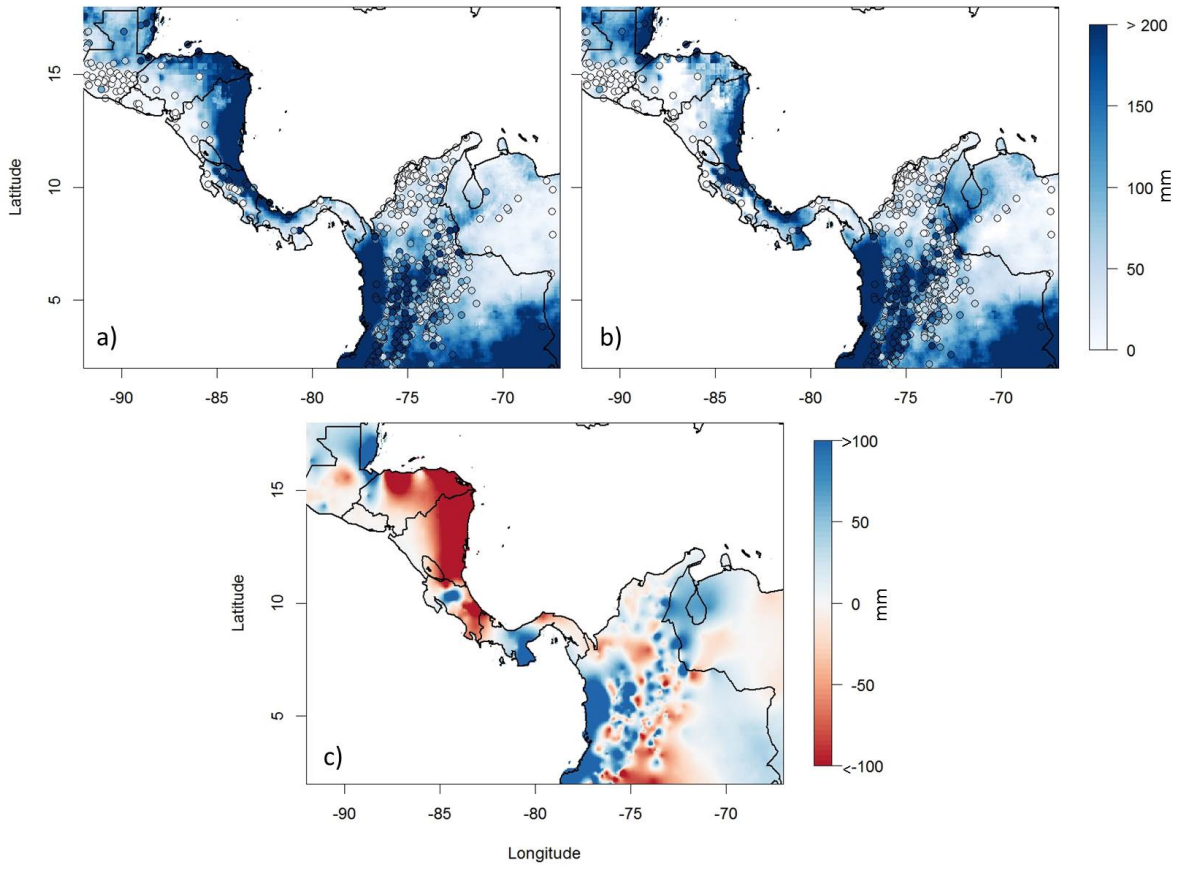

Figure 3.8: a-b) Original and Bayesian median blended CHIRP product for January 2009; station values on same color scale, c) change from original CHIRP due to blending.

Blended - Original). Overall the magnitude of change is small and has coherent spatial structure. As can be seen, there are many localized events in Colombia, implying there is great variability of rainfall even in this representative dry month. Further, note that much of this change is from the estimation of the residual process, ϵ, since CHIRP is a largely unbiased estimator during this month (see Figure 3.2).

Figure 3.9a and b show the 5^{th} and 95^{th} percentiles of the posterior predictive distributions of rainfall for January 2009. The difference between the lower bound and upper bound of this confidence interval is very small, implying that this model is consistent in its estimation process. The uncertainty of the kriging process is quantified using the median parameters from the poste-

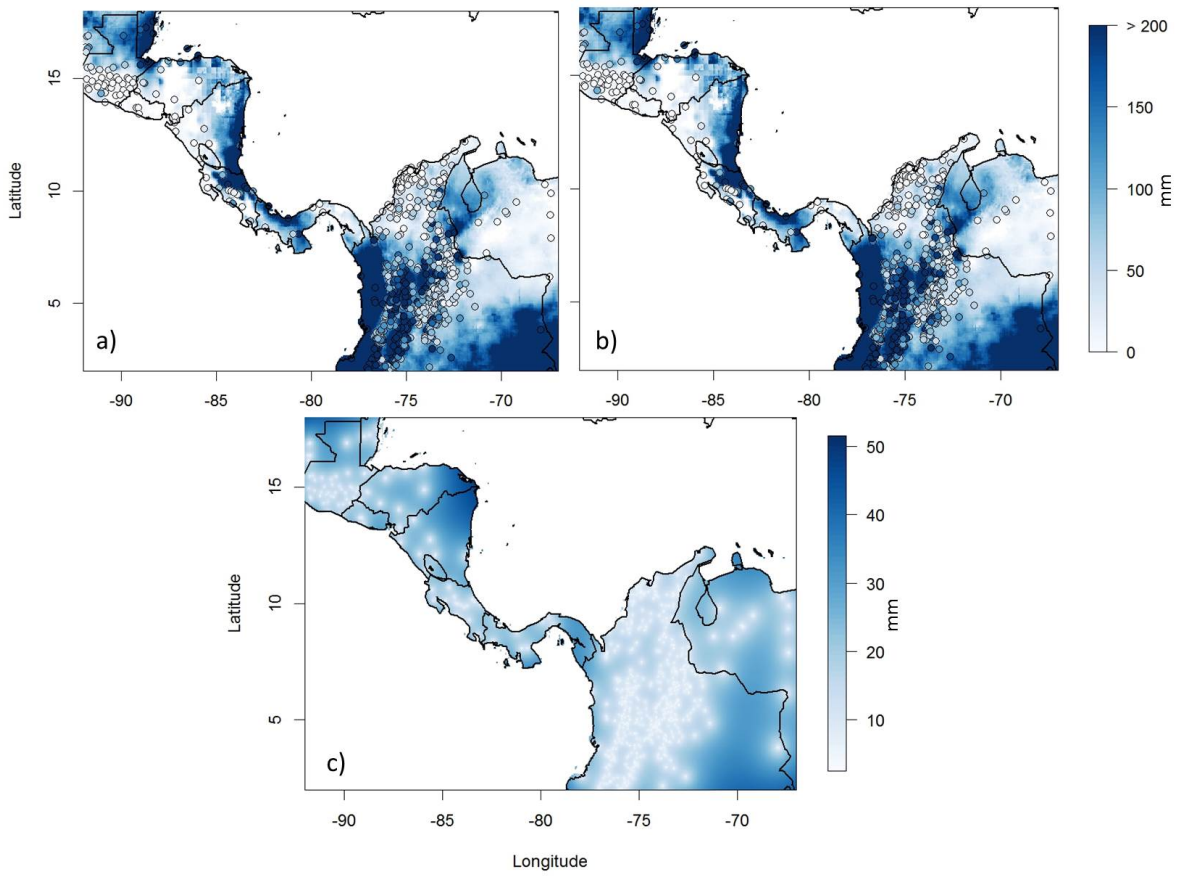

Figure 3.9: a-b) 5^{th} and 95^{th} percentile of Bayesian predictive samples for January 2009.

rior distributions (Figure 3.9c). This shows the uncertainty is greatest at locations furthest from observation locations; similarly, the uncertainty is lowest at locations closest to observations.

July 2009

Spatial maps of the original CHIRP and blended estimates are shown in 3.10a and b – it can be seen that the blended estimates are higher over Nicaragua and Colombia, again consistent with topography. This may be easier to note in Figure 3.10c, which shows the difference between the blended estimates and original CHIRP (Blended − Original). The scale is much greater than for January, due to higher values of the posterior distribution of the parameters β_1 and β_2. The magnitude of increase is on par with the maximum possible increase due to elevation ($\sim$ 280 mm),

along with a CHIRP correction factor of 27% (median of posterior distribution of $\beta_1 = 1.27$). There are very few regions with blended estimates that are lower than the CHIRP values, confirming the consistent underestimation with respect to rain gauge measurements.

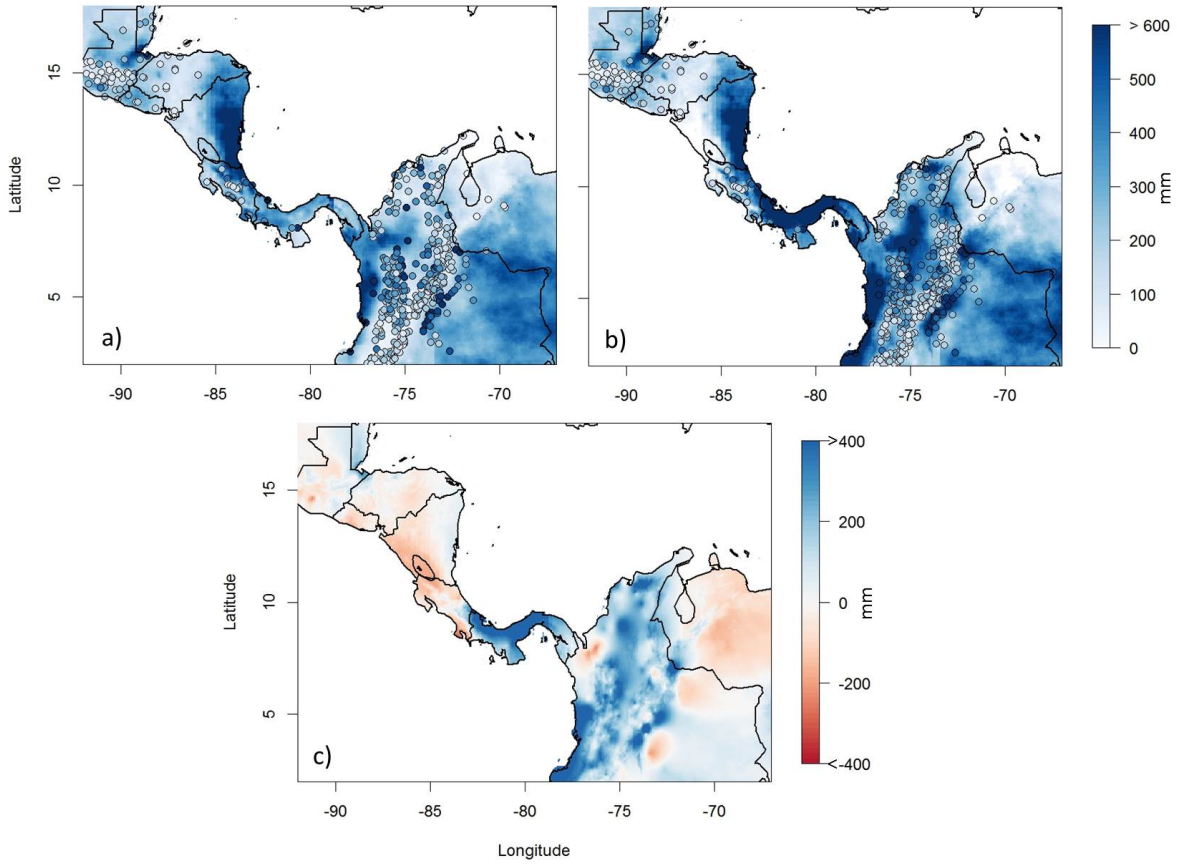

Figure 3.10: Same as Figure 3.8 but for July 2009.

The spatial maps of 5^{th} and 95^{th} percentiles from the predictive posterior distributions (Figure 3.11a-b) also show consistent behavior with elevation. Again, the difference between the lower bound and upper bound of this confidence interval is very small, implying that this model is consistent in its estimation process. The uncertainty of the kriging process is quantified using the median parameters from the posterior distributions (Figure 3.11c). This shows the uncertainty is greatest at locations furthest from observation locations; similarly, the uncertainty is lowest at locations closest to observations.

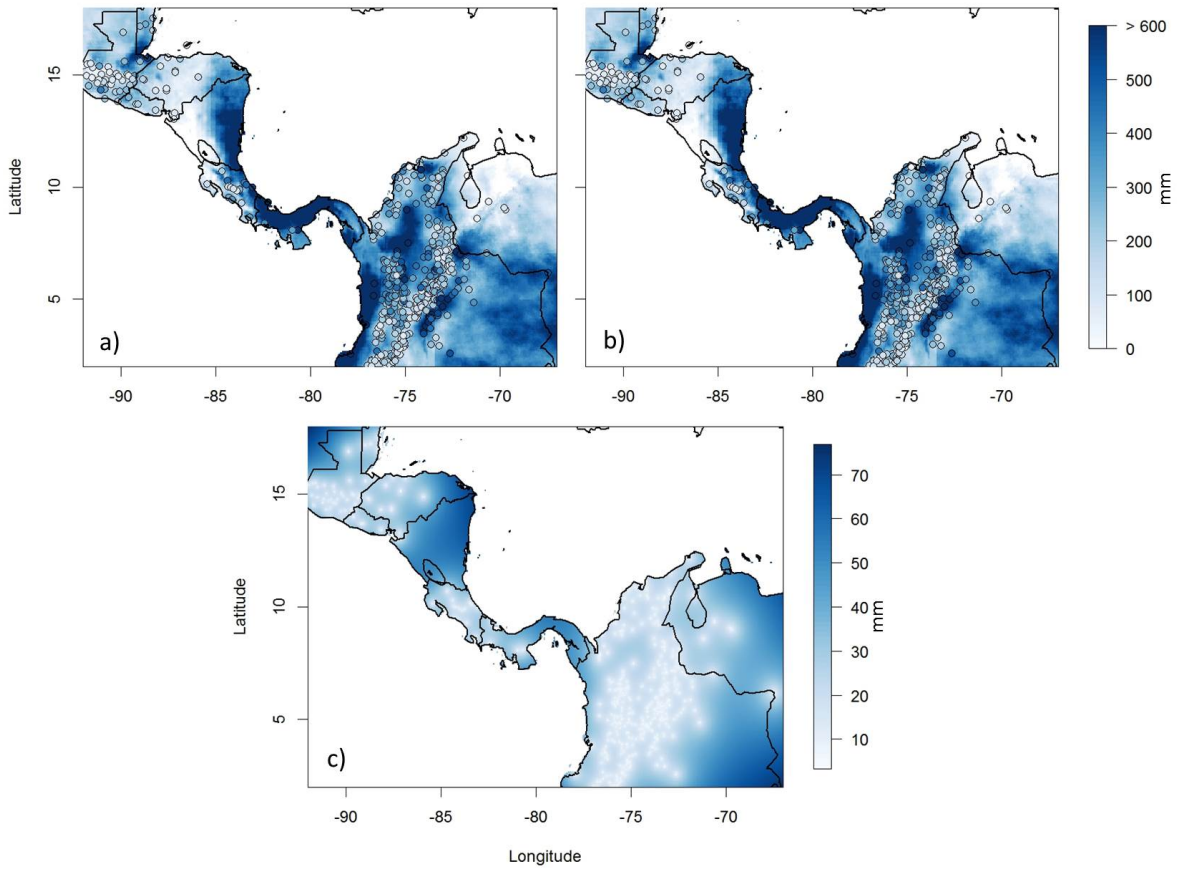

Figure 3.11: Same as Figure 3.9 but for July 2009.

3.4.5 Model Performance on Extreme Events

Good estimation of extreme events is important for hydroclimatic hazard mitigation, especially during wet months. The hazards associated with extreme events consist of floods, landslides, and agricultural overland flow, all of which are destructive to a nation's infrastructure. We define an extreme event at a location whose monthly rainfall exceeds the 85^{th} percentile of all rain gauge measurements for that month. The 85^{th} percentile of all gauged rainfall in July 2009 is 410 mm/month. The locations of these events and the majority of extreme precipitation events occur in the mountainous and coastal regions (Figure 3.12). Figure 3.13a shows the performance of the original CHIRP and 3.13b shows the performance of the blended estimates with respect to extreme events; Table 3.1 summarizes how blending improves performance of the satellite estimate with

respect to extreme events. These statistics are computed from drop-one cross-validation output.

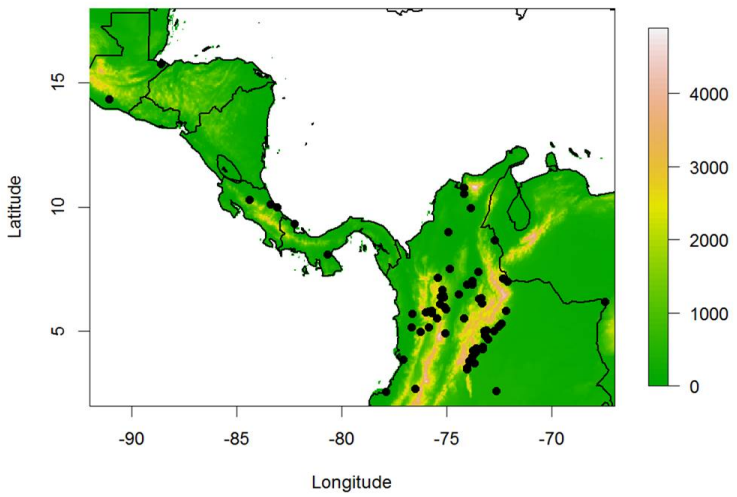

Figure 3.12: Locations of extreme precipitation events for July 2009 (black dots) and elevation shown in color.

	RMSE	MAE	% Bias
Original CHIRP	386 mm	344 mm	−58.8%
Blended Estimate	222 mm	159 mm	−19.2%

Table 3.1: Summary statistics of blending performance on extreme events for July 2009.

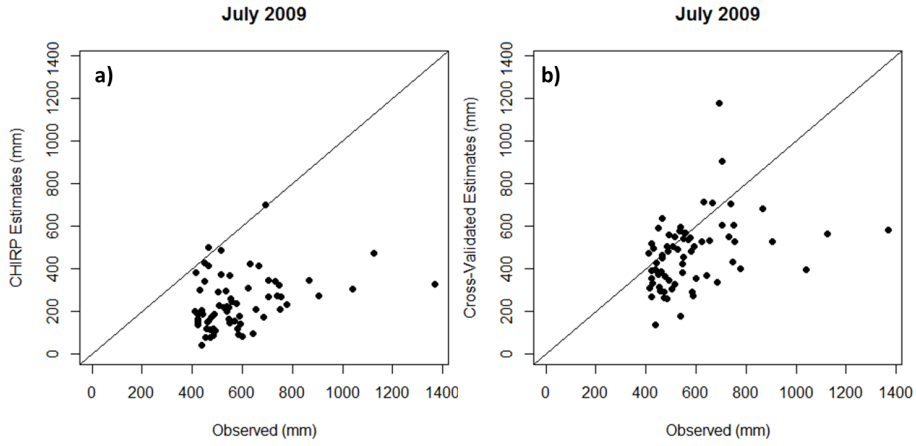

Figure 3.13: Blending performance on extreme events for July 2009.

As can be seen in Figure 3.13 and Table 3.1, the blending method proposed in this chapter has a positive impact on the CHIRP's performance with respect to extreme precipitation events. Validation statistics from Table 3.1 indicate that blending reduces the error and inherent bias of the CHIRP satellite estimate. It should be noted that models for all months in 2009 – the test period – show similar skill with respect to extreme events, thus it is worthwhile to implement this blending method on all available data to produce a more representative gridded time series of monthly precipitation.

3.4.6 Application to Pentad Rainfall Estimation

In this section, we investigate the utility of the blending approach to shorter temporal scales, such as pentad rainfall. This is of interest for near-term blended precipitation estimates for natural hazard mitigation strategies. Blended estimates of precipitation on pentad time scales during the wet season can provide insight into the wetter regions of river basins, can be used to drive hydrologic models for modeling soil moisture and consequently, the potential for flooding and landslides.

The fourth pentad of August 2009 will be used in this application, as it is one of the wettest pentads of 2009, yielding a maximum observation of 385 mm where the CHIRP reports only 130 mm. The BK blending as applied to the monthly precipitation described in previous sections is applied for this pentad data.

Figure 3.14 shows the posterior distributions of the model parameters. The posterior distribution of β_1 suggests that the CHIRP is biased with respect to the rain gauges (the median value of β_1 is 1.21), consistent with our earlier findings; further, elevation has no effect – the median value of β_2 is 0.00. The posterior distribution of the marginal standard deviation σ and effective range ϕ (Figures 3.14c and d) are consistent with the distributions obtained for July rainfall (Figure 3.3) – in that they have similar shapes. The effective range ϕ has a smaller range for the pentadal scale compared to that of July monthly rainfall (Figure 3.3) – this is appropriate because at finer time scales the true rainfall process is much noisier than for monthly sums and magnitudes are smaller. The consistency of the posterior distributions for the kriging parameters confirms the portability

of the BHM.

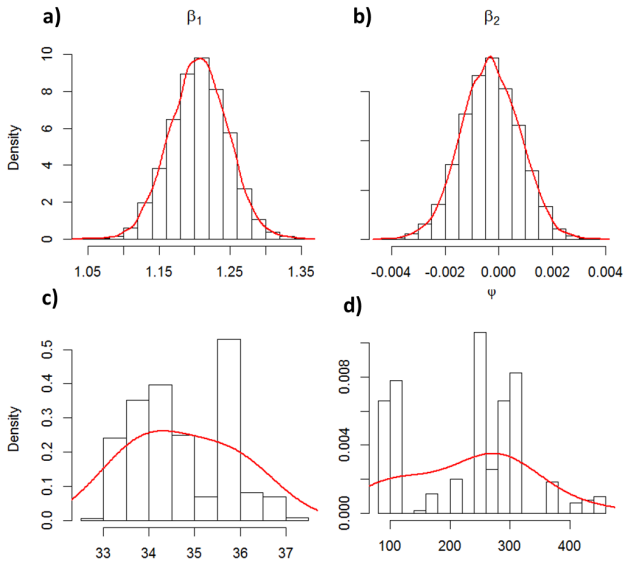

Figure 3.14: Same as Figure 3.2 but for the fourth pentad of August, 2009.

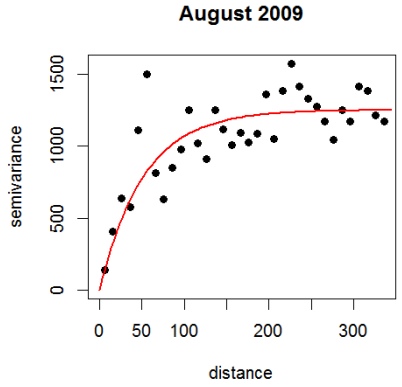

Figure 3.15: Same as Figure 3.4 but for the fourth pentad of August 2009.

Figure 3.15 shows the empirical variogram (black dots) and the theoretical variogram (red line) as estimated by the median parameters of the spatial model from the posterior distributions. The median value based on the BK model indicates kriging parameter values such that the theoretical variogram matches the empirical variogram quite well. These results are consistent to those

on the monthly time scale, suggesting this model is portable and consistent.

To test the predictive capability of the model in this context, we performed drop-one, drop-25%, and drop-50% cross-validation as was done with the monthly precipitation earlier. Figure 3.16 shows the scatterplot for drop-one cross-validation; Figure 3.17 shows boxplots of skill measures – RMSE, MAE, and % Bias for the drop-25% cross-validation, and similarly Figure 3.18 for drop-50% cross-validation. For comparison, the statistics of the original CHIRP product are shown as red dots.

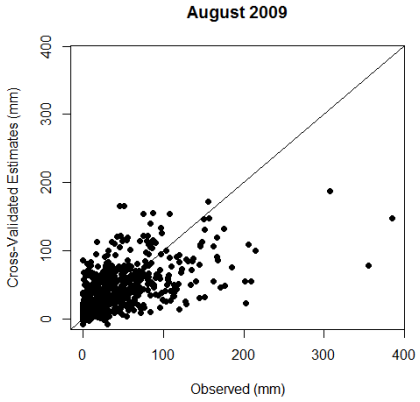

Figure 3.16: Same as 3.5 but for the fourth pentad of August 2009.

For this wet pentad, drop-one cross-validation produces results that eliminate bias and slightly reduce error. RMSE for drop-one cross-validation is 38 mm, that of the original CHIRP product is 40 mm. Similarly, MAE is 25 mm for drop-one cross-validation, that of the original CHIRP product is 26 mm. The resulting bias from drop-one cross-validation, however, is only 0.1% while the original CHIRP product has inherent bias of -21.4%.

While the median values of RMSE and MAE from the cross-validation procedures are similar to the original CHIRP, there is a significant reduction in % Bias. The pentad considered is one of the wettest, thus when 25% and 50% of observations are dropped from the model-fitting process its predictive ability is compromised. Considering this, the performance of the blending method is very good.

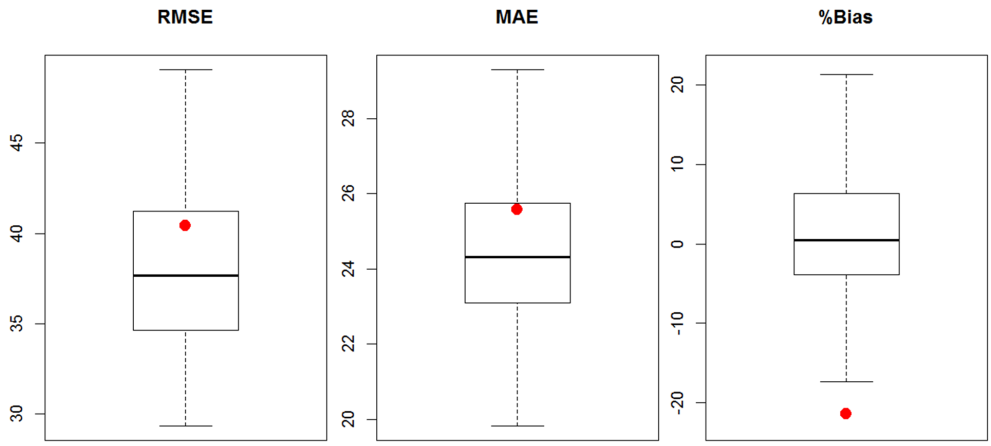

Figure 3.17: Same as 3.6 but for the fourth pentad of August 2009.

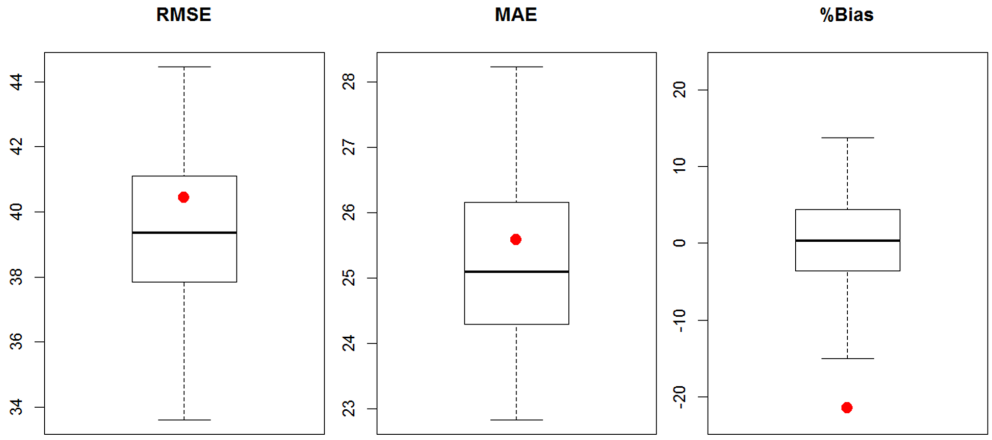

Figure 3.18: Same as figure 3.7 but for the fourth pentad of August 2009.

Figure 3.19 shows the spatial map of original CHIRP for the pentad (3.19a) and the median estimate of the predictive posterior distribution from the Bayesian blending (3.19b). Estimates over much of Nicaragua and Colombia are higher indicating the model has predicted extreme precipitation in these regions. This increase in rainfall over the majority of the study region is due

to the β_1 coefficient which has a median value of 1.21 (Figure 3.14a), thus the CHIRP product is increased everywhere.

The difference between these two estimates is shown in Figure 3.19c. It can be seen that the regions of increase in the pentad CHIRP product have a spatial pattern consistent with the original CHIRP estimate and topography, and there are about equal grid points with the blended estimates less than the CHIRP values – this is consistent with the CHIRP product's over- and underestimation of low and high magnitude precipitation events, respectively.

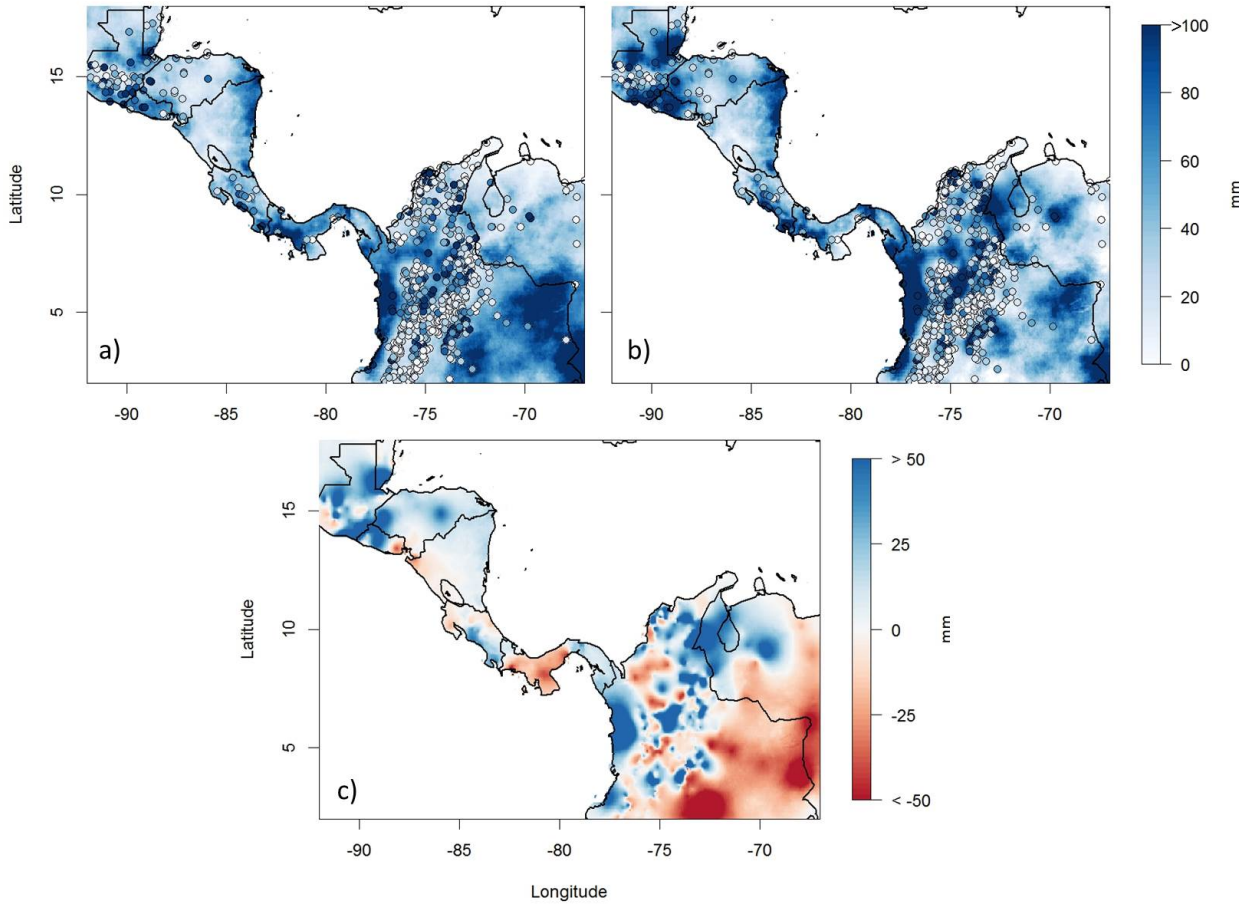

Figure 3.19: Same as Figure 3.8 but for the fourth pentad of August 2009.

The robust confidence interval of the posterior predictive samples is shown in Figure 3.20a-b. Again, the difference between the 5^{th} and 95^{th} percentile values imply a consistent estimation

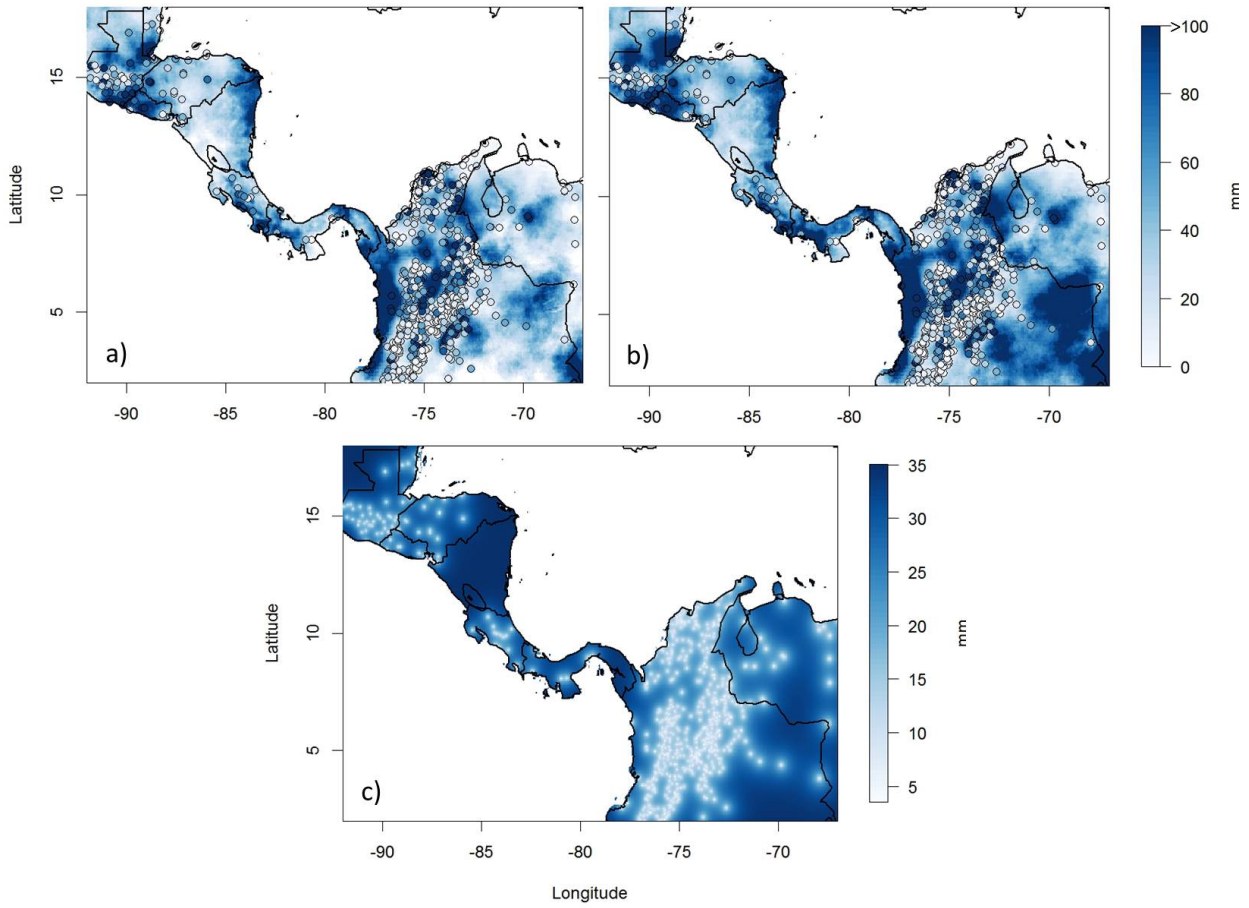

Figure 3.20: Same as Figure 3.9 but for the fourth pentad of August 2009.

process. The uncertainty in the kriging estimates is consistent with that of the monthly rainfall.

To test the ability of the method in estimating extremes, as before we define an extreme event at a location whose rainfall exceeds the 85^{th} percentile of all rain gauge measurements for the pentad – in this case 78 mm. The locations of these events and the majority of extreme precipitation events occur in the mountainous and coastal regions (Figure 3.21).

Figure 3.22a shows the performance of the original CHIRP and 3.22b shows the performance of the blended estimates with respect to extreme events; Table 3.2 summarizes how blending improves performance of the satellite estimate with respect to extreme events. These statistics are computed from drop-one cross-validation output.

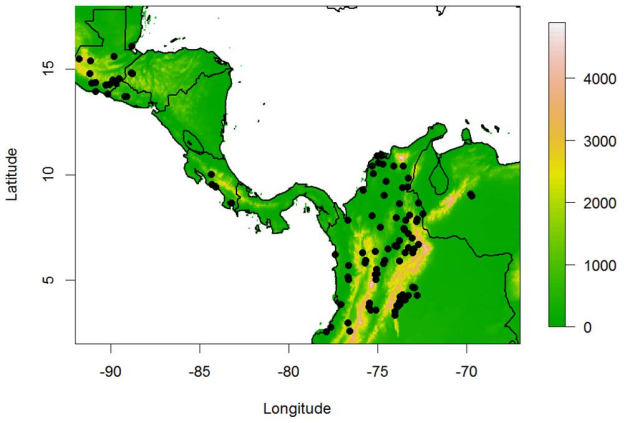

Figure 3.21: Same as Figure 3.12 but for the fourth pentad of August 2009.

	RMSE	MAE	% Bias
Original CHIRP	90 mm	76 mm	−60.5%
Blended Estimate	75 mm	59 mm	−39.7%

Table 3.2: Same as Table 3.1 but for the fourth pentad of August 2009.

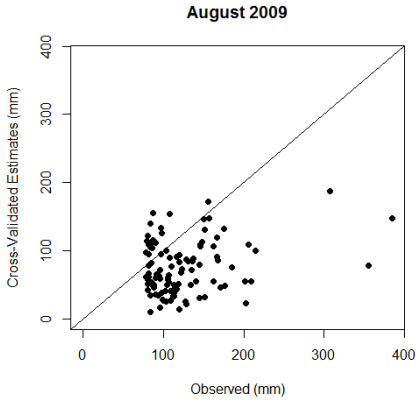

Figure 3.22: Same as 3.13 but for the fourth pentad of August 2009.

As can be seen in Figure 3.22 and Table 3.2, the blending method proposed in this chapter is portable from monthly to pentadal scales and has a positive impact on the CHIRP's performance with respect to extreme precipitation events. Validation statistics from Table 3.2 indicate that blending reduces the error and inherent bias of the CHIRP satellite estimate. In light of these

results, it is encouraged to implement this blending technique for all available data in order to produce a more representative gridded time series of precipitation.

3.5 Summary and Conclusions

A novel approach for blending satellite-derived precipitation estimates with rain gauge measurements has been proposed and validated within the Central American region. This blending method is a Bayesian Hierarchical Model wherein the precipitation process at a location is assumed to be a realization of a (normal) distribution and the mean estimates of the distribution are modeled as a linear function of covariates (satellite-derived estimates and elevation). The residuals from this linear function are modeled as a spatial process using ordinary kriging. The posterior distributions of all the parameters of the hierarchical model are obtained via MCMC. The posterior distribution of the parameters, thus the predictive posterior distribution of the precipitation estimates, provide a complete representation of the variability and a quantification of the uncertainty in the parameters. These distributions can be used to produce probability density functions of precipitation estimates at every grid cell, along with associated confidence intervals.

The model was applied to January and July 2009 precipitation – a dry and wet month, respectively – which showed very good model fits – as expected the model is an exact fit at the observations. Drop-one, and well as drop-25% and -50%, cross-validation predictions of the model all showed good skill especially in eliminating all of the bias prevalent in satellite-derived estimates. Spatial estimates of blended precipitation showed robust and consistent improvements in that satellite estimates were enhanced appropriately based on the precipitation values of the surrounding gauges. The model was also tested with equally skillful performance on pentad time scales where the rainfall process is intensely erratic and driven by topography. The posterior distribution of the model parameters showed the relative importance of the covariates in the wet and dry months – confirming the known underestimation of satellite-derived precipitation. The Bayesian hierarchical model proposed here can produce robust spatial precipitation maps with estimates of uncertainty in near-real-time that will be of immense use in hydrologic and hazard models to provide risk

estimates of various hazards in space that can be used in mitigation efforts. Furthermore, this method offers a complementary suite to the standard methods based on kriging and less traditional local polynomials (Verdin et al. 2013a).

One of the drawbacks of the blending approach proposed here is that it is performed on a spatial snapshot – i.e., individual month or pentad. However, time series of precipitation (both satellite and gauge) are available and thus incorporating the temporal variability can provide greater accuracy. This would involve fitting a separate GLM for each gauged location using all the data available in time and fitting a spatial process for each model parameter. Thus, estimation of precipitation at any location would involve first obtaining the estimates of model parameters from the respective spatial process models and then combining in the hierarchy to obtain the precipitation estimates. The hierarchy could be non-Bayesian (Verdin et al. 2013c) or latent Gaussian process models (Kleiber et al. 2012) or a Bayesian approach used in other applications particularly of modeling spatial extremes (Cooley et al. 2007, Cooley & Sain 2010). The use of temporal variability can help improve the near-real-time blended estimates by incorporating climate drivers that modulate the precipitation variability.

Chapter 4

A Hierarchical Modeling Approach for Time Series Blending

4.1 Introduction

The previous two chapters have considered methods for blending satellite and ground observations that utilize a snapshot approach, where only the information for a single year is considered. Previous efforts (see Chapters 2 and 3) have shown that adjusting the satellite estimate to account for inherent bias produces a more representative estimate, thus enhancing the performance of the Gaussian process model of satellite error. That being said, it is reasonable to assume the bias of the satellite estimate is not constant in space, rather a spatial process that is constant in time. To this end, the blending method proposed in this chapter considers the full time series of data to obtain a more robust estimate of the bias correction factor, which is assumed to vary over the domain. To this end, a bias correction factor is obtained at all locations with near-complete time series data and a Gaussian process model estimates the bias correction factor for all grid cells of the satellite estimate.

Implementing a hierarchical Gaussian process model for estimating model parameters is a relatively new approach to the spatial estimation problem. This method provides a robust framework for parameter estimation in that a model's parameters are allowed to vary spatially, thus better quantifying the physical process it models. A hierarchical approach to spatial parameter estimation has been implemented for the modeling and quantification of precipitation extremes (Cooley et al. 2007, Sang & Gelfand 2008, Schliep et al. 2009, Cooley & Sain 2010), and soil parameter estimation (Šimůnek et al. 1998, Schaap et al. 2001, Vrugt et al. 2002), to name a few.

The proposed method for blending satellite and ground observations of rainfall is a new and unique application of spatial parameter estimation. This approach is novel in the sense it considers the full time series of data at a point to obtain parameters that describe the relationship between in situ rain gauge measurements and high-resolution satellite-derived estimates of rainfall. The hierarchical framework allows for the structured treatment of model parameters that together describe a spatio-temporal process. The study region and data are described in the next section, followed by an in depth description of the methodological framework, and finally the results section.

4.2 Study Region & Data

The details of the study region and data sets are described in greater detail in Chapter 1, an abstract of which is provided below.

4.2.1 Study Region

The study region covers the Central American region (see Figure 4.1). The dry season in the region extends from December through early May and the effects of both Pacific and Atlantic tropical depressions are felt during the wet season that extends from mid-May through November. A major contributor to the prolonged wet season is the inter-tropical convergence zone (ITCZ) – an asymmetric band of convection which encircles the globe. The position of the ITCZ is non-stationary due to the seasonal shift of the trade winds.

Another cause for intense spatial variability of rainfall within the study region is its complex geography, where the elevation ranges from sea level to nearly 5000 meters (as can be seen in Figure 4.1) with numerous mountain ranges spanning Central America and Colombia. Costa Rica and Guatemala have steep elevation gradients along the coasts, making them vulnerable to extreme weather spawning from both Pacific and Atlantic tropical depressions. Coastal mountain ranges cause an abrupt rising of warm, wet air. As this tropical air rises, it cools rather quickly, thus releasing the moisture from the air as precipitation, a major cause for the variability in extreme precipitation events in the region.

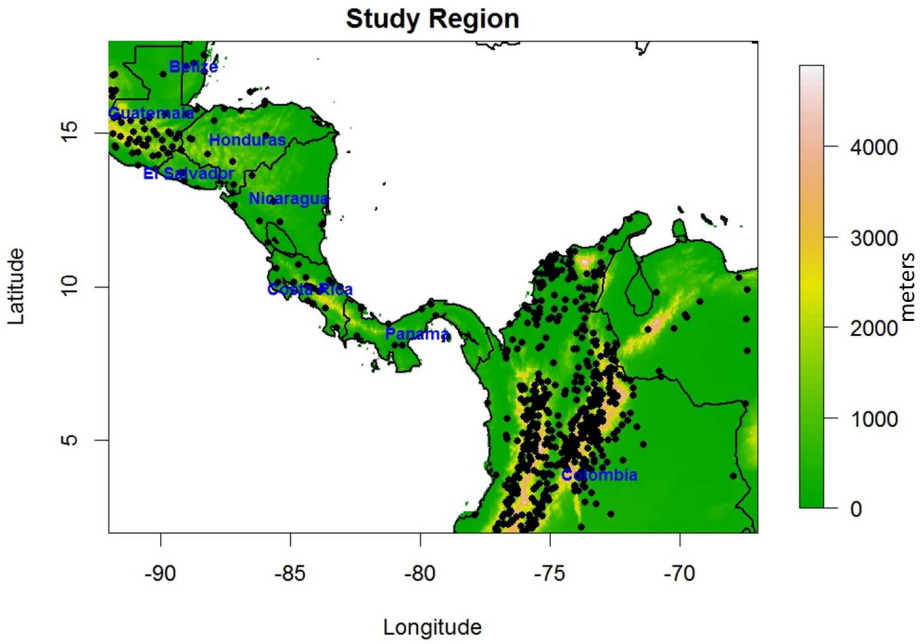

Figure 4.1: Study region geography, rain gauge measurement locations, and country names.

Given the spatial and temporal variability of rainfall in this region coupled with vulnerable socio-economic conditions, droughts and floods can have a debilitating impact on these societies. It has been observed that a single extreme wet precipitation event – whether hurricane, tropical storm, or prolonged precipitation – may result in flooding, landslides, and potentially millions of dollars in damages, thus setting the national infrastructure back many years. Conversely, a prolonged dry spell may be equally devastating to a growing nation, causing widespread crop failures. Thus, a comprehensive understanding of rainfall variability is of critical importance for resources planning and management.

4.2.2 Data

Two 32-year time-series precipitation datasets are utilized in this research – gauge measurements and satellite-derived estimates. Rain gauge measurements are point values that tend to be clustered in populous and lower elevation areas and sparsely scattered throughout the domain; the

satellite-derived estimates are high-resolution – 0.05°x0.05° (approx. 5 km) – areal averages for five day periods (i.e. pentads). The satellite-derived precipitation estimate used in this research is the CHIRP (Climate Hazards group InfraRed Precipitation) covering the first pentad (five-day sum) of January, 1981 through the last pentad of December, 2012. There are six pentads per month, regardless of how many days are in a given month. For example, the sixth pentad of October will consistently be the sum of the last six days, October 26-31. The sixth pentad of February, however, will be the sum of either three or four days, depending on leap year status. Pentads may also be summed to obtain monthly totals of precipitation. It is widely accepted that satellite estimates of rainfall are better at capturing the spatial structure of precipitation, but fail in estimating magnitudes, especially of extreme events – this is because the satellite estimates are derived estimates of precipitation (Xie & Arkin 1997).

The CHIRP estimate is produced using the CHPCLIM (Climate Hazards group Precipitation Climatology), a completely unbounded global long term pentadal climatology of precipitation of the same resolution (0.05°x0.05°), as well as the infrared radiation (IR) satellite product. The CHPCLIM for the appropriate pentad is multiplied by the most recent IR percent-anomaly to produce the CHIRP estimate, shown below:

$$CHIRP_{i,j,k} = \%IR_{i,j,k} \cdot CHPCLIM_k \tag{4.1}$$

In the above equation, $CHIRP_{i,j,k}$ represents the satellite-derived precipitation estimate for pentad k of month j in year i; $\%IR_{i,j,k}$ represents the infrared radiation satellite estimate percent-anomaly for pentad k of month j in year i; and $CHPCLIM_k$ represents the precipitation climatology for pentad k.

The station data (shown in Figure 4.1) was collected from partner sources of the UCSB Climate Hazards Group; quality control and pentadal-aggregation techniques were implemented by the CHG to produce time series data of the same temporal extent as the available CHIRP products (the first pentad of January, 1981 through the last pentad of December, 2012).

4.3 Proposed Blending Methods

A hierarchical modeling approach is proposed for blending time series of precipitation from satellite-derived estimates and rain gauge observations for the 32-year period 1981-2012. To account for seasonality, the model is developed separately for each month or pentad.

We define the following variables used in the method: $Y_{sat}(s,t)$ and $Y_{obs}(s,t)$ are the satellite-derived estimates (i.e. CHIRP) and observed precipitation, respectively, at locations $s = 1...N_s$, where N_s is the number of stations with non-missing data for a given pentad or month, and at time $t = 1...N_t$, where N_t is the number of time points for each time frame (i.e. pentad, month).

In the first level of the hierarchy, for a given location and time frame,

$$Y_{obs}(s,t) \sim N(\mu(s,t), \sigma^2(s,t))$$
$$\mu(s,t) = \beta(s)Y_{sat}(s,t) + \epsilon(s,t) \tag{4.2}$$

This linear model is fitted at each gauge location for any given time frame using all available data for that time frame from all years. Thus a set of s regression coefficients β (one per location) is obtained along with the residual series $\epsilon(s,t)$.

In the next hierarchical step, spatial Gaussian process (i.e. Co-kriging) models are fitted to the regression coefficients and residual series, with easting, northing, and elevation as covariates; covariance between the observations are described by an exponential variogram function with parameters τ^2, σ^2, and ϕ. τ^2 is the nugget effect (the result of dissimilarity between points at infinitesimally small lag), σ^2 is the marginal variance of the Gaussian process, and ϕ is the effective range (the distance at which covariance reaches plateau). Other covariates may be included in the first hierarchical step, resulting in as many β parameters as covariates and, in the second step, resulting in as many spatial models as β parameters.

The spatial model will provide estimates of β at any desired location s with its spatial attributes. However, the spatial model of ϵ provides the same estimate at any desired location s regardless of time, since the spatial model is static (easting, northing, and elevation do not change in time). This poses a problem when estimating the residual ϵ at different time points. To address

this problem, we fitted a spatial model $\epsilon(s,t)$ for each time frame separately. Thus, a separate spatial model is fitted for each time frame and each desired year.

To obtain a blended estimate $\hat{Y}(s_i,t)$ at a desired location s_i and time t, the spatial models are used to obtain estimates of the regression coefficient and the residual, $\beta(s_i)$ and $\epsilon(s_i,t)$, respectively. The blended estimate is:

$$\hat{Y}(s_i,t) = \beta(s_i)Y_{sat}(s_i,t) + \epsilon(s_i,t) \tag{4.3}$$

The spatial models are exact estimators, thus at the gauge locations the blended estimates will match the observations. Therefore, to evaluate this blending method, three increasingly-stressful cross-validation methods are performed. We apply this blending method to estimate monthly precipitation for January and July, 2009 – dry and wet season monthly totals, respectively – and also on the fourth pentad of August, 2009 – a representative wet season pentad. This is to enable comparison with the blending methods proposed and evaluated in previous chapters.

4.4 Results

The proposed hierarchical model is applied to monthly total precipitation – as well as a representative wet pentad – for 2009. This research focuses on 2009 due to newly available data for Colombia provided by partner sources of the Climate Hazards Group. Model fitting and drop-one cross-validated estimates of root mean square error (RMSE), mean absolute error (MAE), and percent bias (% Bias) are computed for each time period. Further analysis such as drop-25% and -50% cross-validation and performance on extreme precipitation events are shown for representative dry (January) and wet (July) months. Analysis will focus on overall performance but, more importantly, the effect that blending has on the CHIRP's representativeness of extreme precipitation events. For consistency, an extreme event is defined at a location where rainfall exceeds the 85^{th} percentile of all events on record. To illustrate the portability of this blending method to a shorter time scale, a representative wet pentad for 2009 is also analyzed. One of the wettest pentads of 2009 occurred in the midst of the region's wet season, during the fourth pentad

of August, with total rainfall of nearly 400 mm.

Analysis of results is broken into the following categories: model fitting, validation, and blending estimates on the satellite grid. As mentioned above, this analysis will focus on January and July, 2009 – representative dry and wet months, respectively – followed by pentadal analysis for the fourth pentad of August, 2009 – a representative wet pentad.

4.4.1 Monthly Totals

Model Fitting

A linear model (Equation 4.2) was fitted at each location using the monthly satellite and observed precipitation, obtaining $\beta(s)$ and $\epsilon(s,t)$ for each location. Spatial models are then separately fitted for these two variables for each month. Figure 4.2 shows the empirical (black dots) and theoretical (red line) variograms of both Co-kriging models (β and ϵ) for both January and July. The theoretical variograms capture the data quite well for all of them, with the variograms of the error field ϵ exhibiting better fit. The spatial structure of the error field is the result of the linear model, which accounts for much of the bias in the satellite estimates, thus leaving small scale variability which is best captured by the spatial model variograms. The satellite estimates poorly capture the gauge observations (see Figures 4.3 & 4.4) – somewhat better in January – while the blended estimates from the spatial models are near-exact and thus fit the observations much better. Table 4.1 summarizes R^2, RMSE, MAE, and % Bias for all months in 2009.

As can be seen from Table 4.1, the original CHIRP estimate has poor representativeness of the gauge measurements, with R^2 values ranging from 0.39 to 0.78; its RMSE and MAE statistics are quite large, with ranges of 48 mm to 154 mm and 35.5 mm to 92.3 mm, respectively; there is also large inherent bias in the CHIRP estimate, with a range of -53.1% to 22.1%. The blended estimates from the model, however, tends to have high R^2, very low error statistics, and almost no percent bias. This is due to Kriging being a near-exact estimator for observations included in the model fitting process. The blended estimates are not exact because there are two Co-kriging models, both

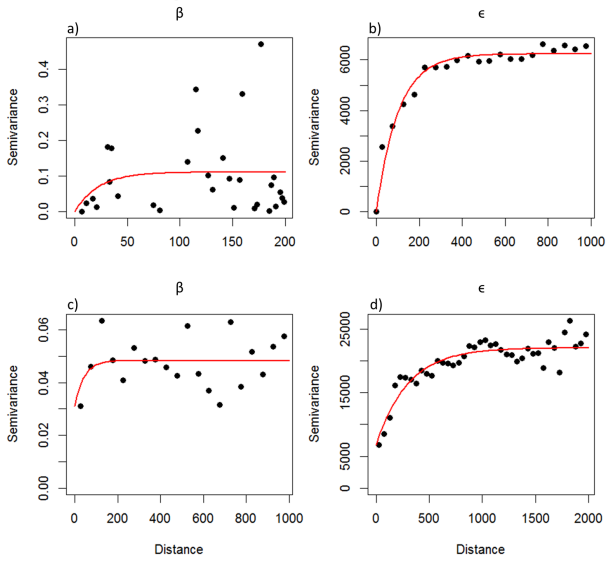

Figure 4.2: Empirical (black) and theoretical (red) variograms for January β (a) and ϵ (b), and for July β (c) and ϵ (d).

	CHIRP R^2	Model R^2	CHIRP RMSE	Model RMSE	CHIRP MAE	Model MAE	CHIRP % Bias	Model % Bias
Jan	0.66	0.95	75.8	28.1	45.2	18.3	8.3	0.8
Feb	0.64	0.70	154.0	126.0	55.9	63.4	-53.1	-9.4
Mar	0.66	0.88	78.2	41.8	35.5	25.1	-38.6	1.3
Apr	0.56	0.74	54.4	41.4	36.4	25.9	17.7	6.5
May	0.49	0.79	113.0	60.3	92.3	41.9	22.1	0.3
Jun	0.50	0.83	88.2	52.0	72.2	41.9	-3.4	2.1
Jul	0.39	0.85	128.0	63.5	74.4	38.4	-7.3	0.0
Aug	0.66	0.85	67.9	44.1	53.6	35.1	0.8	2.0
Sep	0.66	0.94	79.6	34.1	63.1	25.0	9.5	2.2
Oct	0.70	0.83	94.0	62.8	64.7	40.1	10.8	-0.1
Nov	0.78	0.90	69.6	46.0	53.9	34.8	-10.9	-0.1
Dec	0.66	0.92	48.0	20.6	36.3	14.1	2.5	0.7

Table 4.1: Summary statistics reflecting model-fitting performance.

with non-negligible nugget effect. It follows that the proposed blending method greatly improves the representativeness of the CHIRP product for all months in the 2009 test period.

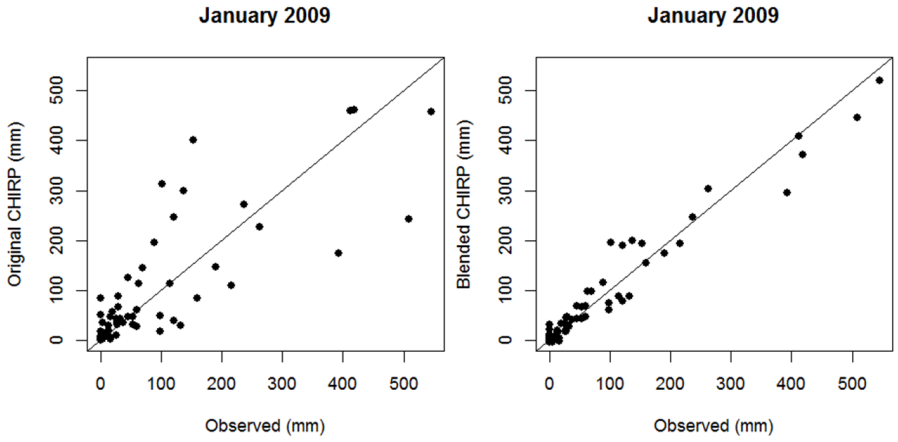

Figure 4.3: Original CHIRP estimates (left) and blended estimates for January 2009.

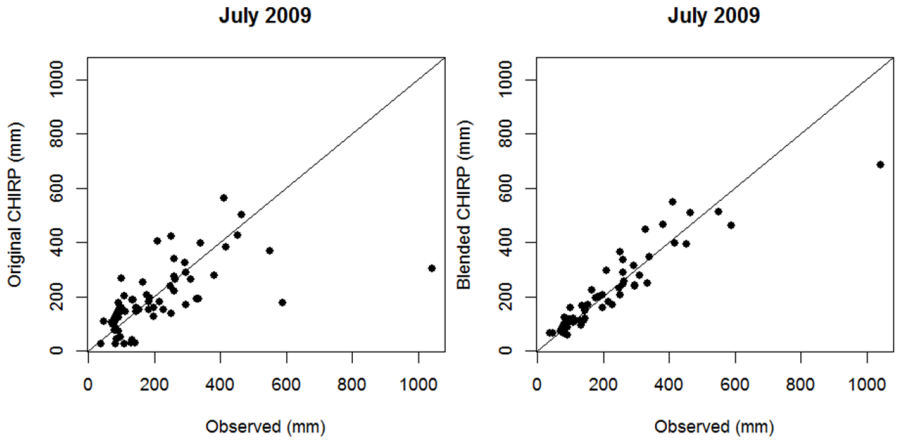

Figure 4.4: Same as Figure 4.3 but for July 2009.

Model Validation

Since the blending method is a near-exact estimator for observations included in the model-fitting process, cross-validation needs to be employed for testing the predictive capability of this blending method. To this end, drop-one, drop-25%, and drop-50% cross-validation procedures are performed. Drop-one cross-validation is a traditional approach for stressing the predictive power of a model, wherein one observation (β or ϵ, separately) and its covariates are dropped from the model-fitting process and the dropped point is subsequently estimated. This process is repeated for every available observation, in turn creating a vector of independently predicted points that is compared with the known observations. Figures 4.5 and 4.6 show the results from drop-one cross-validation for January and July 2009, respectively. Summary statistics for all months in 2009 are shown in Table 4.2, with the statistics of the original CHIRP product included for comparison.

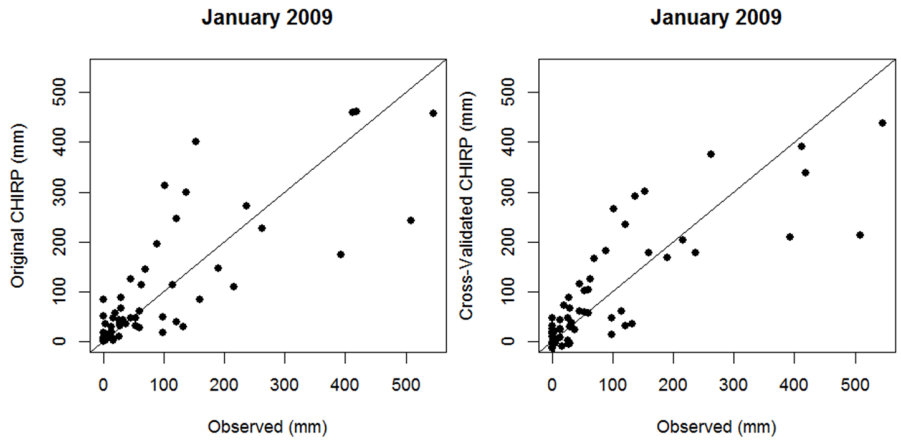

Figure 4.5: Original CHIRP estimates (left) and cross-validated estimates for January 2009.

The months where this method shows little to no improvement – and possibly a degradation of performance – are those where the original CHIRP has good performance. It follows that this blending method should only be used for time periods where drop-one cross-validation shows an improvement from the original CHIRP. Clearly the user must not automate this methodology, as implementation for February 2009 (see Table 4.2) will reduce the overall performance of the CHIRP

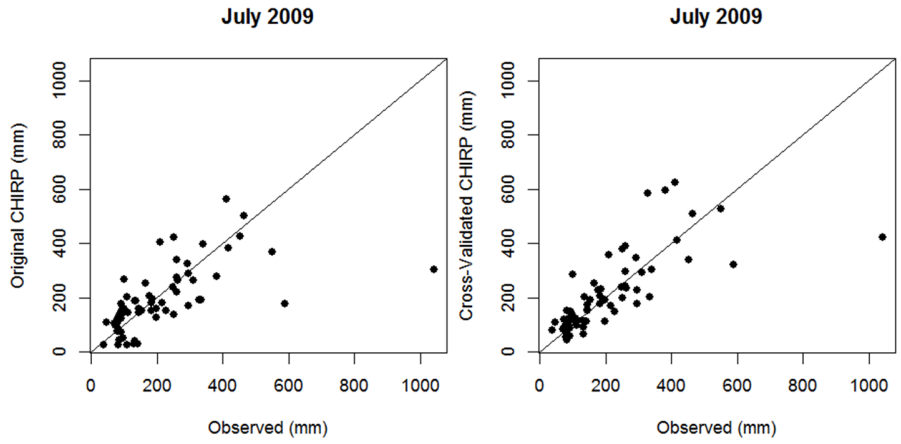

Figure 4.6: Same as Figure 4.5 but for July 2009.

	CHIRP R^2	Model R^2	CHIRP RMSE	Model RMSE	CHIRP MAE	Model MAE	CHIRP % Bias	Model % Bias
Jan	0.66	0.68	75.8	71.3	45.2	47.6	8.3	1.6
Feb	0.64	0.28	154.0	162.0	55.9	85.6	-53.1	-0.2
Mar	0.66	0.79	78.2	48.0	35.5	27.0	-38.6	5.1
Apr	0.56	0.48	54.4	58.2	36.4	34.5	17.7	1.8
May	0.49	0.58	113.0	88.8	92.3	69.8	22.1	-0.6
Jun	0.50	0.43	88.2	96.8	72.2	72.1	-3.4	-0.4
Jul	0.39	0.53	128.0	113.0	74.4	66.3	-7.3	2.4
Aug	0.66	0.68	67.9	66.5	53.6	51.3	0.8	0.0
Sep	0.66	0.68	79.6	75.9	63.1	59	9.5	3.4
Oct	0.70	0.67	94.0	89.5	64.7	59.7	10.8	-0.3
Nov	0.78	0.75	69.6	71.7	53.9	55.9	-10.9	-1.1
Dec	0.66	0.64	48.0	49.0	36.3	34.7	2.5	4.6

Table 4.2: Summary statistics reflecting drop-one cross-validated model performance.

estimate. Conversely, this blending method will yield great improvement for the majority of the months shown in Table 4.2.

To further stress the predictive capability of the blending method, drop-25% validation is carried out, wherein a random sample of 25% of all observations are dropped, the model is fitted on the rest of the observations, and predictions of the dropped points are made. This process is repeated 500 times – for each repetition, root mean square error (RMSE), mean absolute error

(MAE), and percent bias (% Bias) are calculated and shown as boxplots. Drop-50% is also carried out for the test period, which employs the same methodology as drop-25%, except 50% of all observations are dropped at random, thus intensely stressing the predictive capability of the proposed blending method.

Figure 4.7 shows results from drop-25% validation for July 2009; for comparison, statistics of the original CHIRP product are shown as red dots. Similarly, Figure 4.8 shows results from drop-50% validation for July 2009. July 2009 was selected because we aim to illustrate the effectiveness of blending, especially on extreme events, for time scales within the region's wet season. Also, this month is selected to keep consistent with the previous chapters of this research.

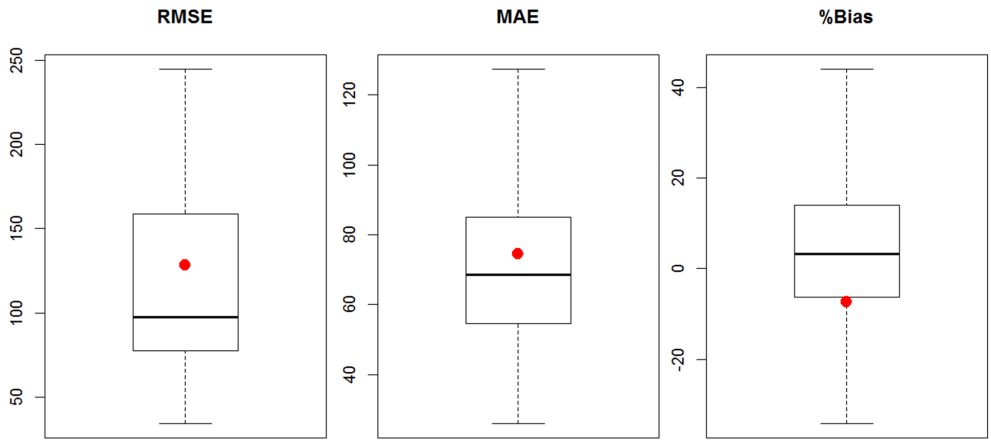

Figure 4.7: Prediction skill measures from drop-25% cross-validation for July 2009. Red dots correspond to estimates from original CHIRP.

It can be seen that, for any month where the cross-validated model shows an improvement, the proposed blending method holds up under the intense stresses of model validation. For both drop-25% and -50% scenarios, the median error statistics (RMSE and MAE) are less than that of the original CHIRP, implying the blending method will improve upon the satellite estimate. Similarly, the inherent bias of the original CHIRP product is reduced to nearly zero for both

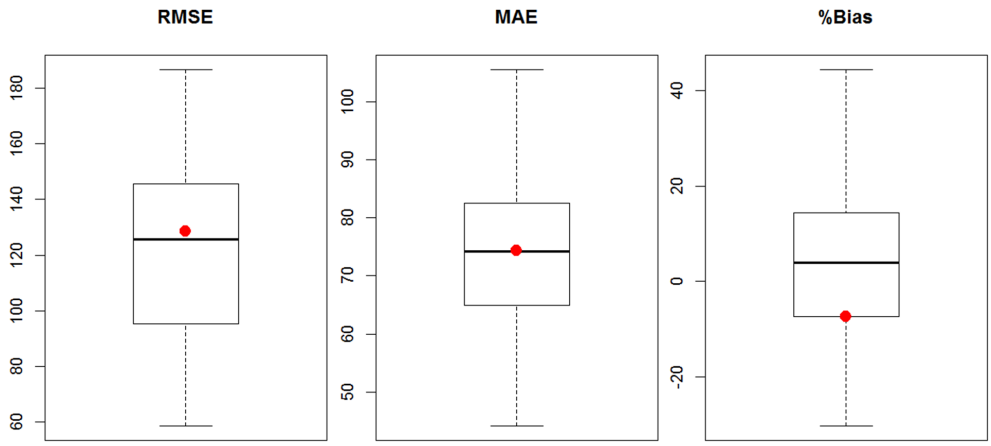

Figure 4.8: Same as 4.7 but from drop-50% cross-validation.

intense validation scenarios, which is encouraging of the skill of the blending method.

Spatial Estimation

Estimates of β and ϵ, as well as blended estimates of precipitation, are made on the grid for all monthly totals in 2009. For illustration this section will focus only on January and July 2009, representative dry and wet months.

January 2009

Figure 4.9 shows the spatial maps of $\hat{\beta}$ and $\hat{\epsilon}$. As expected, the areas of mountainous terrain in Colombia, Pacific Guatemala, and Pacific Costa Rica have a bias correction factor ($\hat{\beta}$) greater than 1, and up to nearly 2, indicating that the CHIRP product underestimates rain gauge measurements in high elevation areas of complex terrain. Meanwhile, the residual field ($\hat{\epsilon}$) indicates that much of the region has a residual field very close to zero, aside from two key regions – Caribbean Nicaragua and Belize, where the CHIRP over- and underestimates rain gauge measurements, respectively. The $\hat{\epsilon}$ process does a good job in matching the values of the observations in these regions, as will be seen in Figure 3.8. The uncertainties in these estimates are quantified in Figure 4.10. It is apparent the

standard error is lowest in regions surrounding observations, and is relatively constant in regions of sparse data.

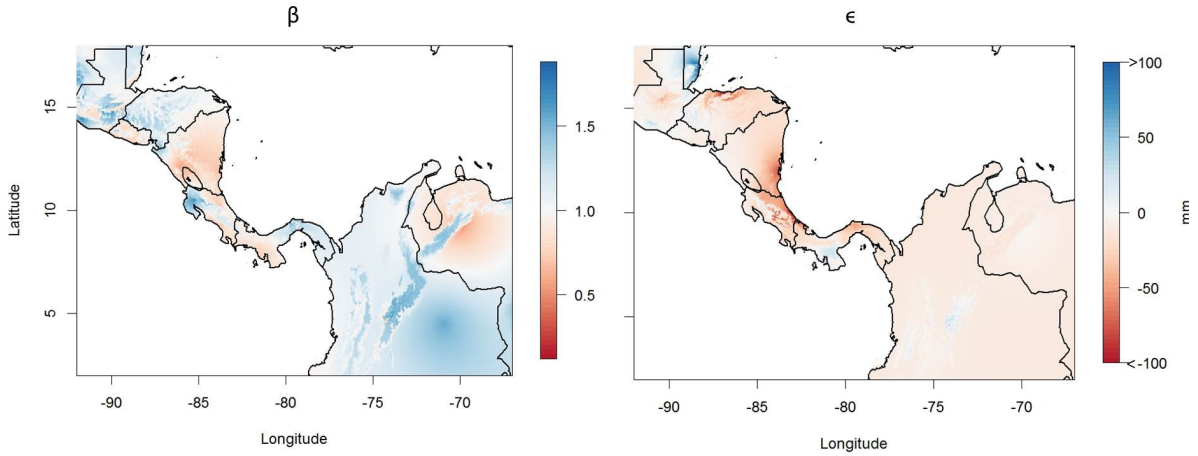

Figure 4.9: Estimates of β (left) and ϵ (right) for January 2009.

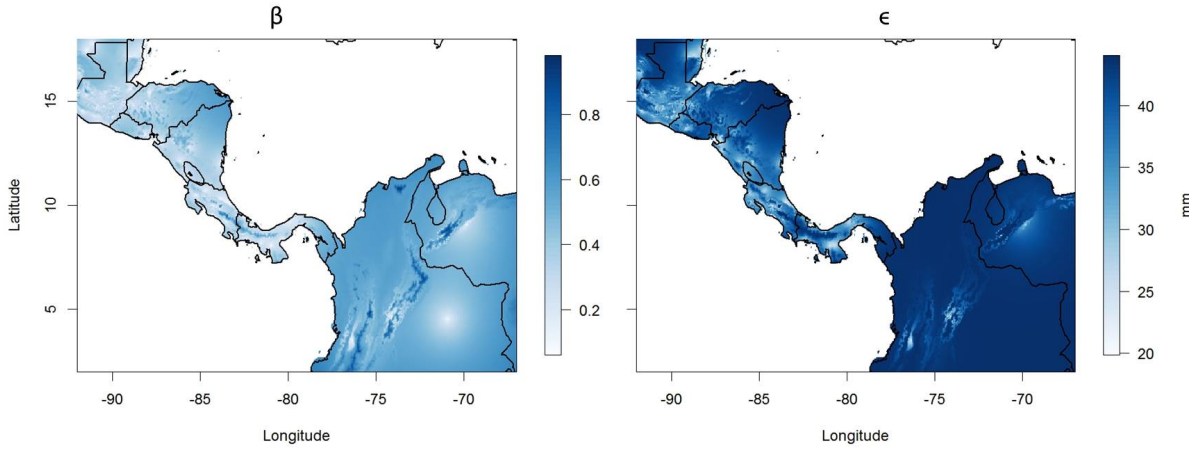

Figure 4.10: Standard error associated with estimation of β (left) and ϵ (right) for January 2009.

The original CHIRP, the blended estimates, and the change in CHIRP (defined as Blended - CHIRP) are shown in Figure 4.11. It is not apparent that blending has much effect for this month, which is due to the fact that the original CHIRP is a relatively good estimator of rainfall for dry months (see Table 4.1). In analyzing the change in CHIRP figure, it is obvious that the CHIRP

overestimates precipitation on the Caribbean coasts of much of Central America, Colombia, and Venezuela. Belize and much of Colombia see an increase in the CHIRP's estimate, implying the CHIRP underestimates rainfall in these regions. These changes are best illustrated in Figure 4.11c, which shows the change in CHIRP due to blending.

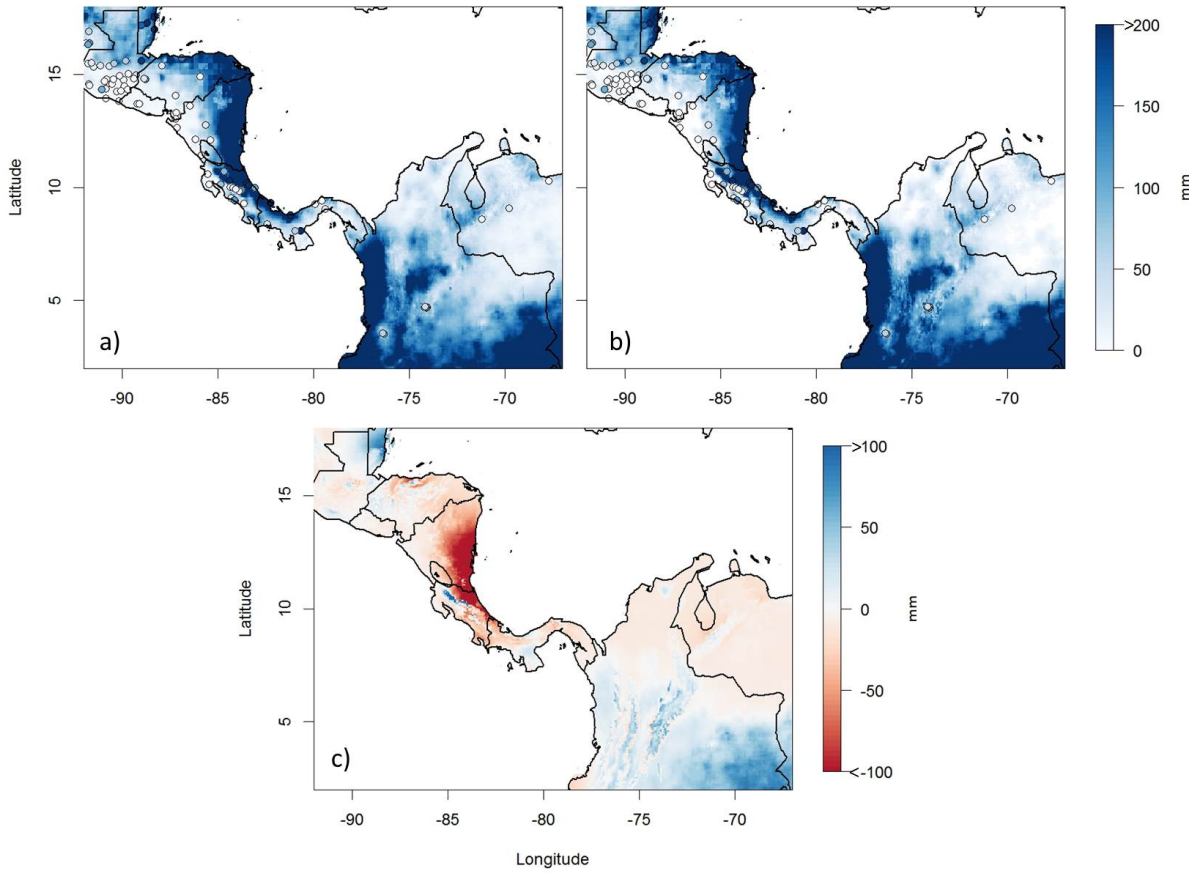

Figure 4.11: a) Original CHIRP estimates, b) blended estimates, and c) change in CHIRP for January 2009. a-b) Rain gauge measurements on same scale for comparison.

July 2009

Shown in Figure 4.12 are spatial maps of $\hat{\beta}$ and $\hat{\epsilon}$. The values of $\hat{\beta}$ are greater than 1 for much of Central America, as well as the mountainous regions of Colombia and Venezuela, indicating underestimation by the CHIRP product. Much of Panama and the flat regions of Colombia and

Venezuela have values of $\hat{\beta}$ less than 1, indicating the CHIRP product overestimates in these regions. The residual field ($\hat{\epsilon}$) indicates that all of Panama and much of Colombia have positive residuals, such that the CHIRP product underestimates in this region. Much of Central America and Venezuela, however, show negative residuals due to the CHIRP product overestimating in these regions. The uncertainties in these estimates are quantified in Figure 4.13. It is apparent the standard error is lowest in regions surrounding observations, and is relatively constant in regions of sparse data.

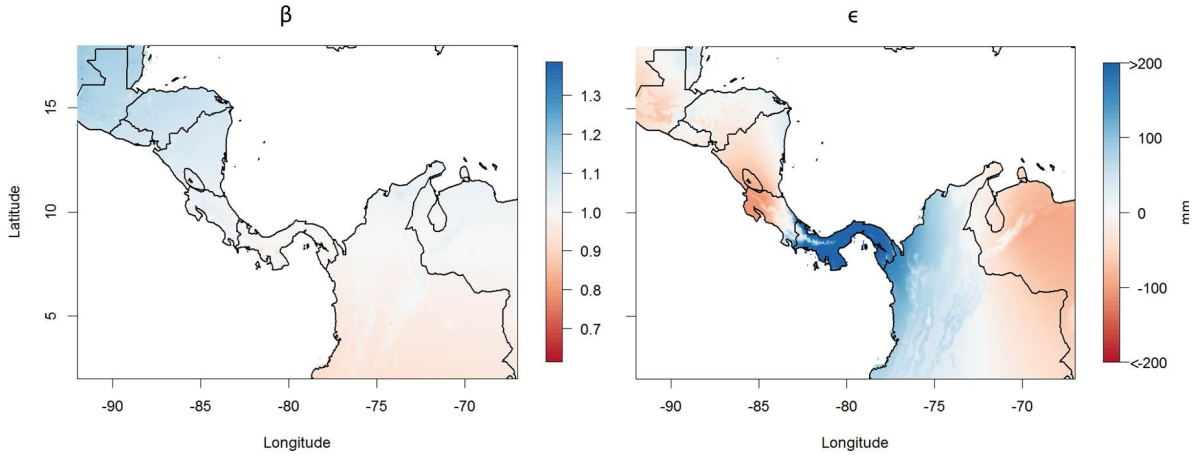

Figure 4.12: Same as Figure 4.9 but for July 2009.

Figure 4.14 shows the original CHIRP, the blended estimates, and the change in CHIRP for July 2009. For this wet season month it is much more obvious that the blended estimate differs from the original CHIRP. The blended estimate reports a lower amount of rainfall in regions of overestimation such as Pacific Costa Rica, Nicaragua, and Venezuela. Similarly, in regions of underestimation, such as Belize, Panama, and much of Colombia, the blended estimate reports more rainfall than the original CHIRP. These modifications are best illustrated in 4.14c, which shows the change in CHIRP due to blending. It is apparent that much of the blended satellite estimate has been adjusted to better match the rain gauge measurements, which are shown as circles with consistent color scale.

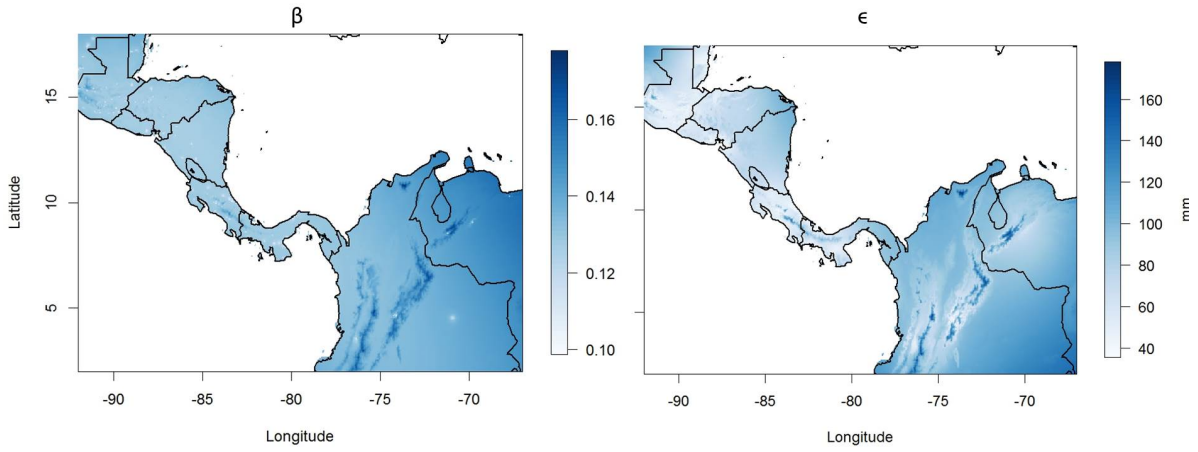

Figure 4.13: Same as Figure 4.10 but for July 2009.

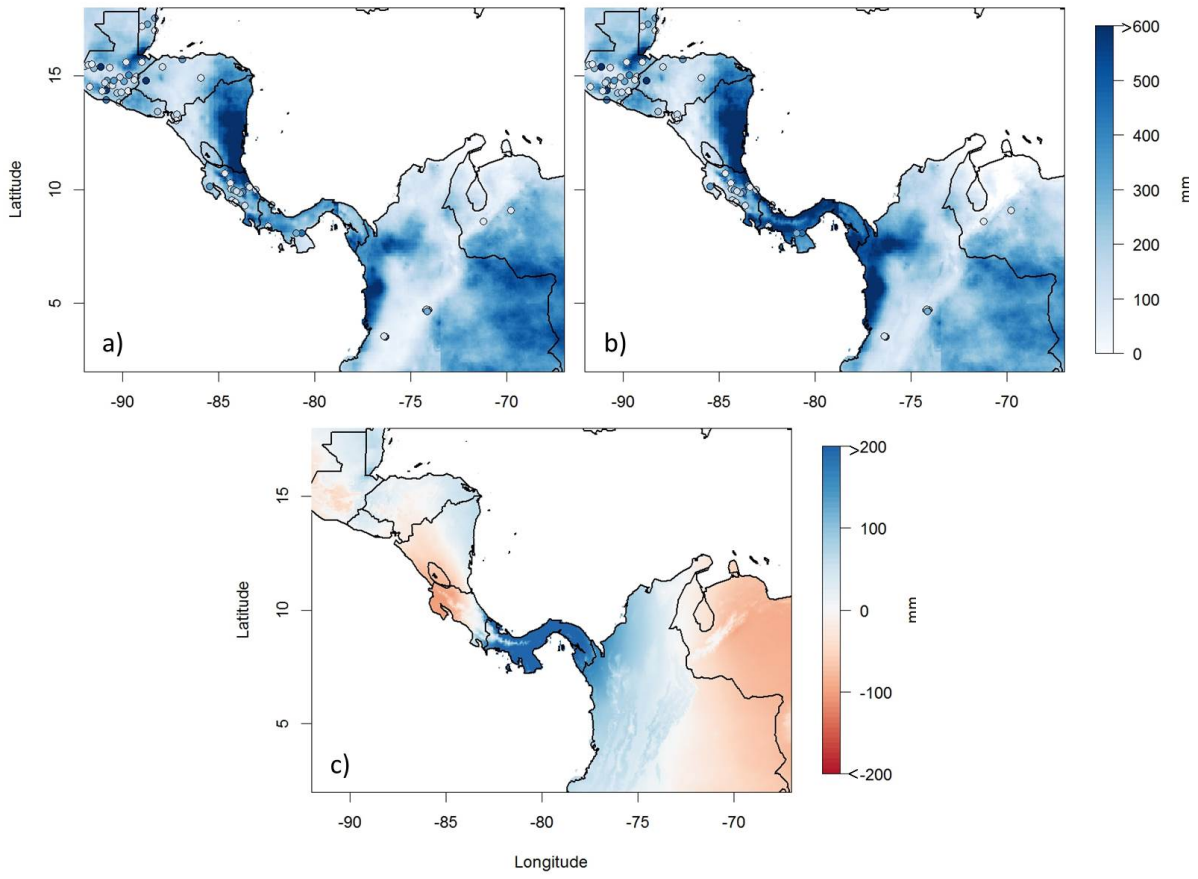

Figure 4.14: Same as Figure 4.11 but for July 2009.

4.4.2 Performance on Extreme Events

Good estimation of extreme events is important for hydroclimatic hazard mitigation, especially during wet months. The hazards associated with extreme events consist of floods, landslides, and agricultural overland flow, all of which are destructive to a nation's infrastructure. We define an extreme event at a location whose monthly rainfall exceeds the 85^{th} percentile of all rain gauge measurements for that month. Of the stations considered in this research (those with a 2/3 complete time series), the 85^{th} percentile for 2009 is 334 mm/month. The locations of these events and the majority of extreme precipitation events occur in the mountainous and coastal regions (Figure 4.15). No extreme stations are found in Colombia because the time series associated with these stations are incomplete. Figure 4.16 shows the performance of the original CHIRP and the blended estimates from the drop-one cross-validated blending model, respectively, for extreme events; Table 4.3 summarizes how this blending method improves performance of the CHIRP satellite estimate with respect to extreme events. These statistics are computed from drop-one cross-validation output.

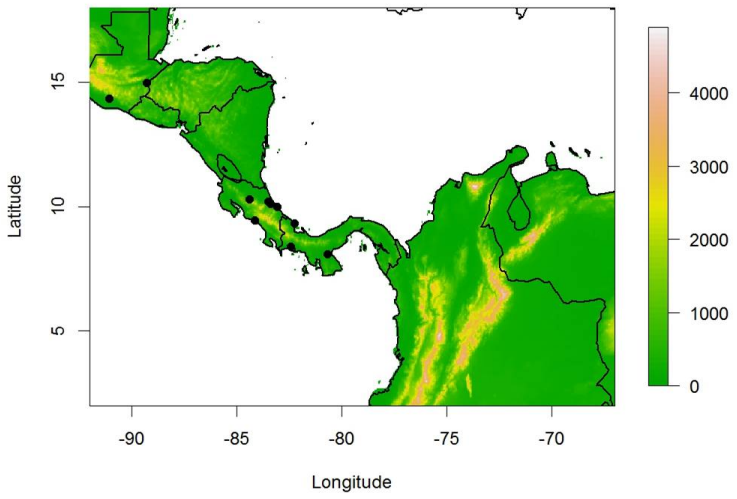

Figure 4.15: Location of extreme precipitation events for July 2009 (black dots) and elevation shown in color.

As can be seen in Figure 4.16 and Table 4.3, the blending method proposed in this chapter has

	RMSE	MAE	% Bias
Original CHIRP	283.5 mm	187.6 mm	-27.8%
Blended Estimate	240.7 mm	166.9 mm	-14.5%

Table 4.3: Summary statistics of blending performance on extreme events for July 2009.

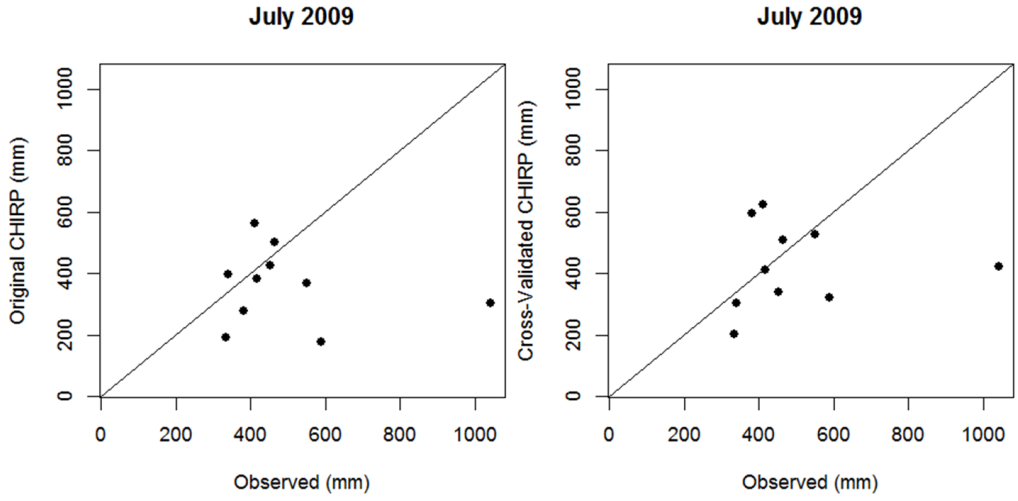

Figure 4.16: Original CHIRP product (left) and blended estimates from cross-validated output for extreme events in July 2009.

a positive impact on the performance of the CHIRP product with respect to extreme precipitation events. Validation statistics from Table 4.3 indicate that the error and bias associated with the original CHIRP product are reduced, even when the blending method is implemented in predictive mode. That being said, for a month where drop-one cross-validation maintains integrity, it is worthwhile to implement this blending method on all available data to produce a more representative gridded time series of monthly precipitation.

4.4.3 Application to Pentad Rainfall Estimation

In this section, we investigate the utility of the blending method to shorter temporal scales, such as pentad rainfall. This is of interest for near-term blended precipitation estimates for natural hazard mitigation strategies. Blended estimates of precipitation on pentad time scales during the

wet season can provide insight into the wetter regions of river basins, can be used to drive hydrologic models for soil moisture and, consequently, the potential for flooding and landslides.

The fourth pentad of August 2009 will be used in this application, as it is one of the wettest pentads of 2009, yielding a maximum observation of 385 mm where the CHIRP reports only 130 mm. The two blending methods as applied to monthly precipitation described in previous sections are applied to this pentadal data. At each observation location, the linear model was fitted using all the pentad rainfall over the 32-year period. Thus the variables β and ϵ are obtained based on potentially 32 observations and the spatial models are subsequently fitted. Figure 4.2 shows the empirical (black dots) and theoretical (red line) variograms of the spatial models of β and ϵ, respectively. Similar to the variograms of the monthly precipitation, that of ϵ exhibits less variability and the theoretical variogram fits the data very well compared to that of β. Figure 4.18 shows that CHIRP estimates are quite poor and underestimate the extreme events. The blending method is a near-exact estimator, thus captures the observations accurately.

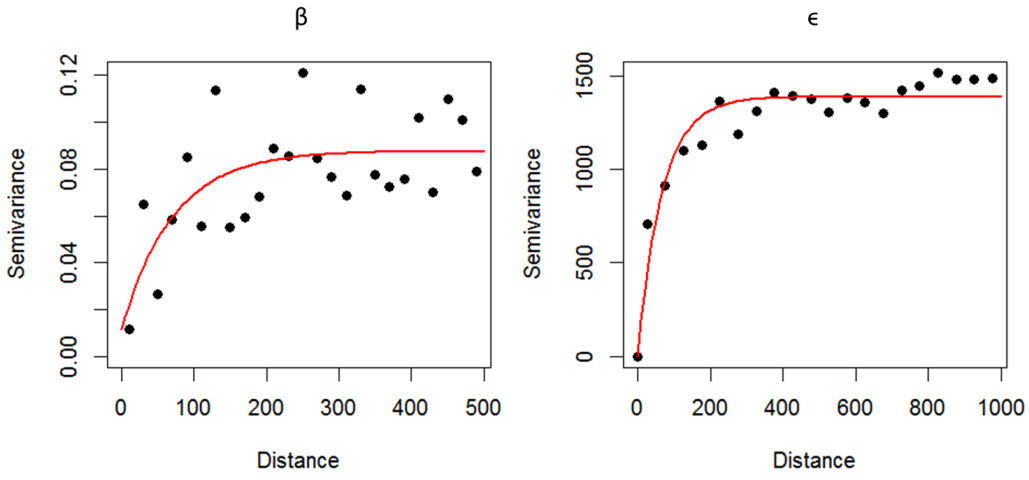

Figure 4.17: Same as Figure 4.2 but for the fourth pentad of August 2009.

As seen earlier, the CHIRP product overestimates regions of no rainfall while underestimating extreme precipitation. For this pentad, the maximum observed rainfall is 385 mm, though the

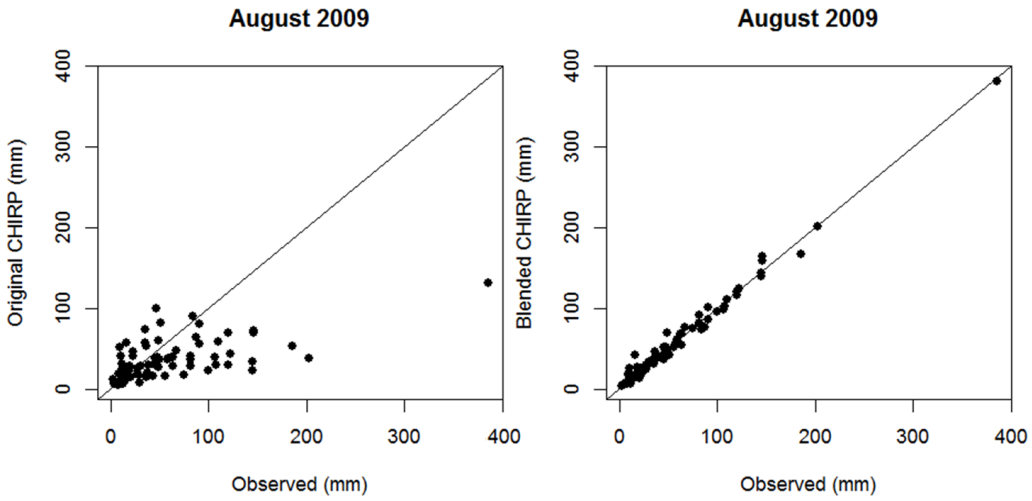

Figure 4.18: Same as Figure 4.3 but for the fourth pentad of August 2009.

CHIRP product reports only 1300 mm of rain the grid cell corresponding to this extreme event. Table 4.4 reports the relative performance of the original CHIRP, as well as the blending method.

	RMSE	MAE	% Bias
Original CHIRP	54.0 mm	34.4 mm	-37.2%
Blended Estimate	7.2 mm	5.0 mm	1.8%

Table 4.4: Same as Table 4.1 but for the fourth pentad of August 2009.

Drop-one cross-validation analysis consistent with that of the monthly totals section was carried out on this representative wet pentad. These predictions are shown as a scatterplot in Figure 4.19, and summary statistics are reported in Table 4.5.

	RMSE	MAE	% Bias
Original CHIRP	54.0 mm	34.4 mm	-37.2%
Blended Estimate	56.0 mm	36.6 mm	1.5%

Table 4.5: Same as Table 4.2 but for the fourth pentad of August 2009.

Under drop-one cross-validation stresses, the error statistics of the blending method is com-

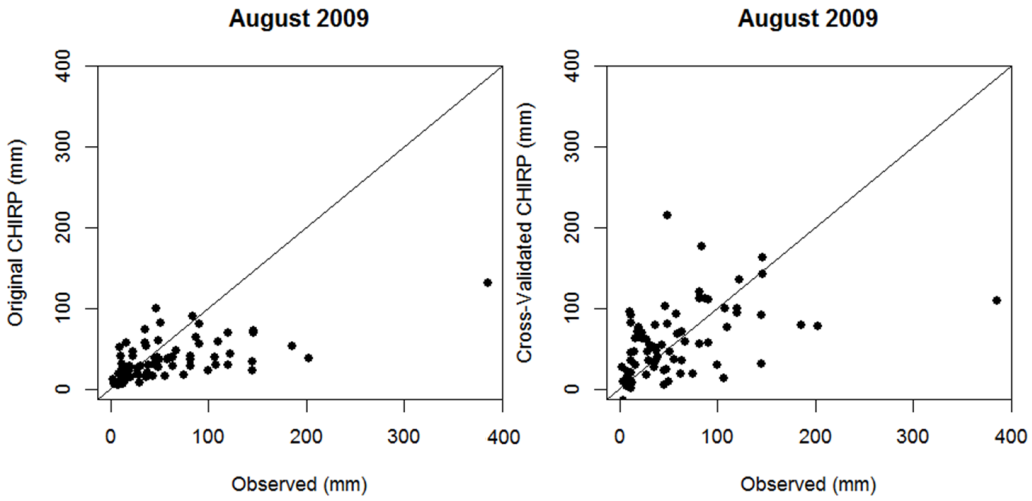

Figure 4.19: Same as Figure 4.5 but for the fourth pentad of August 2009.

parable to that of the original CHIRP product. That being said, this blending method eliminates the large inherent bias of the original CHIRP product even in predictive mode, which is encouraging of the model's predictive skill. Both drop-25% and drop-50% cross-validation techniques were also applied to this pentadal scale, again consistent with the monthly totals section, and statistics such as RMSE, MAE, and % Bias are shown in boxplot form in Figures 4.20 and 4.21. For completeness, the appropriate statistics for the original CHIRP product are shown as red points.

Figures 4.20 and 4.21 show that the blending method performs quit well under the predictive stresses of these validation measures, consistently reducing the RMSE and % Bias of the satellite estimates when compared to the original CHIRP product. It can be seen, however, that the MAE of the original CHIRP product is less than that of the intense validation scenarios. This is due to the blending method's tendency to reduce bias, as well as the noisy spatial structure of rainfall on a pentadal scale.

Implementing this blending method in predictive mode yields spatial maps of $\hat{\beta}$ and $\hat{\epsilon}$, shown in Figure 4.22. From the spatial map of $\hat{\beta}$ it can be seen that the CHIRP product consistently underestimates rainfall magnitude for much of the domain, especially in mountainous areas, due

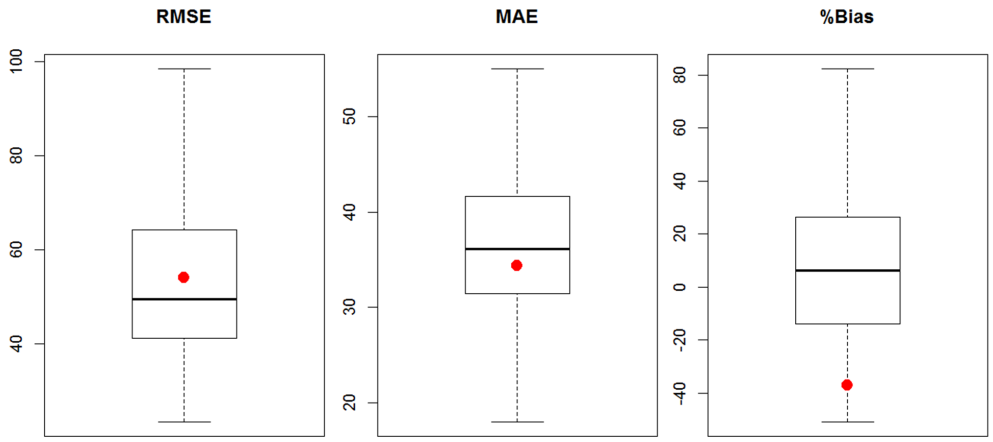

Figure 4.20: Same as Figure 4.7 but for the fourth pentad of August 2009.

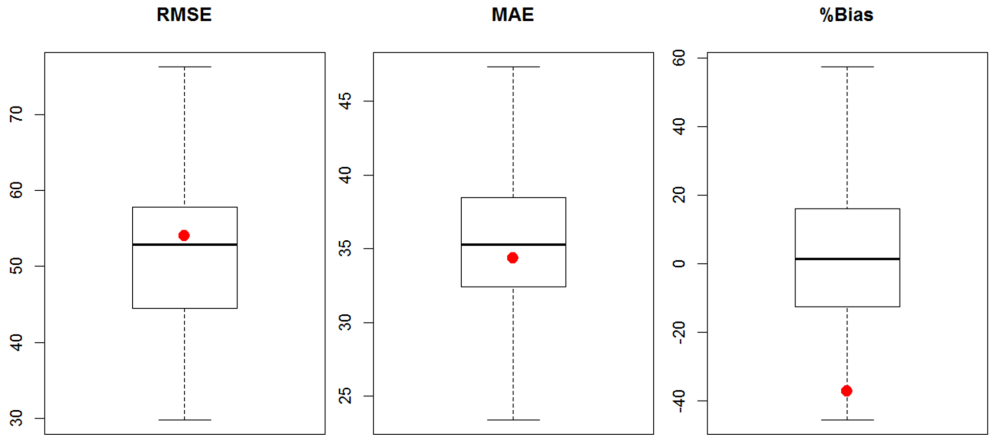

Figure 4.21: Same as Figure 4.8 but for the fourth pentad of August 2009.

to the bias correction factor being greater than 1. There are local deviations to this trend, such as El Salvador, coastal Colombia, and Venezuela. The spatial structure of the residual field $\hat{\epsilon}$ is consistent with that of the $\hat{\beta}$ field – regions with $\hat{\beta} > 1$ have positive residuals and vice versa. The uncertainties in these estimates are quantified in Figure 4.23.

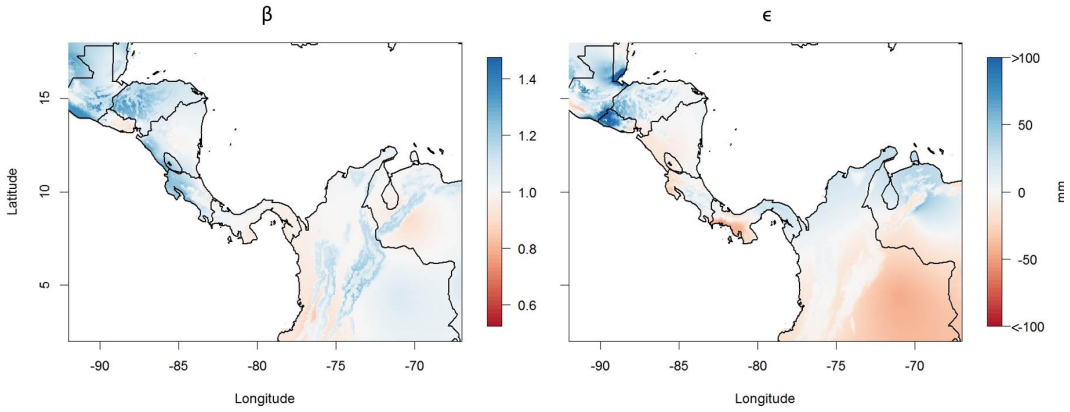

Figure 4.22: Same as Figure 4.9 but for the fourth pentad of August 2009.

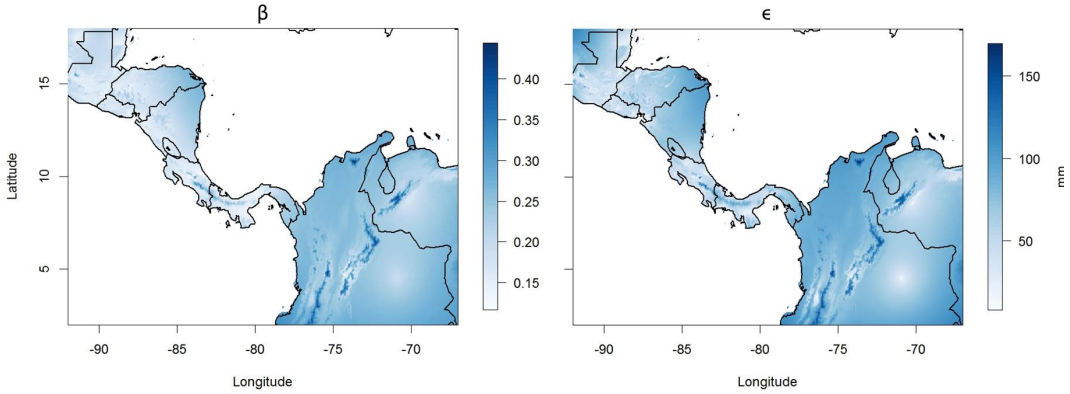

Figure 4.23: Same as Figure 4.10 but for the fourth pentad of August 2009.

The original CHIRP and the blended estimates for this pentad are shown in Figure 4.24a and b. Consistent with blending in a wet season time frame, it is obvious that the blended estimate differs from the original CHIRP. The blended estimate reports a lower amount of rainfall in regions of overestimation such as Pacific Costa Rica and Panama, as well as inland Colombia and Venezuela. Similarly, in regions of underestimation, such as Belize, Guatemala, Honduras, and coastal Colombia and Venezuela, the blended estimate reports more rainfall than the original CHIRP. These modifications are best illustrated in 4.24c, which shows the change in CHIRP due to blending. It is apparent that much of the blended satellite estimate has been adjusted to better

match the rain gauge measurements, which are shown as circles with consistent color scale.

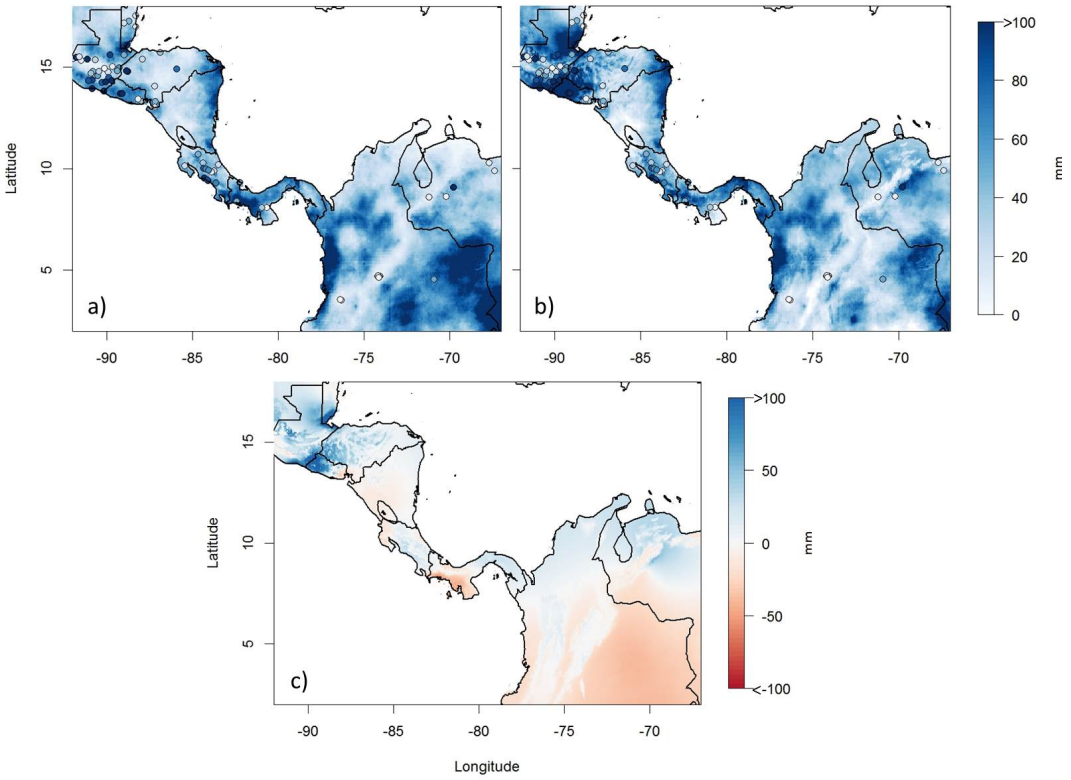

Figure 4.24: Same as Figure 4.11 but for the fourth pentad of August 2009.

To test the ability of this blending method in estimating extremes, as before we define an extreme event at a location where rainfall exceeds the 85^{th} percentile of all rain gauge measurements for the pentad – in this case 107 mm. The locations of these events and the majority of extreme precipitation events occur in the mountainous and coastal regions (Figure 4.25).

Figure 4.26 shows the performance of the original CHIRP as well as output from the blending method with respect to extreme events. Table 4.6 summarizes how the blending method improves the performance of the satellite estimate with respect to extreme events. These statistics – as well as the blended scatterplot – are computed from drop-one cross-validation output.

As can be seen in Figure 4.26 and Table 4.6, the blending method proposed in this chapter is portable from monthly to pentadal scales. It has a positive impact on the CHIRP's performance

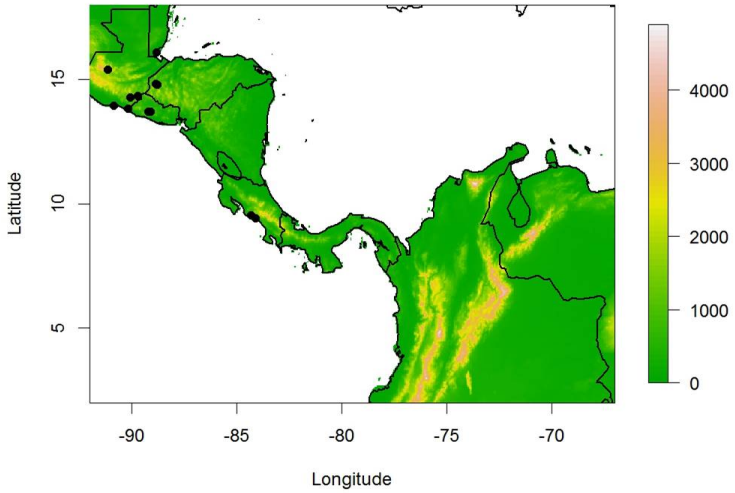

Figure 4.25: Same as Figure 4.15 but for the fourth pentad of August 2009.

	RMSE	MAE	% Bias
Original CHIRP	120.2 mm	107.0 mm	-66.4%
Blended Estimate	100.7 mm	66.6 mm	-38.2%

Table 4.6: Same as Table 4.3 but for the fourth pentad of August 2009.

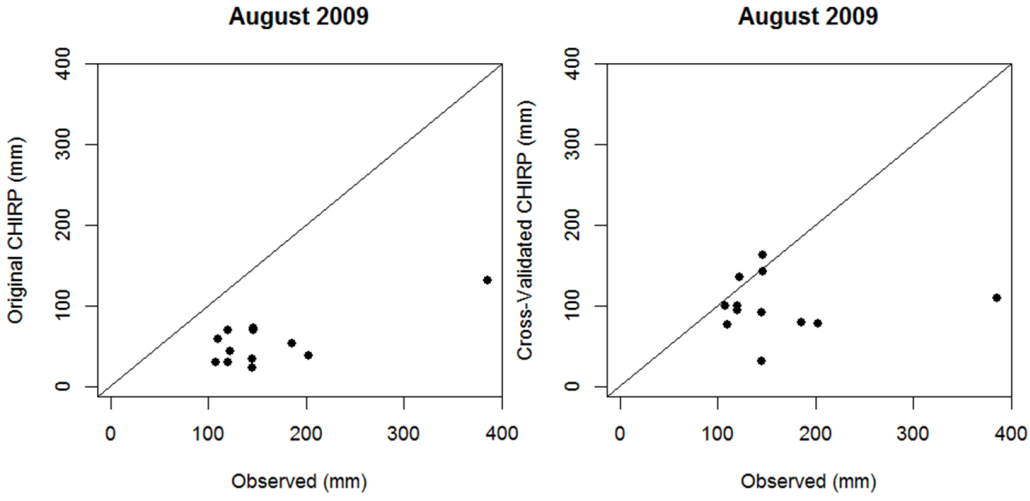

Figure 4.26: Same as Figure 4.16 but for the fourth pentad of August 2009.

with respect to extreme events – validation statistics show that blending reduces the error and inherent bias of the CHIRP satellite estimate, especially of extreme precipitation events. In light of these results, it is encouraged to implement this blending method for all available data in order to produce a more representative gridded time series of precipitation.

4.5 Summary and Conclusions

In this chapter, a unique approach for blending satellite-derived precipitation estimates with in situ rain gauge measurements has been proposed, validated, and applied. This blending technique is a novel approach due to its assumption that satellite-derived precipitation estimates have spatially-varying magnitudes of bias. The benefit of this is a more robust bias correction, resulting in an error field with coherent spatial structure. It has been shown that this technique produces a more representative satellite estimate with respect to rain gauge measurements.

The model fitting process – at the monthly time scale – is most impressive in its error estimation process, which suggests the model will greatly reduce the inherent bias of the original CHIRP product. Validation shows that, even when 50% of the observations are dropped from the model fitting process, the predictive ability of the model is a robust estimator of the true rainfall process. At the pentadal scale, the validation processes produced error statistics relatively consistent with the original CHIRP product, while completely eliminating the great inherent bias. The blended CHIRP's representativeness of extreme events at the pentadal scale is encouraging in its efficacy. It has been shown that this method reduces the error and bias of the original CHIRP product at both monthly and pentadal time scales, especially of extreme events. Therefore it is worthwhile to implement this blending technique, especially during the study region's wet season.

This chapter offers yet another unique approach within the proposed suite of blending methods for solving the problem defined in this research. With these resources, it is possible to retroactively improve satellite estimates of precipitation, thus providing a stronger cornerstone for drought and flood management via forecasting. Blended satellite products resulting from the methods of this research may be used to provide better estimates of initial conditions for hydrologic mod-

els. This research has proven that blending reliable rain gauge measurements with high-resolution precipitation estimates will yield a more representative high-resolution product.

Future work will include implementing this process in a Bayesian Hierarchical framework, which will provide full distributions of the bias correction factor at this fine spatial resolution, which also offers insight into the uncertainty that comes with these estimates. Providing distributions of the error at all grid cells will in turn produce a product with better representation of extreme precipitation events. Another potential blending method to be explored is using the ratio Y_{obs}/Y_{sat} as the dependent variable of an interpolation scheme. While this ratio is similar to the bias correction factor described in this chapter, it is assumed this would account for all the bias as well as the residual, thus correcting the CHIRP in one quick step.

Chapter 5

Summary & Conclusions

This research provides a suite of methods for blending satellite-derived precipitation datasets with in situ rain gauge measurements to assist in providing more accurate gridded precipitation datasets for scientists aimed at forecasting disastrous precipitation events in regions vulnerable to hydroclimatic hazards. These methods are not limited to the study region investigated in this research, but can be applied to any reasonably sized region around the globe. Similarly, it has been shown that these methods are effective not only at coarse (i.e. monthly rainfall) but at fine (i.e. pentadal rainfall) temporal resolution as well, which is useful for obtaining better estimates of initial conditions, driving hydrologic models, and near real-time forecasting.

The second chapter of this research provides a traditional, simplistic approach to blending using both Co-kriging and K-Nearest Neighbor Local Polynomial (LP) models. Because it has been shown that satellite estimates of rainfall are better at capturing the spatial structure of precipitation than rain gauge measurements, the only major assumption for these models is the spatial structure of the satellite error (defined as $r(s) = Y_{obs}(s) - Y_{sat}(s)$) is coherent and the strengths of the respective models are able to fully explain the covariance of the process using location and elevation. It has been shown that this approach to blending is effective in correcting the inherent bias of the satellite estimate, which is especially large during the wet season. Co-kriging is a global spatial model, thus considers all observations within the study region to produce a smoothed map of satellite error that very nearly matches the observations exactly. This model is best suited for regions with a vast network of rain gauge measurement stations, due to the model's

tendency of estimating the expected value of the process. In regions with great variability due to local convection, the expected value of the satellite error will be very close to zero. Therefore this method will only correct the satellite estimate in regions with consistently-sampled rain gauge data. Conversely, Local Polynomial provides local functional estimation, thus considers only a fraction of all observations within the neighborhood of an estimation point, which tends to be much more efficient in representing local nonlinearities in spatial processes. The estimate at an unknown location is obtained by fitting a polynomial of order p using only the $K = \alpha N$ observations within the neighborhood (i.e. convex hull) and using the covariates at the unknown points as predictors. However, this method is susceptible to extrapolation and edge effects – especially for polynomial order 2 or more – thus is not suggested for use in regions with a sparse data network.

A natural extension to the approach described in the second chapter is offered in the third chapter as a unique application of Bayesian Hierarchical Modeling. A hierarchical model is defined, where the true rainfall process at a point is a realization of a normally distributed function defined by $Y_{obs}(s) \sim N(\mu(s), s^2(s))$. The mean of this is obtained by fitting a traditional linear regression model (i.e. $\mu(s) = \beta_1 Y_{sat}(s) + \beta_2 elev(s) + \epsilon(s)$) at each point, thus correcting for bias correction via the β_1 parameter and an elevation-dependent intercept term $\beta_2 elev(s)$, effectively producing a more coherent spatial structure of the residuals $\epsilon(s)$, which are subject to a zero-mean Gaussian process model with parameters σ^2 (marginal variance) and ϕ (effective range). Defining prior distributions on these parameters and implementing a Markov Chain Monte Carlo (MCMC) sampling routine, entire distributions for these parameters are obtained, which allows for the quantification of uncertainty in the true rainfall estimates. Similar to the method described in the second chapter, a kriging scheme is applied to the residual field, thus similar setbacks are associated with this model, namely estimating the expected value (i.e. zero) of the process in regions of sparse data. That being said, the mean function will correct the satellite estimates of the entire domain before implementing the kriging process, and therefore will correct the satellite estimate in these data-sparse regions.

The fourth chapter of this research presents a method similar in nature to that of chapter three but provides a more accessible framework. This method was spurred from the realization that the

bias correction factor (i.e. β_1) obtained from the hierarchical modeling approach is realistically non-uniform over any given study region. To this end, it makes sense to define a hierarchical framework that considers the full time series of data at a point, thus assessing the bias in the satellite estimate by fitting a linear model of the form $Y_{obs}(s,t) = \beta(s)Y_{sat}(s,t) + \epsilon(s,t)$ and extracting the values of $\beta(s)$ and $\epsilon(s,t)$. A Gaussian process model is fitted to the values of $\beta(s)$ using location and elevation as covariates, which produces a time-independent spatial map that will correct bias for the same time frame of any year (i.e. July total rainfall). Because there are year-specific residuals from the linear model used to obtain the values of $\beta(s)$, it follows to implement a separate Gaussian process model to estimate the spatial structure of these residuals. It has been shown that, for many of the months in the test period, this is a robust method for blending ground and satellite observations of rainfall. That being said, before implementing this method an analysis of the drop-one cross-validation performance should be carried out, as the predictive capability of the method is compromised when the satellite estimate is already a good, unbiased estimator of the true rainfall process, typically during the dry season. To address this variable performance, implementing this hierarchical blending method within a Bayesian framework will improve performance of this method, as it quantifies the uncertainty in the parameters as well as the estimated process.

Overall, the method defined in the second chapter requires very little computing power, yet is limited to regions with dense data networks, due to the drawbacks of both models described previously. However, this method is easily implemented in automated mode, iterating through each year and time step separately, thus producing a more representative gridded time series of precipitation. The method of the third chapter performs better than that of the second chapter – reducing both error and bias, especially of extreme precipitation events – but requires much more processing time. This exponential increase in processing time coupled with the snapshot approach (i.e. the posterior distribution of parameters for July 2009 may be notably different than those for July 2010, etc.) limits the automation of this process to those with access to a supercomputer. This realization is what motivated the approach presented in the fourth chapter of this research. While it has been shown that the method described in the fourth chapter shows variable performance,

those time frames where the satellite estimate is a poor estimator of the true rainfall process will benefit greatly from this method. Careful attention to the relative performance of the original satellite estimate is of key importance.

This study has been motivated by the need for a more accurate gridded time series of precipitation data in order to further the efforts of scientists aimed at forecasting disastrous precipitation events within regions vulnerable to hydroclimatic hazards. The methods described in this research are proven effective in that they reduce the error as well as the bias associated with the original satellite estimate, thus improving the representativeness of these gridded time series products. Application of any of these blending methods to the history of the CHIRP satellite estimate will provide a more stable platform for hydroclimatic hazard mitigation and forecasting.

Bibliography

Abe, M. "Counting Your Customers" One by One: A Hierarchical Bayes Extension to the Pareto/NBD Model. Marketing Science, 28(3):541–553, March 2009. ISSN 0732-2399. doi: 10.1287/mksc.1090.0502. URL http://mktsci.journal.informs.org/cgi/doi/10.1287/mksc.1090.0502.

Adler, RF; Huffman, GJ; Chang, Alfred; Ferraro, Ralph; Xie, Pingping; Janowiak, John; Rudolf, Bruno; Schneider, Udo; Curtis, Scott; Bolvin, David; Gruber, Arnold; Susskind, Joel; Arkin, Philip, and Nelkin, Eric. The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present). Journal of Hydrometeorology, December:1147–1167, 2003. URL http://journals.ametsoc.org/doi/abs/10.1175/1525-7541(2003)004%3C1147:TVGPCP%3E2.0.CO;2.

Aelion, C M; Davis, H T; Liu, Y; Lawson, a B, and McDermott, S. Validation of Bayesian Kriging of Arsenic, Chromium, Lead, and Mercury Surface Soil Concentrations Based on Internode Sampling. Environmental Science & Technology, 43(12):4432–8, June 2009. ISSN 0013-936X. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2755059&tool=pmcentrez&rendertype=a

Allenby, Greg M and Rossi, Peter E. Marketing models of consumer heterogeneity. Journal of Econometrics, 89:57–78, 1999.

Apaydin, Halit; Sonmez, F Kemal, and Yildirim, Y Ersoy. Spatial interpolation techniques for climate data in the GAP region in Turkey. Climate Research, 28:31–40, 2004.

Barsugli, Joseph J.; Nowak, Kenneth; Rajagopalan, Balaji; Prairie, James R., and Harding, Benjamin. Comment on When will Lake Mead go dry? by T. P. Barnett and D. W. Pierce. Water Resources Research, 45(9):1–6, September 2009. ISSN 0043-1397. doi: 10.1029/2008WR007627. URL http://www.agu.org/pubs/crossref/2009/2008WR007627.shtml.

Berliner, L. Mark; Wikle, Christopher K., and Cressie, Noel. Long-Lead Prediction of Pacific SSTs via Bayesian Dynamic Modeling. Journal of Climate, 13(22):3953–3968, November 2000. ISSN 0894-8755. doi: 10.1175/1520-0442(2001)013¡3953:LLPOPS¿2.0.CO;2. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0442%282001%29013%3C3953%3ALLPOPS%3E2.0.CO%3

Biau, Gerard; Zorita, Eduardo; von Storch, Hans, and Wackernagel, Hans. Estimation of Precipitation by Kriging in the EOF Space of the Sea Level Pressure Field. Journal of Climate, 12: 1070–1085, 1999.

Block, Paul and Rajagopalan, Balaji. Interannual Variability and Ensemble Forecast of Upper Blue Nile Basin Kiremt Season Precipitation. Journal of Hydrometeorology, 8(3):327–343, June 2007. ISSN 1525-755X. doi: 10.1175/JHM580.1. URL http://journals.ametsoc.org/doi/abs/10.1175/JHM580.1.

Borsuk, Mark E.; Higdon, David; Stow, Craig a., and Reckhow, Kenneth H. A Bayesian hierarchical model to predict benthic oxygen demand from organic matter loading in estuaries and coastal zones. Ecological Modelling, 143(3):165–181, September 2001. ISSN 03043800. doi: 10.1016/S0304-3800(01)00328-3. URL http://linkinghub.elsevier.com/retrieve/pii/S0304380001003283.

Bracken, Cameron; Rajagopalan, Balaji, and Prairie, James. A multisite seasonal ensemble streamflow forecasting technique. Water Resources Research, 46(3), March 2010. ISSN 00431397. doi: 10.1029/2009WR007965. URL http://doi.wiley.com/10.1029/2009WR007965.

Chen, Jinsong and Hubbard, Susan. Estimating the hydraulic conductivity at the South Oyster Site from geophysical tomographic data using Bayesian techniques based on the normal linear regression model displays variation Oyster Site. Water Resources Research, 37(6):1603–1613, 2001.

Clark, Martyn P. and Slater, Andrew G. Probabilistic Quantitative Precipitation Estimation in Complex Terrain. Journal of Hydrometeorology, 7:3–22, 2006.

Cleveland, William S; Devlin, Susan J, and Cleveland, S. Locally Weighted Regression : An Approach to Regression Analysis by Local Fitting. 83(403):596–610, 2010.

Cooley, Daniel and Sain, Stephan R. Spatial Hierarchical Modeling of Precipitation Extremes From a Regional Climate Model. Journal of Agricultural, Biological, and Environmental Statistics, 15(3):381–402, March 2010. ISSN 1085-7117. doi: 10.1007/s13253-010-0023-9. URL http://link.springer.com/10.1007/s13253-010-0023-9.

Cooley, Daniel; Nychka, Douglas, and Naveau, Philippe. Bayesian Spatial Modeling of Extreme Precipitation Return Levels. Journal of the American Statistical Association, 102 (479):824–840, September 2007. ISSN 0162-1459. doi: 10.1198/016214506000000780. URL http://www.tandfonline.com/doi/abs/10.1198/016214506000000780.

Craven, P and Wahba, G. Smoothing noisy data with spline functions. Numerische Mathematik, 1978. URL http://link.springer.com/article/10.1007/BF01404567.

Cressie, Noel. The Origins of Kriging. Mathematical Geology, 22(3):239–252, April 1990. ISSN 0882-8121. doi: 10.1007/BF00889887. URL http://link.springer.com/10.1007/BF00889887.

Cui, Haiyan; Stein, Alfred, and Myers, D O N E. EXTENSION OF SPATIAL INFORMATION , BAYESIAN KRIGING AND UPDATING OF PRIOR VARIOGRAM. Environmetrics, 6(December 1993):373–384, 1995.

Curtis, Scott and Adler, Robert. ENSO indices based on patterns of satellite-derived precipitation. Journal of Climate, 13:2786–2793, 2000. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0442(2000)013%3C2786%3AEIBOPO%3E2.0.CO%3B2.

Dai, Aiguo; Wigley, T.M.L.; Boville, B.A.; Kiehl, J.T., and Buja, L.E. Climates of the Twentieth and Twenty-First Centuries Simulated by the NCAR Climate System Model. Journal of Climate, 14:485–519, 2001. URL http://journals.ametsoc.org/doi/full/10.1175/1520-0442(2001)014%3C0485:COTTAT%3E2.0.CO%3B2.

Desbarats, AJ; Logan, CE; Hinton, MJ, and Sharpe, DR. On the kriging of water table elevations using collateral information from a digital elevation model. Journal of Hydrology, 255:25–38, 2002. URL http://www.sciencedirect.com/science/article/pii/S0022169401005042.

Finley, Andrew O; Sang, Huiyan; Banerjee, Sudipto, and Gelfand, Alan E. Improving the performance of predictive process modeling for large datasets. Computational Statistics & Data Analysis, 53(8):2873–2884, June 2009. ISSN 0167-9473. doi: 10.1016/j.csda.2008.09.008. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2743161&tool=pmcentrez&rendertype=a

Furrer, Reinhard; Sain, Stephan R.; Nychka, Douglas, and Meehl, Gerald A. Multivariate Bayesian analysis of atmosphereocean general circulation models. Environmental and Ecological Statistics, 14(3):249–266, July 2007. ISSN 1352-8505. doi: 10.1007/s10651-007-0018-z. URL http://link.springer.com/10.1007/s10651-007-0018-z.

Gallagher, Kerry; Charvin, Karl; Nielsen, Soren; Sambridge, Malcolm, and Stephenson, John. Markov chain Monte Carlo (MCMC) sampling methods to determine optimal models, model resolution and model choice for Earth Science problems. Marine and Petroleum Geology, 26(4):525–535, April 2009. ISSN 02648172. doi: 10.1016/j.marpetgeo.2009.01.003. URL http://linkinghub.elsevier.com/retrieve/pii/S0264817209000075.

Goovaerts, P. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. Journal of Hydrology, 228(1-2):113–129, February 2000. ISSN 00221694. doi: 10.1016/S0022-1694(00)00144-X. URL http://linkinghub.elsevier.com/retrieve/pii/S002216940000144X.

Grantz, Katrina; Rajagopalan, Balaji; Clark, Martyn, and Zagona, Edith. A technique for incorporating large-scale climate information in basin-scale ensemble streamflow forecasts. Water Resources Research, 41(10), October 2005. ISSN 00431397. doi: 10.1029/2004WR003467. URL http://doi.wiley.com/10.1029/2004WR003467.

Handcock, MS and Stein, ML. A Bayesian analysis of kriging. Technometrics, 35(4):403–410, 1993. URL http://amstat.tandfonline.com/doi/abs/10.1080/00401706.1993.10485354.

Hartkamp, A Dewi; Beurs, Kirsten De; Stein, Alfred, and White, Jeffrey W. Interpolation Techniques for Climate Variables Interpolation Techniques for. NRG-GIS, 99(01), 1999.

Helsel, D.R. and Hirsch, R.M. Statistical methods in water resources. Elsevier, New York, 1995. URL http://books.google.com/books?hl=en&lr=&id=jao4o5X1pvgC&oi=fnd&pg=PP2&dq=Statistical+Met
http://books.google.com/books?hl=en&lr=&id=jao4o5X1pvgC&oi=fnd&pg=PP2&dq=Statistical+methods-

Hoef, JM Ver and Frost, K.J. A Bayesian hierarchical model for monitoring harbor seal changes in Prince William Sound, Alaska. Environmental and Ecological Statistics, 10:201–219, 2003. URL http://link.springer.com/article/10.1023/A:1023626308538.

Holdaway, Margaret R. Spatial modeling and interpolation of monthly temperature using kriging. Climate Research, 6:215–225, 1996.

Huffman, George J.; Adler, Robert F.; Rudolf, Bruno; Schneider, Udo, and Keehn, Peter R. Global Precipitation Estimates Based on a Technique for Combining Satellite-Based Estimates, Rain Gauge Analysis, and NWP Model Precipitation Information. Journal of Climate, 8:1284–1295, 1995.

Huffman, George J.; Bolvin, David T.; Nelkin, Eric J.; Wolff, David B.; Adler, Robert F.; Gu, Guojun; Hong, Yang; Bowman, Kenneth P., and Stocker, Erich F. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. Journal of Hydrometeorology, 8(1):38–55, February 2007. ISSN 1525-755X. doi: 10.1175/JHM560.1. URL http://journals.ametsoc.org/doi/abs/10.1175/JHM560.1.

Janowiak, John E; Gruber, Arnold; Kondragunta, C.R.; Livezey, Robert E., and Huffman, George J. A Comparison of the NCEP-NCAR Reanalysis Precipitation and the GPCP Rain Gauge-Satellite Combined Dataset with Observational Error Considerations. Journal of Climate, 11:2960–2979, 1998. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0442(1998)011%3C2960:ACOTNN%3E2.0.CO%3B2.

Joyce, RJ and Janowiak, JE. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. Journal of Hydrometeorology, 5:487–503, 2004. URL http://journals.ametsoc.org/doi/pdf/10.1175/1525-7541(2004)005%3C0487%3ACAMTPG%3E2.0.CO%3B2.

Kennedy, Marc C. and O'Hagan, Anthony. Bayesian calibration of computer models. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63 (3):425–464, August 2001. ISSN 1369-7412. doi: 10.1111/1467-9868.00294. URL http://doi.wiley.com/10.1111/1467-9868.00294.

Kleiber, William; Katz, Richard W., and Rajagopalan, Balaji. Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes. Water Resources Research, 48(1):1–17, January 2012. ISSN 0043-1397. doi: 10.1029/2011WR011105. URL http://www.agu.org/pubs/crossref/2012/2011WR011105.shtml.

Kuczera, George; Kavetski, Dmitri; Franks, Stewart, and Thyer, Mark. Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-dependent parameters. Journal of Hydrology, 331(1-2):161–177, November 2006. ISSN 00221694. doi: 10.1016/j.jhydrol.2006.05.010. URL http://linkinghub.elsevier.com/retrieve/pii/S0022169406002745.

Lall, U. Recent advances in nonparametric function estimation: Hydrologic applications. Reviews of Geophysics, (July):1093–1102, 1995. URL http://onlinelibrary.wiley.com/doi/10.1029/95RG00343/full.

Lall, Upmanu; Moon, Young-Il, and Bosworth, Ken. Locally weighted polynomial regression: Parameter choice and application to forecasts of the Great Salt Lake. Reports, Paper 243, 1995. URL http://onlinelibrary.wiley.com/doi/10.1029/2004WR003782/full.

Le, Nhu D and Zidek, James V. Interpolation with Uncertain Spatial Covariances: A Bayesian Alternative to Kriging. Journal of Multivariate Analysis, 43(2):351–374, November 1992. ISSN 0047259X. doi: 10.1016/0047-259X(92)90040-M. URL http://linkinghub.elsevier.com/retrieve/pii/0047259X9290040M.

Li, Ming and Shao, Quanxi. An improved statistical approach to merge satellite rainfall estimates and raingauge data. Journal of Hydrology, 385(1-4):51–64, May 2010. ISSN 00221694. doi: 10.1016/j.jhydrol.2010.01.023. URL http://linkinghub.elsevier.com/retrieve/pii/S0022169410000715.

Liu, Yuqiong and Gupta, Hoshin V. Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. Water Resources Research, 43(7):1–18, July 2007. ISSN 00431397. doi: 10.1029/2006WR005756. URL http://doi.wiley.com/10.1029/2006WR005756.

Loader, C. Locfit: An Introduction1. (x), 1997. URL http://www.stat.uni-muenchen.de/ leiten/Lehre/Material/GLM_0708/Tutorium/locfit.pdf.

Loader, Clive. Local regression and likelihood, volume 42. Springer, New York, November 1999. doi: 10.2307/1270956. URL http://www.jstor.org/stable/1270956?origin=crossref http://cyber.sibsutis.ru:82/Monarev/docs/nauka/PROBABILITY/MVsa_Statistics and applications/Loader C. Local Regression and Likelihood (Springer, 1999)(305s).pdf.

Omre, Henning. Bayesian KrigingMerging Observations and Qualified Guesses in Kriging. Mathematical Geology, 19(1):25–39, 1987. URL http://link.springer.com/article/10.1007/BF01275432.

Prairie, JR and Rajagopalan, Balaji. Statistical nonparametric model for natural salt estimation. Journal of Environmental Engineering, (January):130–138, 2005. doi: 10.1061/(ASCE)0733-9372(2005)131. URL http://ascelibrary.org/doi/pdf/10.1061/(ASCE)0733-9372(2005)131%3A1(130).

Prairie, JR and Rajagopalan, Balaji. Modified K-NN model for stochastic streamflow simulation. Journal of Hydrologic Engineering, (August):371–378, 2006. URL http://ascelibrary.org/doi/abs/10.1061/(ASCE)1084-0699(2006)11:4(371).

R Development Core Team, . R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2011. URL http://www.R-project.org/. ISBN 3-900051-07-0.

Raftery, Adrian E.; Gneiting, Tilmann; Balabdaoui, Fadoua, and Polakowski, Michael. Using Bayesian Model Averaging to Calibrate Forecast Ensembles. American Meteorological Society, 133(5):1155–1174, May 2005. ISSN 0027-0644. doi: 10.1175/MWR2906.1. URL http://journals.ametsoc.org/doi/abs/10.1175/MWR2906.1.

Rajagopalan, Balaji and Lall, Upmanu. Locally weighted polynomial estimation of spatial precipitation. Journal of Geographic Information and Decision Analysis, 2(2):44–51, 1998. URL http://civil.colorado.edu/ balajir/my-papers/lpoly.pdf.

Rajagopalan, Balaji; Grantz, Katrina; Regonda, Satish; Clark, Martyn, and Zagona, Edith. Ensemble streamflow forecasting: Methods and applications, 2005.

Regonda, Satish Kumar; Rajagopalan, Balaji; Clark, Martyn, and Zagona, Edith. A multimodel ensemble forecast framework: Application to spring seasonal flows in the Gunnison River Basin. Water Resources Research, 42(9), September 2006. ISSN 00431397. doi: 10.1029/2005WR004653. URL http://doi.wiley.com/10.1029/2005WR004653.

Sahu, Sujit K. and Mardia, Kanti V. A Bayesian kriged Kalman model for short-term forecasting of air pollution levels. Journal of the Royal Statistical Society: Series C (Applied Statistics), 54(1):223–244, January 2005. ISSN 0035-9254. doi: 10.1111/j.1467-9876.2005.00480.x. URL http://doi.wiley.com/10.1111/j.1467-9876.2005.00480.x.

Sang, Huiyan and Gelfand, Alan E. Hierarchical modeling for extreme values observed over space and time. Environmental and Ecological Statistics, 16(3):407–426, January 2008. ISSN 1352-8505. doi: 10.1007/s10651-007-0078-0. URL http://link.springer.com/10.1007/s10651-007-0078-0.

Schaap, Marcel G.; Leij, Feike J., and van Genuchten, Martinus Th. Rosetta: a Computer Program for Estimating Soil Hydraulic Parameters With Hierarchical Pedotransfer Functions. Journal of Hydrology, 251(3-4):163–176, October 2001. ISSN 00221694. doi: 10.1016/S0022-1694(01)00466-8. URL http://linkinghub.elsevier.com/retrieve/pii/S0022169401004668.

Schliep, Erin M.; Cooley, Daniel; Sain, Stephan R., and Hoeting, Jennifer a. A comparison study of extreme precipitation from six different regional climate models via spatial hierarchical modeling. Extremes, 13(2):219–239, December 2009. ISSN 1386-1999. doi: 10.1007/s10687-009-0098-2. URL http://link.springer.com/10.1007/s10687-009-0098-2.

Tebaldi, Claudia and Sansó, Bruno. Joint projections of temperature and precipitation change from multiple climate models: a hierarchical Bayesian approach. Journal of the Royal Statistical Society: Series A (Statistics in Society), 172(1):83–106, January 2009. ISSN 09641998. doi: 10.1111/j.1467-985X.2008.00545.x. URL http://doi.wiley.com/10.1111/j.1467-985X.2008.00545.x.

Thyer, Mark; Renard, Benjamin; Kavetski, Dmitri; Kuczera, George; Franks, Stewart William, and Srikanthan, Sri. Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis. Water Resources Research, 45(12):1–22, December 2009. ISSN 00431397. doi: 10.1029/2008WR006825. URL http://doi.wiley.com/10.1029/2008WR006825.

Towler, Erin; Rajagopalan, Balaji, and Summers, R Scott. Using Parametric and Nonparametric Methods to Model Total Organic, Alkalinity, and pH after Conventional Surface Water Treatment. Environmental Engineering Science, 26(8):1299–1308, 2009.

Towler, Erin; Rajagopalan, Balaji; Gilleland, Eric; Summers, R Scott; Yates, David, and Katz, Richard W. Modeling hydrologic and water quality extremes in a changing climate: A statistical approach based on extreme value theory. Water Resources Research, 46(11), 2010a.

Towler, Erin; Rajagopalan, Balaji; Summers, R. Scott, and Yates, David. An approach for probabilistic forecasting of seasonal turbidity threshold exceedance. Water Resources Research, 46(6), June 2010b. ISSN 00431397. doi: 10.1029/2009WR007834. URL http://doi.wiley.com/10.1029/2009WR007834.

Trenberth, KE and Caron, JM. The Southern Oscillation Revisited: Sea Level Pressures, Surface Temperatures, and Precipitation. Journal of Climate, 13:4358–4365, 2000. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0442(2000)013%3C4358%3ATSORSL%3E2.0.CO%3B2.

Vrugt, Jasper a.; Bouten, Willem; Gupta, Hoshin V., and Sorooshian, Soroosh. Toward improved identifiability of hydrologic model parameters: The information content of experimental data. Water Resources Research, 38(12):48–1–48–13, December 2002. ISSN 00431397. doi: 10.1029/2001WR001118. URL http://doi.wiley.com/10.1029/2001WR001118.

Šimůnek, J; van Genuchten, Martinus Th.; Gribb, Molly M., and Hopmans, Jan W. Parameter estimation of unsaturated soil hydraulic properties from transient flow processes. Soil and Tillage Research, 47:27–36, 1998. URL http://www.sciencedirect.com/science/article/pii/S0167198798000695.

Wikle, CK. Hierarchical Bayesian models for predicting the spread of ecological processes. Ecology, 84(6):1382–1394, 2003. URL http://www.esajournals.org/doi/abs/10.1890/0012-9658(2003)084%5B1382:HBMFPT%5D2.0.CO%3B2.

Xie, Pingping and Arkin, PA. Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. Journal of Climate, 9:840–858, 1996. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0442(1996)009%3C0840:AOGMPU%3E2.0.CO;2.

Xie, Pingping and Arkin, PA. Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. Bulletin of the American Meteorological Society, 78(11):2539–2558, 1997. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0477(1997)078%3C2539%3AGPAYMA%3E2.0.CO%3B2.

Xie, Pingping and Arkin, PA. Global monthly precipitation estimates from satellite-observed outgoing longwave radiation. Journal of Climate, 11:137–164, 1998. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0442(1998)011%3C0137:GMPEFS%3E2.0.CO%3B2.