MINIMIZING BIASES IN RADAR PRECIPITATION ESTIMATES

A Dissertation

by

DOUGLAS BRENT MCROBERTS

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	John W. Nielsen-Gammon
Committee Members,	Gerald R. North
	Courtney Schumacher
	Steven M. Quiring
Head of Department,	Ping Yang

December 2014

Major Subject: Atmospheric Sciences

ABSTRACT

The demand for real-time drought information in recent years led to the development of a suite of objective drought indicators that relies on the high-resolution Stage IV precipitation estimates that are produced each day by the National Weather Service in near real-time. The drawback to using the Stage IV dataset for this purpose is the presence of numerous biases in the estimates, which lead to erroneous assessments of drought conditions. Among the types of biases in the Stage IV dataset are

1. Underestimation of precipitation due to beam blockage.
2. Range-dependent errors that originating from the measurement of reflectivity above the surface.
3. Mean-field biases resulting from radar calibration and measurement errors.

A three stage bias correction procedure is developed and evaluated for minimizing the biases, methods used to produce an improved, bias-adjusted Stage IV precipitation dataset. The original Stage IV data are initially corrected by a beam blockage identification procedure and Kriging interpolation to replace the precipitation values in grid cells affected by blockage. Next, range-dependent and mean field biases are identified and corrected by use of a statistical model based on the vertical profile of reflectivity in mixed-phase precipitating systems. The last bias quantification procedure estimates and removes a two-dimensional field of residual biases using available gauges as an assumed unbiased estimate of the ground truth.

Data withholding testing showed the bias-adjusted Stage IV dataset to have a significant reduction in the overall bias relative to the original Stage IV precipitation dataset. This includes a reduction in the overall bias at each of the three major steps. The bias-adjusted Stage IV dataset will be utilized in the drought indicators to enable a better objective assessment of real-time drought conditions.

TABLE OF CONTENTS

Page

Page

Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

LIST OF TABLES

TABLE Page

CHAPTER I

INTRODUCTION

I.1. Motivation

Drought is one of the costliest forms of natural disaster in the United States, causing an estimated $6 billion – $8 billion dollars annually, and in recent decades, drought trails only tropical cyclones in terms of economic loss (Elliott et al. 2013). Adjusting for inflation, Smith and Katz (2013) estimated that droughts and heat waves caused $210 billion in damages between 1980-2011, which does not even include the devastating 2012 drought that may have led to one-year losses totaling $30 billion (Elliott et al. 2013). Drought has an almost universally negative impact on agriculture, soil moisture, and water resources, and in most of the United States, drought conditions are sensitive to local variations in precipitation (McRoberts and Nielsen-Gammon 2012). The goal of this work is to provide an unbiased quantification of drought at the local level for the purposes of monitoring and planning.

The United States Drought Monitor (USDM; Svoboda et al. 2002) is a single map that provides a national overview of drought intensity on a weekly basis, with a focus on short-term drought impacts related to agriculture and long-term impacts related to hydrology. Though it is meant for regional-level and state-level drought monitoring, the USDM has been used in recent years as an assessment of drought conditions by decision makers at the county and sometimes even city levels (Quiring 2009). The USDM is a popular tool among the leaders of local government, who prefer to make their decisions based on snapshots of drought intensity, i.e., maps, figures, and tables, tailored specifically to their municipality rather than taking a more holistic approach to understanding the spatiotemporal characteristics of drought conditions (Dow et al. 2009). This type of approach by local decision makers over the years has forced the USDM to evolve into a one-stop shop for assessment of drought at these smaller spatial scales.

Local precipitation variations are the primary driver of drought (and non-drought) conditions in most locations, and rain gauge data provide the best assessment of precipitation in a single location. However, the usefulness of gauges is directly related to the temporal availability and spatial density of available data, so smaller-scale drought monitoring capabilities are limited in times with few observations and in regions where gauges are sparse. Other than gauges, most of the initial indicators used by the USDM were at spatial scales too broad for local interpretation, such as climate division or state averages of common drought indices such as the Palmer Drought Severity Index or Standardized Precipitation Index (Svoboda et al. 2002).

Recently, smaller-scale drought monitoring has been enhanced with the use of a multisensor approach that combines gauge data with radar and satellite data in a Multisensor Precipitation Estimator (MPE) algorithm (Seo and Breidenbach 2002; Lawrence et al. 2003; Fulton 2005). The National Weather Service (NWS) River Forecast Centers (RFCs) have been producing MPEs at high spatial resolution for more than a decade. The MPEs are independently generated at each RFC with a base field of radar precipitation estimates that are quality controlled by gauge data, and when gauge and radar data are limited, the MPE algorithm uses satellite precipitation estimates.

Each day, the MPE data are mosaicked into a single product containing 24-hour precipitation totals across the continental United States and Puerto Rico. Each single, daily mosaic is known as a Stage IV precipitation estimate (Lin and Mitchell 2005) and is available on the Hydrologic Rainfall Analysis Project (HRAP) coordinate system, which is a polar stereographic projection with a spatial resolution of roughly 4 km × 4 km. The Stage IV precipitation dataset is spatiotemporally complete across the continental United States with daily data available from the Advanced Hydrologic Prediction Service website (water.weather.gov/precip) dating back to 1 January 2005 and dating back to 2002 upon request from the National Center for Atmospheric Research (NCAR).

I.2. High-Resolution Drought Monitoring Products

The availability of the Stage IV precipitation estimates and the need for higher resolution drought monitoring led an initiative to develop an experimental high-resolution drought monitoring tool for the state of Texas, called the MPE drought estimator (MPEDE; McRoberts and Nielsen-Gammon 2012). The real-time assessment of drought by the MPEDE products was made possible through the use of the daily Stage IV precipitation mosaics, readily available on the AHPS website. Fig. 1.1 is the MPEDE Standardized Precipitation Index (SPI) for the 12 months period ending on 31 December 2013.

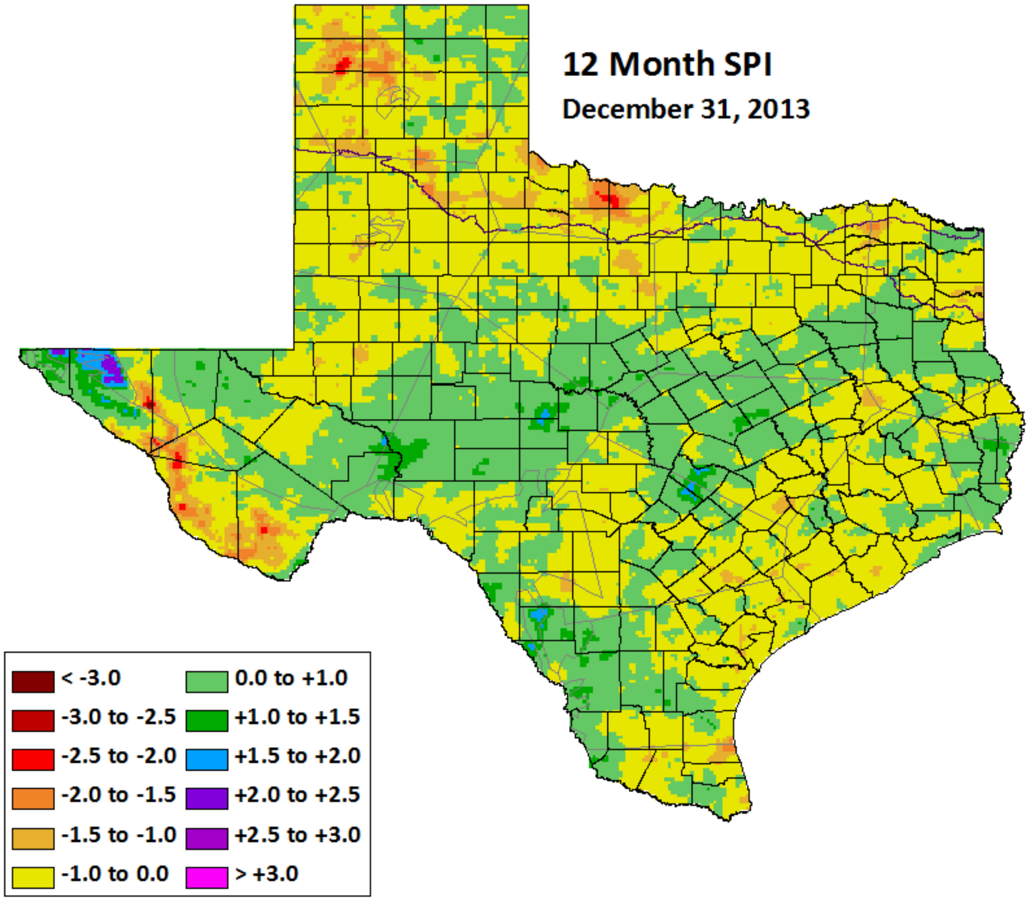


Fig. 1.1. The 12-month MPEDE SPI for the period ending 31 December 2013, available on the Texas A&M University MPEDE webpage.

Most drought monitoring indices classify drought intensity through comparison of a current observation of a meteorological parameter(s), e.g., accumulated precipitation, with past observations of the same parameter. For a given time of year and temporal scale, historical data are used to construct a probability distribution function (PDF) so that assessment of drought using current observations is possible. The SPI fits observed historical precipitation values to a Gamma or Pearson Type III distribution function, which is subsequently mapped onto a Gaussian distribution. The SPI quantifies a precipitation value as the number of standard deviations above or below a normally distributed PDF with a mean of zero and standard deviation of zero (McKee et al. 1993), i.e., an SPI value of -1 indicates a negative precipitation anomaly of one standard deviation.

The Stage IV estimates only date back a little more than a decade, so the MPEDE made use of the NWS Cooperative Observer Program (COOP) network of gauges for historical data (McRoberts and Nielsen-Gammon 2012). The historical CDFs computed at each COOP gauge used in the MPEDE products were determined by the computation of L-moment ratios, which non-parametrically describe the shape and scale of a distribution of values (Hosking and Wallis 1997; McRoberts and Nielsen-Gammon 2012). Through the use of regional frequency analysis (described in great detail by Hosking and Wallis 1997) and a thorough testing procedure of candidate parametric distributions, it was determined that a Pearson Type III (P3) distribution was a suitable parameterization of the L-moment ratios, particularly for characterizing the lower tail, which is critical for an objective assessment of drought severity. Computation of the P3 parameters at clusters of COOP gauge with sufficient historical records was followed by interpolation of the parameters to each HRAP grid cell (McRoberts and Nielsen-Gammon 2012).

For a given temporal scale, real-time assessment of the MPEDE drought severity at an HRAP grid cell compares a normalized Stage IV precipitation estimate to its historical CDF (McRoberts and Nielsen-Gammon 2012). The Stage IV estimates are normalized by the 1981-2010 Parameter-Elevation Regression on Independent Slopes Model (PRISM;

Daly et al. 1994) precipitation normal, which is defined as the P3 “location” parameter at all grid cells. The statistical comparison of the current value to the past distribution of data values at the same location results in a percentile measure, which allows for simple interpretation of drought frequency (Svoboda et al. 2002) that is used in the USDM classification of drought (Table 1.1). A Stage IV precipitation estimate P_0 in the 4th percentile can be expressed as $\mathrm{P}(P \leq P_0) = 0.04$, such that any randomly chosen precipitation value P from the historical CDF only has a 4% chance of being less than or equal to P_0. In everyday language, a P_0 value in the 4th percentile is referred to as a “one in 25 year drought.”

Table 1.1. USDM categories of drought severity as in Svoboda et al. (2002).

Category	**Qualitative severity**	**Percentile range**
D0	Abnormally dry	20th to 30th
D1	Moderate drought	10th to 20th
D2	Severe drought	5th to 10th
D3	Extreme drought	2nd to 5th
D4	Exceptional drought	≤ 2nd

The MPEDE experimental products were accepted as a drought monitoring tool for Texas by the USDM community of authors and contributors, which led to interest in expansion to other parts of the United States. A proposal submitted to and accepted by the United States Department of Agriculture (USDA), in collaboration with the state climate offices of Indiana (Purdue University) and North Carolina (NC State University), created an initiative to spatially expand and improve the MPEDE products. Based on the overall scope of the project, the MPEDE products were rebranded as a High Resolution Drought Trigger Tool (HIRDTT), a name that emphasizes the impetus of the project. In the context of drought, “trigger” is an objective indicator of drought reaching a pre-specified index value that requires a response. The overarching goal of HIRDDT is to provide non-experts with the confidence and capability to make local decisions based on an appropriate objective indicator.

I.3. Biases in Stage IV Precipitation Estimates

The task of improving the quality of the Stage IV precipitation data used as the real-time inputs to the HIRDTT products was designated to the Texas state climate office (housed at Texas A&M University). There are numerous deficiencies in the Stage IV estimation procedure that result in biases in the real-time precipitation values that are used as inputs to the MPEDE algorithm. The Stage IV precipitation estimates have improved since their initial release back in 2002 thanks to advances in the MPE algorithms at RFCs and the recent implementation of dual polarization Doppler radars by the NWS. However, biases still plague Stage IV precipitation estimates and the resulting HIRDTT products require expert knowledge of these biases for a reasonable estimate of drought severity in some regions. For example, Stage IV 36-month precipitation estimates used as input to an HIRDTT Standardized Precipitation Index product show clear non-physical discontinuities at the edges of the Shreveport, LA radar (KSHV) coverage area that might lead a non-expert to underestimate the severity of drought is some regions at an intermediate distance from the radar location (Fig. 1.2). An expert user will likely recognize the biases present in Fig. 1.2, but this recognition still leaves assessment of quantitative drought severity surrounding the KSHV radar to be a speculative exercise at best.

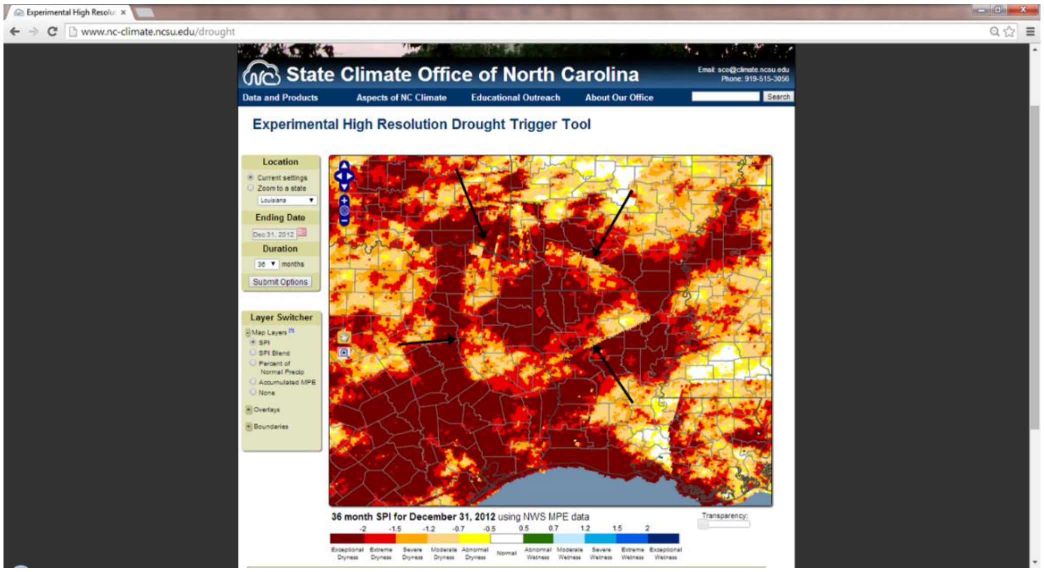


Fig. 1.2. Screenshot of the 36-month SPI for the period ending 31 December 2012 taken from the NC State University HIRDTT webpage. The marker denotes the location of the KSHV radar and arrows the discontinuities existing at the edges of the KSHV radar domain.

A more focused goal of this work is to improve the Stage IV precipitation estimates through objective quantification of biases. Biases in radar precipitation estimates (including Stage IV estimates), can be broken into different components, which are

1. azimuthal-dependent biases $B(d, \theta)$,
2. mean-field biases $B(r)$,
3. range-dependent biases $B(d)$, and
4. two-dimensional biases $B(\Phi, \Lambda)$,

where d is the distance from the closest radar (range), θ is the angular direction from the radar location (azimuth), r refers to a given radar, and Φ is the latitude and Λ is the longitude for a given location. The first four biases are systematic and minimization of these biases will allow for a Stage IV precipitation estimate (P_0) to more accurately depict the true surface precipitation (P_T). In equation form, the goal of this work is to minimize the first four bias terms on the right-hand-side of

$$\frac{P_0}{P_T} = 1 + B(d,\theta) + B(r) + B(d) + B(\Phi,\Lambda), \tag{1.1}$$

so that the ratio of P_0 to P_T is approximately one for a given time period.

A methodology will be developed in subsequent chapters to objectively quantify the different types of biases. This quantification requires knowledge of the mechanisms responsible for different types of biases and the resulting spatiotemporal properties of these biases in the Stage IV dataset. The final goal of this work is to produce improved Stage IV precipitation estimates through minimization of the four bias terms on the right-hand-side of (1.1).

I.4. Outline

I.4.a. Chapter II

A literature review details the basics of radar reflectivity measurements and transforming these reflectivity values into radar precipitation estimates. This is followed by an overview of the operational procedures for processing the radar returns to produce the Stage IV precipitation estimates. An understanding of the mechanisms leading to biases in radar precipitation estimates is provided. A discussion of the recent upgrade of radars by the NWS to dual polarization and the impact on Stage IV precipitation estimates follows.

I.4.b. Chapter III

This chapter will detail a methodology for correcting azimuthal-dependent biases $B(d, \theta)$, which are due to blockage of a transmitted radar beam by a fixed, non-meteorological obstacle. Beam blockage results in a systematic under-estimation of precipitation in a given azimuth θ that is dependent on the range d of the obstacle. Identification of regions with Stage IV estimates contaminate by beam blockage is followed by a correction procedure for data in the affected locations.

I.4.c. Chapter IV

A methodology will be presented to correct mean-field biases $B_{M\text{-}F}$ and range-dependent biases $B(d)$ in the Stage IV precipitation estimates. This is accomplished through creation of a new statistical model that characterizes $B_{M\text{-}F}$ and $B(d)$ as a smooth, continuous function given a set of spatially irregular and discrete bias data points. The azimuthal corrections were done prior to the mean-field and range-dependent bias corrections in order to reduce the statistical noise in the bias data points. The resulting bias function will be used for correction of the Stage IV precipitation estimates.

I.4.d. Chapter V

The final minimization procedure will characterize two-dimensional biases $B(\Phi, \Lambda)$ remaining in Stage IV precipitation estimates (P_3) that have been corrected for the other three types of biases. For a given P_3 precipitation field, a continuous two-dimensional bias field will be constructed using gauge data and this field will be applied to the P_3 data to produce a Stage IV precipitation estimate P_f that minimizes the four types of biases in (1.1).

I.4.e. Chapter VI

The final chapter will provide a summary of the bias correction procedures and possible future projects resulting from this dissertation.

CHAPTER II

LITERATURE REVIEW

II.1. Overview of Radar Precipitation Estimates

In the mid-1990s, 158 Weather Surveillance Radar-1988 Dopplers (WSR-88Ds) were installed as part of the Next Generation Weather Radar (NEXRAD) program (Crum and Alberty 1993), including 142 radars providing coverage across the continental United States (Rogalus and Ogden 2011). The Stage IV dataset in this study includes data originating from 104 WSR-88D locations in the central and eastern United States (Fig. 2.1).

Each WSR-88D has an accumulation algorithm used to estimate precipitation totals on an hourly time scale (Fulton et al. 1998). The WSR-88D precipitation estimates are used as an initial inputs to the Stage IV algorithm at nine of the twelve National Weather Service (NWS) River Forecast Centers (RFCs; Fig. 2.2), which produce hourly, high-resolution radar precipitation estimates (Habib et al. 2009). The remaining three RFCs are located in the mountainous western United States, outside of the region of this study. Because the complex terrain leads to unreliable radar-based precipitation estimates, those three RFCs use gauges to produce the hourly precipitation estimates (Henkel and Peterson 1996). Each hour, the analyses produced at the twelve individual RFCs are combined into a single Stage IV mosaic that covers all of the continental United States and Puerto Rico (Zhang et al. 2011).

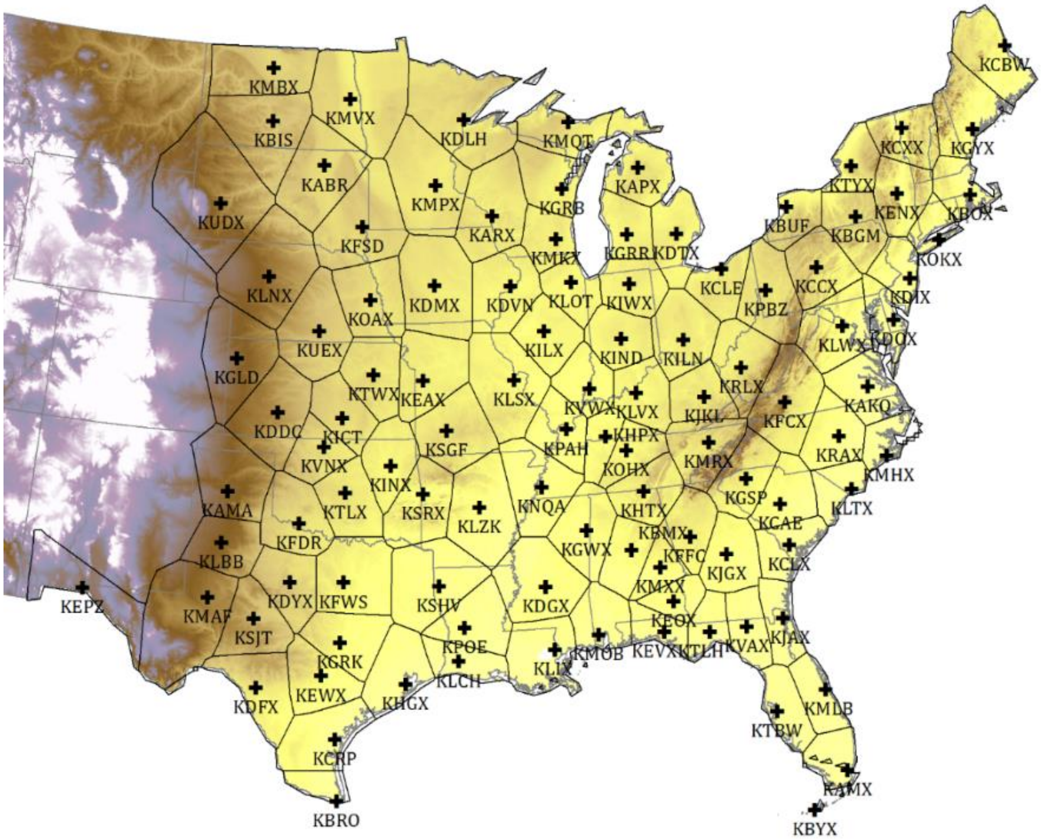


Fig. 2.1. Locations (crosshairs) and regions covered (polygons with black outlines) for the 104 WSR-88Ds in the central and eastern United States.

An overview of the WSR-1988 Doppler (WSR-88D) System estimation of precipitation and the algorithm that produces Stage IV precipitation estimates will be provided. This will be preceded by a discussion of radar reflectivity measurements and the computation of radar precipitation estimates. Following a description of the algorithm, there will be a thorough description of error sources in both gauge measurements and radar precipitation estimates and the negative impact these errors have on the Stage IV precipitation dataset, which includes biases that need to be corrected.

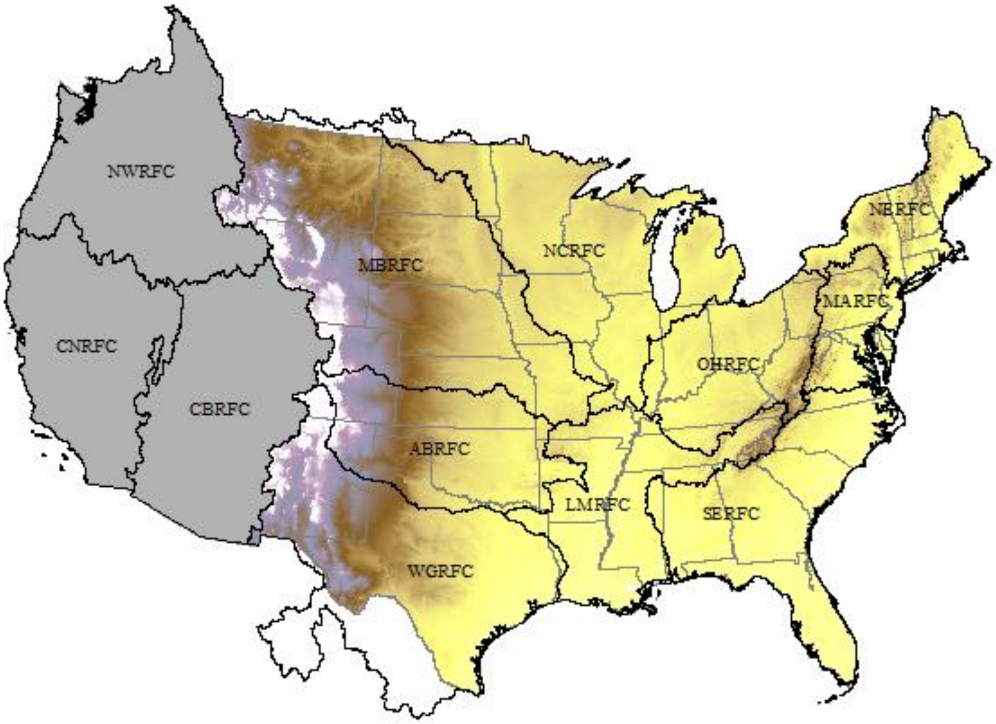


Fig. 2.2. The 12 River Forecast Center boundaries in the Continental United States, with the three western RFCs that don't use WSR-88D as initial inputs to the Stage IV algorithm (gray shading).

II.2. Properties of WSR-88D Measurements

II.2.a. Radar Reflectivity

Weather radars transmit frequent pulses of radiation with wavelength λ from an antenna that are reflected back to the radar by objects in the atmosphere. The radar reflectivity factor (Z) is a quantitative measure of the power returned to the radar from the transmitted radiation. The amount of power returned depends on the backscattering properties of both meteorological and non-meteorological targets in the path of the transmitted pulse (Rinehart 2004). Each transmitted pulse is a three-dimensional beam with power P_t, with a return power (P_r) from each target that is dependent on the scattering cross-section (σ) of the target

$$P_r = P_t \frac{g^2 \lambda^2 \sigma}{64 \pi d^4}, \tag{2.1}$$

where d is the distance from the target to the radar, and g is the antenna gain (Rinehart 2004). The antenna gain g is the power of the transmitted signal, and the relative gain at anywhere within the beam with strength g is the ratio g/g_0, where g_0 is the gain at the axis of the beam, where the transmitted power is the strongest. By convention, the azimuthal beamwidth is considered the width of the beam where $g \geq 0.5$ (Rinehart 2004), or the half-power beamwidth ($\theta_{1/2}$; Fig, 2.3). Any dimensionless quantity X that is the ratio of two values ($G = g/g_0$ for example) can be converted to the decibel (dB), where

$$\mathrm{dB}X = 10 \log_{10}(X). \tag{2.2}$$

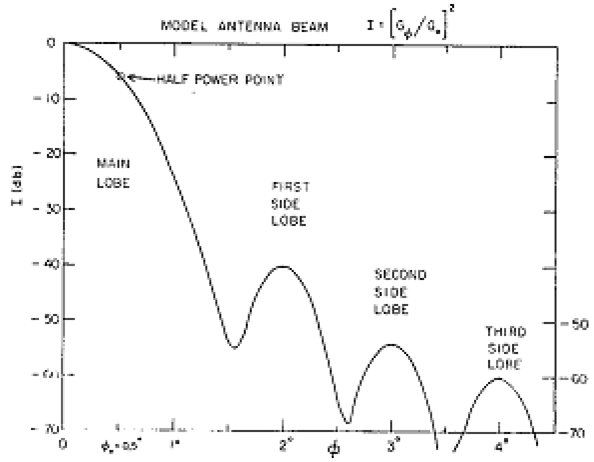


Fig. 2.3. Power of the radar beam as a function of azimuthal distance from the beam axis for a beam with $\theta_{1/2}$, originally in Donaldson (1964).

This conversion is normally done for quantities that vary over several orders of magnitude, and is why the half-power beamwidth $\theta_{1/2}$ is often referred to as the -3 dB beamwidth. Assuming that backscattering is dominated by the Rayleigh region ($D << \lambda$), the cross-sectional area of a spherical target is

$$\sigma = \frac{\pi^5 |K|^2 D^6}{\lambda^4}, \tag{2.3}$$

which simplifies (2.1) to

$$P_r = \frac{C |K|^2 z}{d^2}, \tag{2.4}$$

where C is a constant dependent on the properties of the transmitted radiation, D is the diameter of an object assumed to be a spherical raindrop, $|K|^2$ is the refractivity of liquid water, and z is the radar reflectivity factor (Battan 1973). Which in terms of size distribution with units in $\mathrm{mm^6m^{-3}}$ is given by

$$z = \int_0^\infty D^6 N(D)\, dD\,, \tag{2.5}$$

where *N(D)* is the number density of raindrops for a given diameter (Rinehart 2004). The radar reflectivity factor z (2.5) assumes Rayleigh scattering of the transmitted radiation by spherical liquid particles with $D << \lambda$ (Battan 1973). Because the range of z is extensive for precipitation, z is converted to dimensionless unit Z and is expressed as the decibel of Z (dBZ). The value of Z is the ratio of z to z_0, where $z_0 = 1\ \mathrm{mm^6m^{-3}}$, expressed as

$$\mathrm{dB}Z = \log_{10}\left(\frac{z}{1\mathrm{mm^6m^{-3}}}\right)\,. \tag{2.6}$$

WSR-88Ds transmit waves in the S-band, with λ ranging from 10 to 11.1 cm (Villarini and Krajewski 2010a). S-band waves are effective at detecting precipitation and limiting attenuation, the scattering and absorption of radiation by atmospheric objects that can include clouds, precipitation, and hail, in addition to non-meteorological objects such as insects and birds (Anagnostou et al. 1998). Attenuation in the atmosphere decreases as the wavelength of the radiation increases, though increasing the wavelength of transmitted radiation will decrease the detection of smaller objects.

Each WSR-88D measures radar reflectivity in volume scan sweeps at several tilt angles ranging from 0.5° to 20° from an antenna that is located several meters above the surface of the radar site (Klazura and Imy 1993). The radar reflectivity measured at the lowest elevation angle (0.5°) is the base reflectivity (Z_b) whereas the composite reflectivity (Z_c) represents the largest Z from all the elevation angles. The WSR-88D has a built-in algorithm that switches the radar antenna from clear air mode to precipitation mode when Z_c indicates the existence of precipitation (Fulton et al. 1998). The time needed for the WSR-88D to complete a full multi-elevation volume scan decreases from 10 min in clear air mode to 6 min in precipitation mode (Vignal and Krajewski 2001). Four additional volume scan sweeps are added at higher elevation angles in precipitation mode to the 5 elevation angles included in clear air mode. The faster rotation speed of the WSR-88D during precipitation mode allows for better temporal resolution of Z and reduces attenuation by non-precipitation objects that can be detected in clear air mode (Fulton et al. 1998).

II.2.b. Radar Rainfall Rate

WSR-88Ds merge the Z values from the four lowest elevation angles into a single hybrid scan reflectivity when the radar is in precipitation mode to estimate accumulated precipitation (Legates 2000). The estimates of precipitation are on a 1 km × 1° azimuthal degree polar grid, out to a range of 230 km from each radar site. The choice of scan used in the HSR is determined at each grid cell by the tilt angle that captures the "optimum" altitude of 1 km above the ground, though some adjustments have been made to account for surface elevation changes (Fulton et al. 1998). The default construction of the hybrid scan reflectivity value uses the fourth highest tilt out to a range of 20 km from the radar, the third highest tilt from 20 km to 35 km, the highest second tilt from 35 km to 50 km, and the lowest tilt beyond 50 km (Smith et al. 1996).

The radar estimates precipitation at each grid cell from the liquid water content present, which is related to the size distribution of raindrops in the atmosphere (Marshall and Palmer 1948). The rainfall rate R for a given drop size distribution is

$$R = \frac{\pi}{6}\int_0^\infty D^3 N(D) v(D)\, dD\,, \tag{2.7}$$

where ν is the terminal velocity of a spherical raindrop with diameter D (Seo 2010). Measured Z (2.6) and assumed R (2.7) are both expressed empirically in terms of raindrop distribution size, so these parameters can be combined to derive a precipitation estimate from a given value of the hybrid scan reflectivity (Smith et al. 1996). This Z-R relationship is

$$Z = aR^b\,, \tag{2.8}$$

where a and b are coefficients that can vary depending upon rainfall regime (Battan 1973). The WSR-88D default coefficients are a = 300 and b = 1.4, which represents a compromise between stratiform and convective rainfall events, though a = 250 and b = 1.2 is used as an alternative in locations dominated by tropically driven precipitation (Fulton et al. 1998).

II.3. WSR-88D Data Processing

RFCs began generating regional mosaics of hourly radar precipitation estimates in 1993, with most using WSR-88D Stage II data as the initial radar-based input (Young et al. 2000); the exceptions were the California-Nevada, Colorado Basin, and Northwest RFCs (Fig. 2.2; Henkel and Peterson 1996). In an effort to improve the Stage II data, the National Severe Storms Laboratory (NSSL) began generating next-generation quantitative precipitation estimates (Q2 for short) data in 2006 as part of the National Mosaic and Q2 (NMQ) project (Vasiloff et al. 2007). Both the Stage II and Q2 algorithms use radar reflectivity measurements, with Q2 moving away from the radar-centric Stage II approach toward a multi-sensor approach that combines radar, satellite, model, and gauge data (Vasiloff et al. 2007; Glaudemans et al. 2008; Chen et al. 2013). Between 2007 and 2011, the North Central, Ohio, Southeast, and West Gulf RFCs transitioned from using Stage II data to Q2 data as the initial radar-based input for the RMPAs (Zhang et al. 2011).

II.3.a. Stage II Processing Algorithm

Each WSR-88D contains a set of built-in algorithms called the Precipitation Processing System (PPS), which produces radar precipitation estimates based on Z measurements and the specified Z-R relationship (Fulton et al. 1998). The PPS algorithm within each WSR-88D produces Stage I data used as one the initial inputs to the regional mosaics at most of the RFCs (Fig. 2.2). The PPS goes through several sub-algorithms, beginning with the determination of R from Z that eventually leads to accumulated precipitation products (Rogalus and Ogden 2007). The PPS algorithm determines an initial rainfall rate R_i from the WSR-88D Z-R relationship, using Z_m as the primary input (Fulton et al. 1998). During a precipitation event, R is integrated over an entire hour to produce an radar-based quantitative precipitation estimate *(QPE*) for hour h, given by

$$QPE(h) = \sum_{i=1}^{n_s} R^i T^i \ , \tag{2.9}$$

where n_s is the number of scans and for each scan T^i is the time period of the scan.

The PPS algorithm eliminates hourly outliers and performs adjustments using available hourly precipitation data from available gauges. The real-time gauge data is used to compute a multiplicative mean-field bias correction for the entire field with a Kalman filter adjustment algorithm (Smith and Krajewski 1991). The mean-field bias is based on collocated, non-zero radar-gauge pairs with a requirement that at least three of these pairs are present at a given time for the adjustment to occur. The WSR-88D PPS generates adjusted Stage I products on a 2 km × 1° azimuthal degree polar grid using adjacent radial pairs of R on the 1 km × 1° azimuthal degree polar grid (Fulton et al. 1998). An additional output of each WSR-88D is an hourly digital precipitation array that is produced each hour on the HRAP grid (Kalinga and Gan 2012).

The digital precipitation array from each WSR-88D is further processed by the individual RFCs to produce Stage II data, which undergoes more rigorous quality control with a goal to provide an optimal radar-gauge precipitation estimate. The Stage II process used hourly gauge observations and the Stage I data (Breidenbach et al. 1998), the latter of which has numerous error sources. A Kalman filter approach calculates the mean-field bias at each radar in an approach similar to Stage I, except that the mean-field bias is computed using a "memory" of non-zero precipitation values from the recent past at gauges within the radar domain (Seo 1998). The number of hours in the memory span (*m*) can be adjusted based on the number of available non-zero hourly gauge measurements and is typically inversely proportional to the number of non-zero precipitation measurements. The memory span *m* can extend up to a maximum one year at radars where data are sparse (Breidenbach et al. 1998). Using gauge data going back *m* hours, the Stage II mean-field bias at a given WSR-88D uses a Kalman filter modified from the Stage I approach to incorporate the memory span (Seo 2002).

After correcting the mean-field bias of the digital precipitation, the Stage II algorithm performs local adjustments at each HRAP grid cell using non-zero gauge measurements, considered to be the ground truth during this step (Breidenbach et al. 1998). An "optimal" estimate at each HRAP grid is a linear combination of the mean-field bias-adjusted estimate and non-zero precipitation values from nearby gauges that minimizes expected error variance (Fulton et al. 1998). The relative weight of the mean-field bias-adjusted values used in the optimal estimate decreases as distance of the nearest reporting gauge increases (Breidenbach et al. 1998). The estimate will be an exact match to a non-zero gauge measurement at an HRAP cell containing a non-zero precipitation measurement. The result of the Stage II algorithm is an array of HRAP grid cells with bias-adjusted radar data merged with gauge data (Habib et al. 2009).

II.3.b. Next-Generation Quantitative Precipitation Estimation Algorithm

As part of the NMQ initiative, Q2 products are generated by the NSSL using radar data from WSR-88D, Terminal Doppler Weather Radar, and Canadian operational radars (Zhang et al. 2009). The radar data and satellite data are combined with rain gauge data and input from numerical weather models (Glaudemans et al. 2008) to produce radar precipitation estimates at 1 km resolution over CONUS. Single-radar hybrid scan reflectivity fields are constructed using the Z at each location from the volume scan at the lowest altitude with acceptable data (Zhang et al. 2011). The hybrid scan reflectivities (HSR) produced by the individual radars are combined to create a regional mosaic ($RHSR$)

$$RHSR = \frac{\sum_i {w_d}^i {w_h}^i HSR^i}{\sum_i {w_d}^i {w_h}^i}, \tag{2.10}$$

using each radar i, where w_d is a horizontal weighing function defined by

$$w_d = \exp\left(-\frac{d^2}{L^2}\right) \tag{2.11}$$

and w_h is a vertical weighing function defined by

$$w_d = \exp\left(-\frac{h^2}{H^2}\right) \tag{2.12}$$

with scaling factors L and H (Zhang et al. 2009).

Table 2.1. List of Z-R relationships, taken from Zhang et al. (2009).

Precipitation type	***Z*-*R* relationship**
Convective	$Z = 300R^{1.4}$
Stratiform	$Z = 200R^{1.6}$
Tropical	$Z = 230R^{1.25}$
Snow	$Z = 75R^{2.0}$

The NMQ algorithm produces 3-dimensional reflectivity mosaics using 31 vertical levels (Zhang et al. 2011), providing a vertical profile of reflectivity (VPR) at each NMQ grid cell (1 km^2). The Q2 algorithm then uses the VPR at each grid cell to determine the Z-R relationship, using four different Z-R relationships based on the precipitation type (Table 1; Zhang et al. 2009), providing a major advantage over single-radar processing (Story 2011). A radar-only Q2 product is produced every 2.5 min, with local gauge-corrected Q2 products produced at the top of every hour (Zhang et al. 2011). Additive radar-gauge biases are calculated and interpolated onto the NMQ grid by using inverse distance weighting, which is based on the distance of each gauge from a given grid cell (Kim et al. 2009).

II.4. Mosaicking of Radar Precipitation Estimates

After the WSR-88D network was installed, production of the regional mosaics at most RFCs used the Stage III processing algorithm. At each RFC, the Stage II data from each individual radar was combined into a single product covering the entire RFC area of responsibility (Breidenbach et al. 1998). The Stage III processing algorithm was utilized by RFCs in the central and eastern United States. Because of unreliable radar returns due to the complex topography, the western RFCs (Fig. 2.2) utilized a different processing algorithm. Mountain Mapper combines gauge observations and the PRISM dataset from Oregon State University (Vasiloff et al. 2007) to produce gridded precipitation estimates for the regional mosaics.

During the summer of 1996, the Arkansas-Red Basin RFC transitioned from the Stage III algorithm to a processing algorithm called Process 1, which uses Stage II data to spatially interpolate gauge estimates (Young et al. 2000) and is now an updated version called as Process 3. RFCs still using Stage III processing by the year 2000 transitioned to the Multi-Sensor Precipitation Estimator (MPE) processing algorithm by 2003, an approach that uses satellite, radar, and gauge data to produce an optimal regional mosaic (Habib et al. 2009). It is not uncommon for people to refer to generically refer to the regional mosaics as Stage III data, but in this study Stage III refers specifically to the once-used processing algorithm.

II.4.a. Stage III Processing

At each RFC, the Stage III was a mosaicking of the Stage II precipitation estimates from the radars into a regional product (Fulton et al. 1998). Multiple radars cover most HRAP grid cells, in which case a choice was made between using the mean value of all the overlapping radar values or the value derived from radar showing the maximum reflectivity (Breidenbach et al. 1998). In addition to mosaicking of the Stage II data, RFC personnel had the ability to manually remove anomalous gauge or radar data that might have led to significant errors (Kalinga and Gan 2012).

II.4.b. Multi-sensor Precipitation Estimator (MPE)

Deficiencies in the Stage III precipitation estimation algorithm led to the development of the MPE algorithm, which has been used by RFCs as far back as 2003 to produce regional mosaics (Yilmaz et al. 2005). In addition to the use of available gauges and the Stage I precipitation estimates, the MPE procedure uses Geostationary Operational Environmental Satellite (GOES) precipitation estimates that are available every 15 minutes at 4 km resolution (Scofield and Kuligowski 2003). Satellite remote sensing of precipitation provides estimates in mountainous and remote regions where radar and gauge coverage is limited (Villarini et al. 2009). Where multiple radars provide adequate coverage, the MPE algorithm is able to select the individual radar that provides the best coverage at each HRAP grid cell (Marzen and Fuelberg 2005) with a mosaicking procedure that uses the precipitation estimate from the radar with the lowest unobstructed sampling volume (Seo 2002).

A major advantage of the MPE algorithm when compared to the Stage III algorithm is the delineation of the effective radar coverage for each radar that use several years of Stage II data to produce a radar precipitation climatology (Habib et al. 2009). The delineation offers an improved mean-field bias correction factor for each WSR-88D since the use of radar data is limited to regions where the radar is able to consistently detect precipitation (Seo 2002). MPE has an enhanced memory span feature, with the ability to store the mean-field bias for 10 different time spans (Marzen and Fuelberg 2005). The mean-field bias correction from the shortest time span with a predetermined minimum number of gauge/radar pairs is then applied to the Stage I data (Fulton 2005). In addition, the MPE program has a local bias correction algorithm that can adjust the value of the multiplicative bias determined by the mean-field bias correction at each HRAP grid cell based on nearby gauge observations (Seo 2002).

II.4.c. Process 3 (P3)

Process 1 (now P3) was developed by the Tulsa District of the US Army Corps of Engineers and was implemented in 1996 by the ABRFC, which covers an area that includes all of Oklahoma and at least part of every surrounding state (Young et al. 2000). Each hour, an initial radar mosaic covering all of the ABRFC is created on the HRAP grid, averaging the digital precipitation array estimates from each WSR-88D that overlaps a particular grid cell (Schmidt et al. 2000). At each HRAP cell containing at least one available hourly gauge measurement, a gauge-to-radar ratio is computed by dividing the hourly gauge value by the initial radar-only mosaic. All the available hourly gauge measurements from a given time are then used to create a triangulated irregular network (TIN; Glaudemans et al. 2008). At HRAP cells not containing gauge data, the gauge-to-radar ratio is computed using a distance-weighted interpolation of available gauge data from the TIN (Young et al. 2000). The final regional mosaic produced by the ABRFC multiplies the initial radar-only estimate at each HRAP grid cell by the local gauge-to-radar ratio. The P3 approach is effective in the ABRFC because the Oklahoma Mesonet provides a spatially dense network compared to most other parts of the United States (Young et al. 2000).

II.4.d. Mountain Mapper

The Mountain Mapper algorithm was developed due to the inaccuracies of radar precipitation estimates in the mountainous western United States caused by the rapid changes in topography (Schaake et al. 2004). Mountain Mapper uses the 1 km resolution PRISM monthly precipitation climatologies to compute hourly normals (P_n) for a given calendar month at each available, non-zero gauge. At each of these gauges, the percent of normal precipitation (PoN) is computed as the ratio the of gauge value (P_G) to the P_n value (see 3.2). All the available PoN data are then interpolated onto the NMQ grid using inverse distance weighting (Zhang et al. 2011). The final step in the Mountain Mapper algorithm is multiplying the PoN at each grid cell by its PRISM hourly normal to get a resulting precipitation field on the HRAP grid (Henkel and Peterson 1996).

II.5. Production of Stage IV Precipitation Estimates

The regional QPE mosaics produced by each RFC are mosaicked into hourly QPE products that cover all of the continental United States, which are known as Stage IV precipitation estimates (Zhang et al. 2011). At any given time, the individual RFCs determine the mosaic that best represents precipitation in their region based on experience and precipitation regime (Habib and Qin 2013) and produce 1-hour and 6-hour regional mosaics that have passed extensive quality control procedures. Every day, the NWS produces a daily Stage IV mosaic of the 1200 UTC -1200 UTC precipitation totals on the

HRAP grid that is a summation of the 24 hourly mosaics (Lin and Mitchell 2005). However, the Stage IV analyses contain a large number of biases that result from inaccuracies in the measurement of precipitation, which will be discussed in the following section.

II.6. Precipitation Data Inaccuracies

II.6.a. Spatial Scales of Gauges and HRAP Grid Cells

Attempts to correct the inaccuracies of radar-based precipitation estimates typically involve the use of gauge data. There are major differences in both the spatial and temporal sampling between the two measurement types that can make adjustments of radar precipitation estimates difficult when using gauge measurements (Jayakrishnan et al. 2004). Gauge measurements continuously sample an area only on the order of 100 cm^2 while radar estimates typically cover an area on the order of 10 km^2 (Ciach and Krajewski 1999). Joss and Waldvogel (1990) showed that precipitation rates can vary by one or two orders of magnitude within a single storm. Though gauges provide in situ measurements of precipitation, the density of most existing gauge networks is too sparse to capture spatial variability in a way that is possible using WSR-88D precipitation estimates (Jayakrishnan et al. 2004).

The sampling frequency of gauge data is much different than that of radar sampling, so even comparisons of ideal radar and gauge measurements can differ significantly because precipitation is being recorded at different intervals (Steiner et al. 1999). WSR-88D precipitation estimates are based on instantaneous samples repeated at intervals of several minutes, though rainfall intensity can very significantly between two sampling intervals (Austin 1987). Gauges provide continuous, in situ measurements of precipitation, but totals are typically recorded at much longer time intervals between measurements than radar estimates.

II.6.b. Radar Precipitation Estimation Error Sources

Errors in radar precipitation estimates can arise from a number of sources, which can be classified as incorrect measurements of Z, uncertainties in the Z-R relationship, and variations in the vertical profile of Z that lead to range-dependent errors (Austin 1987; Smith et al. 1996; Baeck and Smith 1998; Vignal and Krajewski 2001). An assessment is provided here of the various factors that lead to errors in WSR-88D precipitation estimates. Knowledge of these error sources is important in the context of this study, because systematic errors can result in biases in the Stage IV precipitation estimates, even with thorough quality control measures.

II.6.b.1) Beam Blockage

Beam blockage at the lowest tilt angles can be a significant source of errors (Baeck and Smith 1998) that can lead to significant biases in the Stage IV precipitation estimates. Although it is primarily a function of azimuth (θ), if blockage occurs, it begins at a range d_b that is a function of θ which makes beam blockage a two-dimensional problem. These errors appear in the Stage IV data as a systematic underestimation of precipitation in an azimuthal sector (θ), beyond the range d_b of the obstacle. Beam blockage is a major issue because the lowest radar tilt angles, for which beam blockage is most likely to occur, provide the best estimate of surface precipitation if unobstructed (Smith 1998). The WSR-88D algorithm (Fulton et al. 1998) has the capability of correcting for known obstructions by using higher tilt angles, but this can lead to significant azimuthal-dependent biases (Andrieu et al. 1997; Young et al. 1999). Therefore, even correcting for known blockage features for a given radar can still lead to a systematic underestimation of precipitation for an azimuthal sector θ.

Beam blockage commonly occurs in mountainous regions as a beam traveling from a lower elevation encounters the ground. The complex topography of the western United States (Westrick et al. 1999) is the reason all bias corrections in this study are limited to the central and eastern United States. Other than complex terrain, the main sources of beam blockage in the Stage IV precipitation dataset are tall buildings, trees, water towers, and cell towers near the radar location (Fabry et al. 1992; Holleman 2006; Overeem et al. 2009). The size of the sector θ can give some insight into the type of obstruction that is responsible for the blockage. Beam blockages due to terrain appear as fatter “slices” than the blockages due to point structures such as tall buildings, which appear as thinner “spikes” in the Stage IV precipitation data.

II.6.b.2) Other Reflectivity Measurement Errors

The accuracy of WSR-88D precipitation measurements depends on the stability of the radar hardware components, e.g., the antenna, transmitter, and receiver (Harrison et al. 2000). Hardware calibration is done separately at each WSR-88D radar site. Miscalibration of instrumentation has been shown to produce significant errors in Z

measurements; Smith et al. (1996) suggested differences in calibration between adjoining radars can lead to differences in *Z* measurements as high as 30% in overlapping areas. The PPS algorithm that computes the WSR-88D precipitation estimates has correction functions to account for this and other hardware malfunctions, such as missing Z measurements (Fulton et al. 1998).

Ground clutter results when a transmitted beam intercepts a target on the ground, such as a permanent building, leading to enhanced returned power over that produced by airborne targets (Harrison et al. 2000). Ground clutter is usually limited to the lowest WSR-88D tilt angle within 20 to 30 miles of the radar site and is unique to each radar site (Chrisman et al. 1995). Sources of ground clutter can include stationary targets such as buildings and trees, moving biological targets such as insects, bats, and birds, sun strobes, and electronic interference (Zhang et al. 2011). Each WSR-88D utilizes a predefined clutter suppression map derived empirically through radar returns during times of fair weather; the clutter suppression map can be updated as often as needed using off-line procedures when the WSR-88D is in clear air mode (Fulton et al. 1998). The clutter suppression algorithm reduces the returned power whose radial velocity is near zero in known ground clutter areas when the radar is in precipitation mode (Chrisman et al. 1995).

Anomalous propagation (AP) of a WSR-88D transmitted beam occurs when the beam is directed to the ground and returns non-zero reflectivity values for areas not receiving precipitation (Krajewski and Vignal 2001). The occurrence of AP is typically associated with cold season temperature inversions or outflow boundaries from thunderstorms, during which super-refraction of the beam directs it toward the ground (Smith et al. 1996). The WSR-88D PPS uses higher elevation tilt angles when AP is detected, though the detection algorithm is more successful in situations where real precipitation does not exist (Fulton et al. 1998).

Attenuation can become a significant problem during heavy precipitation events and can lead to a significant decrease or even total loss of backscattered power from a transmitted WSR-88D beam, resulting in erroneously low Z measurements (Harrison et al. 2000). Precipitation occurring at the radar site can lead to wet radome attenuation, in which both the transmitted and incoming radiation can be attenuated by water collected on the radome, a dome that protects the WSR-88D antenna (Legates 2000).

II.6.b.3) Uncertainty in Z-R Relationship

Even assuming perfect Z measurements, errors in rainfall estimation can be caused by an inaccurate Z-R relationship (Smith and Krajewski 1991). If these errors persist, it could lead to a systematic mis-estimation of precipitation and biases in the Stave IV precipitation data. The WSR-88D PPS algorithm uses a spatially and temporally constant Z-R relationship (Fulton et al. 1998), which assumes the drop size distribution is also constant though it is known that large variations can occur within a single precipitation event and from storm to storm (Austin 1987). Uijlenhoet et al. (2003) showed that that ideal coefficient of a in (2.8) varies from 200 to 450 during the passage of a squall line. Accurately characterizing the Z-R relationship is difficult because of the within-storm variations of drop size distribution and because Z is proportional to the sixth power of the drop diameters (Harrison et al. 2000).

The Q2 algorithm determines the Z-R coefficients for individual grid cells based on the VPR and is produced every 5 minutes at 1 km resolution, which are major improvements over the WSR-88D PPS algorithm (Kim et al. 2009). However, the Q2 computation of R from the Z-R relationship is based only on four choices of *a* and *b* to represent the environmental DSD at a given grid cell (Zhang et al. 2009). The WSR-88D PPS algorithm default coefficients of $a = 300$ and $b = 1.4$ are viewed as the best compromise between stratiform and convective precipitation events, though Battan (1973) listed a total of 69 possible Z-R relationships, which are based on differences in geography and precipitation regime. Even with similar precipitation regimes, the ideal Z-R relationship may depend on the measurement and data analysis techniques used (Rogalus and Ogden 2011).

II.6.b.4) Vertical Profile of Reflectivity

Range-dependent biases in radar precipitation estimates can be attributed to variations in precipitation intensity with height, known as the VPR effect. The curvature of Earth's surface means that the height of the radar beam at the lowest tilt angles increase approximately as the square of the distance from the radar site (Rinehart 2004).

$$h = \sqrt{d^2 + \left(\frac{16}{9} + d\sin\phi\right) r_E} - \frac{4}{3} r_E \, , \tag{2.13}$$

where r_E is the Earth's radius and ϕ is the tilt angle of the radar beam.

This leads to discrepancies between estimates of *R* and the precipitation rate occurring at the surface at a given location (Vignal and Krajewski 2001). Numerically, the VPR (see equation 4.4) is the ratio of reflectivity at height *h* (Z_h) to the reflectivity at the surface (Z_0) The vertical variability of precipitation and the range-dependence of the height Z_h at which reflectivity measurements are taken are the primary mechanisms leading to range-dependent biases in radar precipitation estimates (Bellon et al. 2005), which includes the Stage IV dataset.

An overestimation of precipitation can occur with the presence of solid hydrometeors such as hail, sleet, and melting snow that have higher reflectivity values due to the presence of water on ice (Legates 2000). Each of the four tilt angles used in the construction of the hybrid scan of reflectivity at any given WSR-88D exhibit a significant climatological peak in Z measurements in regions where the beam intercepts the melting layer, called the radar bright band (Baeck and Smith 1998). Bright-band effects can lead to a significant overestimation of precipitation during the mid-latitude cold season at a range of 50-100 km, where the second tilt angle used in the hybrid scan frequently intersects the melting layer (Smith et al. 1996). The WSR-88D PPS algorithm places an upper limit on Z measurements used to compute rainfall rates, known as a “hail cap” (default of 53 dBZ). This upper limit reduces contamination from hail by changing the reflectivity of all returns initially exceeding the hail cap to the reflectivity of the hail cap (Fulton et al. 1998).

Attempts to correct range-dependent errors due to the VPR effect have used both empirically-derived and model VPRs to determine the ratio of estimated precipitation at different tilt angles to the ground (Villarini and Krajewski 2010b for an extensive review). Vignal et al. (2000) found using a correction based on the spatial variability of VPR provided the most effective correction for radar precipitation estimates. However, any approach that uses spatial variability of the VPR requires some knowledge of the vertical distribution of reflectivity, information that is not available in the Stage IV dataset. Therefore any attempt to model the VPR in this study must do so as a function of range (i.e., Ciach et al. 2007; Krajewski et al. 2011) rather than height.

II.6.b.5) Other Range-Dependent Error Sources

The beam transmitted by the WSR-88D maintains a constant solid angle as it travels away from the radar site, so the beam widens as the distance from the radar increases (Legates 2000). This leads to an effect called partial beam filling, which can average out small, intense rain features imbedded in thunderstorms and lead to an underestimation of Z measurements (Villarini and Krajewski 2010b). At long ranges, these beams can overshoot cloud tops and completely miss a precipitation event that is being captured at the surface by gauge measurements (Smith et al. 1996). Both of these issues can cause serious degradation of long-range Z measurements (Fulton et al. 1998). The MPE delineation of each radar site's effective coverage area has diminished the influence of errors caused by partial beam filling and the overshooting of cloud tops. At a given HRAP grid cell, the MPE delineation assigns the WSR-88D that best minimizes range-dependent errors according to the radar climatologies (Seo 2002). However, there is still a general underestimation of precipitation in the Stage IV dataset for grid cells that are near the radar domain boundaries shown in Fig. 2.1.

Strong winds below the beam can produce systematic errors on shorter time scales by horizontally displacing a measured raindrop, an effect that has yet to be directly accounted for by the Q2 improvement of Stage II data (Zhang et al. 2011). If precipitation occurs in a region with a strong prevailing wind, there could be a systematic spatial displacement of radar precipitation estimates from the gauge measurements. Below beam evaporation can lead to an overestimation of precipitation (Krajewski and Smith 2002), while coalescence of raindrops below the beam can lead to the underestimation of precipitation by the WSR-88D (Legates 2000). The frequency of these below beam effects increases with distance from the radar as the vertical separation between the beam and the ground is increased.

II.6.c. Gauge Measurements

In reality, gauge measurements are irregularly spaced, point estimates of precipitation that suffer from a number of deficiencies (Sevruk 1991) and cannot be directly treated as the ground truth for precipitation data (Ciach and Krajewski 1999). Also, gauge networks are usually too sparse to properly capture the spatial variability of precipitation (Jayakrishnan et al. 2004). Inaccuracies in precipitation measurements taken from a single gauge be attributed to a number of sources: inhomogeneities in the precipitation record at the site, biases in the precipitation measurements, errors resulting in flawed gauge readings, and random errors (Groisman and Legates 1994).

Inhomogeneities in the precipitation record at a gauge site can appear as a discontinuity in time series of accumulated, e.g., annual precipitation or as a gradual change over time that cannot be explained by variations in climate (Peterson and Easterling 1994). Abrupt discontinuities are caused by instrumentation changes, exposure changes, and station moves. Changes in the environment surrounding the station may result in an artificial trend in a time series of precipitation values (Groisman and Legates 1994). The accuracy of gauge measurements is most affected by under-catch due to horizontal winds, the wind effect, which leads to a systematic underestimation of precipitation (Neff 1997). Changes in instrumentation at a gauge site can result in a change in the wind effect and can be responsible for inhomogeneities in the precipitation record (New et al. 2000).

Standard, non-recording precipitation gauges issued by the NWS are 20 cm in diameter, are elevated about 1 m above ground level, do not come with wind shields, and are susceptible to under-catch (Yang et al. 1998). Past studies have quantified the under-catch due to the wind effect by comparing above-ground, wind-exposed gauges to buried pit gauges (Neff 1977; Groisman and Legates 1994; Yang et al. 1998; Duchon and Essenberg 2001) and the underestimation of precipitation in the above-ground gauges relative to the collocated pit gauges ranged from 2% to as much as 10%. Under-catch increases with increased wind speeds and with a larger fraction of smaller drops and can be as high as 50% with snow (Nespor and Sevruk 1999).

The recording, tipping bucket gauge is another common type of gauge that produces inaccurate measurements due to the wind effect (because its orifice is similar), but the tipping bucket gauge also has a number of other deficiencies (Kalinga and Gan 2012). A tipping bucket gauge collects water in a bucket until it is full and then drains the water into a collection funnel, with the rainfall rate computed from the size of the bucket and number of tips within a prespecified period of time (Habib et al. 2001). When rainfall rates approach 300 mm hr^{-1}, water can accumulate faster than the drainage capacity of the tipping bucket gauge, so these gauges systematically underestimate precipitation during heavy rainfall events (Tapiador et al. 2011). During convective events, tipping bucket gauges can suffer from hardware malfunctions, transmission interruptions, and power failures (Habib et al. 2001).

Splashing, evaporation, and wetting, which occurs when precipitation adheres to a gauge and evaporates before being measured and additional reasons for the systematic underestimation of gauge measurements (Habib et al. 2001). Improper maintenance can lead to the orifice of a tipping bucket gauge to become out-of-level, and this tilt increases the magnitude of the wind effect and can lead to double tips (Sieck et al. 2007). The accumulation of particulates such as dust, leaves, insects, and bird droppings can clog standard gauges and adversely affect the tipping mechanism on tipping bucket gauges (Kalinga and Gan 2012). Random gauge errors include damage to gauges, inaccurate recording by tipping bucket gauges, and inaccurate observer reporting of non-recording measurements (Groisman and Legates 1994). In order to minimize the inaccuracies of gauge measurements, Ciach and Krajewski (1999) recommend the installation of two or more gauges of different designs at each measurement site to provide redundancy in the measurements.

Despite the numerous potential issues with individual gauges, any network of gauges should be free of spatially dependent biases since each gauge provides an independent measurement of precipitation. Though the Stage IV estimation algorithm uses gauge data to determine and correct for spatial biases in the radar precipitation estimates, there are biases that may aggregate over time that are not accounted for by the RFC correction procedure. Since the corrections are done on time scales of 24 hours or less (Lin and Mitchell 2005), it is difficult to account for subtle biases that show up more prominently in long-term data. However, gauge data aggregated over a period of time can be useful in evaluating and correcting these long-term spatial biases. These corrections can be done both in a single dimension, i.e. range, or can be applied to a given two-dimensional field of Stage IV precipitation data. Chapters IV and V will detail methods for using gauge data to minimize these long-term spatial biases.

II.7. WSR-88D Upgrade to Dual Polarization

The biggest development in the past few years has been the upgrade from single-polarization to dual-polarization (D-P) capabilities for the WSR-88D network. The upgrade started in 2011 and was completed for all WSR-88Ds in the continental United States by the end of June 2013 (Crum et al. 2013). The theoretical potential for D-P to reduce errors in radar precipitation measurements has been studied for more than a decade (Jameson 1991; Ryzhkov and Zrnic 1995; Doviak et al. 2000). The recent upgrades to D-P in the WSR-88D network has allowed for evaluation of D-P radar precipitation estimates, but because the upgrade was so recent, peer-reviewed publications have on the subject have been limited (Vasiloff 2012; Cunha et al. 2013; Ryzhkov et al. 2014).

Before the D-P upgrade, single-polarization (S-P) radars transmitted pulses with horizontally (H) polarized waves (Fulton et al. 1998). When a S-P radar beam encounters an object, it is only able to retrieve a one-dimensional measurement of that object, whether meteorological or non-meteorological. In addition to the H polarization, a D-P radar beam has the capability of transmitting a radar beam that is polarized in the vertical (V) direction. A WSR-88D with D-P capability will alternate between H and V polarization of transmitted beams, which allows to retrieve a two-dimensional measurement of an intercepted object (Doviak et al. 2000). The additional measurement capabilities of the D-P WSR-88D allows for a more accurate discrimination of hydrometeors from non-hydrometeor and characterization of hydrometeor type (Cunha et al. 2013).

The D-P radar technology has the capability of computing more variables related to the distribution of hydrometeors, which ought to improve radar precipitation estimation over the traditional single-polarization Z-R relationship. The differential reflectivity (Z_{DR}) measures the reflectivity difference of the H and V transmitted pulses in decibels, with the equation

$$Z_{DR} = 10\log\left(\frac{Z_H}{Z_V}\right). \tag{2.14}$$

Combining Z_{DR} with Z makes identification of hail more straightforward, since hail is more spherical and thus has a smaller Z_{DR} than rain drops (Straka et al. 2000). In wintry precipitation events, Z_{DR} can help delineate rain ($Z_{DR} > 1$ dB) from snow ($Z_{DR} < 0.5$ dB). The specific differential phase (K_{DP}) measures the difference in the propagation of the H and V transmitted beams and values of K_{DP} are highest in areas of heavy precipitation (Straka et al. 2000), and has a value

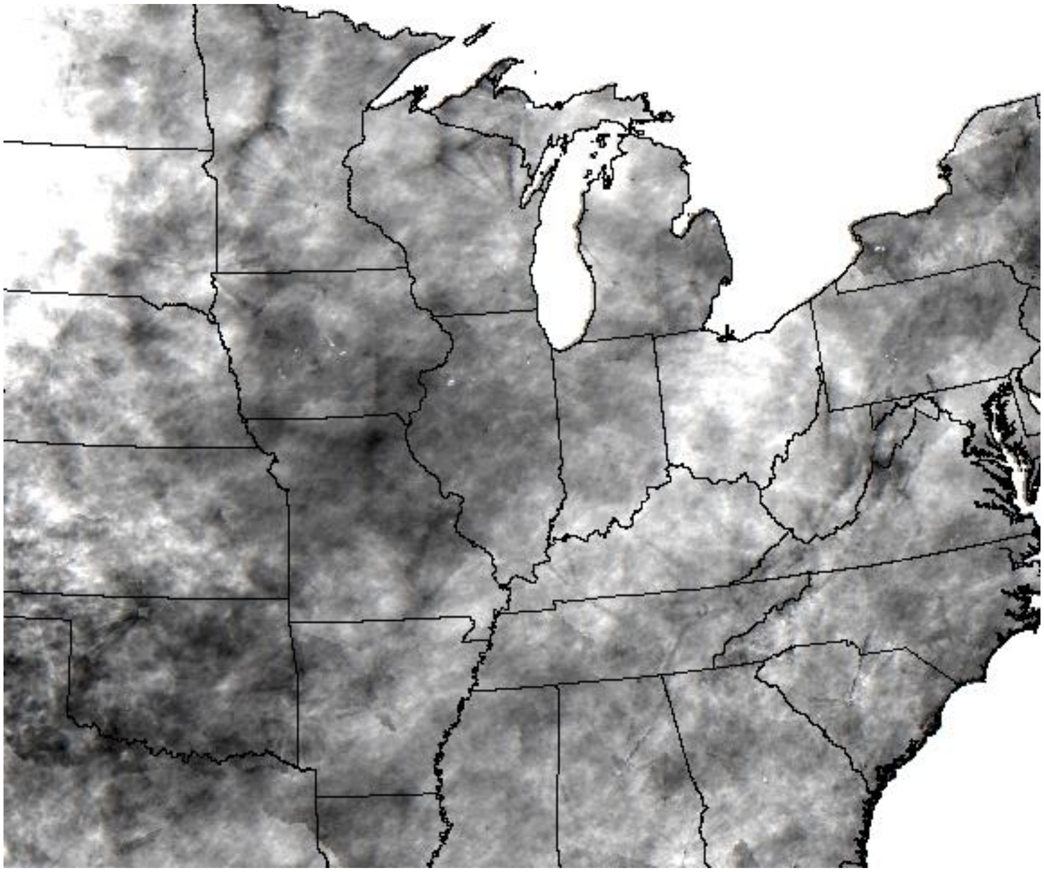
$$K_{DP} = \frac{1}{2} \frac{\phi_{DP}(d_2) - \phi_{DP}(d_1)}{d_2 - d_1}, \tag{2.15}$$

where ϕ_{DP} is the differential phase

$$\phi_{DP} = \phi_H - \phi_V . \tag{2.16}$$

The differential phase measures the difference in the propagation of an H wave relative to a V wave, with a larger differential expected as precipitation intensity increases due to attenuation of the signal. Ryzhkov et al. (2014) argue that a rainfall estimation algorithm based on the specific attenuation is preferable to using Z, Z_{DR}, and K_{DP} as the estimation algorithm is less sensitive to knowledge of the drop size distribution.

The computation of these new parameters in the WSR-88D algorithm has improved D-P precipitation estimation relative to the S-P estimates (Vasiloff 2012; Cunha et al. 2013; Ryzhkov et al. 2014), particularly in lighter precipitation events. Despite all the improvements to the D-P reflectivity measurements relative to the S-P measurements, the transmitted beams are still returning backscattered power from some height h above the surface that is dependent on range d. Even with a perfect assessment of the measured reflectivity, there will still be biases in the radar precipitation estimates that are dependent on range, assuming there is not a vertically uniform VPR. The Stage IV 12-month PoN precipitation ending on 30 June 2014 (Fig. 2.4a) is based on data collected entirely after the upgrade to D-P at all 104 radars in the central and eastern United States. The number of biases in Fig. 2.4a visually appears to be fewer than those in Fig. 2.4b, the Stage IV 12-month estimate with an ending date of 30 June 2011; however biases are still visible in Fig. 2.4a.



(a)

Fig. 2.4. Stage IV 36-month PoN precipitation ending (a) 30 June 2014 and (b) 30 June 2011. The minimum value (black) is *PoN* = 50% and the maximum value (white) is *PoN* = 150% in both maps.

(b)

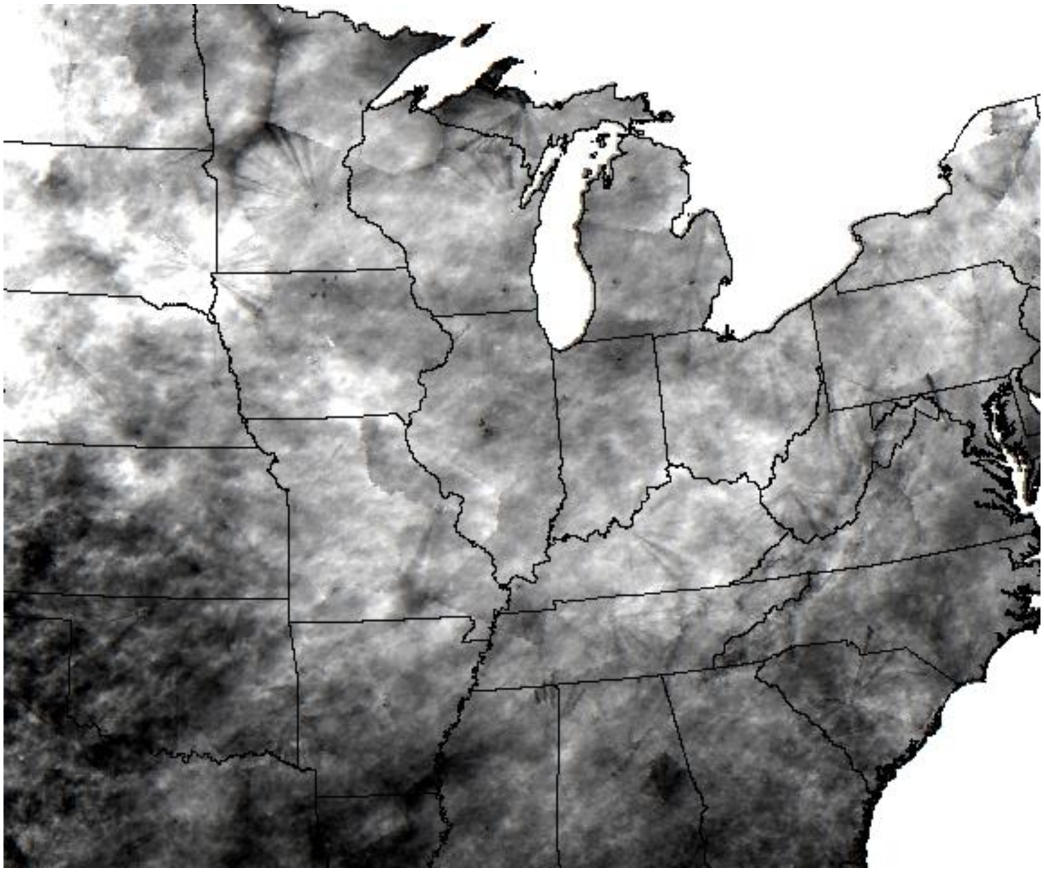


Fig. 2.4. Continued.

The first of the two types of biases that are the most prominent in the long-term Stage IV precipitation estimates is beam blockage. In the third (next) chapter, a beam blockage detection and correction methodology will be detailed. The fourth chapter will address the second and third types of bias prevalent in the Stage IV precipitation data, which are mean-field biases and range-dependent biases. The fifth chapter will detail a methodology for correcting two-dimensional biases remaining following the bias corrections done in chapters three and four.

CHAPTER III

PROCEDURE TO DETECT AND CORRECT BEAM BLOCKAGE

III.1. Overview

This chapter will detail the correction of biases that appear in the Stage IV precipitation dataset due to obstacles that block that path of the transmitted beam. In this chapter, the term obstacle is used for a fixed, non-meteorological target that inhibits propagation of the radar beam. The reduction in radar-estimated precipitation associated with true beam blockage in a given azimuth (θ) should be uniform beyond the range of the obstacle (d_b), aside from differences due to detection of hydrometeors at different altitudes. Given our conceptual equation of bias sources for radar precipitation estimates (P_0), equation (1.1) for azimuthal biases is

$$\frac{P_0}{P_T} = B\left(d, \theta\right). \tag{3.1}$$

Correction for the azimuthal-dependent biases caused by radar beam blockage will be done for regions objectively identified by a three-step beam blockage detection procedure. The first step of the detection process will flag HRAP grid cells within azimuths with anomalously low Stage IV precipitation values relative to grid cells in neighboring azimuths. The second step is to identify patterns of flagged grid cells consistent with beam blockage, which is done independently for each radar domain and time period. The third step in the detection algorithm is a temporal consistency check at each radar domain to prevent the spurious flagging of beam blockage in azimuths and to identify the onset or termination of beam blockage due to apparent man-made or man-controlled obstacles such as towers or trees. The precipitation values for the grid cells identified as being blocked by the beam blockage detection algorithm will be adjusted using neighboring, non-blocked grid cells to produce a dataset of beam blockage-adjusted Stage IV precipitation estimates (P_1), with the subscript 1 used since these data are adjusted for a single type of bias source.

III.2. Data

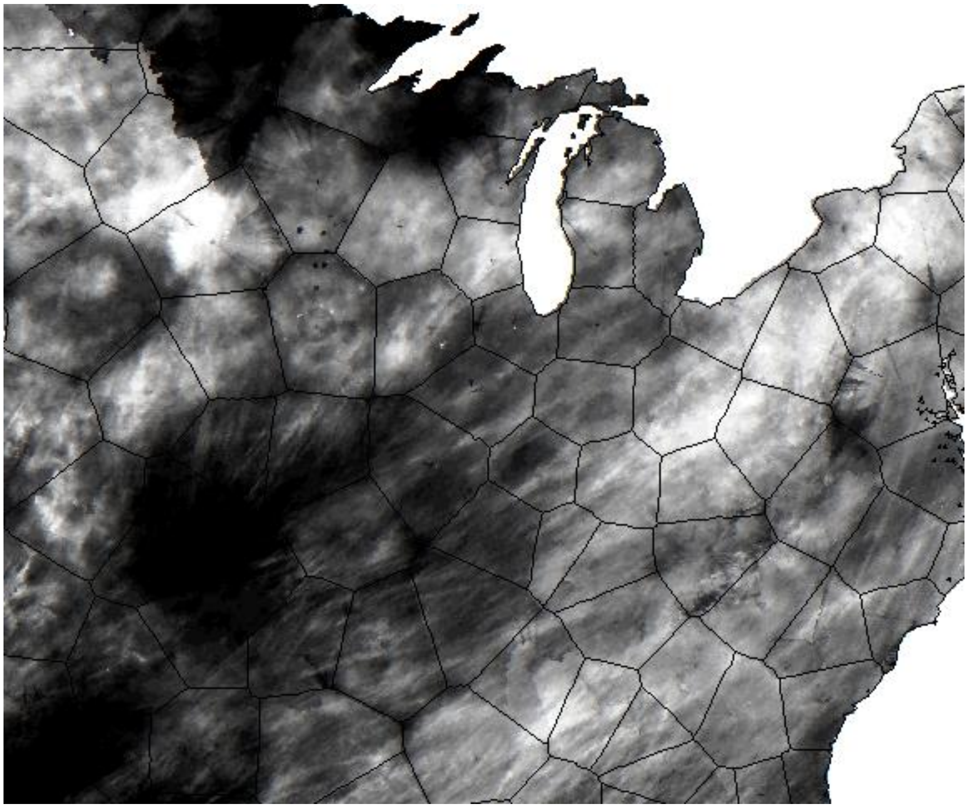
III.2.a. Computing Range

The location coordinates for both the radar-gauge pair biases and Stage IV PoN precipitation are the latitude and longitude at the center of the HRAP grid cell. The distance d of any from a nearby WSR-88D radar r is computed using the formula for the great circle on a spherical earth is

$$d = r_E \cos^{-1}\left[\cos\left(\Lambda - \Lambda_r\right)\cos\Phi\cos\Phi_r + \sin\Phi\sin\Phi_r\right], \tag{3.2}$$

where Φ is latitude, and Λ is longitude. Each HRAP grid cell was assigned to the WSR-88D in closest proximity, meaning the radar domains are Thiessen polygons given the network of available radars (Fig. 2.1). Assuming complete azimuthal coverage, the number of grid cells within the two ranges d_1 and $d_2 = d_1 + \Delta d$ increases with the square of the range. For example, the number of grid cells in an annulus centered at 100 km will have four times the number of grid cells as an annulus centered at 50 km.

Visual inspection of long-term Stage IV precipitation fields (i.e., Fig. 3.1b) demonstrates that maximum range of most beam blockage features are located at or very near a Theissen polygon boundary line. Additionally, there are numerous examples of discontinuities in the Stage IV precipitation fields occurring at the edges of the Theissen polygons (Figs. 3.1a and 3.1b). Therefore, the choice to assign each HRAP grid cell to the nearest single radar appears visually to be a reasonable one.



(a)

Fig. 3.1. The Stage IV (a) 1-month and (b) 36-month PoN precipitation ending on 31 December 2012. The minimum value (black) is *PoN* = 25% and the maximum value (white) is *PoN* = 250% in (a). The minimum value (black) is *PoN* = 50% and the maximum value (white) is *PoN* = 150% in (b).

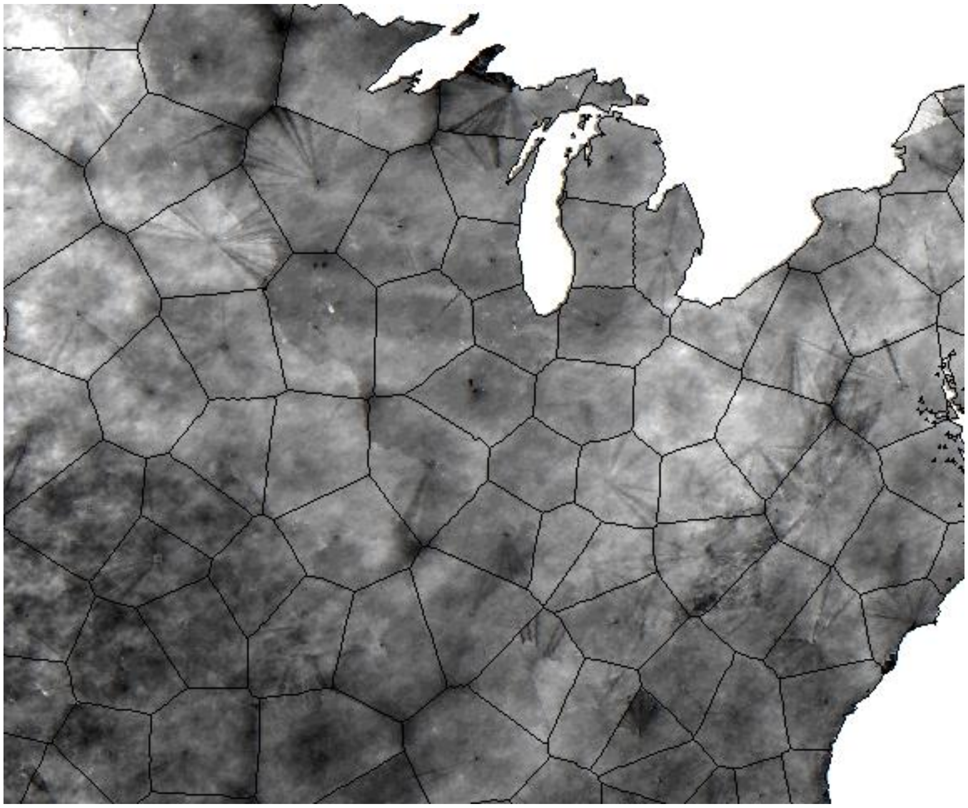
(b) 

Fig. 3.1. Continued.

III.2.b. Percent of Normal Precipitation

Beam blockage in the Stage IV precipitation dataset will be identified using 36-month PoN precipitation (P_N) data. The Stage IV data used in this study are from the period 1 January 2005 – 31 December 2012. The ending date of each accumulation period falls on the last calendar day of a month. In this chapter there are 61 overlapping periods for which Stage IV 36-month PoN precipitation was computed, with the first period ending on 31 December 2007. A time step Δt = 1 month exists between computed 36-month PoN precipitation values.

The stage IV PoN at a grid cell is $P_N = P_N(d_H, \theta_H)$, with the subscript H specifically referring to the grid cell coordinates relative to the nearest radar location. The computation of PoN precipitation is straight-forward

$$P_N = \frac{P_0}{P_n}, \tag{3.3}$$

where P_0 is the radar precipitation estimate and P_n is the climatological precipitation normal. PoN allows for direct comparison of the radar precipitation estimates between grid cells with knowledge of the expected spatial variability in the precipitation climate. Precipitation normals for the period 1981-2010 have been generated by the PRISM group (Daly et al. 1994) at 800m resolution using a digital elevation model, and are available on the HRAP grid in the continental United States. It is worth noting that the PRISM normals do not use radar inputs, so the normals data are free of any biases that show up in radar precipitation estimates.

The use of 36-month periods, as opposed to shorter time periods, decreases the likelihood of both Type I and Type II errors in the identification of azimuthal sectors contaminated by beam blockage (Fig. 3.1). If using data from a shorter time scale, random variability may allow real swaths of local precipitation minima to be misidentified as beam blockage (Type I error). Additionally, the use of shorter time periods may result in failing to identify areas actually affected by beam blockage (Type II error) because the ratio of spatial variability to precipitation amount is too large. For example, identifying beam blockage in the Southeast US during a single summer season could be difficult if the only precipitation mechanism is isolated convection resulting from daytime heating. A time period of 36 months ensures a representative sample of actual precipitation at all grid cells within each radar domain. However, it is also important to have a time period short enough to adequately identify any sources of temporal changes in the beam blockage properties at each radar, and 36 months was thought to be a sufficient upper limit. These sources include changes in the radar network configuration, the construction or tearing man-made or man-controlled obstacles, and any changes in the Stage IV processing scheme discussed in the previous chapter.

III.3. Radar Geometry

III.3.a. Computing HRAP Grid Cell Azimuth Angles

Computation of the azimuth angle θ_H for an HRAP grid cell uses the point at the center of the HRAP grid cell and the radar *r*. In degrees, the value of θ at the center of the HRAP grid cell is

$$\theta_H = \frac{180^\circ}{\pi}\left[\arctan\left(\frac{\cos\Phi_H \sin\left(\Lambda_H - \Lambda_r\right)}{\sin\Phi_H \cos\Phi_r - \cos\Phi_H \sin\Phi_r \cos\left(\Lambda_H - \Lambda_r\right)}\right)\right], \tag{3.4}$$

where (Φ_H, Λ_H) are the latitude-longitude coordinates of the grid cell center and (Φ_r, Λ_r) are the coordinates of the WSR-88D, all in radians. However, each grid cell has a width $W_H = 4$ km and this cannot be ignored since the azimuthal width $d\theta_H$ taken up by each grid cell has a functional dependence on range. The azimuthal width $d\theta_H$ at d_H is approximated to be the ratio of W_H to the circumference of a circle at that range, given by

$$d\theta_H = \frac{W_H}{2\pi d} \times 360^\circ . \tag{3.5}$$

III.3.b. Azimuthal Sectors

The half-power beam width $\theta_{1/2}$ of an operational WSR-88D is 1° (Fulton et al. 1998; Fig. 2.3), so each radar domain will be divided into 360 non-overlapping “azimuthal sectors,” with each sector having an azimuthal span of 2 $\theta_{1/2}$ = 1°. Each azimuthal sector has a width of exactly 1° and the azimuthal endpoints of each azimuthal sector will be whole degrees, e.g., the 90° sector will span 90° to 91°. All azimuthal sectors within a given domain will originate at the site of the WSR-88D and extend radially to the edge of the radar domain (d_{max}). A span of ranges in a given azimuthal sector will be referred to as a “radial span.” For example, the area in the 90° sector from 30 km to 60 km from the radar location is the 30 km – 60 km radial span of the 90° sector.

In the beam blockage detection algorithm, the initial objective is to determine whether or not each of these azimuthal sectors may be blocked. If possible blockage is detected for a given azimuthal sector, the second objective is to determine at which range d_b the blockage occurs. The basic geometry of the azimuthal sectors dictates that the geographical width of each azimuthal sector (W_θ) increases with distance from the radar location. Replacing $d\theta_H$ with 1° and W_H with W_θ in (3.5) and then solving for W_θ, we get

$$W_\theta = \frac{\pi d}{180^\circ} . \tag{3.6}$$

Given that the Stage IV 36-month PoN precipitation data is used to detect beam blockage, each HRAP grid cell needs to be assigned to the azimuthal sectors. However, the HRAP grid cells have a constant width of $W_H = 4$ km that is independent of range. Given the complexity of the geometry, the task of assigning HRAP grid cells to azimuthal sectors is not as straight-forward as simply assigning each grid cell to a single azimuthal sector based on the value of θ_H found in (3.4). The azimuthal width $d\theta_H$ is greater than 1° for d_H values less than 229.1 km, so in these grid cells beam blockage in any of two or more azimuthal sectors may lead to precipitation values that are biased low. Therefore, grid cells cannot be restricted to a single azimuthal sector in the beam blockage detection procedure. For an HRAP grid cell to be used in the beam blockage detection algorithm for a given azimuthal sector, the grid cell must span at least half of the azimuthal sector at range d_H.

III.4. Beam Blockage Detection Procedure

This section describes a flagging procedure for detection of areas in each radar domain that are affected by beam blockage. The foundation of this section is that at a constant range and a time period of 36 months, PoN precipitation modeled purely as a function of azimuth for all possible azimuths tends toward a smooth and continuous curve. For an azimuthal sector to be identified as having beam blockage, it must have the following properties flagged.

1. Sufficiently lower PoN precipitation than predicted by the models.
2. Continuous radial span of sufficiently lower than expected PoN precipitation.

3. Temporal consistency of the continuous radial span having lower than expected PoN precipitation.

The first stage in detecting regions of beam blockage is to assign a "low precipitation" flag $F_1(d_H, \theta_H)$ to individual HRAP grid cells with PoN precipitation values that are significantly less than the model curve. This initial flagging ($F_1 = 0$ if not flagged, $F_1 = 1$ if flagged) is done independently for each grid cell. The second stage is to assign a "range continuity" flag $F_2(\theta)$ for azimuthal sectors at each radar with a sufficient number and radial consistency of grid cells flagged in the first stage. The third stage is a quality control procedure designed to ensure that the azimuthal sectors flagged in the second stage have sufficiently low 36-month PoN precipitation values relative to neighboring sectors with $F_2 = 0$. If an azimuthal sector fails the third stage quality control test, its F_2 value is changed from one to zero.

The first three stages of the beam blockage detection algorithm are done independently for all 61 of the 36-month periods in the timeframe of this study. At this point each flag is a function of both azimuthal sector θ and time t, such that "range continuity" flag $F_2 = F_2(\theta, t)$. The fourth stage checks the temporal consistency of the of the F_2 flags for each azimuthal sector. A time series of F_2 flags is constructed, and for temporal segments of the time series with a sufficient consistency of F_2 flags, each time t within that segment is assigned a "temporal continuity" flag $F_3(\theta, t) = 1$ and at all other times $F_3(\theta, t) = 0$. For a given azimuthal sector θ, each 36-month period t is considered to have beam blockage detected if $F_3(\theta, t) = 1$.

III.5. “Low precipitation” Flagging Procedure

III.5.a. Separating PoN Data into Annuli

The flagging of HRAP grid cells with anomalously low precipitation values is the first stage in the beam blockage detection procedure. Ideally, this would be accomplished by modeling Stage IV 36-month PoN precipitation purely as a function of azimuth with no range dependence of the data. However, each grid cell has a unique d_H value, which makes it impossible to create a beam blockage detection model using only data with a constant range. In order to model the Stage IV PoN values in each WSR-88D domain strictly as functions of azimuth angle, the HRAP grid cells need to be grouped in a way that provides both sufficient azimuthal sampling and minimizes contamination from range-dependent biases.

The first step in the “low precipitation” flagging process is to divide each radar domain into non-overlapping annuli centered at the radar location. Each annulus is bounded by an inner ring with radius r and an outer ring with radius $R = r + d\rho$, where $d\rho$ is the radial span of the annulus. The division of each radar domain into annuli allows each HRAP grid cell to be grouped with other grid cells having a similar distance to the radar location. The choice of $d\rho$ is critical because if it is too small, azimuthal sampling for the model at a given annulus may be insufficient and if $d\rho$ is too large, the model may be contaminated by range-dependent biases. A universal choice of $d\rho = 10$ km was decided as an acceptable compromise between the two aforementioned and competing factors. Fig. 3.2 shows the annuli for the Minneapolis, MN (KMPX) and Slidell, LA (KLIX) radar domains. The

KMPX domain (Fig. 3.2a) has complete azimuthal coverage out to a range of roughly 110 km, whereas the KLIX domain (Fig 3.2b) only has complete coverage out to a range of about 30 km. For a given 36-month period, the Stage IV PoN data are grouped by annuli and ordered by azimuth angle.

(a)

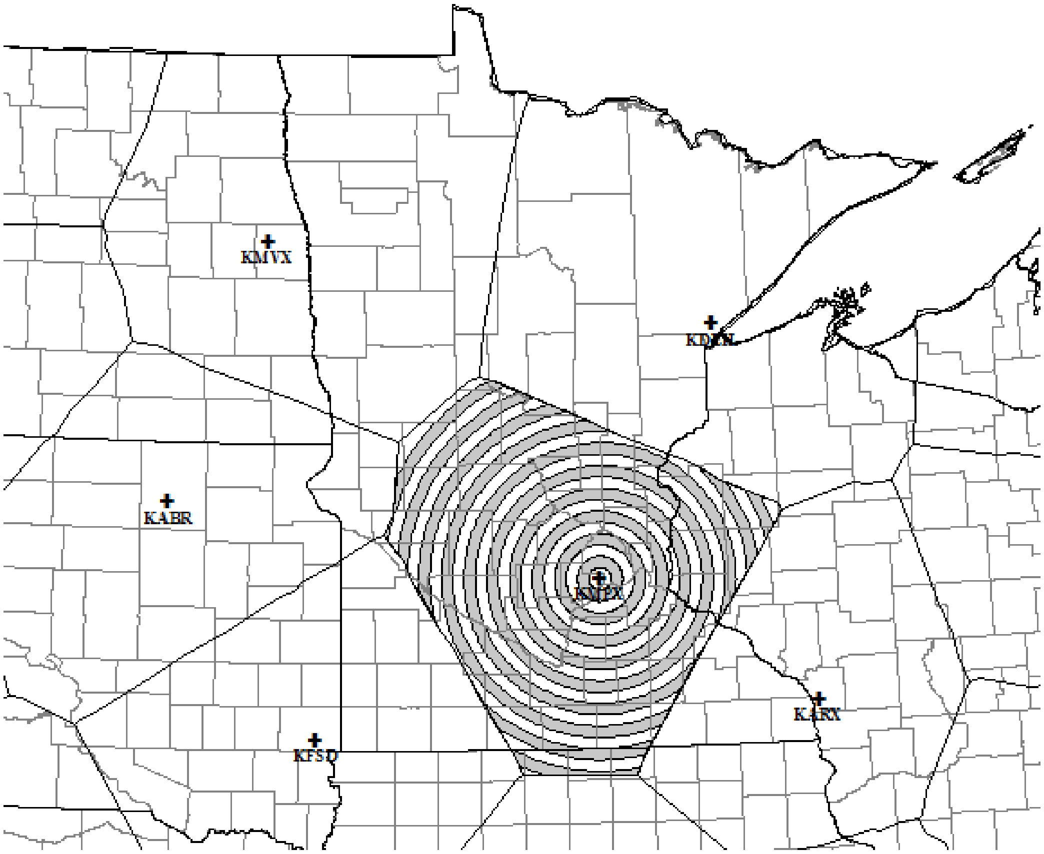
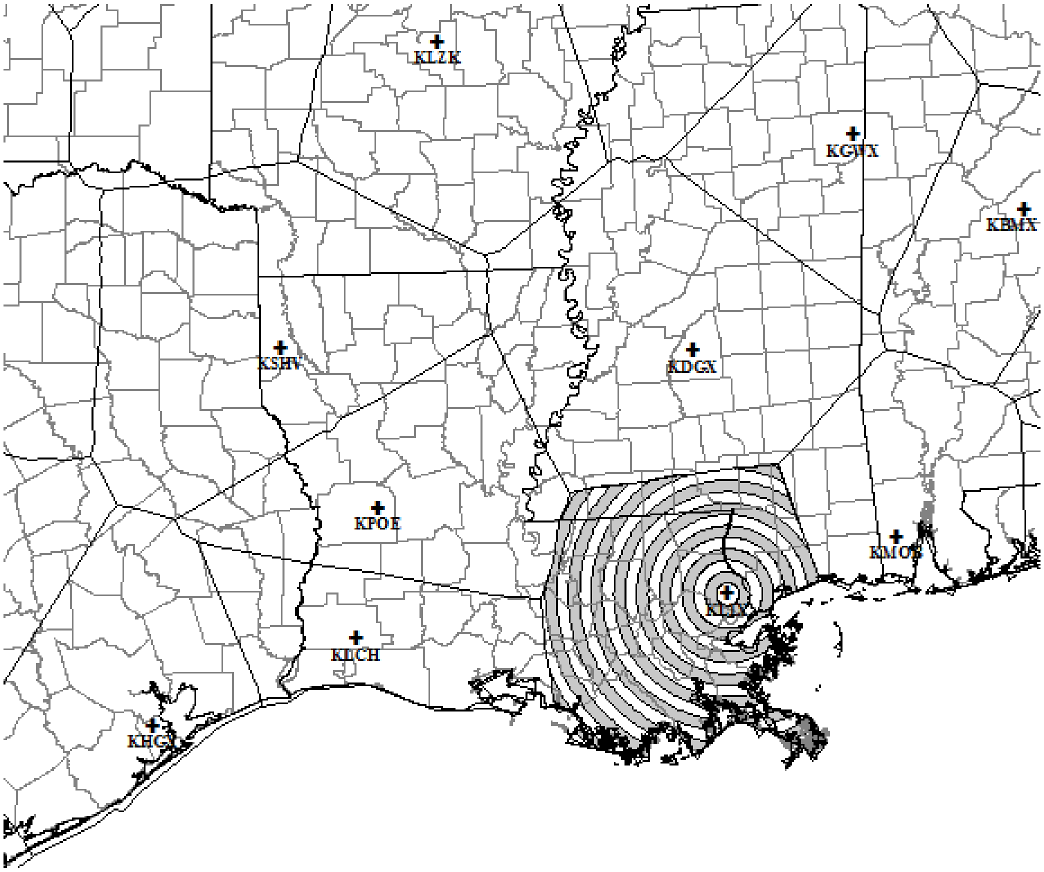


Fig. 3.2. The WSR-88D locations (crosshairs) and the annuli (circles) for the (a) KMPX and (b) KLIX radar domains.



(b)

Fig. 3.2. Continued.

III.5.b. Modeling Percent of Normal as Function of Azimuth Angle

For a given radar and annulus, the Stage IV 36-month PoN precipitation is modeled using a low-order Fourier series function $f(\theta)$ in the flagging of "low precipitation" HRAP grid cells. This model form was chosen to ensure the function is periodic over the 360° span of the annulus. Another desirable property of the Fourier series is that it will provide a realistic model form for all possible azimuthal distributions of PoN precipitation. This includes an exact fit of $f(\theta)$ for an annulus with azimuthally uniform field of PoN precipitation values or for a field with linear spatial variations.

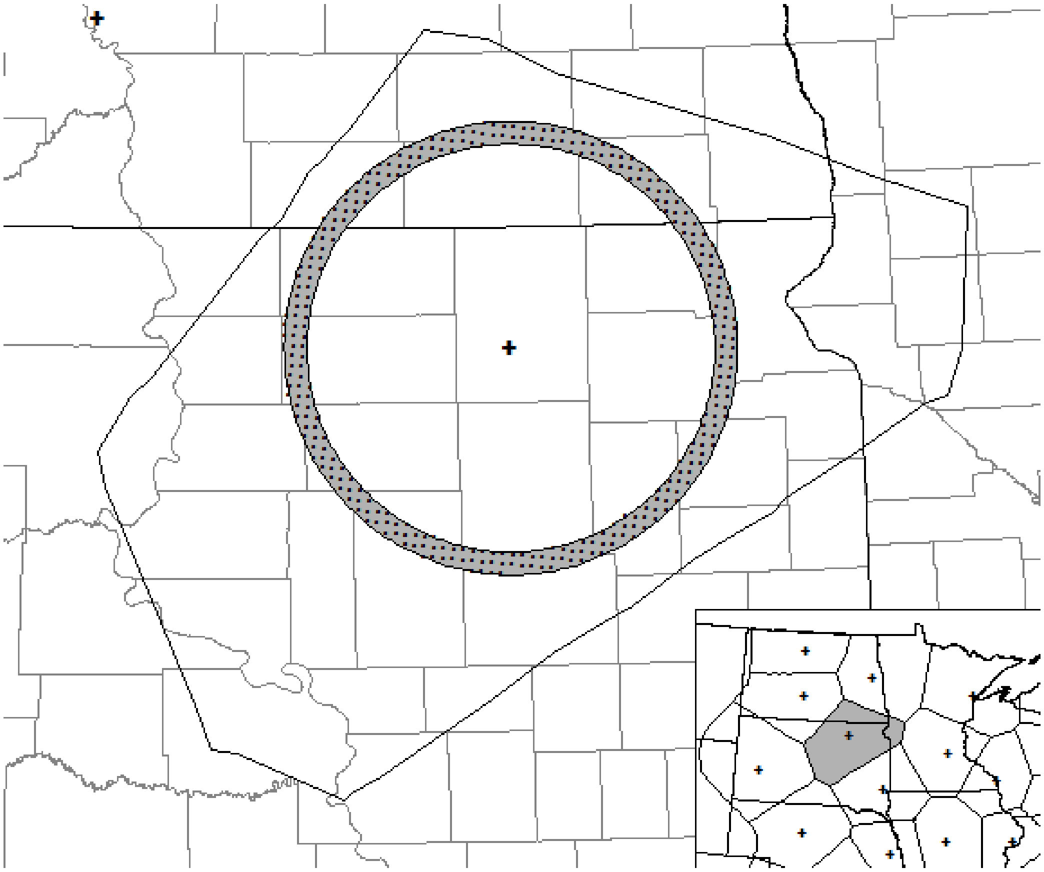
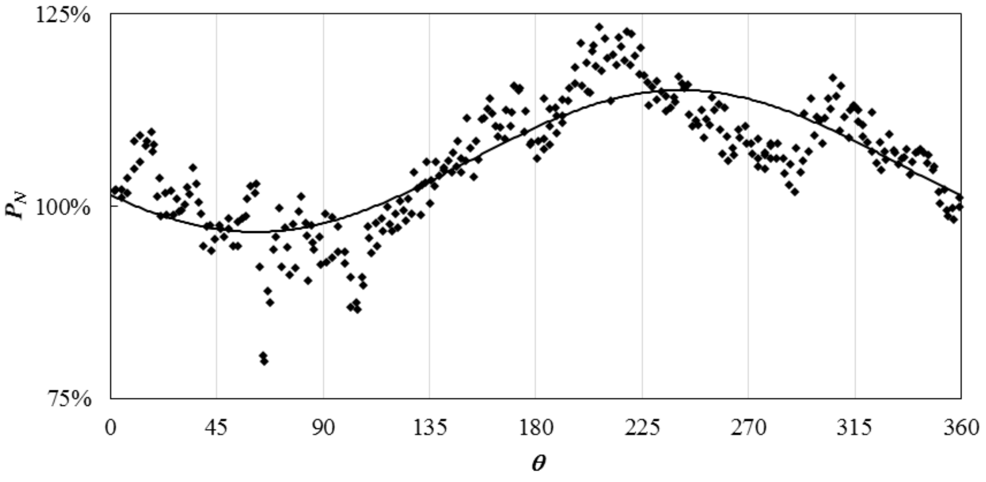
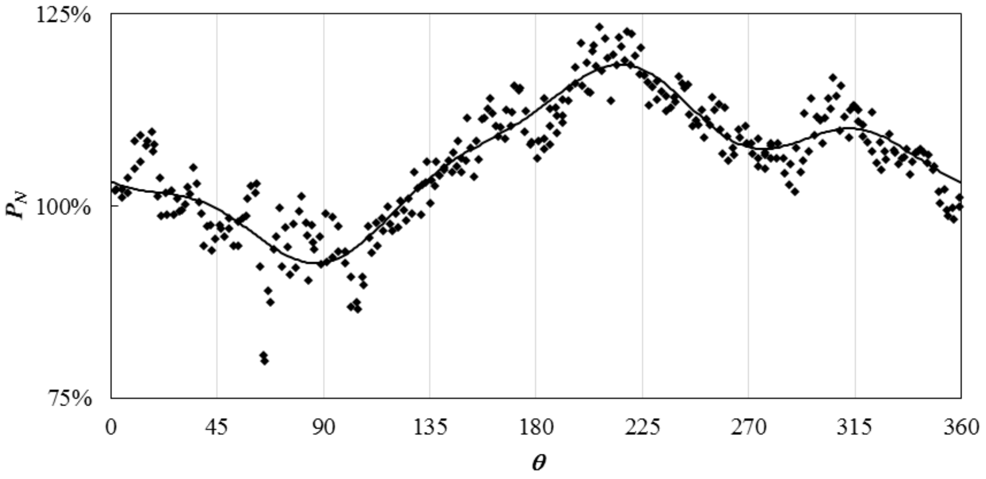
(a) 

Fig. 3.3. (a) The annulus (gray ring) and HRAP grid cells (black dots) for the 90-100 km radial span of the KABR radar domain, centered at the crosshairs. The Stage IV 36-month PoN precipitation data, ending 31 December 2012, as a function of azimuth (black diamonds) in the annulus with a Fourier series fits (black line) using $k = 1$ (b), $k = 4$ (c), and $k = 16$ (d).

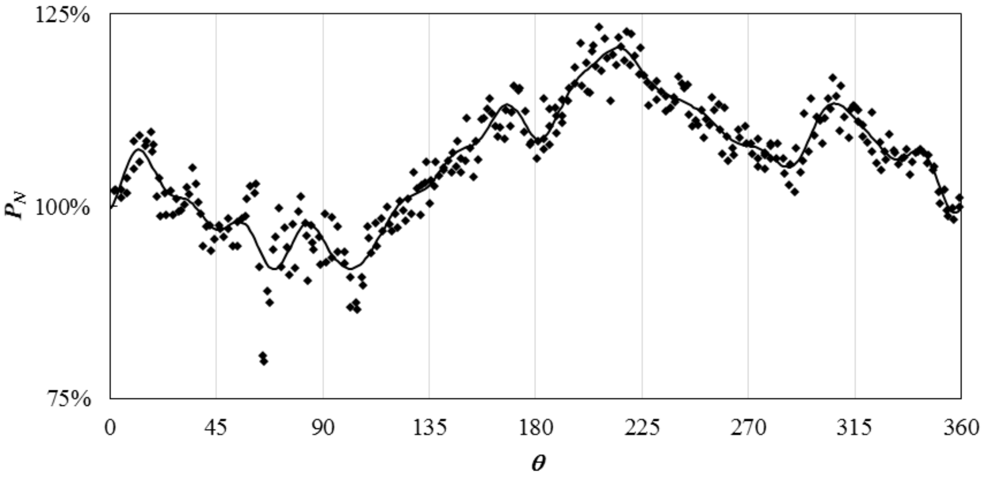


(b)



(c)

Fig. 3.3. Continued.



(d)

Fig. 3.3. Continued.

The wavenumber *k* for the Fourier series fit needed to both be adequate for depicting the actual azimuthal variability of rainfall, which is improved with an increase in *k*, but resistant to any overfitting to beam-blockage artifacts that may occur if the wavenumber is too large. Fig. 3.3 is the Fourier series function of the annulus with a radial span of 90 km – 100 km at the Aberdeen, SD (KABR) radar for the 36-month period ending 31 Dec 2012, using *k* =1, *k* = 4, and *k* = 16. Based on visual inspection of Fig. 3.3 and numerous other examples, *k* = 4 is an appropriate wavenumber for the Fourier series to model the azimuthal variations of PoN precipitation in an annulus without overfitting (Fig. 3.3c).

The form of the Fourier series is

$$f(\theta) = c + \sum_{n=1}^{k} a_n \cos(n\theta) + \sum_{n=1}^{k} b_n \sin(n\theta), \tag{3.7}$$

where

$$c = \frac{1}{n_N}\sum_{i=1}^{n_N} P_i \, , \tag{3.8}$$

is the mean using all n_N Stage IV PoN precipitation data points (P_i). For each order n,

$$a_n = \sum_{i=1}^{n_N} \left(P_i - c\right)\cos\left(n\theta\right), \tag{3.9}$$

and

$$b_n = \sum_{i=1}^{n_N} \left(P_i - c\right)\sin\left(n\theta\right). \tag{3.10}$$

III.5.c. Flagging Algorithm

In a given radar domain, the Fourier series modeling is done independently for each annulus. The “low precipitation” flagging assigns $F_1(d_H, \theta_H) = 1$ to each HRAP grid cell within an annulus having large negative residuals to the model fit $f(\theta)$ in (3.7). The magnitude of each residual i is measured by the squared residual (r_i^2), which is

$$r_i^2 = \left(P_i - f\left(\theta_i\right)\right)^2 . \tag{3.11}$$

After computing r_i^2 for each P_i, the grid cells are ranked and arranged in descending order by the magnitude of the squared residual. The flagging process compares the magnitude of r_i^2 at each grid cell i to the mean squared residual (MSR) for all the grid cells ranked higher. For a grid cell with a rank of i, the value of MSR_i is

$$MSR_i = \frac{r_i^2}{MSR}, \quad (3.12)$$

where N is the total number of Stage IV PoN data points in the annulus used to compute $f(\theta)$. The assumption is that if the ratio R_i of the squared residual for a grid cell i (3.11) to the MSR of all the other grid cells with a higher rank (3.12) is sufficiently large, then the Stage IV PoN precipitation value is an outlier that may have been caused by beam blockage. The ratio R_i in equation form is

$$R_i = \frac{r_i^2}{MSR}, \quad (3.13)$$

and takes into account and normalizes the natural variability of the residuals in the annulus. If the residual r_i is negative and the ratio R_i is greater than a pre-determined threshold ratio value R_0, the grid cell is flagged as being blocked. The threshold ratio value $R_0 = 3.75$ was determined based on visual inspection of the resulting blockages using several different R_0 test values.

The step-wise algorithm for the "low precipitation" flagging of individual HRAP grid cells within an annulus is as follows.

1. Set $f(d_H, \theta_H) = 0$ for all N grid cells.
2. Create a Fourier series model fit $f(\theta)$ using all N grid cells within the annulus.
3. Compute the value r_i^2 for each of the N grid cells.
4. Rank and arrange the grid cells by magnitude of r_i^2 in descending order.

5. Set $i = 1$.
6. Compute R_i.
7. Is $R_i \geq R_o$? If yes, proceed to Step 8. If no, proceed to Step 10.
8. If $r_i < 0$, set $F_1(d_H, \theta_H) = 0$ for the HRAP grid cell with rank i.
9. Increase i by one and go back to Step 5.
10. "Low precipitation" flagging algorithm is complete. Store the number of grid cell processed by the algorithm (N_1).

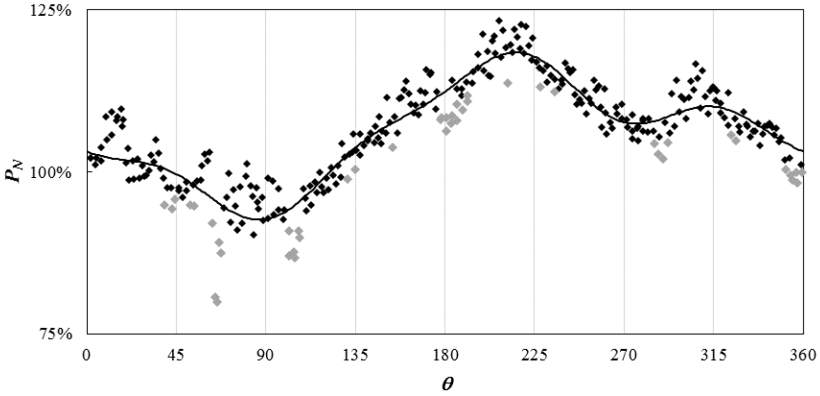


Fig. 3.4. The same as Fig. 3.3b, but with HRAP grid cells detected by the flagging algorithm (gray diamonds) in the 90 km – 100 km annulus in the KABR radar domain.

Fig. 3.4 shows the same Stage IV 36-month PoN precipitation from the KABR radar domain modeled using a Fourier series with k =4 shown in Fig. 3.2c with the grid cells flagged by the aforementioned algorithm. Fig. 3.5 shows the Stage IV 36-month PoN and the flagged HRAP grid cells for all annuli in the KABR radar domain. Visual inspection shows there is likely blockage in the PoN field in azimuthal sectors around 65° and this is well-detected for all annuli (Fig. 3.5). The results of the "low precipitation" flagging procedure will be analyzed to detect radial spans with a consistency of F_1 values in the "range continuity" flagging procedure.

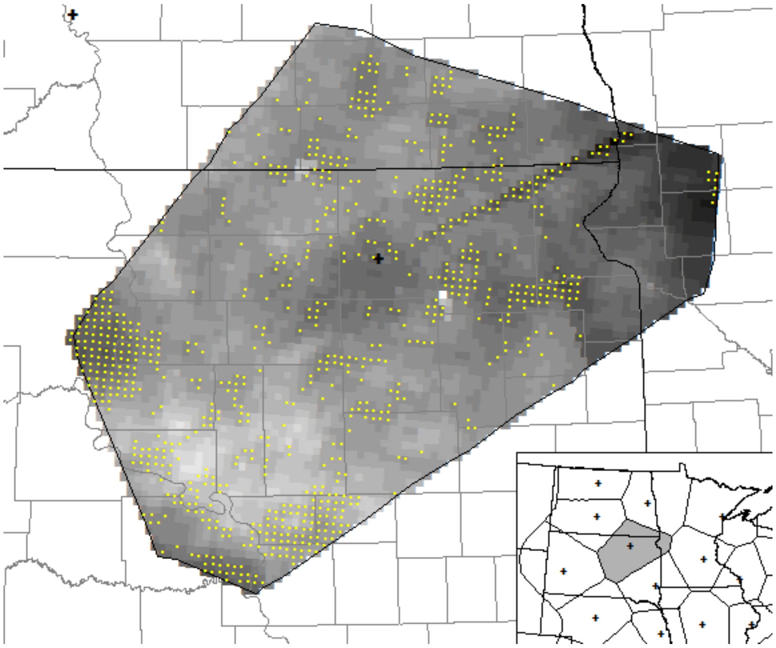


Fig. 3.5. The Stage IV 36-month PoN precipitation, ending 31 December 2012 for the KABR radar domain (contoured field); minimum value (black) is *PoN* = 50% and the maximum value (white) is *PoN* = 150%. The HRAP grid cells detected by the flagging algorithm are included (yellow squares).

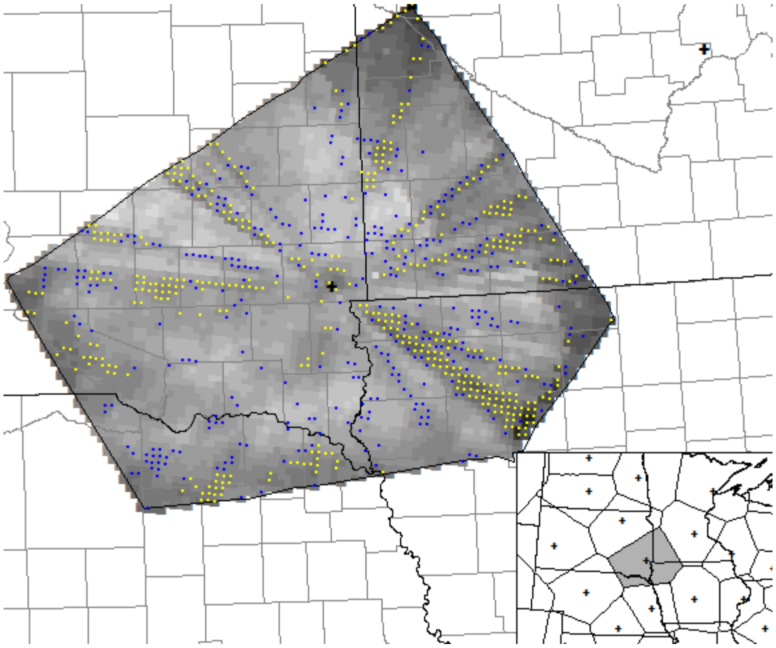


Fig. 3.6. The Stage IV 36-month PoN precipitation, ending 31 December 2012 for the Sioux Falls, SD (KFSD) radar domain (contoured field); minimum value (black) is *PoN* = 50% and the maximum value (white) is *PoN* = 150%. The HRAP grid cells detected by the initial flagging algorithm ($k = 4$; yellow squares) and the secondary flagging algorithm ($k = 12$; blue squares) are included.

A viable solution found to combat this issue was to rerun the flagging algorithm in each annulus to compute a new Fourier series. This version of the algorithm builds the model $f(\theta)$ using only Stage IV 36-month PoN precipitation data whose residual was not flagged in the initial run. The second run of the algorithm models the remaining PoN precipitation data using a Fourier series with $k = 12$ (e.g., Fig. 3.6). The assumption for this second run is that it has the capability of capturing beam blockage features in azimuths where the variability was too complex for a proper assessment using a lower-wavenumber Fourier series fit. To reduce the variance of the second Fourier series, all grid cells with a rank ≤ N_1 were not used in its construction, even those with positive residuals and an $F_1 = 0$.

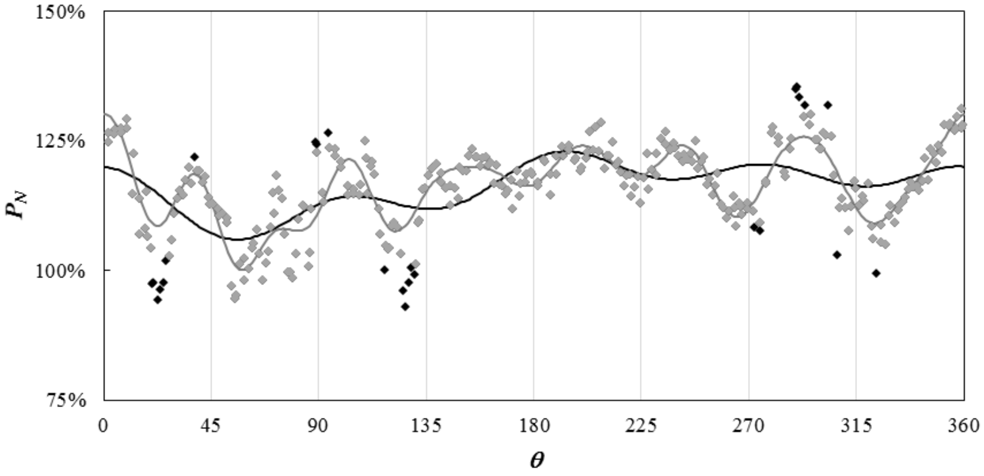


Fig. 3.7. The KFSD Stage IV 36-month PoN precipitation data, ending 31 December 2012, as a function of azimuth (all diamonds) in the 90-100 km annulus with the initial Fourier series fits (black line). The grid cells with $R_i \geq R_0$ (black diamonds) are removed for the second flagging algorithm Fourier series fit (gray line).

Fig. 3.7 shows the Fourier series model fit for each HRAP grid cell flagging algorithm run for the KFSD annulus with a radial span of 90 km – 100 km, using 36-month PoN precipitation data from 31 December 2012. Of note, is the difference in the two models for azimuth angles between 135° and 180°. In the second run of the "low precipitation" flagging algorithm, the model more accurately characterizes the azimuthal variability. This allows the second run of the algorithm to flag HRAP grid cells around 145° as being potentially contaminated by beam blockage.

III.6. “Range Continuity” Flagging Procedure

In a given radar domain, azimuthal sectors with Stage IV 36-month PoN precipitation data contaminated by beam blockage will ideally have all HRAP grid cells flagged in the “low precipitation” flagging procedure beyond a range d_b where the radar beam intercepts an apparent obstacle. However, other types of errors and the reality of the natural spatial variability of precipitation add noise to the data, even at the 36-month time scale, and prevent proper “range continuity” flagging of these grid cells. The noise in the data may also lead to the misidentification of beam blockage within a given annulus (Fig. 3.4 for example), but these grid points tend to be random with little consistency with range in a given sector. In Fig. 3.4, the azimuthal sectors around 65° that visually appear to suffer from beam blockage in the PoN precipitation data are consistently, though not always, flagged. The guiding principle of flagging azimuthal sectors in the “range continuity” flagging procedure is a consistency of grid cells with $F_1 = 1$ in the radial span beyond the point where the radar beam encounters an apparent obstacle. In addition to the noise in the data, the basic geometry of the radar domains present issues for the consistency of flagging within an azimuthal sector.

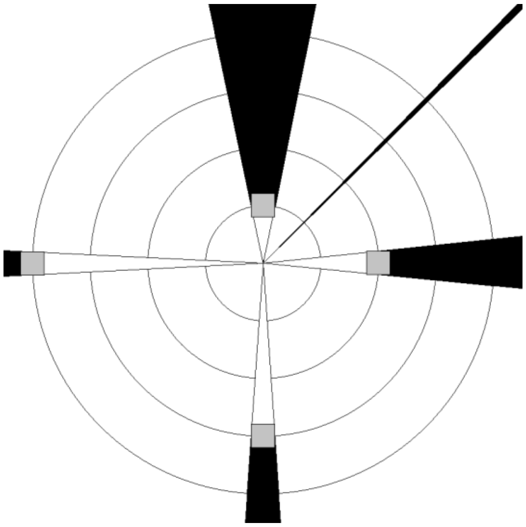
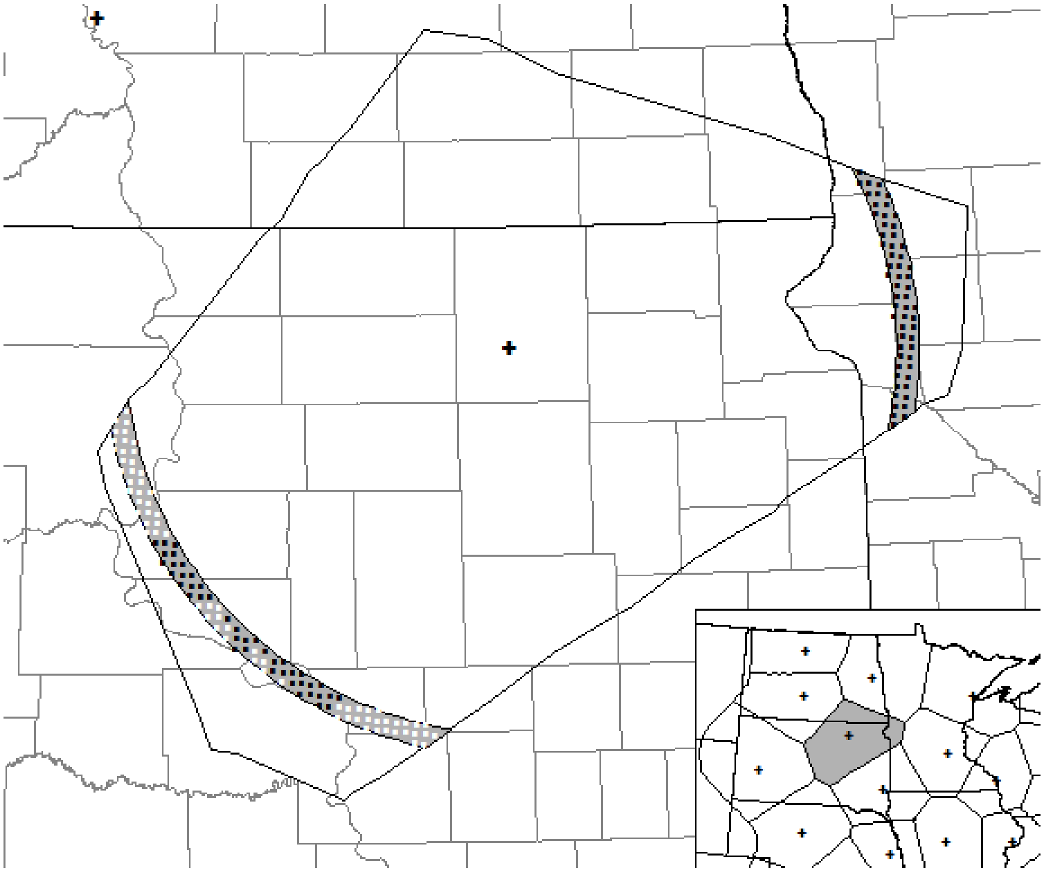


Fig. 3.8. A conceptual model of the azimuthal coverage (whit-e slices) and blockage (black slices) of HRAP grid cells (gray boxes) for ranges of 10 km (top), 20 km (right), 30 km (bottom), and 40 km (left). For reference, an azimuthal sector of width 1° was included at an azimuth of 45°.

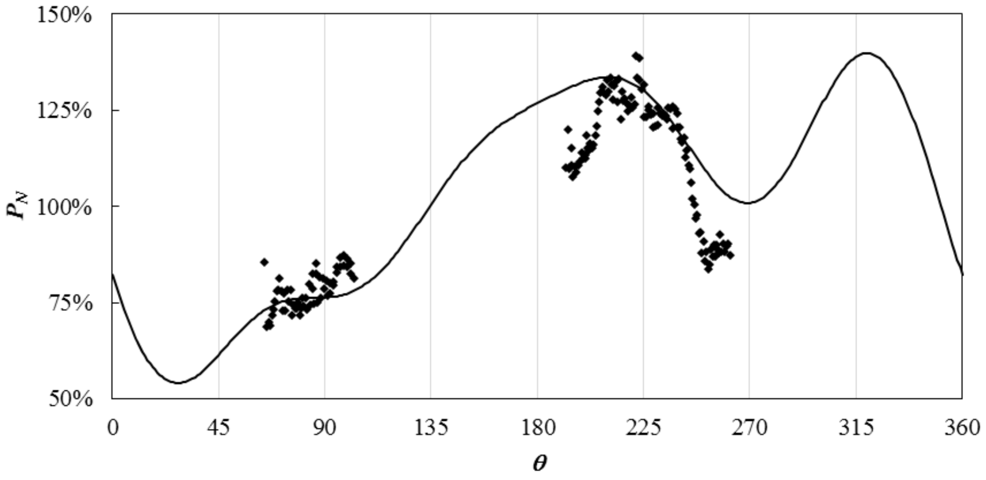
III.6.a. Geometric Limitations of "Range Continuity" Flagging Procedure

The first factor to take into consideration in the "range continuity" flagging procedure is the large azimuthal width $d\theta_H$ of grid cells close to the radar location. The azimuthal width of the HRAP grid cells is displayed graphically in Fig. 3.8 for the outer rings of the first four annuli out to 40 km. Real beam blockage caused by a point source within a grid cell at close ranges may be smoothed out if adjoining sectors within the same grid cell are not affected. The issue of large $d\theta_H$ values arises when grid cells at close ranges are flagged (F_1 = 1) because it is difficult to assess the true azimuthal width of "low precipitation" given a single grid cell. However the impact of this issue should be minimal, because the azimuthal width of true beam blockage occurring very close to the radar can be determined by grid cells with larger *d* values. In azimuthal sectors affected by beam blockage, there should be a large number of grid cells with $F_1 = 1$ beyond the range of initial blockage.



(a)

Fig. 3.9. (a) The annulus (gray ring) and HRAP grid cells (black dots) for the 170-180 km radial span of the KABR radar domain, centered at the crosshairs. (b) The Stage IV 36-month PoN precipitation data, ending 31 December 2012, as a function of azimuth (black diamonds) in the annulus with the initial Fourier series fits (black line).



(b)

Fig. 3.9. Continued.

The second and more troublesome concern is that for most radar domains, there are Fourier series fits for annuli at far ranges that have poor azimuthal coverage. For example, the annulus with a radial span of 170 km – 180 km at the KABR radar domain (Fig. 3.9a) only has two azimuthal segments; one from about 65° to about 100° and the other from about 190° to about 260°. The lack of azimuthal coverage is due to the irregular shape of the radar boundary polygon and leads to the Fourier series providing a poor fit for the available data, particularly for the segment in the western part of the domain (Fig. 3.9b). The poor fit leads to the spurious detection of beam blockage for about 40% of the grid cells in the western segment of the KABR radar domain, simply because the model fit $f(\theta)$ is so much better for the eastern segment (Fig. 3.9).

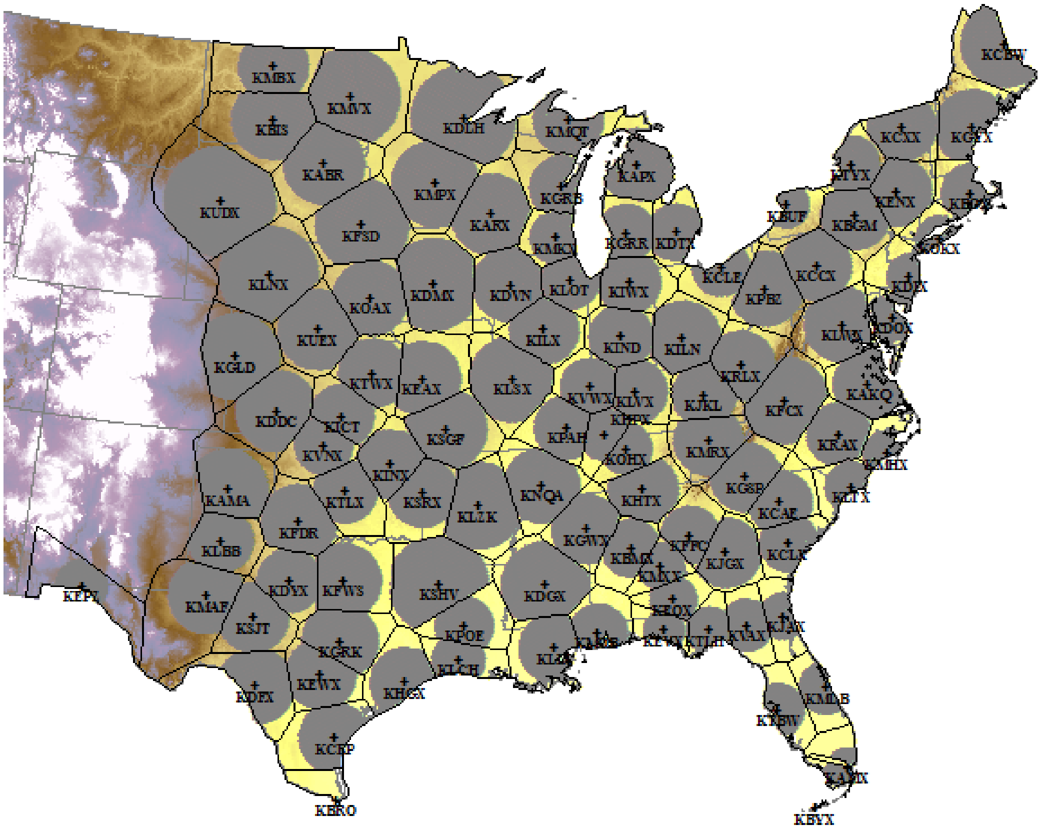


Fig. 3.10. Locations (crosshairs) and regions covered (polygons with black outlines) for the 104 WSR-88Ds in the central and eastern United States. The HRAP grid cells which belong to annuli with at least 180° of azimuthal coverage are included.

Therefore, a restriction was put in place so that the initial algorithm in the "range continuity" flagging procedure will only flag radial spans covering annuli with at least 180° of azimuthal coverage. The range interval of the outermost annulus meeting this coverage criteria will vary from radar domain to radar domain. For a given azimuthal sector, the range of the outer ring meeting the coverage criteria is assigned to a variable $d_x(\theta)$ if the range is less than d_{max}; otherwise $d_x = d_{max}$. Fig. 3.10 shows the spatial coverage of the grid cells examined in this first step for each of the 104 radar domains in the central and eastern United States. Using the criterion of at least 180° rather than some larger threshold such as 270° or 360° ensures that the beam blockage detection algorithm can be used on most coastal radars, such as Melbourne, FL (KMLB). Only Brownsville, TX (KBRO) and Key West, FL (KBYX) do not have at least one annulus meeting this criterion.

III.6.b. Flagging Algorithm

The guiding principle in the “range continuity” flagging procedure is that true beam blockage in a given azimuthal sector should extend to all ranges beyond the point of blockage $d_b(\theta)$. If blockage is identified in an azimuthal sector, the key is to determine a value of $d_b(\theta)$ where the blockage initiates. In a given azimuthal sector, the grid cells are ranked and arranged in ascending order by range d for all ranges in the radial span $d_0 - d_x$, with d_0 referring to the radar location. For every azimuthal sector, each HRAP grid cell has a an F_1 value, and each grid cell with $F_1 = 1$ is assessed as a potential beam blockage initiation point. In each azimuthal sector, a potential $d_b(\theta)$ is identified an must satisfy the following two “range continuity” criteria based on grid cells from $d_b - d_x$, which are

1. the number of flagged grid cells $n_b(\theta)$ in the radial span $d_b - d_x$ is greater than some pre-specified minimum threshold value (n_b), and
2. the ratio of flagged grid cells ($F_1 = 1$) to total grid cells in the radial span d-d_x is greater than some pre-specified minimum threshold value (R_b).

Based on visual inspection of several possible combinations, $N_b = 12$ and $R_b = 0.6$ were chosen as the ideal threshold values.

If there is more the one value of d that satisfies the "range continuity" criteria within a given azimuthal sector, the statistical significance of the blockage for each these ranges will be computed. For each range d meeting the two criteria, all the F_1 values within the sector azimuthal sector are divided into two groups. The first group (g_0) contains n_0 grid cells from ranges $d_0 - d$, with a computed mean of F_1 values given by $\overline{g_0}$; the second group (g_x) contains n_x grid cells from ranges $d - d_x$ with a computed mean of F_1 values given by $\overline{g_x}$. At each potential blockage initiation range, the value $dg = \overline{g_x} - \overline{g_0}$ is computed, where $\overline{g_x} > \overline{g_0}$ and $\overline{g_x} \geq R_b$ according to the second rule of the "range continuity" criteria. For a given range d, the statistical significance of the blockage is represented by the z-score z_d of the difference dg, given as

$$z_d = \frac{\overline{g_x} - \overline{g_0}}{\sqrt{\left(\frac{\overline{g_0}n_0 + \overline{g_x}n_x}{n_0 + n_x}\right)\left(\frac{1}{n_0} + \frac{1}{n_x}\right)}} . \tag{3.14}$$

The range d with the maximum value of z_d within a given azimuthal sector that also satisfies the two "range continuity" criteria is assigned to $d_b(\theta)$.

III.6.c. Sequence for "Range Continuity" Flagging Procedure

The "range continuity" flagging of azimuthal sectors, which is done for sectors with a radial span containing a sufficient number of HRAP grid cells with $F_1 = 1$, has several steps. The sequence of this two-step procedure is dictated by the geometrical limitations of the irregularly-shaped radar domains.

III.6.c.1) Flagging Close and Intermediate Ranges

Given the limitations to using grid cells in the radial span beyond d_x in a given radar domain, the first step in the "range continuity" flagging procedure is to detect potential beam blockage initiation in azimuthal sectors at close and intermediate ranges. At a given radar domain, each azimuthal sector θ with a radial span meeting the two "range continuity" flagging criteria are assigned a value of $F_2(\theta) = 1$. In addition, the minimum range $d_b(\theta)$ of each sector flagged is stored and beam blockage is assumed to cover the entire azimuthal sector in the range interval $d_b(\theta) \leq d \leq d_{max}(\theta)$.

III.6.c.2) Beam Blockage Flagging at Far Ranges

The next step in the "range continuity" flagging procedure is to determine if beam blockage initiates in ranges beyond d_x. This procedure is more important for radar domains as the value of d_x diminishes for two main reason: 1) there are likely more grid cells beyond d_x when the value is small and 2) the radar beam is more likely to encounter an obstacle at closer ranges given the height of the beam. At this point, we take a step back and reexamine the "low precipitation" flagging for HRAP grid cells beyond d_x. By convention, each annulus examined will have less than 180° coverage, so the "low precipitation" flagging algorithm used for close and intermediate ranges needed tweaking to account for the lack of coverage.

Data from HRAP grid cells in azimuthal sectors with a value $F_2 = 1$ are removed in the computation of the Fourier series at each annulus since these data are deemed contaminated by beam blockage. The removal of the previously flagged data to compute the Fourier series model was done to provide a more accurate fit (Fig. 3.11 for example). Unlike the computation of the Fourier series, the data from the blocked sectors are used in the flagging of individual HRAP grid cells that is done through analysis of the model residuals (Steps 3 – 7 in the "low precipitation" flagging algorithm). Using the data from azimuthal sectors flagged prior to the current flagging procedure ensures that new grid cells receiving a flag value $F_1 = 1$ have similarly "low precipitation" to grid cells already assumed to have beam blockage.

Following the flagging of "low precipitation" grid cells, the "range continuity" flagging procedure looked at azimuthal sectors with $F_2 = 0$ to determine if there was at least one radial span now meeting the two "range continuity" flagging criteria. If these criteria were met in azimuthal sector θ, the sector was given a value of $F_2(\theta) = 1$, and the range $d_b(\theta)$ was determined as the range with the maximum value of z_d, computed using (3.14).

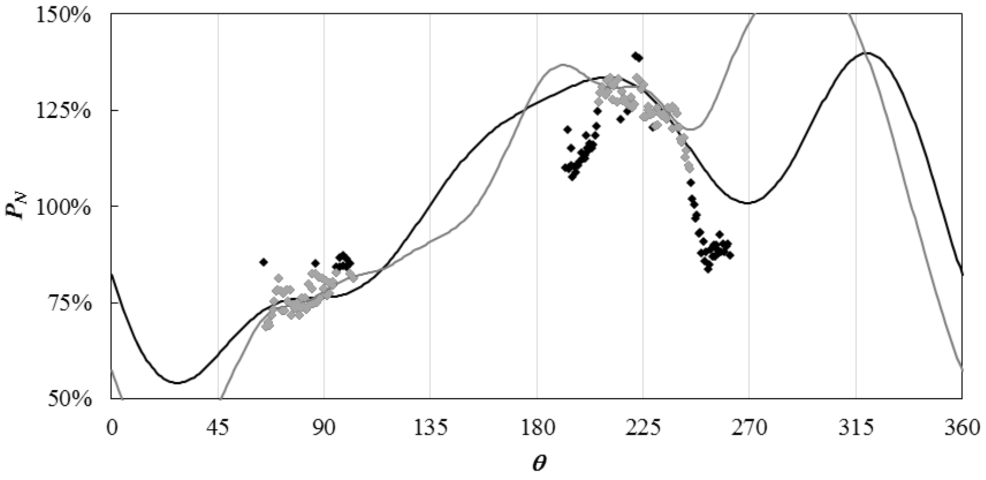


Fig. 3.11. The KABR Stage IV 36-month PoN precipitation data, ending 31 December 2012, as a function of azimuth (all diamonds) in the annulus with the initial Fourier series fits (black line). The grid cells detected in the initial "low precipitation" and "range continuity" flagging algorithms (black diamonds) are removed for the second "low precipitation" flagging algorithm Fourier series fit (gray line).

III.6.d. Results of the "Range Continuity" Flagging Procedure

Fig. 3.12 shows the results of the "range continuity" flagging procedure for azimuthal sectors at the KABR radar domain for Stage IV 36-month PoN precipitation ending 31 December 2012. The one group of sectors around 65° that visually has beam blockage is well detected by the algorithm, but there appears to be spurious detection of beam blockage in other sectors. Fig. 3.13 has the results of the beam blockage detection algorithm for the KFSD domain using Stage IV PoN precipitation data from the same time period. Spurious detection of beam blockage is less of an issue based on the results of the algorithm in the KFSD domain, which identifies the visually apparent beam blockage features.

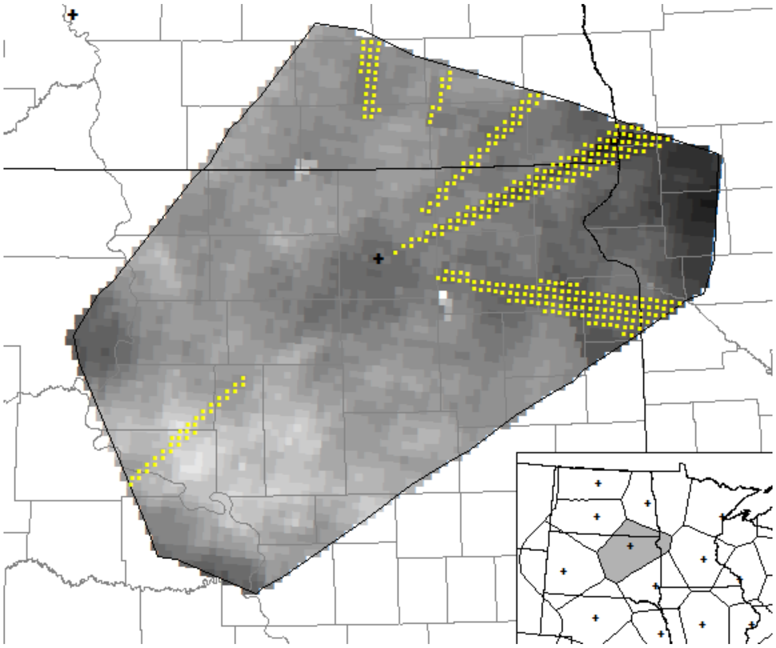


Fig. 3.12. The Stage IV 36-month PoN precipitation, ending 31 December 2012 for the KABR radar domain (contoured field); minimum value (black) is *PoN* = 50% and the maximum value (white) is *PoN* = 150%. The HRAP grid cells in azimuthal sectors meeting the criteria of the beam blockage detection algorithm are included (yellow circles).

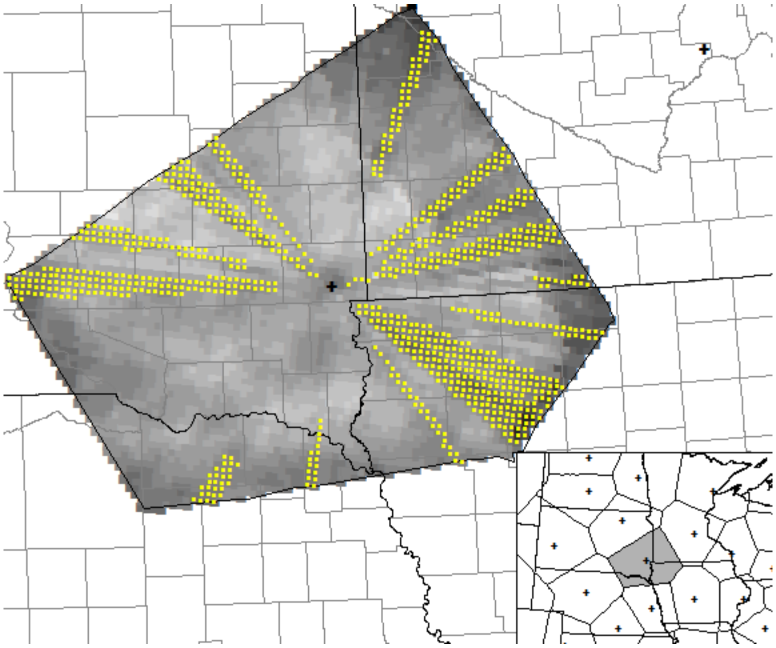


Fig. 3.13. Same as Fig. 3.12, but for the KFSD radar domain.

Fig. 3.14 shows the results of the beam blockage detection algorithm for most of the radar domains in the central and eastern United States, with any boundaries purposefully omitted to make visual inspection of the performance a much easier task. Most of the visually apparent beam blockage features in the Stage IV 36-month PoN precipitation are detected with little in the way of false detection. However, the values of variables such as R_o, N_b, and R_b in the algorithm were chosen so that 1) the ratio of Type I errors to Type II errors is greater than one and 2) both types of errors were minimized. False detection is preferred to non-detection of true blockage in the algorithm as the results of the "range continuity" flagging procedure will go through a quality control (QC) test that will result in some sectors with $F_2 = 1$ being changed to $F_2 = 0$.

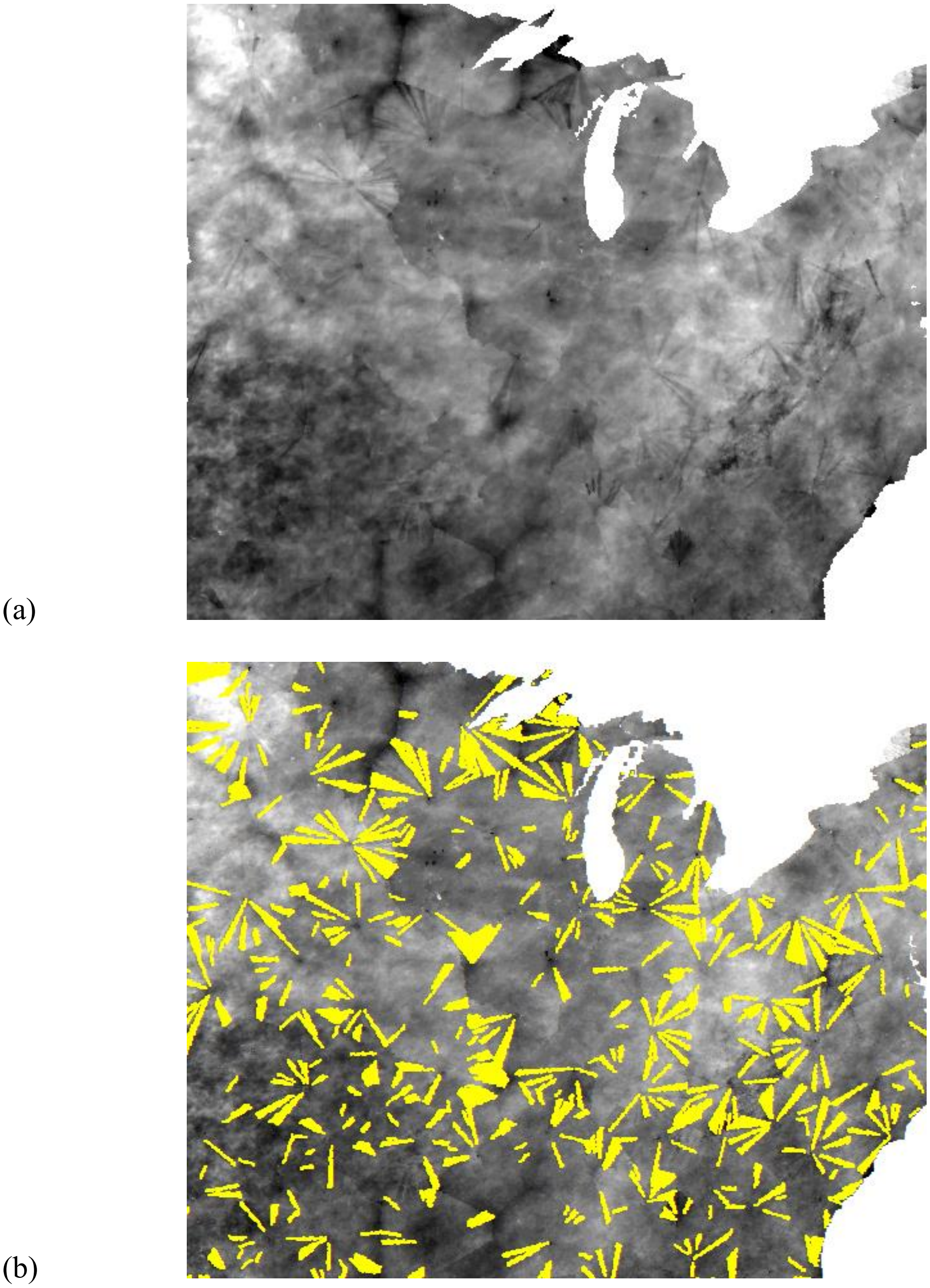


Fig. 3.14. (a) The Stage IV 36-month PoN precipitation, ending 31 December 2012; minimum value (black) is *PoN* = 50% and the maximum value (white) is *PoN* = 150% in both maps. (b) The same as (a) but azimuthal sectors flagged in the "range continuity" flagging procedure included (yellow shading).

III.7. Quality Control Test for Flagged Azimuthal Sectors

A simple QC test will ensure that azimuthal sectors with $F_2 = 1$ have Stage IV 36-month PoN precipitation values that are actually less than neighboring and sectors with $F_2 = 0$. In the context of this QC test, an azimuthal "slice" will be defined as two or more adjoining sectors flagged in the "range continuity" flagging procedure. Each radar domain was broken up into the 10 km annuli for the QC test to minimize range-dependent errors in the comparison of PoN precipitation values from neighboring sectors.

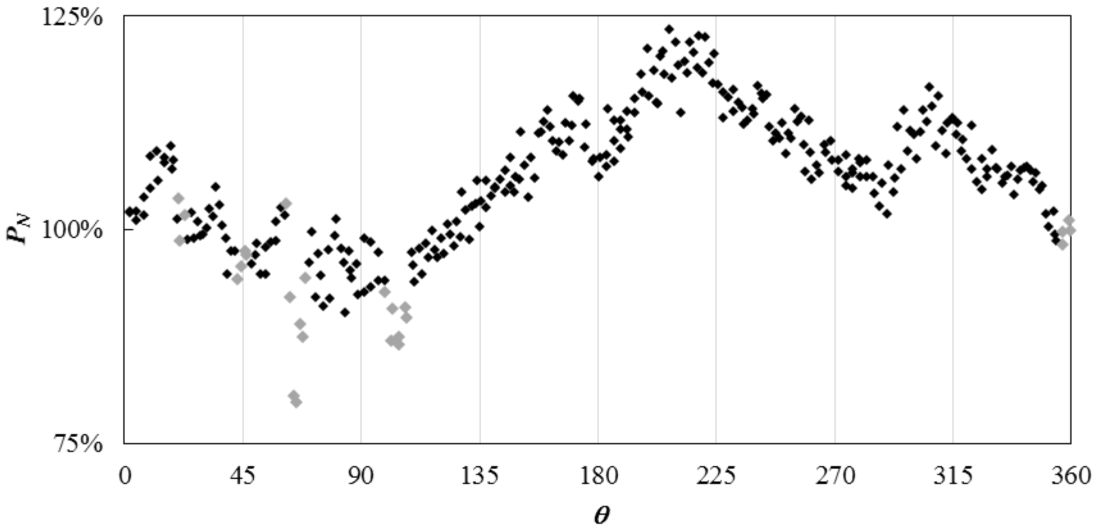


Fig. 3.15. The KABR Stage IV 36-month PoN precipitation data, ending 31 December 2012, as a function of azimuth (all diamonds) in the 90 km – 100 km annulus. The grid cells indicated by the beam blockage detection algorithm are highlighted by gray diamonds.

At an annulus within a given radar domain, the grid cells are ordered by θ_H and the algorithm looks for a sequence of blocked cells. Fig. 3.15 is the Stage IV 36-month PoN precipitation at the KABR radar domain for the annulus with a radial span of 90 km – 100 km (Fig. 3.3a) and has 4 separate slices in which beam blockage has been indicated by the "range continuity" flagging procedure. The two azimuthal slices near 65° and 100° have a "textbook" blockage signature that we would expect to see in the PoN precipitation data within an annulus. However, the other 3 slices, around 20°, 45°, and 355°, don't appear to have any distinction from the neighboring data either visually (Fig. 3.12) or statistically (Fig. 3.15).

The QC test metric for each azimuthal slice of grid points considered blocked is simply the ratio of the mean PoN precipitation of the blocked cells ($F_2 = 1$) to the mean PoN precipitation of the two bounding non-blocked cells with ($F_2 = 0$). The algorithm scans for a couplet where a grid cell $i = 1$ has a value of $F_2 = 0$ and grid cell $i = 2$ has a value of $F_2 = 1$. Each ordered grid cell with $F_2 = 1$ is passed over until another grid cell $i = n$ with with $F_2 = 0$ is found. Then, the QC test ratio for this sequence (R_S) is computed using

$$R_S = \frac{2\sum_{i=2}^{n-1} P_N^{\,i}}{(n-2)\left(P_N^{\,i} - P_n^{\,i}\right)}, \tag{3.15}$$

where P is PoN precipitation.

For a given sector, computation of R_S for each grid cell deemed blocked by the "range continuity" flagging algorithm is followed by looking at the R_S values along the radial span $d_b(\theta) - d_{max}(\theta)$. A key decision for analysis of the R_S values is determining a proper threshold R_Q such that grid cells in a given sequence with $R_S \leq R_Q$ are considered truly contaminated by beam blockage and thus pass the QC test. In Fig. 3.15, $R_S = 0.87$ for the sequence near 65°, whereas the other 3 seqences of grid cells have R_S values of 1.01, 0.99, and 0.99. Three different possibilities for the threshold value of R_S were explored: 0.95 (Fig. 3.16), 0.98 (Fig. 3.17), and 1.00 (Fig. 3.18). For a given sector, we will determine the number of grid cells $n_{QP}(\theta)$ where $R_S \leq R_Q$ and the number of grid cells $n_{QF}(\theta)$ where $R_S \leq R_Q$. An azimuthal sector θ will pass the QC check if $n_{QP}(\theta) \geq n_{QF}(\theta)$ and $n_{QP}(\theta) > N_Q$. After careful consideration, $N_Q = 10$ was chosen as an appropriate threshold.

Figs. 3.16 – 3.18 show examples of azimuthal sectors passing the QC test in the KABR and KFSD radar domains. Based on these two radars and a thorough visual inspection of other key radars with PoN significantly visibly affected by beam blockage (Fig. 3.14a), a threshold value of $R_Q = 0.98$ appears visually as properly verifying and rejecting sectors deemed blocked in the "range continuity" flagging procedure. Azimuthal sectors with an initial value of $F_2 = 1$ will continue to be flagged if passing the QC test; otherwise a value of $F_2 = 0$ will be assigned.

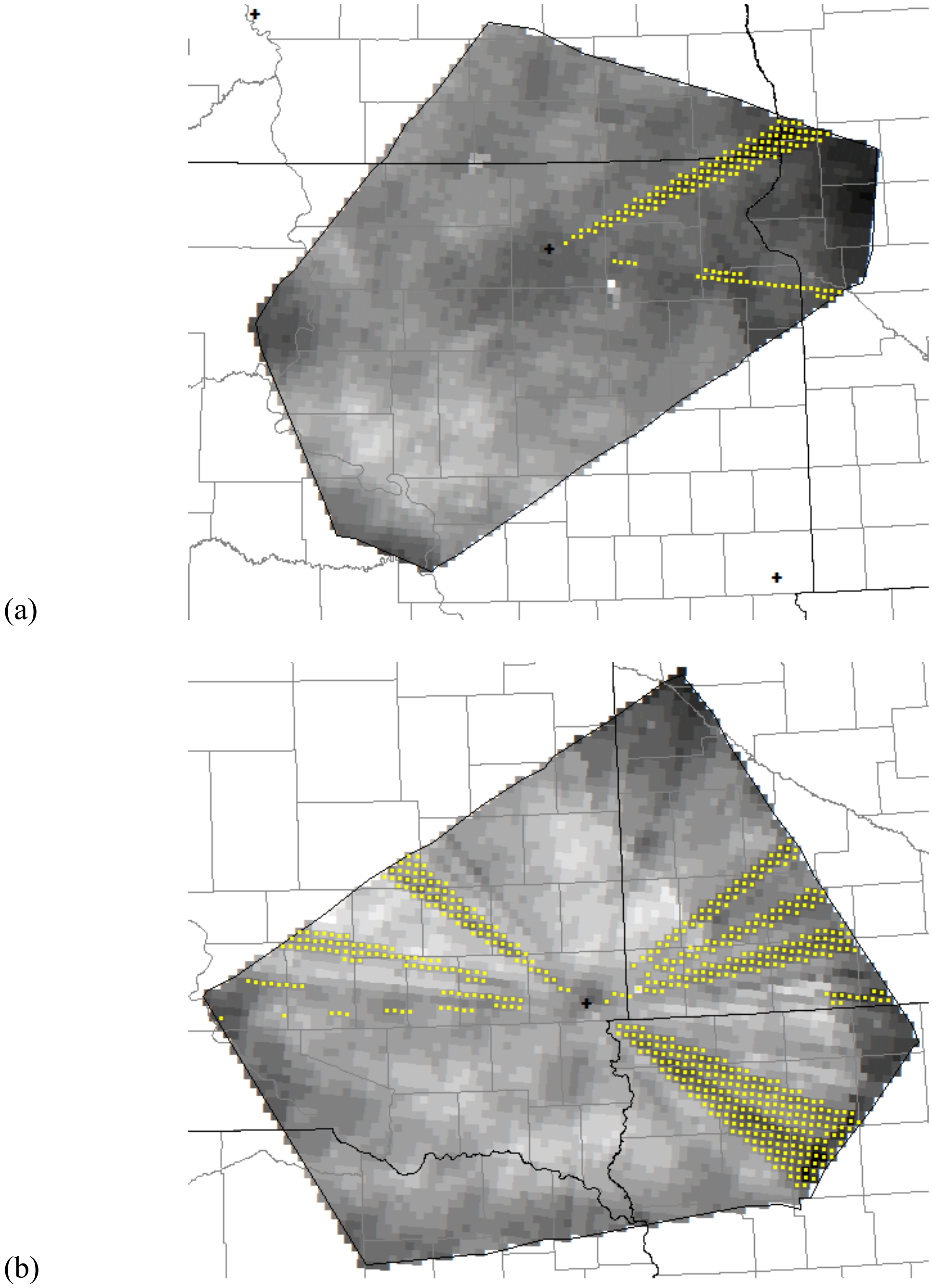


Fig. 3.16. The Stage IV 36-month PoN precipitation, ending 31 December 2012 for the (a) KABR and (b) KFSD radar domains; minimum value (black) is *PoN* = 50% and the maximum value (white) is *PoN* = 150%. The HRAP grid cells (yellow dots) in azimuthal sectors meeting the criteria of the QC test algorithm are included, using $R_Q = 0.95$.

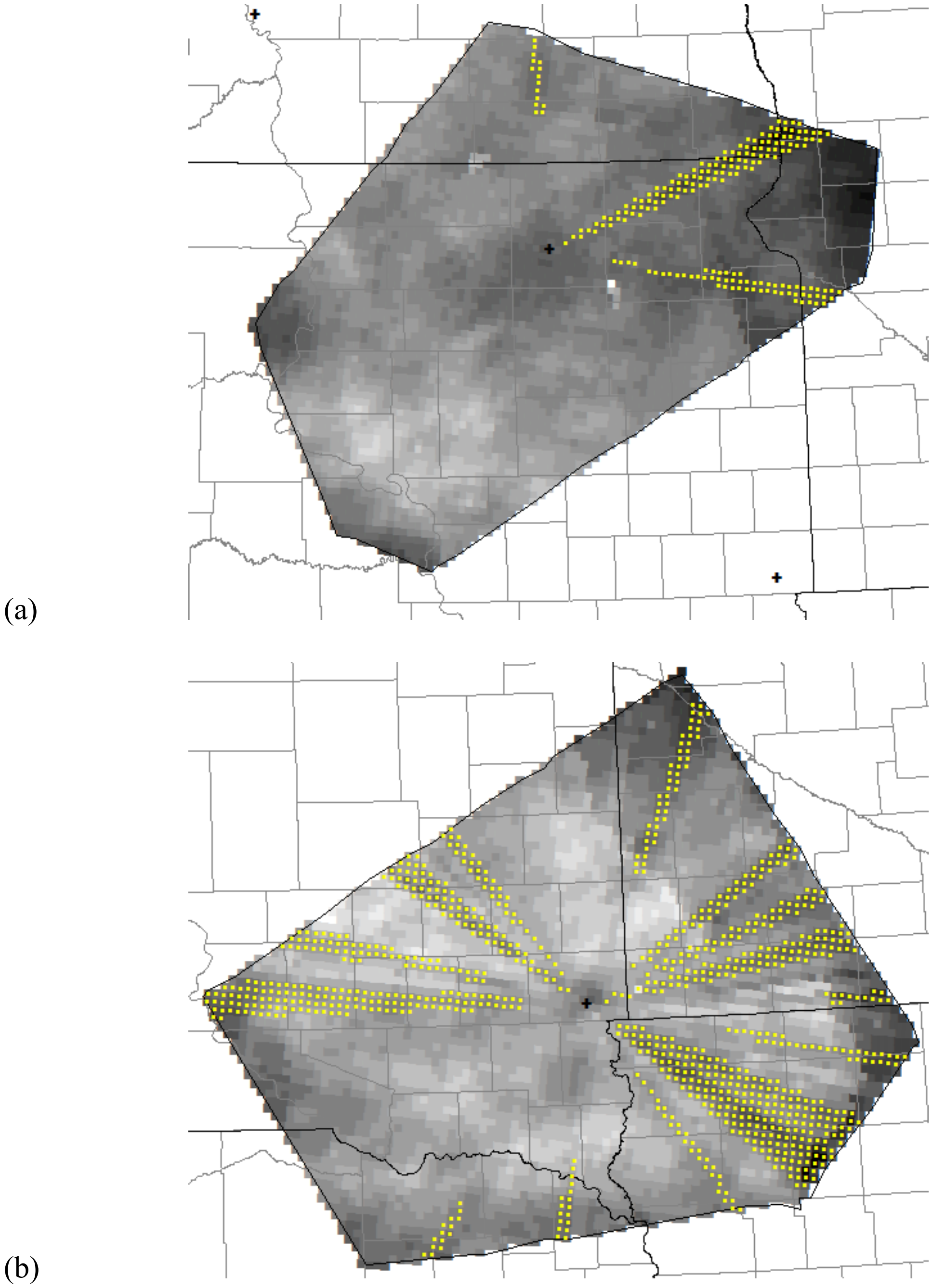


Fig. 3.17. Same as Fig. 3.16, using a threshold value of $R_Q = 0.98$.

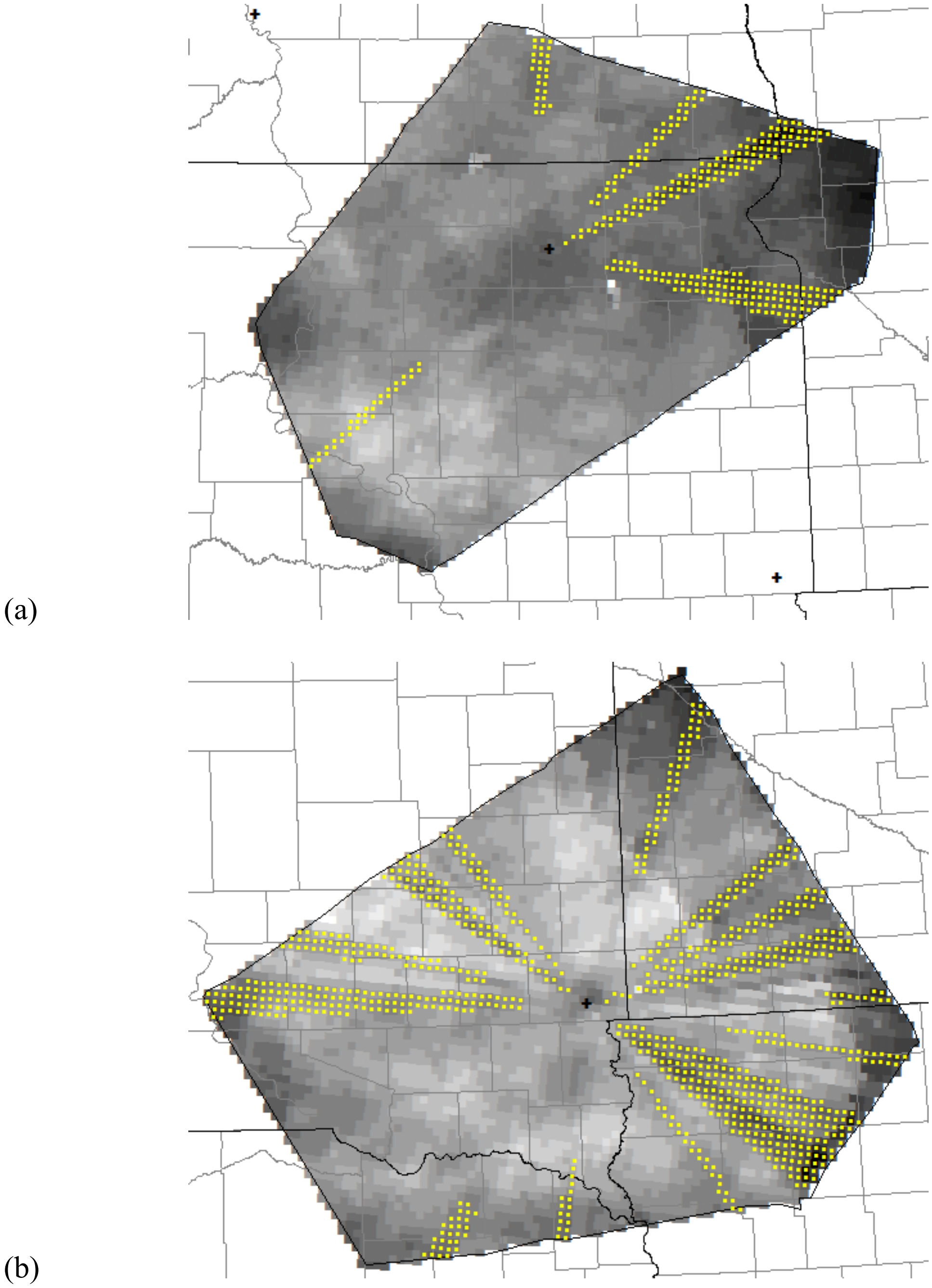


Fig. 3.18. Same as Figs. 3.16 and 3.17, using a threshold value of $R_Q = 1.0$.

III.8. Temporal Consistency of Beam Blockage

III.8.a. Characteristics of Monthly Time Series of 36-Month Beam Blockage Detection Data

Up to this point, detection of beam blockage features in each radar domain has been described for a single 36-month period. This section examines the temporal consistency of beam blockage for azimuthal sectors detected as blocked ($F_2 = 1$) during the timeframe of this study (2005-2012). The goal of this section is to remove any inconsistencies that may be present in a time series of F_2 data points and to identify any significant changepoints in detected blockage (Fig. 3.19).

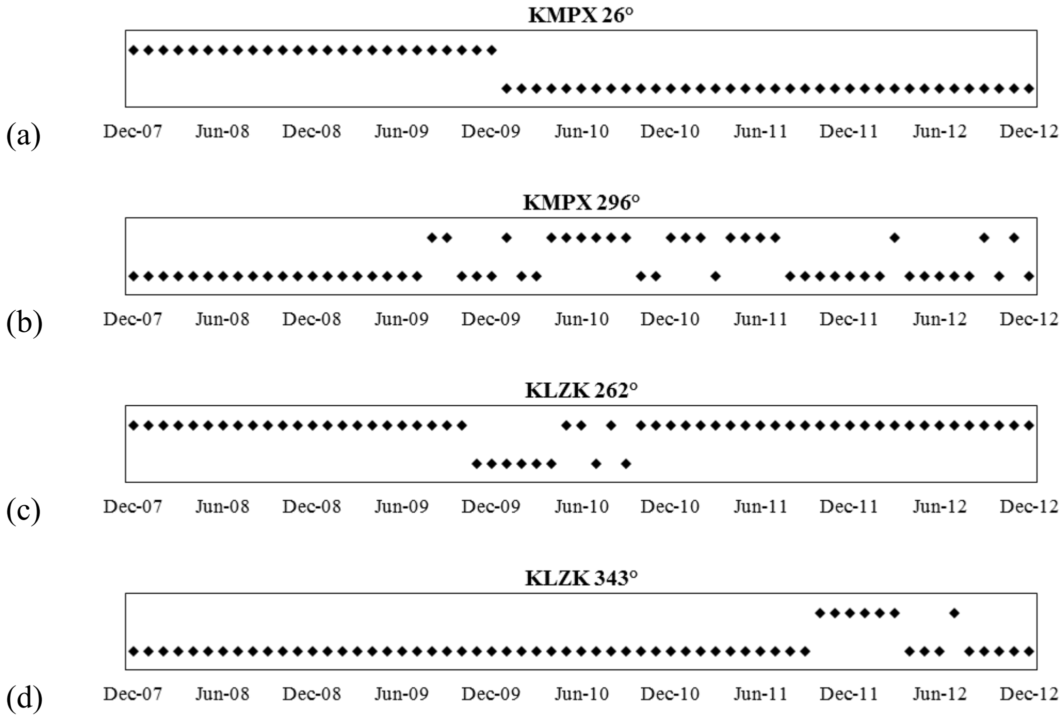


Fig. 3.19. Time series of F_2 values for the period December 2007 – December 2012. The results for the (a) KMPX 26°, (b) KMPX 296°, (c) KLZK 262°, and (d) KLZK 343° azimuthal sectors are shown.

III.8.b. Detecting Changepoints in Time Series of the "Range Continuity" Flags

Each azimuthal sector at each radar is being evaluated independently for the presence of a changepoint in a time series of F_2 data (Fig. 3.19). The time series is examined sequentially from beginning to end for a changepoint, examining each possible partitioning of the F_2 data points into two groups, given that each group has at least two data points. The metric used to determine the suitability of a partition t_0 as a changepoint is the statistical significance of the difference in the means between the two groups, with group 1 (g_1) to the left of the partition and group 2 (g_2) to the right of the partition t_0. Similar to z_d given by (3.14), the statistical difference (z_{abs}) between the mean of group 1 ($\overline{g}_1$) and the mean of group 2 ($\overline{g}_2$) is

$$z_{abs} = \left| \frac{\overline{g}_2 - \overline{g}_1}{\sqrt{\left(\frac{\overline{g}_1 n_1 + \overline{g}_2 n_2}{n_1 + n_2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \right| . \tag{3.16}$$

A changepoint will partition the data into one group with mostly or all F_2 values of zero and the other with mostly or all F_2 values of one, with $\overline{g}_2 - \overline{g}_1$ having a possible range of [-1, 1]. The first criterion for a partition of F_2 values in a time series to be considered a potential changepoint is that z_{abs} exceeds a minimum threshold value z_c determined to be $z_c = 4.0$ (e.g., Fig. 3.20a). The second criterion is conditional depending on the means of the two groups, and is

a. if $\overline{g}_2 - \overline{g}_1 > 0$, then both $\overline{g}_1 < 0.5$ and $\overline{g}_2 > 0.5$ or

b. if $\overline{g_2} - \overline{g_1} < 0$, then both $\overline{g_1}$ is > 0.5 and $\overline{g_2} < 0.5$.

If more than one z_{abs} value exceeds z_c and meets the second, conditional criterion the partition t_0 with the maximum value of z_{abs} will be chosen as the changepoint of the time series.

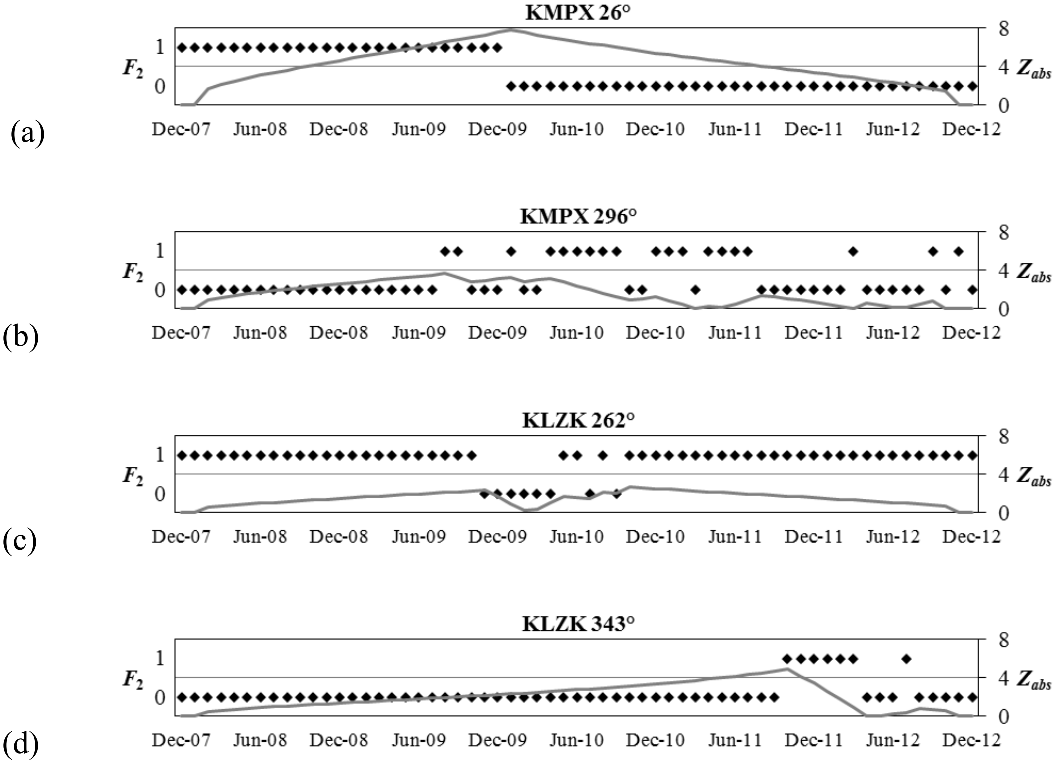


Fig. 3.20. Same as Fig. 3.19, with a time series of Z_{abs} values included (gray lines).

III.8.c. Assigning "Temporal Consistency" Flags

At each radar domain, the time series of F_2 data points is analyzed for changepoints in all azimuthal sectors. Given the reality that values of both $F_2 = 0$ and $F_2 = 1$ co-exist in the time series of some azimuthal sectors where changepoints were not detected, there needs to be an objective determination of a single, constant beam blockage status. If the sum of all the F_2 values exceeds ½ of the number of F_2 data points ($n = 61$, so $\frac{1}{2}n = 30.5$), the azimuthal sector θ is considered blocked throughout the entire period; otherwise $d\theta$ is considered free of beam blockage. Therefore, any given sector $d\theta$ with no changepoint is assigned a single "temporal consistency" flag $F_3(\theta)$ value ($F_3 = 0$ if not blocked; $F_3 = 1$ if blocked; see Fig 3.21c) for all times. For sectors with a changepoint identified in the "temporal consistency" procedure, the value of F_3 is a function of time, such that $F_3 = F_3(\theta, t)$.

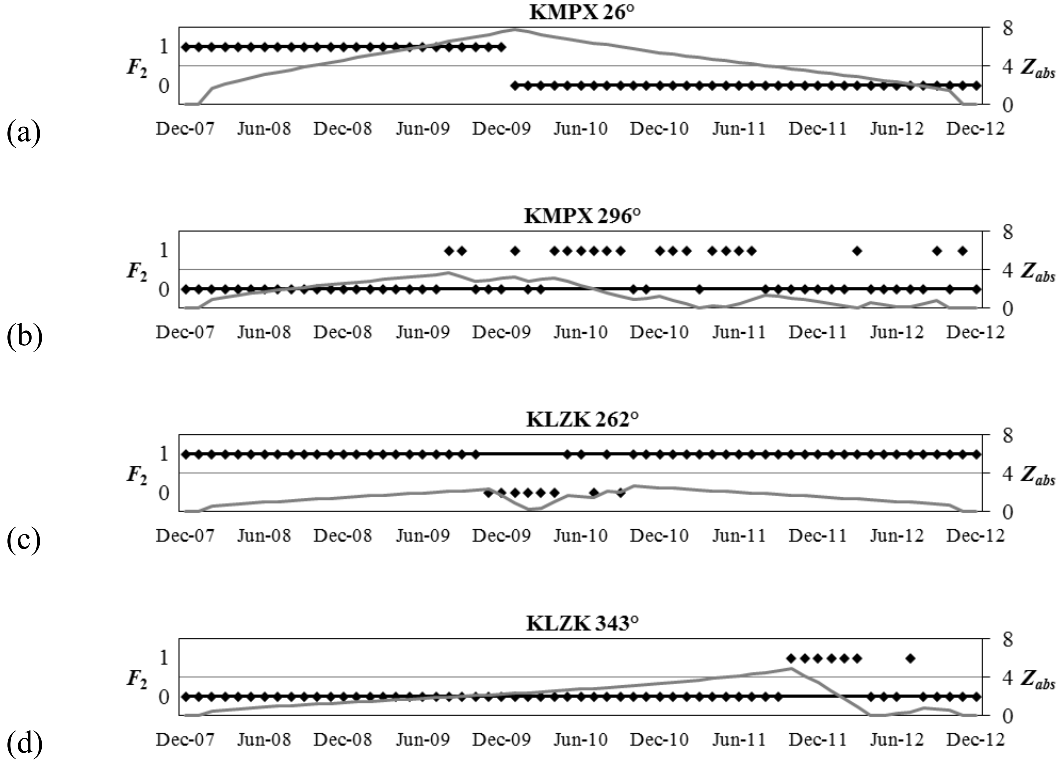
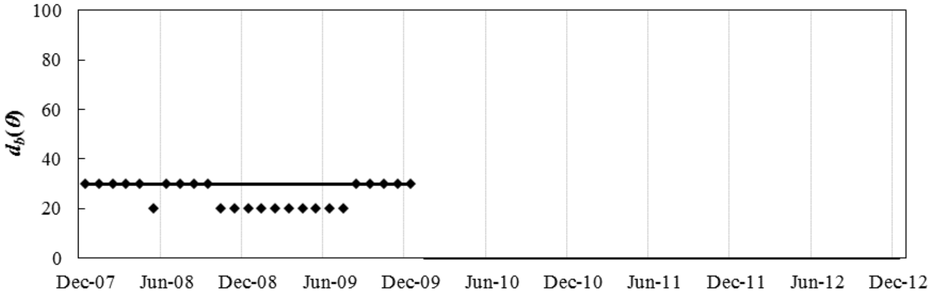
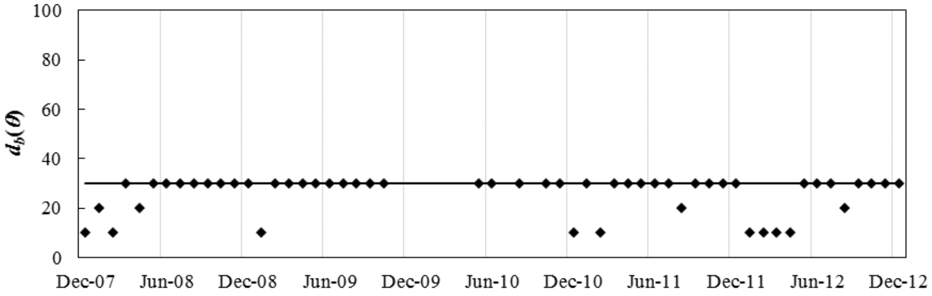


Fig. 3.21. Same as Fig. 3.20, with values of F_3 included (straight black lines).

Assuming that the "temporal consistency" flag for azimuthal sector θ is $F_3 = 1$, there may be inconsistency in the values of d_b, the range at which beam blockage is assume to initiate (in a spatial sense). Given a span of times from t_0 to $t_0 + \Delta t$, the range $d_b(\theta, t)$ with the highest frequency of occurrence is designated as the single range d_b at which beam blockage is assumed to initiate, with the reality that the obstacle is fixed in nature. Assignment of the "temporal consistency" flag F_3 is the final say for the beam blockage status of a particular azimuthal sector given a time series of Stage IV 36-month PoN precipitation data (Fig. 3.22). However, the application of these results is not limited to the 36-month time scale and the applicability to shorter time scales needs to be defined, which is done in the following section.



(a)



(b)

Fig. 3.22. The time series of independently-found $d_b(d\theta)$ values from beam blockage detection algorithm (black diamonds) and the final value of $d_b(d\theta)$ (black line). The results for the (a) KMPX 26° and (b) KLZK 262° azimuthal sectors are shown.

III.9. Beam Blockage Correction

III.9.a. "Overall" Beam Blockage Designation

The correction for beam blockage in the Stage IV data will be done using the azimuthal sectors identified with the 36-month PoN precipitation beam blockage flagging procedure with values of $F_3 = 1$. However, given that the time stamp of an F_3 data point is at the end of a 36-month period, the estimated changepoint for real beam blockage in this study is t_0 – 18 months, where t_0 is the changepoint specified by the three-step beam blockage flagging procedure. Each calendar month in the 18 months prior to changepoint t_0 will be assigned an "overall" beam blockage flag F that is the same value as the F_2 value for all times $t \geq t_0$. In essence, the changepoint in the F time series is the F_3 changepoint t_0 shifted to the left by 18 months. For azimuthal sectors with no changepoints, the single value of F_3 is assigned as a universal "overall" flag F.

For a given azimuthal sector, if beam blockage exists in a radial span $d_b(\theta)$-$d_{max}(\theta)$ for a given time period, the data will be flagged as missing and will be replaced by interpolated data that uses neighboring, unblocked HRAP grid cells to fill in these gaps. It is important to note the interpolation will be done using PoN precipitation as opposed to accumulated precipitation values to account for any spatial variations of the climatological precipitation that may exist.

The identification of beam blockage was done using only 36-month PoN precipitation, but the results of the detection procedure must be applied to all possible accumulation periods, e.g., 1-month, 6-month, 12-month, 24-month, etc. The application of the beam blockage results will be done at the shortest possible time period in this study, which is the 1-month accumulation period. For HRAP grid cells in an azimuthal sector considered “blocked” (*F* = 1), the original Stage IV 1-month PoN precipitation value is considered missing and a new estimated value is found through interpolation of values from neighboring “non-blocked” grid cells. The interpolated value is then transformed to a 1-month precipitation total using an inverse form of (3.3). For a given accumulation period, i.e., 12-months, the interpolation procedure is used for to estimate 1-month PoN and total precipitation values for each calendar month with a flag $F = 1$. Continuing with the 12-month example, all twelve 1-month totals will be added together to get a 12-month total, which can be used to compute other precipitation metrics (*PoN*, departure from normal, etc.).

III.9.b. Overview of Ordinary Kriging

Rather than adjusting the PoN precipitation values of grid cells contaminated by beam blockage using some metric measuring the magnitude of blockage, these grid cells will be considered missing and neighboring grid cells will be used to produce an estimate of precipitation using an established interpolation method. Geostatistical interpolation techniques estimate a missing value $\hat{Z}$ by assigning weights w_i to n neighboring grid cells with sampled Z_i values such that the estimate $\hat{Z}$ is

$$\hat{Z} = \frac{\sum_{i=1}^{n} w_i Z_i}{\sum_{i=1}^{n} w_i}. \tag{3.17}$$

Inverse distance weighting (IDW) is a commonly used interpolation technique, but the weights are found with a function based completely on one-dimensional distance, so estimates may be skewed if the sampled points are spatially heterogeneous. Kriging is a geostatistical linear interpolation method that is an alternative to IDW that accounts for the spatial heterogeneity of samples used to compute the estimates. Kriging interpolation weights the neighboring grid cells according to spatial covariance, a measure of the similarity between two values of variable as a function of geographical distance, or lag (usually dentoed as *h*), between points (Bohling 2005a). The spatial covariance *C*(*h*) for a set of *n* unique pairs of grid points separated by lag *h* is

$$C(h)=\frac{1}{n}\sum_{i=1}^{n}\left(Z(x_i)-\overline{Z_i}\right)\left(Z(x_i+h)-\overline{Z_{i+h}}\right), \tag{3.18}$$

where $\overline{Z_i}$ is the mean for the "tail" pair member and $\overline{Z_{i+h}}$ is th mean for the "head" pair member. An alternative measure to spatial covariance as a measure of similarity as a function of lag is the semivariance $\gamma(h)$ which for a given lag *h* is

$$\gamma(h)=\frac{1}{2}\left[\frac{1}{n}\sum_{i=1}^{n}\left(Z(x_i)-Z(x_i+h)\right)^2\right], \tag{3.19}$$

which is typically done for a nominal lag interval. Obviously, the semivariance is zero when the lag *h* is zero, but is assumed a "nugget" value if $\gamma(h) >> 0$ for values approaching zero (Bohling 2005a). The values of $\gamma(h)$ are plotted against *h* on a semivariogram, which is shown comceptually in Fig. 3.23 (Karl and Maurer 2010). The semivariogram data is fit to an existing semivariogram model $g(h)$ using the value c of the "sill" and the lag a at which the sill is located (Bohling 2005a). The corresponding model covariance $C(h)$ can be computed using $g(h)$ as

$$C(h) = c - g(h). \tag{3.20}$$

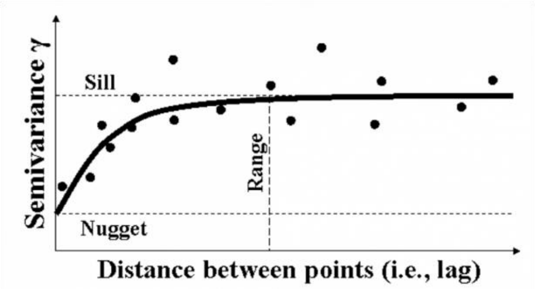


Fig. 3.23. Conceptual semivariogram data and model, taken from Karl and Maurer (2010). Reprinted from *Ecolo. Inform.*, **5**, Karl, J. W., and B. A. Maurer, Spatial dependence of predictions from image segmentation: A variogram-based method to determine appropriate scales for producing land-management information, 194-202, 2010, with permission from Elsevier.

Given a set of n neighboring grid cells the $n \times n$ model covariance matrix $\mathbf{K}$ is computed for all pairs of neighbors and the n length vector $\mathbf{k}$ of model covariances between each neighbor and the unsampled target grid cell (Bohling 2005b), with each covariance computed using the pair's lag (3.18). In simple Kriging, the the mean μ is assumed known and each weight i assigned to a neighboring grid cell is a vector $\mathbf{w}$ with a relationship to $\mathbf{K}$ and $\mathbf{k}$ as

$$\mathbf{K}\mathbf{w} = \mathbf{k}. \tag{3.21}$$

If the mean μ is assumed not to be known as is the case when interpolating using ordinary Kriging, each matrix is augmented (Bohling 2005b) so that

$$\begin{bmatrix} C_{1,1} & C_{1,2} & \cdots & C_{1,n} & 1 \\ C_{2,1} & C_{2,2} & \cdots & C_{2,n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{n,1} & C_{n,2} & \cdots & C_{n,n} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} C_{1,0} \\ C_{2,0} \\ \vdots \\ C_{n,0} \\ 1 \end{bmatrix}. \tag{3.22}$$

III.9.c. Estimating Data at HRAP Grid Cells Contaminated by Beam Blockage

The Stage IV 1-month accumulated precipitation and PoN precipitation values at grid cells deemed blocked for a given radar-month are interpolated using an ordinary Kriging procedure that use n = 12 neighboring points. To reduce the effects of range-dependent biases, only grid cells from the same radar annulus are used as neighboring points, with an equal number of “unblocked” neighbors ($n_{1/2}$ = 6) on either side of a “blocked” point used in the Kriging procedure. The empirical semivariance data in the estimation procedure uses a nominal lag interval of 1 km. and the model semivariograms are constrcted separately for each radar annulus.

The ordinary Kriging procedure for correcting the 1-month Stage IV PoN data at a grid cell contaminated by blockage is best explained using a straightforward example. Fig. 3.24 is the Stage IV PoN data for January 2012 for the 100 km – 110 km annulus at the KABR radar domain.

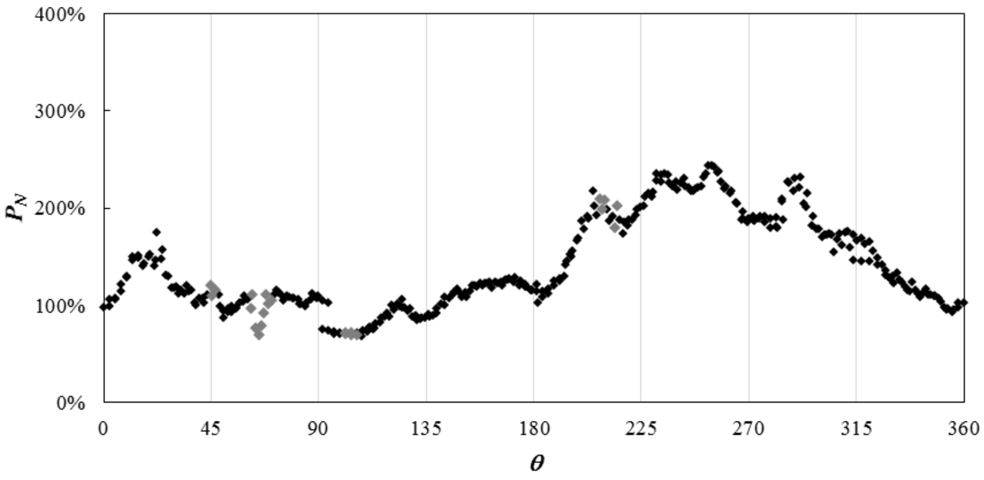


Fig. 3.24. The Stage IV 1-month PoN precipitation data for January 2012 at the KABR radar domain in the 100 km – 110 km annulus as a function of azimuth. “Unblocked” (black diamonds) and “blocked” (gray diamonds) data points are differentiated.

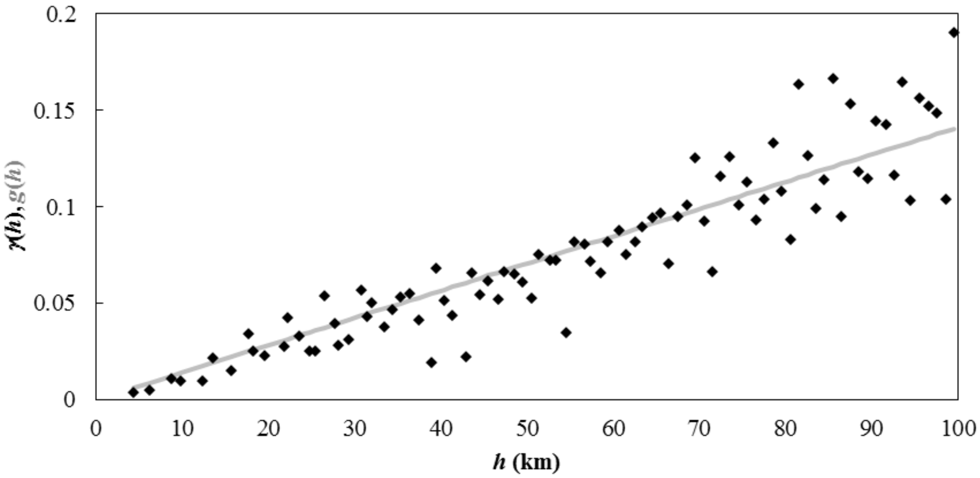


Fig. 3.25. Empirical and model semivariogram constructed using pairs of “unblocked” Stage IV PoN precipitation data points from Fig. 3.24.

The nominal semivariance $\gamma(h)$ is computed for each nominal lag less than 100 km with available data using only pairs for which both data points are considered “unblocked” (Fig. 3.25). Using the available $\gamma(h)$ data, a continuous model semivariogram function $g(h)$ is fit to the data points (also included on Fig. 3.25). Given inspection of several plots of $\gamma(h)$ data for several different radar-months, it was determined that a straight line model fit was most appropriate as a universal $g(h)$ model form. The slope and intercept parameters of the straight line model are found using simple linear regression (SLR), which minimizes the residual sum-of-squares (RSS).

$$\begin{bmatrix}
0.2218 & 0.2156 & 0.2156 & 0.2131 & 0.2080 & 0.2080 & 0.1941 & 0.1908 & 0.1885 & 0.1857 & 0.1822 & 0.1803 & 1.0000 \\
0.2156 & 0.2218 & 0.2131 & 0.2156 & 0.2131 & 0.2094 & 0.1908 & 0.1868 & 0.1857 & 0.1822 & 0.1780 & 0.1772 & 1.0000 \\
0.2156 & 0.2131 & 0.2218 & 0.2156 & 0.2094 & 0.2131 & 0.1885 & 0.1857 & 0.1826 & 0.1803 & 0.1771 & 0.1746 & 1.0000 \\
0.2131 & 0.2156 & 0.2156 & 0.2218 & 0.2156 & 0.2156 & 0.1857 & 0.1821 & 0.1803 & 0.1771 & 0.1734 & 0.1719 & 1.0000 \\
0.2080 & 0.2131 & 0.2094 & 0.2156 & 0.2218 & 0.2131 & 0.1821 & 0.1780 & 0.1771 & 0.1734 & 0.1692 & 0.1685 & 1.0000 \\
0.2080 & 0.2094 & 0.2131 & 0.2156 & 0.2131 & 0.2218 & 0.1802 & 0.1771 & 0.1746 & 0.1719 & 0.1685 & 0.1664 & 1.0000 \\
0.1941 & 0.1908 & 0.1885 & 0.1857 & 0.1821 & 0.1802 & 0.2218 & 0.2156 & 0.2156 & 0.2131 & 0.2080 & 0.2080 & 1.0000 \\
0.1908 & 0.1868 & 0.1857 & 0.1821 & 0.1780 & 0.1771 & 0.2156 & 0.2218 & 0.2131 & 0.2156 & 0.2131 & 0.2095 & 1.0000 \\
0.1885 & 0.1857 & 0.1826 & 0.1803 & 0.1771 & 0.1746 & 0.2156 & 0.2131 & 0.2218 & 0.2156 & 0.2094 & 0.2131 & 1.0000 \\
0.1857 & 0.1822 & 0.1803 & 0.1771 & 0.1734 & 0.1719 & 0.2131 & 0.2156 & 0.2156 & 0.2218 & 0.2156 & 0.2156 & 1.0000 \\
0.1822 & 0.1780 & 0.1771 & 0.1734 & 0.1692 & 0.1685 & 0.2080 & 0.2131 & 0.2094 & 0.2156 & 0.2218 & 0.2131 & 1.0000 \\
0.1803 & 0.1772 & 0.1746 & 0.1719 & 0.1685 & 0.1664 & 0.2080 & 0.2095 & 0.2131 & 0.2156 & 0.2131 & 0.2218 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 0.0000
\end{bmatrix}
\times
\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ w_6 \\ w_7 \\ w_8 \\ w_9 \\ w_{10} \\ w_{11} \\ w_{12} \\ \mu \end{bmatrix}
=
\begin{bmatrix} 0.2080 \\ 0.2022 \\ 0.2043 \\ 0.1995 \\ 0.1941 \\ 0.1955 \\ 0.2032 \\ 0.2022 \\ 0.1971 \\ 0.1963 \\ 0.1941 \\ 0.1902 \\ 1.0000 \end{bmatrix}$$

Fig. 3.26. Matrix equation used to solve for the ordinary Kriging weights for each of the 12 neighboring grid cells used to determine the corrected Stage IV 1-month PoN precipitation for January 2012 at HRAP grid cell 618583.

Using (3.20), all the covariance terms on both the right-hand and left-hand sides of (3.22) are computed from c and the semivariogram function $g(h)$ shown in Fig. 3.25. The value of the sill c is assumed to be the sample variance s^2 for all the data “unblocked” data points in the 100 km – 110 km annulus. The matrix equation used to solve for the weights (each neighbor assigned a number between 1 and 12) and determine the corrected January 2012 Stage IV PoN precipitation value for HRAP grid cell 618583 is shown in Fig. 3.26. Grid cell 618583 has $\theta = 65°$ and is a local minimum in a series of “blocked” grid cells within the 100 km – 110 km annulus. Fig. 3.27 shows the neighboring grid cell weights and PoN precipitation values used in the ordinary Kriging estimation at grid cell 618583.

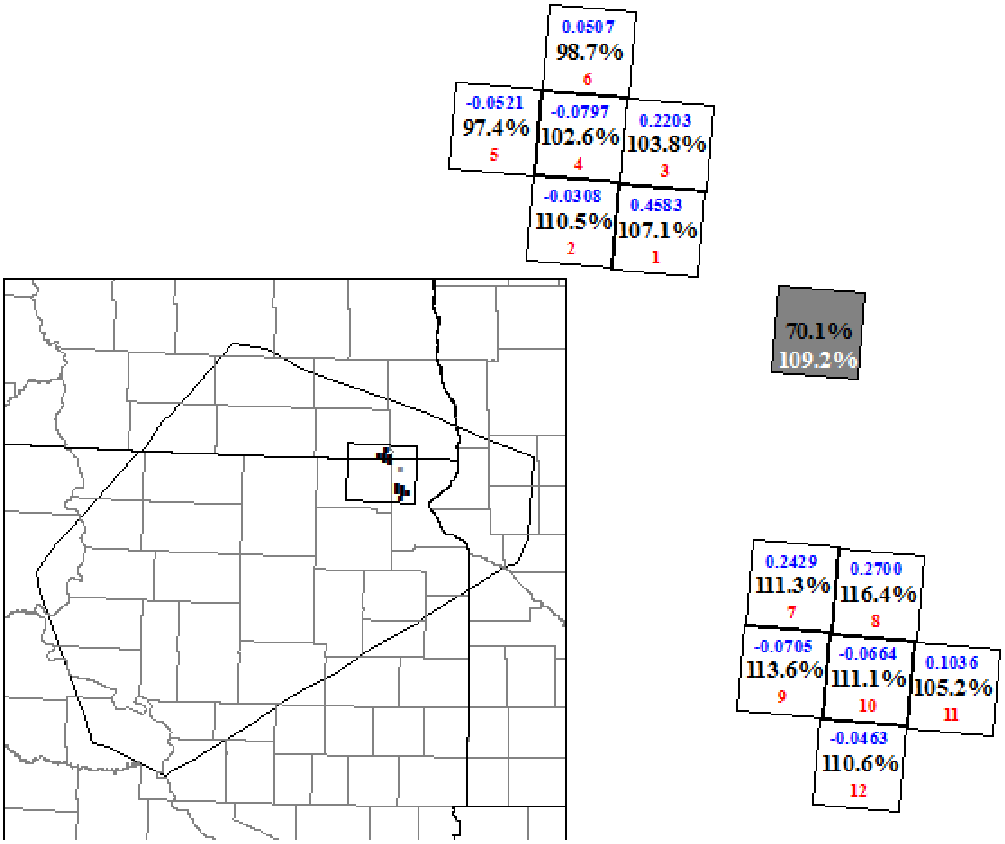


Fig. 3.27. Map of the neighboring gird cells (white boxes) used in the ordinary Kriging estimation of the corrected Stage IV 1-month PoN precipitation for January 2012 at HRAP grid cell 618583 (gray box). The neighbor ID (red number), neighbor interpolation weight (blue number), and Stage IV PoN precipitation values (black numbers) are shown.

Given the spatial configuration of the neighboring grid cells in Fig. 3.27, it is apparent that the ordinary Kriging weights give preferential treatment to the first grid cell encountered in a particular direction, i.e., neighboring grid cells 1,3, 7, and 8. Neighbors with other grid cells lying between their location and the target grid cell (in this case 618583) have negative weights, i.e., 4, 5, 9, and 10. Fig. 3.27 is an example of how Kriging reduces the effect of clustering by giving members of a cluster in closer proximity to the interpolation point a much higher weight than members of the cluster farther away.

The ordinary Kriging procedure was used to estimate Stage IV 1-month precipitation for all “blocked” grid cells during the period January 2005 – December 2012. The resulting precipitation dataset, corrected for beam blockage using the methodology described in this chapter (example in Fig. 3.28), will go through a procedure that corrects for both mean-field and range-dependent biases that will be detailed in the next chapter.

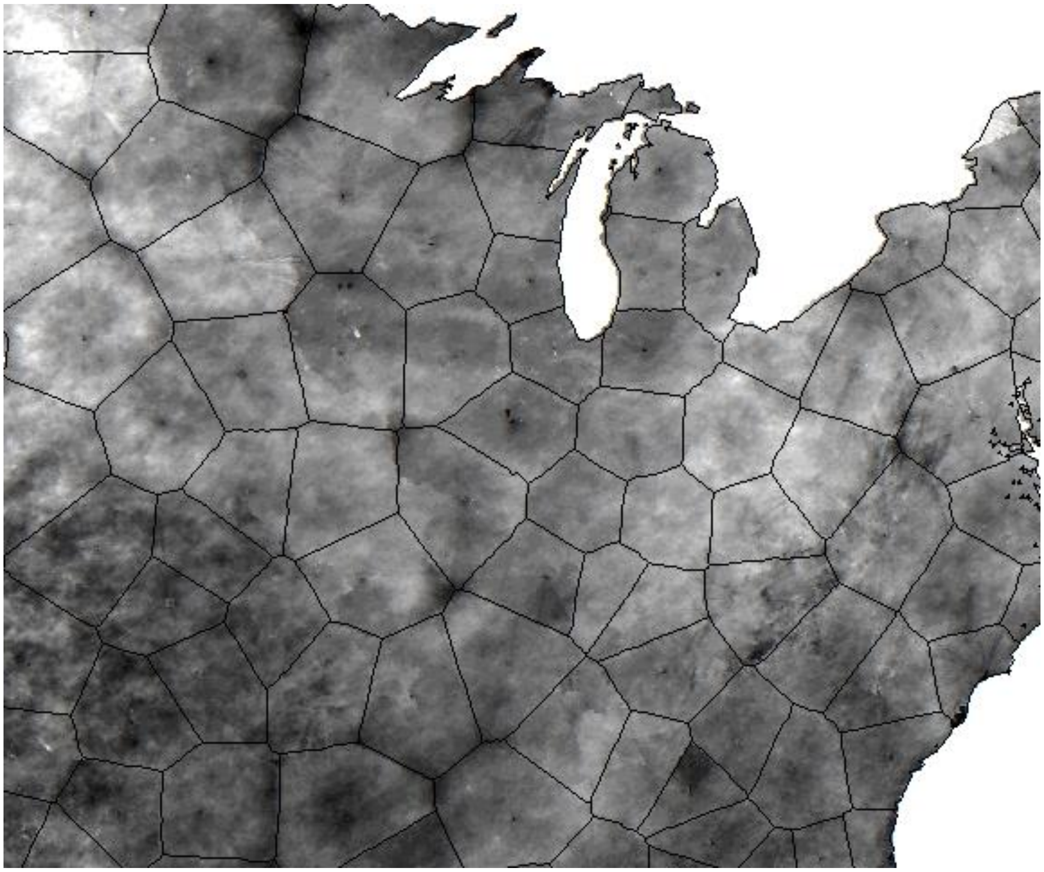


Fig. 3.28. Same as Fig. 3.14a, but with Stage IV 36-month PoN precipitation corrected for beam blockage and the boundaries of radar domains included. The minimum value (black) is $PoN = 50\%$ and the maximum value (white) is $PoN = 150\%$.

III.9.d. Error Reduction in Data Corrected for Beam Blockage

At each grid cell and time t at which blockage was detected, a raw 1-month Stage IV PoN precipitation value (P_{N0}) and 1-month beam-blockage adjusted PoN precipitation value (P_{N1}) exists. Using the 1981-2010 PRISM precipitation normals (P_n), the precipitation total P can be computed for each of the two P_N values using

$$P = P_N \times P_n . \tag{3.23}$$

Because the beam blockage detection and correction methodology presented is done using only Stage IV PoN precipitation data, the performance of the methodology can be evaluated using gauge data (G) as a proxy for the true surface precipitation P_T. At each time t and each gauge i located within a grid cell for which the 1-month Stage IV PoN precipitation was adjusted for beam blockage, the error ε_i was computed as

$$\varepsilon_i = P - G \tag{3.24}$$

for both the raw (P_0) and beam blockage-adjusted (P_1) precipitation values. The error checking is done to ensure the beam blockage corrections led to systematically lower ε_i values in the P_1 data relative to the original Stage IV precipitation dataset.

The two evaluation metrics, the overall root mean square error ($RMSE$) and overall mean bias (MB) were computed for both the P_0 and P_1 data using 1-month data from all possible gauges with 100% data completeness during the period January 2005 – December 2012 (Table 3.1). The overall RMSE in Table 3.1 was computed as

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\frac{1}{n_t}\sum_{i=1}^{n_t}\left(P_{t,i} - G_{t,i}\right)^2}\ , \tag{3.25}$$

where N is the total number of times ($N = 96$) and n_t is the number of collocated gauges available at each time t. The overall mean bias was computed similarly to (3.25) as

$$MB = \frac{1}{N}\sum_{t=1}^{N}\frac{1}{n_t}\sum_{i=1}^{n_t}\left(P_{t,i} - G_{t,i}\right). \tag{3.26}$$

The two evaluation metrics were also computed for longer accumulation periods (Table 3.1). For a given gauge with 100% data completeness for accumulation period *a*, ε_i was computed only if each and every 1-month period *t* within the accumulation period was considered "blocked" at the collocated HRAP grid cell.

Table 3.1 indicates a significant reduction in the *RMSE* for the beam blockage-adjusted data (P_1) relative to the original Stage IV dataset (P_0) for all accumulation periods. Additionally, the overall underestimation of the P_0 dataset relative to gauge data at grid cells considered blocked is improved significantly in the P_1 dataset. For example, the magnitude of the 36-month beam blockage-adjusted overall mean bias (MB_1) is more than 5 times lower than the magnitude of the overall raw bias (MB_0).

Table 3.1. The *RMSE* and *MB* used for comparison of raw Stage IV (subscript 0) and beam blockage-adjusted (subscript 1) precipitation errors for selected accumulation periods.

	1	**3**	**6**	**12**	**18**	**24**	**36**
MSE_0	16.9 mm	38.3 mm	66.0 mm	123.7 mm	186.2 mm	256.1 mm	405.3 mm
MSE_1	16.1 mm	34.8 mm	57.7 mm	99.1 mm	142.7 mm	187.8 mm	267.8 mm
MB_0	-6.7 mm	-18.3 mm	-35.5 mm	-77.6 mm	-122.4 mm	-178.8 mm	-314.9 mm
MB_1	-1.8 mm	-4.0 mm	-7.1 mm	-16.2 mm	-23.9 mm	-35.3 mm	-55.1 mm

CHAPTER IV

MODELING MEAN-FIELD AND RANGE-DEPENDENT BIASES

IV.1. Introduction

The focus in this chapter is on correcting the systematic range-dependent biases in the Stage IV precipitation estimates. Range-dependent biases are those that can be described as strictly as a function of d (3.2) using a continuous function. In addition to correcting biases as a function of range, any correction applied will correct the mean-field bias of a given precipitation field within a radar domain. Given our conceptual equation of bias sources for radar precipitation estimates (1.1), the equation for Stage IV data biases in this chapter is

$$\frac{P_0}{P_T} = B(d) + B_{M-F}\,. \tag{4.1}$$

A procedure will be developed to estimate the range-dependent bias as a function of range, given a set of discrete bias data points. The algorithm includes finding a suitable set of bias data points and developing a parametric model to accurately assess the range-dependent biases. This range-dependent bias model must be robust enough to handle the geographical and seasonal variations in range-dependent biases that occur in central and eastern CONUS. For any given WSR-88D and any time period defined by ending date t and accumulation period a, the resulting range-dependent bias estimation model will be applied to the all the Stage IV P_1 precipitation estimate at the radar, which has already been corrected for errors due to beam blockage. The resulting dataset will contain Stage

IV precipitation estimates (P_3) corrected for beam blockage, mean field biases, and range dependent biases using the model bias values.

Here are some desirable properties that the developed model chosen to estimate the Stage IV range-dependent bias as a function of range should have.

1. The model should be able to represent the true bias as a continuous function of d for all ranges in a radar domain, which includes realistic values for the biases and for the range-dependent bias corrected precipitation data, i.e., no negative values.
2. The model should be able to account for known processes that lead to range-dependent or mean-field biases (see Section II.6).
3. The model should be resistant to overfitting of the available bias data.
4. The model should be resistant to outliers.

This chapter will discuss the properties of the Stage IV biases in the context of how an ideal model should handle the range dependence of biases within a WSR-88D radar domain. A testing procedure called leave-one-out cross validation (LOOC-V) is used to determine the ideal form of the model and to objectively test different model parameterizations and specifications. The outcome was development of a decision-based model that determines whether or not there exists a maximum in the bias data associated with the vertical profile of reflectivity (VPR).

IV.2. Properties of Range-Dependent Biases

The term true surface precipitation (P_T) is the actual amount of liquid water falling on a pre-specified area over a specific time interval (Villarini and Krajewski 2010b; Rinehart 2004). Range-dependent bias corrections of radar precipitation estimates use rain gauge data (P_G) as an approximation for P_T. Although individual gauge measurements may contain errors, gauge networks as a whole are assumed to be unbiased. Any computed Stage IV precipitation bias relative to P_T measured at HRAP grid cell i is

$$B^i = \frac{P_1^i}{P_T^i} - 1 , \tag{4.2}$$

but given that the dataset used in the chapter has already been corrected for beam blockage, the biases in this chapter are conceptually

$$B^i = \frac{P_1^i}{P_T^i} - 1 = \frac{P_0^i}{P_T^i} - B^i\left(d,\theta\right) - 1 . \tag{4.3}$$

The following summarizes sources of errors in radar precipitation estimates that could potentially lead to either mean-field biases or range-dependent biases in Stage IV precipitation estimates.

IV.2.a. Mean-Field Bias Sources

Radar miscalibration and incorrect Z-R relationships may lead to biases are roughly equal in magnitude across an entire radar domain. Miscalibration occurs when the constant C in the equation that measures total backscattered power is in inaccurate because of changes over time to different radar components (Villarini and Krajewski 2010b). Miscalibration is most easily identified by measurements of a storm at the same location by more than one radar. Z-R relationships contribute to mean-field bias when storms with relatively homogeneous drop size distributions (DSD) are incorrectly characterized by given Z-R relationships. Heterogeneities in DSDs may lead to random errors, which are neither mean-field nor range-dependent bias sources.

IV.2.b. Range-Dependent Bias Sources

Range-dependent biases can occur because radars measuring reflectivity at a height h above the surface that increases with increasing range d, with h dependent on both d and the tilt angle ϕ of the radar beam (2.13). In a typical precipitation event, the hydrometeors being measured have different properties at height h than when they reach the surface. The VPR is a dimensionless quantity that is an estimate of the ratio of the measured reflectivity at height h relative to the surface, given as

$$VPR(h) = \frac{Z(h)}{Z_0}. \tag{4.4}$$

Assuming a spatially-invariant VPR throughout a given radar domain, the VPR influence may result in a maximum of (4.4) at an intermediate range related to the radar bright-band. In addition to the VPR influence, the effects of attenuation, beam filling, and overshooting due to the radar beam geometry can lead to a systematic underestimation of P_T as d increases.

IV.3. Point Bias Data

There are two types of bias data points that were deemed useful for construction of a range-dependent bias estimation model. Radar-gauge pair bias data points (B_G) were computed using (4.2) as the ratio between Stage IV precipitation and gauge precipitation accumulated over the same time span at the same location. The second, more novel, type of bias data points (B_N) were computed from Stage IV PoN precipitation data, which were also used to identify regions of beam blockage in the previous chapter. PoN precipitation is used as a proxy of the range-dependent bias as it can characterize spatial variations, but it cannot be used to characterize bias magnitude. PoN is more useful than other dimensionless measures of precipitation, such as each precipitation value divided by the mean-field precipitation, because it eliminates the effects of climatological range-dependent variations of precipitation within the radar domain.

For a given radar, ending date t, and accumulation period of a months, the radar-gauge pair bias data points and PoN-based bias data points will be used in the range-dependent bias estimation model. A testing procedure described in the next section determined the appropriate form of the parametric model. Separate models will be built for both data types using only bias data within the a previous months preceding ending date t. For each parameter in the model, the radar-gauge pair value will be combined with the Stage IV PoN precipitation value into a single parameter value. The relative weights of the two parameter values are maximum-likelihood estimates of the statistical confidence of the two model fits. Validation of the range-dependent bias estimation model will use only the radar-gauge pairs, because unlike Stage IV PoN precipitation, these data points are a direct assessment of Stage IV biases.

IV.3.a. Radar-Gauge Pair Bias Data

The latitude and longitude coordinates of each COOP, WBAN, and CoCoRaHS gauge (Fig. 4.1) were used to pair each gauge with the HRAP grid cell it is located within. For a given radar and time period all gauges with 100% data availability at the monthly time scale were used in the development of range-dependent bias estimation model. In the last section of this chapter, an analysis will determine an appropriate threshold for other time periods, given that the number of gauges with data completeness diminishes as time scales get longer. The radar-gauge pair bias value B_G for a single data point i and time period of length n_d days is

$$B_G^i = \frac{P_1^i}{P_G^{\ i}} - 1 = \frac{\sum_{d=1}^{n_d} P_1^{i,d} w_d}{\sum_{d=1}^{n_d} P_G^{\ i,d} w_d} - 1, \tag{4.5}$$

where d is a single day within the period and w_d is a daily weighting coefficient that equals zero if a gauge measurement is missing and one if it is non-missing.

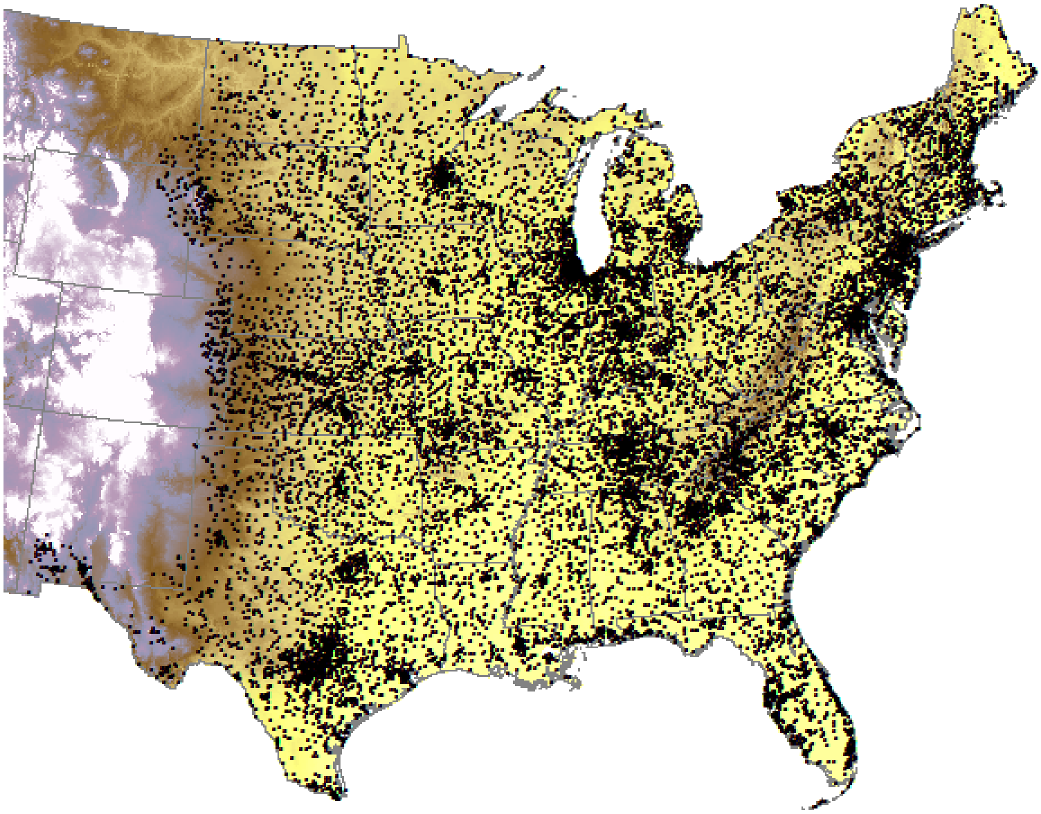


Fig. 4.1. Locations of the available radar-gauge pairs in the central and eastern CONUS.

To avoid dividing by very small numbers, only B_G values computed from pairs with both P_0 and P_G values above some minimum threshold will be included in the range-dependent bias estimation model, which was determined to be 3.0 mm based on inspection of the effect low precipitation values can have on computed bias values (Fig. 4.2). For a given radar-gauge pair bias, the larger of the P_0 and P_G values is assigned to A and the smaller to B, and Fig 4.2 shows the ratio A:B plotted as a function of B.

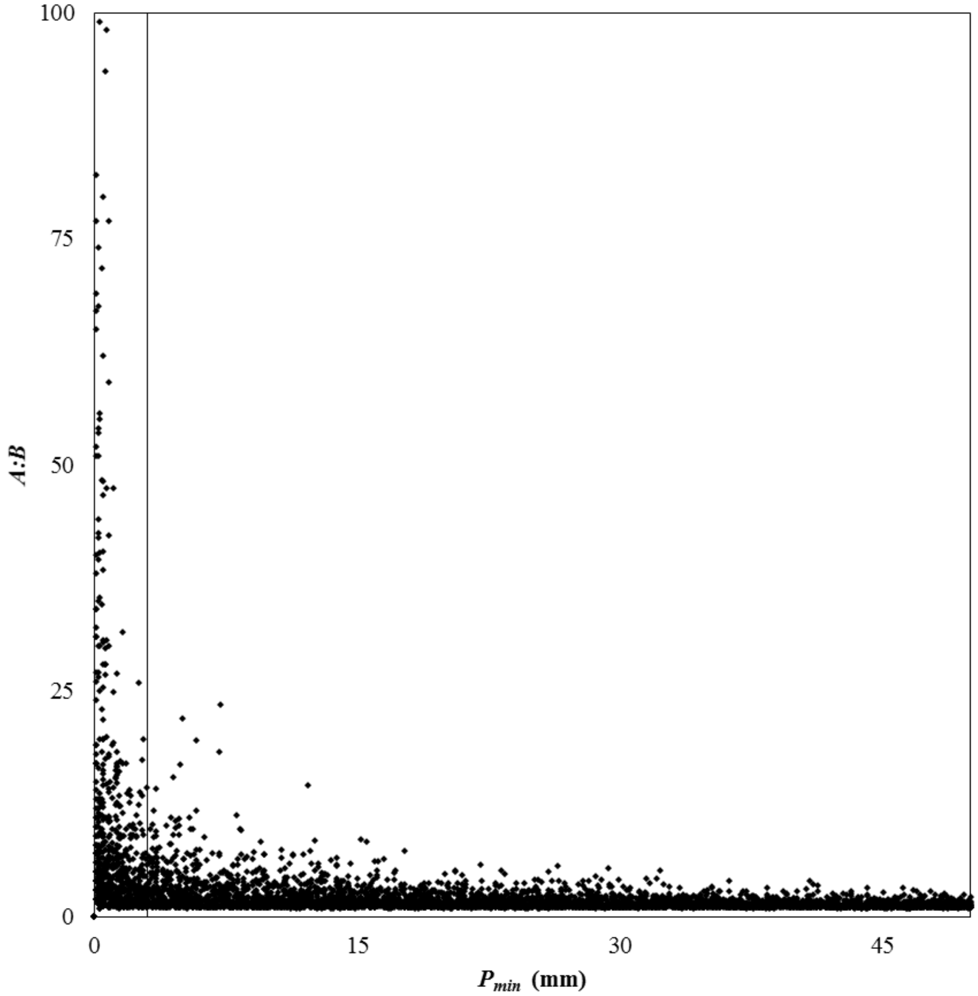


Fig. 4.2. The ratio A:B of the larger value precipitation value (A) to the smaller value precipitation value (B) in a radar-gauge pair as a function of B.

IV.3.b. Stage IV Percent of Normal Precipitation Bias Data

Unlike the irregularly spaced radar-gauge pair bias data, the Stage IV PoN precipitation data has complete two-dimensional spatial coverage across a given radar domain (see Section III.2 for a more detailed description). Therefore, it is possible to obtain generally equal azimuthal sampling for most intermediate range intervals at most ranges. The main assumption is that PoN gives an indication of the spatial variability of bias that does not rely upon the availability of gauges, which makes it quasi-independent from the radar-gauge pair biases. One drawback is that the mean-field bias cannot be computed using strictly PoN precipitation. A secondary assumption guiding the use of Stage IV PoN precipitation as a proxy for range-dependent bias is that averaging all the samples for a given range interval minimizes any azimuthal dependencies.

In a given radar domain, all the HRAP grid cells were arranged in order of increasing distance and then split into n separate bins, each of which will represent a single Stage IV PoN bias data point. The choice was made that the number of bins should equal the number of radar-gauge pair bias data points and that the bins should be constructed without knowledge of azimuthal angles. The bins were constructed to be of roughly equal size, so that the difference in the number of members for any two randomly selected bins j_1 and j_2 has an absolute value of zero or one. All n_j members of bin j have lie within an annulus that is unique to that bin. Fig. 4.3 is an example of the bins constructed for the Little Rock, AR (KLZK) radar domain for the 1-month period ending 31 December 2012.

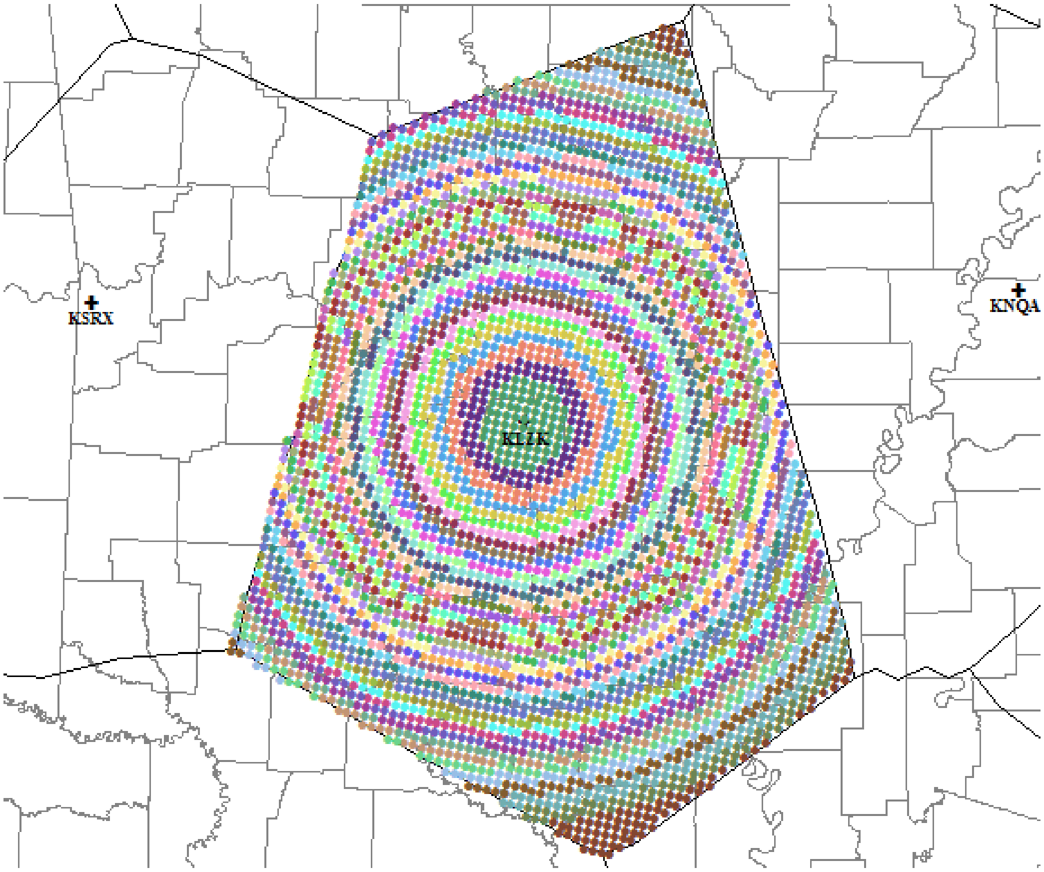


Fig. 4.3. The 56 bins used to create the 1-month S4 PoN precipitation bias data points at the Little Rock, AR (KLZK) WSR-88D for the period ending 31 December 2012.

The single data point from each bin j has both a single range value (d_N) and Stage IV PoN bias value (B_N) computed as an equal-weighted average of all n_j member values

$$B_N{}^j = \frac{1}{n_j} \sum_{i=1}^{n_j} P_N{}^i \, . \tag{4.6}$$

Combining the Stage IV B_N parameters with the radar-gauge pair parameters in the range-dependent bias estimation model requires an adjustment of the B_N value at each bin by an amplitude A. For a given radar domain and set of bias data points for each type

$$A = \frac{\underset{1 \le i \le n}{med}\, B_G}{\underset{1 \le i \le n}{med}\, B_N}, \tag{4.7}$$

using the median of the n data points for each data type. The Stage IV PoN bias data point for bin j is simply the Stage IV PoN bias value computed in (4.6) multiplied by the amplitude A.

IV.4. Bias Estimation Model Testing

IV.4.a. Leave One Out Cross-Validation

The range-dependent bias estimation model is an idealization of bias as a continuous function of range, predicted using only discrete data points at random ranges. Cross-validation (C-V) is a widely used testing procedure for estimating model prediction error for random and irregularly-spaced datasets (Hastie et al. 2005). In K-fold C-V, n sample data points are split into K equal-sized or roughly equal-sized parts. K prediction models are constructed, each time leaving a different part k out as “validation data” while the model is constructed using data points from the remaining $K - 1$ parts, which are referred to as “training data”. The prediction error for each individual model is computed using the data points in the kth validation group. The most commonly used metric for prediction error and the one used in this study is mean squared error (MSE; Hastie et al. 2005). MSE is simply the sum of squared error (SSE) divided by the total number of validation groups. Separating data into a training group and validation group allows one to ensure that the model does not overfit the n sample data points. If the prediction error was estimated using only training data points, the value of the error would get smaller and smaller as more parameters were included in the model (Hastie et al. 2005).

Table 4.1. The overall MSE of the LOOC-V tests done for all radar-months on the radar-gauge pair bias data points. Each test looked at the two different possibilities, each of which is listed.

Test	**Possibility 1**	**Possibility 2**
1	MK2011 model form	**SLR single straight line model form*
	0.1675	0.1480
2	**SLR single straight line model form*	T-S single straight line model form
	0.2205	0.2206
3	SLR single straight line model form	**T-Sw single straight line model form*
	0.22052	0.22049
4	**T-Sw CDF approximately 0.5*	T-Sw CDF = 0.5
	0.220487	0.220491
5	**CoCoRaHS not used*	CoCoRaHS used
	0.2205	0.2209
6	**T-Sw single straight line model form*	T-Sw merged maximum model form
	0.1966	0.1949
7	All partitions $m_{T\text{-}S1} > 0$ and $m_{T\text{-}Sw2} < 0$	Only partitions $MSE_1 > MSE_2$
	0.2040	0.2019

Leave-one-out C-V (LOOC-V) is a special case of *K*-fold C-V where $K = n$, the total number of data points in the complete sample. LOOC-V is approximately unbiased for the true prediction error, whereas cross-validation testing with lower values of *K* may be biased (Hastie et al. 2005). Therefore, LOOC-V testing will be used to determine appropriate choices for the range-dependent bias estimation model, which includes the model form, the number of parameters, the method for determining the parameters, and any restrictions needed to be implemented in the model construction.

IV.4.b. Details of Testing

For a given radar domain, LOOC-V testing will use all *n* COOP and WBAN radar-gauge pair bias data points as validation data points. The CoCoRaHS stations will not be used as validation gauges since the network does not have any automated quality control process, unlike the other two networks (Cifelli et al. 2005). Each LOOC-V test will determine the preferred choice from a set of two possible outcomes, based on the choice with the lower overall MSE value. For instance, one of the initial LOOC-V tests will determine whether the inclusion of CoCoRaHS radar-gauge pair biases in the training group is preferred to their exclusion. Table 4.1 will summarize the overall MSE for each of the two choices.

Since this study is concerned with correcting biases on monthly and longer time scales, the accumulation period $a = 1$ was chosen for the LOOC-V testing procedure. Another assumption is that noise in the bias data points due to random errors is highest at the shortest time scales, and any model form that can handle 1-month biases is more than capable of handling biases at longer time scales. Only radar-gauge pairs with no missing gauge data were used for testing at each of the 104 WSR-88Ds and 96 calendar months in the testing period for a total of 9,984 "radar-months." Further LOOC-V testing on longer accumulation periods will examine if the prediction provided by the training group is improved using radar-gauge pairs from stations with missing gauge data.

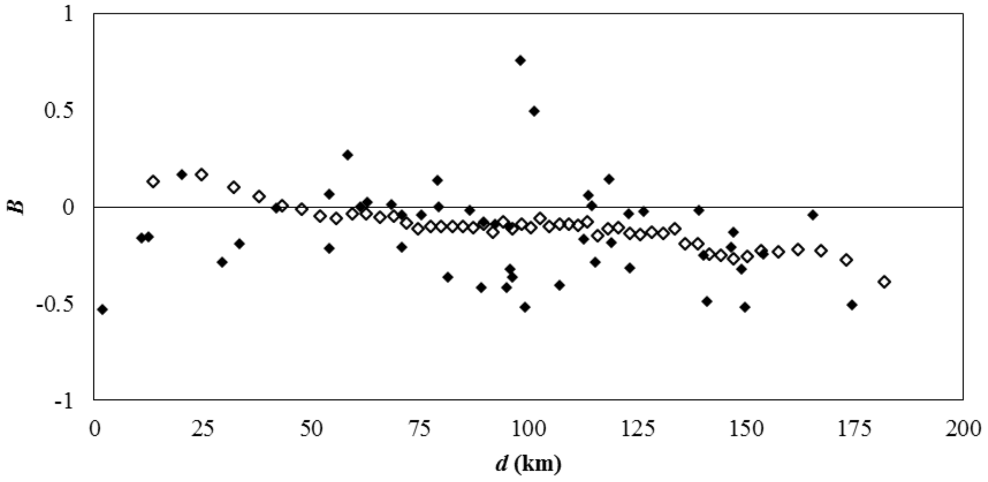


Fig. 4.4. 1-month B_G (solid diamonds) and B_N (open diamonds) data points at the KLZK radar domain for the period ending 31 December 2012. Data included (minus one outlier, which is shown in Fig. 4.6) have bias values between $-1.0 < B < 1.0$.

The Stage IV PoN bias data points will not be included in the LOOC-V testing, with the assumption for a given radar domain and time period these data will have significantly less noise than the radar-gauge pair bias data points. The Stage IV PoN has complete spatial coverage, so the range-interval averaged bias data points will have greater spatial consistency and there is no introduction of random errors due to the gauge data. Therefore, the model form found appropriate for the radar-gauge pair bias data points will be applied to the PoN bias data points. Fig. 4.4 is an example of both types of data points for the Little Rock, AR (KLZK) radar domain for the 1-month period ending 31 Dec 2012. Both types of bias data points for the KLZK December 2012 radar-month show the same the basic dependence on range, with a bit of disagreement at shorter ranges, but there is much less variance and greater spatial consistency with the Stage IV PoN bias data points.

IV.5. Determining Appropriate Model Form

The initial stage of determining the range-dependent bias estimation model form was seeing if there was an existing model form that was appropriate given the properties of our bias data points. The single most important factor leading to range-dependent biases is the VPR. For single storms, spatial non-uniformity of the VPR may cause random biases. For longer accumulation periods the overall VPR influence can cause overestimation of P_T and a maximum in B data at range intervals where the radar beam typically intercepts the bright-band. In this section, LOOC-V testing only used COOP and WBAN stations in the training group to determine an appropriate form of a D-R bias estimation model form.

IV.5.a. Krajewski et al. (2011) VPR Model

There has been a great deal of work done to correct for range dependent biases caused by the VPR, including correction algorithms by Andrieu and Creutin (1995), Seo et al. (2000), Vignal and Krajewski (2001), and Zhang et al. (2008). However, these and most other correction procedures require that availability of reflectivity scans at a number of different tilt angles, information that is not available in the Stage IV precipitation dataset. Krajewski et al. (2011) created a conceptual statistical model (K2011 model) of the "climatological" VPR at a given WSR-88D as a function of height h above the surface, or $VPR(h)$. The objective of the Krajewski et al. (2011) study was to create a model of the range-dependent errors that targets the VPR as the source of uncertainty in radar precipitation estimates, which falls in line with the goal of this section. Also, the conceptual model is useful in the context of correcting the Stage IV dataset as it can created without a priori knowledge of the vertical structure of the reflectivity measurements.

The K2011 model is a continuous model characterized by three primary vertical layers. The lowest K2011 vertical layer is a below-bright-band region where $VPR(h)$ is one. The middle layer is where a bright-band is located and the $VPR(h)$ is assumed greater than one. The top layer is an above-bright-band region where $VPR(h)$ is assumed to decrease linearly with height.

Parameterization of the K2011 model begins with characterizing the enhanced reflectivity associated with the climatological bright-band region. The height of the maximum bright-band enhancement (h_{BB}), the ratio of the reflectivity maximum at h_{BB} relative to h_0 (Z_{max}), and the vertical depth of the bright-band region (e_{bb}) are combined so that within the K2011 bright-band region,

$$VPR\left(h\right) \approx 1 + Z_{max} \exp\left[-\left(\frac{h - h_{bb}}{e_{bb}} \right)^2 \right]. \tag{4.8}$$

The above-bright-band layer in the K2011 model is characterized by a single parameter quantifying the linear decrease in VPR with altitude (s_{bb}) so that

$$VPR\left(h\right) \approx \exp\left[-\frac{\ln 10}{10} s_{bb} \left(h - h_{bb} \right) \right]. \tag{4.9}$$

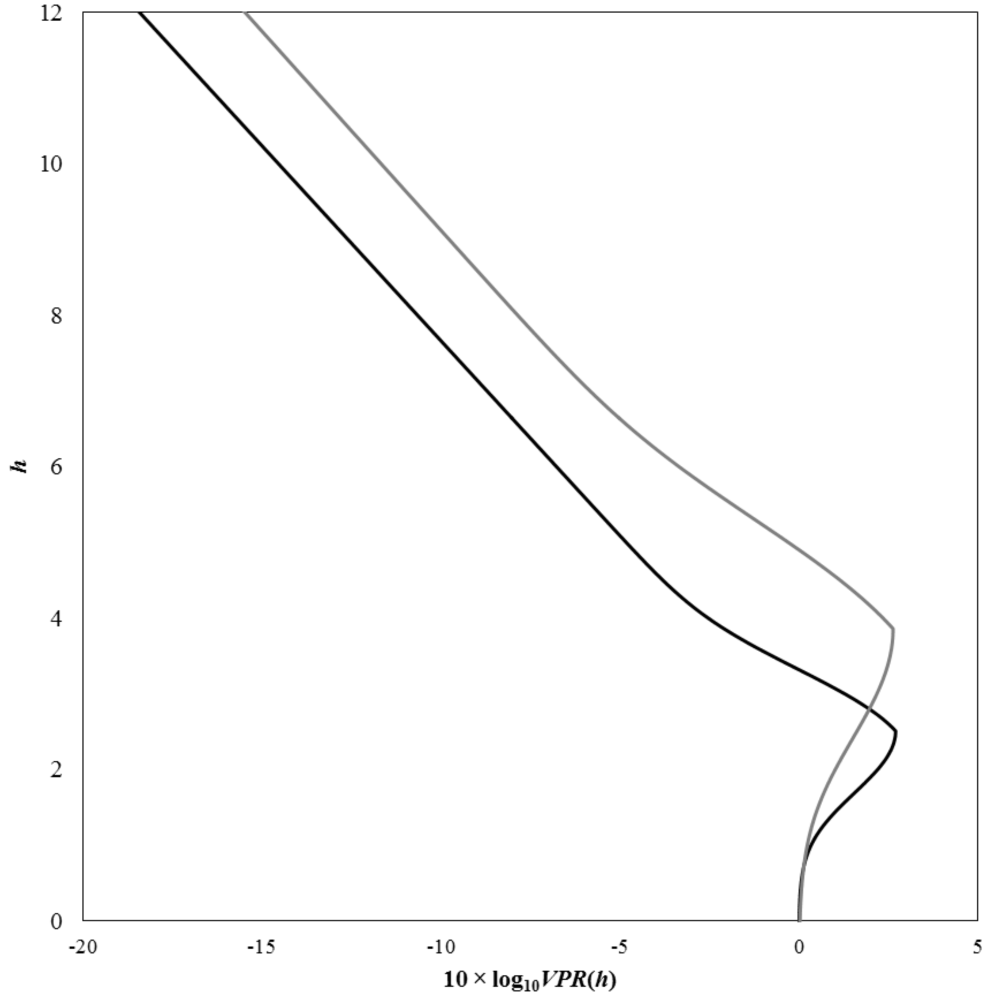


Fig. 4.5. The vertical profile of reflectivity as a function of height in the Krajewski et al. (2011) conceptual model for the 1.48° tilt of the Tulsa, OK (KINK) WSR-88D. The cold season (black line) parameters are $h_{bb} = 2.51$, $e_{bb} = 0.98$, $Z_{max} = 0.86$, and $slop = 1.94$ and the warm season (gray line) parameters are $h_{bb} = 3.86$, $e_{bb} = 1.73$, $Z_{max} = 0.83$, and $slop = 1.90$, taken from Krajewski et al. (2011).

Combining (4.8) and (4.9) into a single equation to describe the magnitude of the VPR influence at any height *h* with the K2011 parameterization is

$$VPR(h) = \left\{ 1 + Z_{max} \exp\left[-\left(\frac{h - h_{bb}}{e_{bb}} \right)^2 \right] \right\} \exp\left[-\delta_h \frac{\ln 10}{10} s_{bb} \left(h - h_{bb} \right) \right], \quad (4.10)$$

where δ_h is a delta function that is zero if $h < h_{bb}$ and one if $h \geq h_{bb}$. Fig. 4.5 shows the Krajewski et al. (2011) VPR for the 1.48° tilt of the Tulsa, OK (KINX) WSR-88D, given separate parameterization for the cold season and warm season. The parameterization was determined by minimizing the sum-of-square differences between the model VPR and empirical VPR data.

One drawback to the K2011 model is that it is not possible to do a direct transformation of the VPR from height coordinates to range coordinates without knowledge of the tilt angles used in the radar precipitation estimates. For each of the three lowest tilt angles, Krajewski et al. (2011) model the VPR influence at distance *d* from the radar by integrating the VPR found in (4.10) throughout the vertical width of the beam, which is dependent on range.

IV.5.b. Modified Krajewski et al. (2011) VPR Model

The K2011 was considered to be a good starting point for the range-dependent bias estimation model given the model form properties we are looking for. However, adjustments to the K2011 model form were necessary if it was to be considered a candidate. The most obvious modification is that the model needs to be a function of range rather than height given the bias data points are in range coordinates. Only transforming to range coordinates might be sufficient if the goal was to find a range-dependent VPR estimation model. However, the K2011 VPR model characterizes biases at ranges between the radar location and the region with enhanced reflectivity due to the bright-band as approximately one. This approximation is not sufficient for a candidate model since it assumes a mean-field bias of one, and this shortcoming was addressed with the addition of a mean-field bias term to the modified Krajewski et al. (2011) model (MK2011 model).

The MK2011 model of range-dependent biases, with a vector of parameters **a** = $\{d_{bb}, \varepsilon_{bb}, \zeta_{max}, \sigma_{bb}, \beta\}$, is given by

$$B(d) = \left\{1 + \zeta_{max} \exp\left[-\left(\frac{d - d_{bb}}{\lambda_{bb}}\right)^2\right]\right\} \exp\left[-\delta_d \sigma_{bb} \left(d - d_{bb}\right)\right], \tag{4.11}$$

where β is mean-field bias adjustment factor, d_{bb} is the distance from the radar of the maximum VPR enhancement, ζ_{max} is the magnitude of the maximum enhancement, λ_{bb} is the horizontal width of the enhancement region, σ_{bb} characterizes the decrease is VPR beyond the bright-band region, and δ_D is a delta function that is zero if $d < d_{bb}$ and one if $d \geq d_{bb}$. Since there is no simple analytical solution, determination of the MK2011 parameters must be done through numerical methods. Because the MK2011 best-fit vector of **a** is found numerically, a reasonable numerical range for the possible values of each parameter was determined (Table 4.2), with a possible 11^5 or 161,051 permutations. In the MK2011 model, the range is normalized by the d_{max} of the radar domain such that the possible range of the bias data points is (0, 1]. The best-fit vector of ***a*** given a set of bias data points is the permutation which minimizes the RSS. Fig. 4.6 is an example of the

MK2011 constructed for the Little Rock, AR (KLZK) radar domain for the 1-month period ending 31 December 2012.

Table 4.2. Numerical possibilities for the five parameters in the MK2011 model form.

Parameter	**Minimum**	**Interval**	**Maximum**
D_{bb}	-1.0	0.3	2.0
ε_{bb}	0.0	0.05	0.5
ζ_{max}	0.0	0.5	5.0
σ_{bb}	0.0	0.3	3.0
β	0.5	0.1	1.5

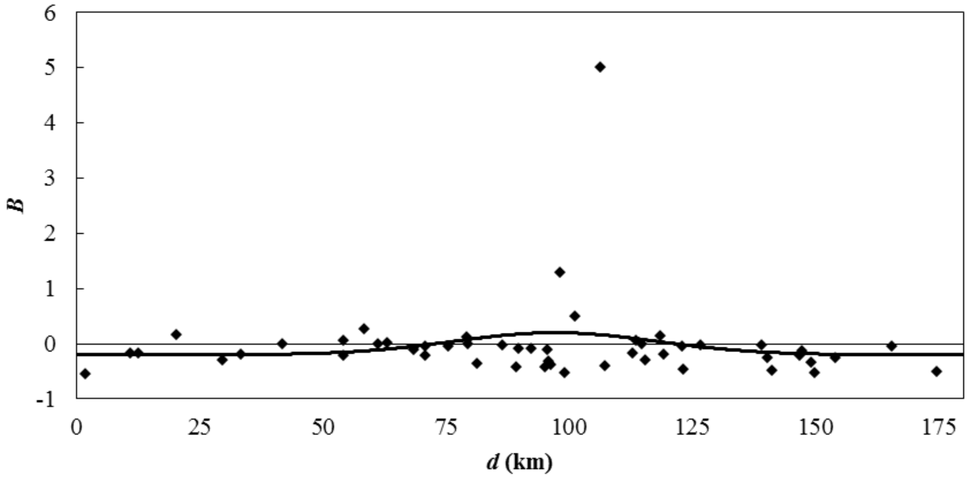


Fig. 4.6. The modified Krajewski et al. (2011) model fit (solid black line) to the radar-gauge pair B_G data at the Little Rock, AR (KLZK) WSR-88D for the 1-month period ending 31 December 2012.

IV.5.c. MK2011 Model Form vs. Single Straight Line Model Form

The performance of the 5-parameter MK2011 model form was compared using the LOOC-V testing procedure to a 2-parameter single straight line model form, which uses SLR for parameter estimation. The two SLR estimated parameters are the slope (m_{SLR}) and y-intercept (b_{SLR}) found through minimization of the RSS. The objective of the testing was to determine if the MK2011 model is overfitting the radar-gauge pair bias data. The computational expense of finding the MK2011 model best-fit **a** meant only one randomly chosen data point was used for validation at each radar-month. Both the best-fit MK2011 parameters and SLR parameters of the training group were determined and each model estimate was computed for the *d* value of the validation data point.

The LOOC-V testing results show a lower overall MSE was lower for the single straight line model than for the MK2011 model (Table 4.1). Both model types determine parameters based on minimization of the RSS, and extreme outliers can lead to large residual values. However, the 5-parameter MK2011 model gives has a greater capability of adjusting its model shape to fit outliers and thus minimize the RSS, which leads to overfitting. Fig. 4.7 is an example of overfitting by the MK2011 model that is handled better by a SLR model for the Norfolk, VA (KAKQ) radar domain for the 1-month period ending 31 October 2008. The removal of only a single radar-gauge pair bias data point from COOP station 444044 (Holland 1E, VA) completely changes the shape of the MK2011 model fit in the KAKQ domain for the same period. In this case, there are 43 available radar-gauge pair biases in Fig. 4.7, so one it would be difficult to attribute the overfitting a lack of training data.

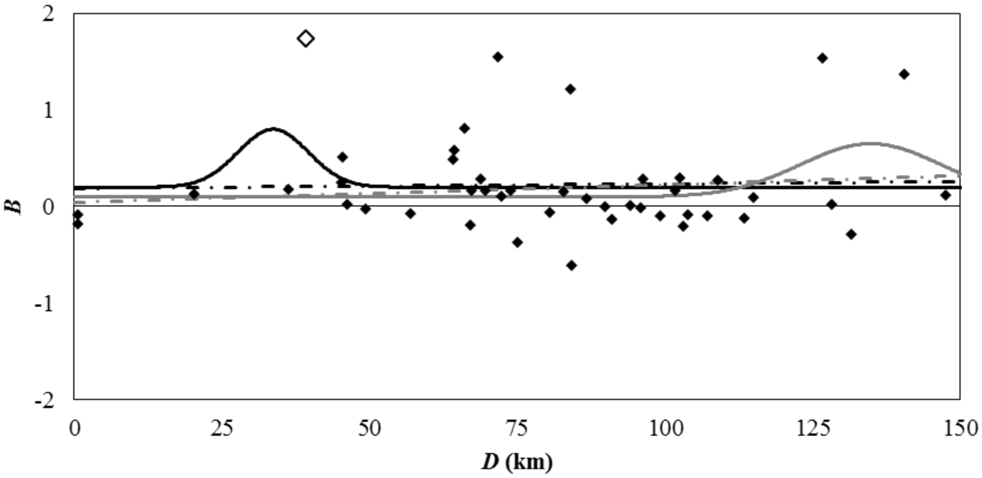


Fig. 4.7. The modified Krajewski et al. (2011) model fit (solid black line) to the radar-gauge pair B_G data at the KAKQ radar domain for the 1-month period ending 31 October 2008; the solid gray line is the MK2011 model fit removing only the open diamond. The SLR model fit (dashed black line) using all the data points and the SLR model fit removing only the open diamond (dashed gray line) are included for comparison.

The SLR model was found to be preferable to the MK2011 model because it has fewer parameters, which makes it more resistant to overfitting. However, a single straight line model form does not have the capability to model a maximum in the VPR, which was a desirable property of the MK2011 model. Looking back at the desired properties of an ideal range-dependent bias estimation model form, both model forms satisfy property 1 (continuous function of *D* for all possible ranges), the MK2011 model form satisfies property 2 (ability to model both range-dependent and mean-field bias processes), and the single straight line model form satisfies property 3 (resistant to overfitting). So one would assume the logical next step is to merge the two model forms into a single model.

However, SLR parameter estimation still doesn't quite satisfy property 4, as it is sensitive to outliers since it is based on minimization of the RSS. Also, the SLR model form is based on parametric statistics and assumes homoscedasticity of the radar-gauge pair biases throughout the radar domain with increasing range. To better satisfy property 4 in a merged model form, an alternative method of single straight line parameter estimation called Theil-Sen was explored and its ability to handle radar-gauge pair bias data relative to SLR was assessed using LOOC-V testing.

IV.5.d. Alternative Estimator of Regression Model Parameters

IV.5.d.1) Theil-Sen Estimation of Parameters

Theil (1950) introduced a method for estimating the slope as an alternative to traditional regression techniques that uses non-parametric statistics. The Theil (1950) slope is the median of the slopes computed from each possible pair of data points. Sen (1968) limited the set of possible slopes to only pairs of points i and j in which $x^i \neq x^j$, which in our case limits pairs to having points from different HRAP grid cells. For a given radar domain, ending date t, and accumulation period a, the data points are initially ordered by increasing range. Candidate Theil-Sen slopes (m^{ij}) are computed for each pair of data points for which $i < j$ as

$$\underset{i<j}{m^{ij}} = \frac{B^j - B^i}{d^{ij}}, \tag{4.12}$$

where $d^{ij} = d^j - d^i$. The Theil-Sen slope ($m_{T\text{-}S}$) is the median of all n_m candidate slopes, written as

$$m_{T-S} = \underset{1 \le i < j \le n}{\mathrm{med}} \left(\frac{B^j - B^i}{d^{ij}} \right). \tag{4.13}$$

The Theil-Sen y-intercept ($b_{T\text{-}S}$) is determined by passing a straight line with slope $m_{T\text{-}S}$ through all the data points. For each point i a y-intercept b^i is computed as

$$b^i = B^i - m_{T-S} d^i \tag{4.14}$$

and $b_{T\text{-}S}$ is the median of all n ordered y-intercepts, written as

$$b_{T-S} = \underset{1 \le i \le n}{\mathrm{med}} \left(B^i - m_{T-S} d^i \right). \tag{4.15}$$

Theil-Sen parameter estimation is more robust than SLR and is able to handle a dataset with up to 29.3% of the data points being corrupt or outliers (Rousseeuw and Leroy 2003), an advantage that is desirable for the irregularly-spaced radar-gauge pair bias data. The LOOC-V testing revealed that the performance of the SLR and Theil-Sen methods was nearly identical with a slight edge to SLR (Table 1.2).

IV.5.d.2) Weighted Theil-Sen Estimation of Parameters

Because the Theil-Sen method was slightly outperformed by the SLR, a minor tweak to the T-S method was made. The traditional Theil-Sen slope estimate is based on equal weighting of each pair for which a slope is computed, regardless of the distance between the two data points. However, the magnitude of each slope computed from a pair of data points i and j is sensitive to the distance between the two points (Fig. 4.8). In our case, the slope of a line is less sensitive to noise in B_G values when the distance between the two points increases. Sievers (1978) suggests that each pair should have a computed weight based on the value d^{ij} instead of using a single universal weight of $1/n_m$. The weight of each pair (w^{ij}) is formulated as

$$\underset{i<j}{w^{ij}} = \frac{d^{ij}}{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} d^{ij}}, \tag{4.16}$$

with the computation of m^{ij} is shown in (4.12).

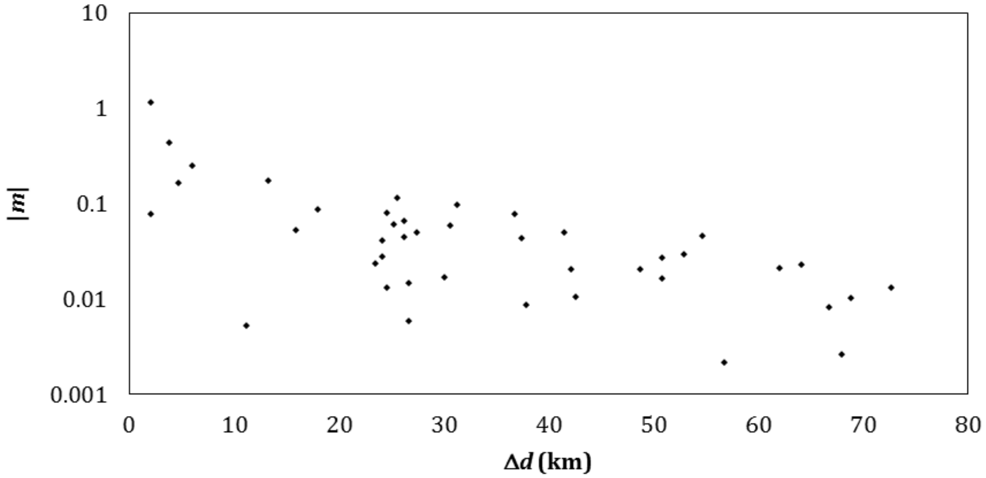


Fig. 4.8. The absolute value of the computed slope (|m|) for each pair of radar-gauge pair bias data points as a function of the distance (Δd) between each pair at the KFDR radar domain for the 1-month period ending 30 November 2005.

Instead of the traditional Theil-Sen method of estimating the slope as the median out of a set of equally-weighted candidates, the weighted Theil-Sen (T-Sw) method introduced by Sievers (1978) estimates the slope ($m_{T\text{-}Sw}$) using a weighted rank procedure. After computing the slope m^{ij} and weight w^{ij} of each pair, the pairs are ordered by increasing slope and each is assigned a rank r^{ij}. The pair with the lowest slope value was assigned $r^{ij} = 1$ and $r^{ij} = n_m$ was assigned to the pair with the largest slope value. A cumulative weighting function W is then introduced with an initial value of zero and for any rank R has a value

$$W = \sum_{k=1}^{R} w^{ij} \ni r^{ij} = k \ . \tag{4.17}$$

The function W is increased until it reaches a value ≥ 0.5 corresponding to rank $r_{1/2}$. The T-Sw slope is then computed as

$$m_{T-Sw} = \frac{1}{2}\left(m^{ij} \ni r^{ij} = r_{1/2} - 1\right) + \frac{1}{2}\left(m^{ij} \ni r^{ij} = r_{1/2}\right) \tag{4.18}$$

by Sievers (1978). The stated goal of the T-Sw is to find the median of the weighted cumulative distribution function, i.e. $W = 0.5$. The Sievers (1978) determination of the median was modified to allow for differential weighting of the two slopes on either side of $W = 0.5$, such that

$$m_{T-Sw} = \left[\left(\frac{1}{2} - W_1\right)\left(m^{ij} \ni r^{ij} = r_{1/2} - 1\right)\right] + \left[\left(W_2 - \frac{1}{2}\right)\left(m^{ij} \ni r^{ij} = r_{1/2}\right)\right], \tag{4.19}$$

where $W_1 < 0.5 \leq W_2$. Table 4.1 indicates the overall MSE in LOOC-V testing was nearly identical for determining $m_{T\text{-}Sw}$, with a slight preference for (4.18) instead of (4.19) and better performance than both SLR and traditional Theil-Sen regression. Therefore, the

Sievers (1978) computation of the single straight line slope parameter is preferred to the unweighted Theil-Sen. The T-Sw y-intercept is found in the same manner as the traditional Theil-Sen method, substituting $m_{T\text{-}Sw}$ for $m_{T\text{-}S}$ in both (4.14) and (4.15).

IV.5.d.3) Comparison of Theil-Sen and Weighted Theil-Sen

The T-Sw method of finding parameters performed marginally better in LOOC-V testing than both SLR and traditional Theil-Sen (Table 4.1). The Altus AFB, OK (KFDR) radar domain for November 2005 is an example of a radar-month where the traditional Theil-Sen and T-Sw methods produced noticeably different values for parameter estimation (Fig. 4.9). The KFDR bias data points prove to be something of a worst-case scenario for the traditional Theil-Sen method, which is that most range intervals are unsampled and data points are clustered into narrow range intervals with large variability. The median slope value $m_{T\text{-}S}$ to be -0.0030 km^{-1}, whereas the value of $m_{T\text{-}Sw}$ was -0.0085 km^{-1}.

IV.5.e. Testing for Inclusion of CoCoRaHS Gauges

The appropriateness of using (or not using) CoCoRaHS gauges was assessed for T-Sw parameter estimation of a single straight line using LOOC-V testing. For a given radar-month, each COOP and WBAN gauge was withheld as a validation point and the T-Sw parameters of two single straight lines were computed, each using a different training group. The first training group only contained the other COOP and WBAN radar-gauge pair biases while the second training group contained all the data points from the first training group in addition to CoCoRaHS radar-gauge pair biases. The results indicate a slightly lower overall MSE using only COOP and WBAN gauges in the training group (Table 4.1).

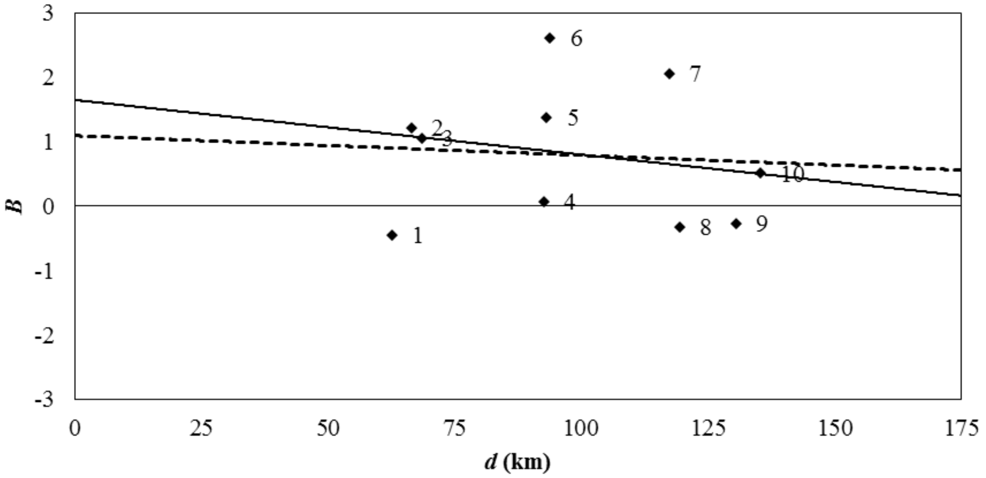


Fig. 4.9. Radar-gauge pair B_G data (solid numbered diamonds) at the KFDR radar domain for the 1-month period ending 30 November 2005. The straight line model fit using the traditional Theil-Sen method (dashed line) and the Sievers (1978) weighted T-S method model fit (solid line) are included.

IV.6. Conditional-VPR (conVPR) Model

IV.6.a. Overview

The Sievers (1978) T-Sw method of finding the slope and intercept of a single straight line satisfies properties 1, 3, and 4 of a preferred range-dependent bias estimation model. However, a single straight line model form will not be able to account for the VPR effect that leads to range-dependent biases, and thus, is not able to account for all significant physical processes (property 2). This section will introduce a new model that combines T-Sw estimation with the MK2011 conceptual model form that approximates the VPR as a function of range and satisfies all the properties of an ideal model. Based on LOOC-V testing of the T-Sw single straight line model form, only COOP and WBAN stations will be used in the merged, decision-based model form called the conditional-VPR (conVPR) model.

The basic concept of a conVPR model is to determine the existence or non-existence of a VPR maximum in the bias data points for a given radar-month. The non-existence of a maximum in the bias data points indicates one of three possibilities, which are

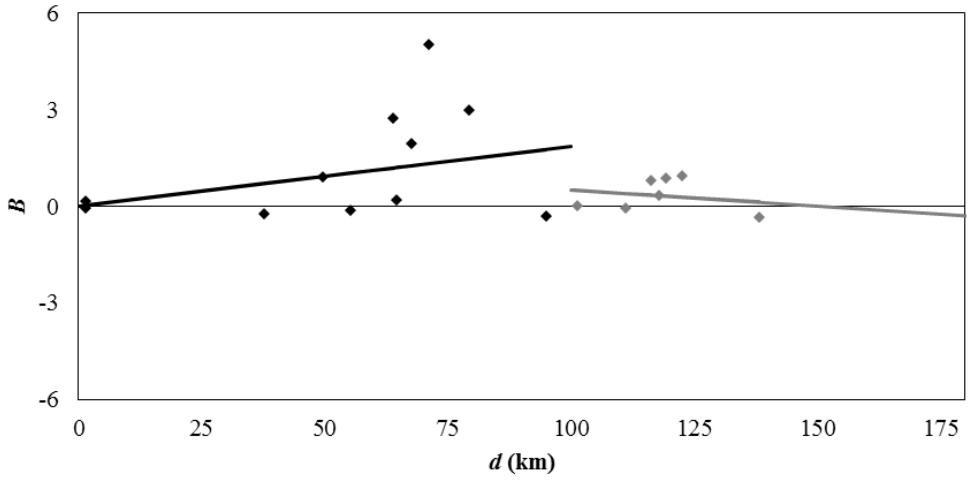
1. the VPR maximum is at a height corresponding to a range beyond the spatial footprint of a given radar domain,
2. a VPR maximum is not present, or
3. the VPR maximum is not statistically significant.

The first possibility likely indicates a primarily convective VPR structure where the height of the VPR corresponds to a range beyond the maximum of the radar domain. Therefore, a model with a single straight line having positive slope is more likely in the warm season in most locations. The second possibility can occur when temperatures throughout the radar domain are below freezing, which means there is no melting layer and thus no VPR maximum.

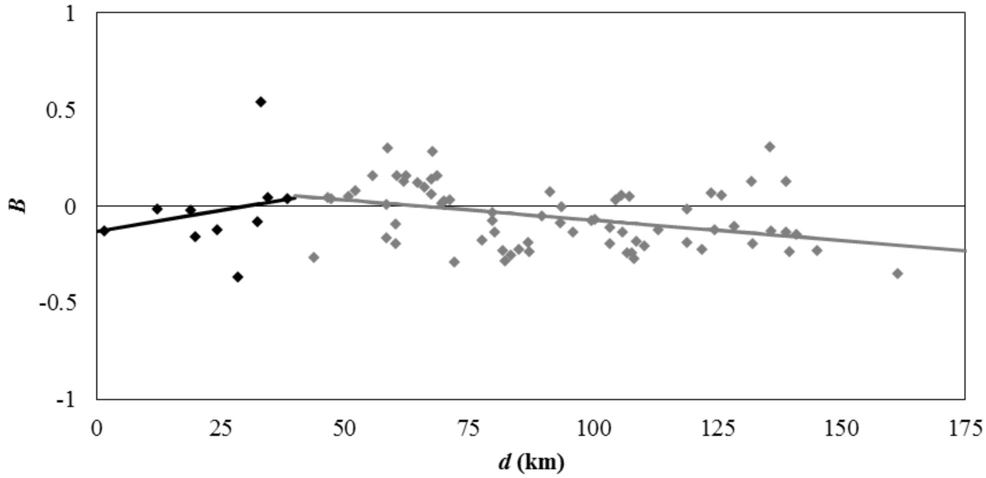
For each radar-month, the conVPR model will be constructed independently for both the radar-gauge pair biases (B_G) and the Stage IV PoN biases (B_N). If the bias data points indicate that a VPR maximum exists, the conVPR model will be a combination of two separate T-Sw single straight lines merged at a single range $d_\cap$, which we will call the T-Sw merged maximum model form with a VPR maximum $B_\cap$. If the bias data points do not indicate a VPR maximum, the conVPR model will contain a single T-Sw straight line found using the Sievers (1978) T-Sw method described in the previous section and we will call this the single conVPR model form.

IV.6.b. Defining a Candidate VPR Maximum

To determine each candidate T-Sw merged maximum model form maximum (B_{max}) for a given radar-month, the bias data points are partitioned several times. Each partition contains two non-overlapping groups such that Group 1 contains bias data points for $0 \leq d < d_p$ and Group 2 from $d_p \leq d \leq d_{max}$, with d_p defined as the changepoint for each partition. This first partition examined is at $d_p = 10$ km and for each subsequent partition, the range at which the partitioning occurred was increased by 10 km, which is consistent with the spatial step in the beam blockage detection procedure. The two criteria for a partition to contain a candidate B_{max} is that the T-Sw slope of Group 1 is positive ($m_{T\text{-}Sw1} > 0$) and the T-Sw slope of Group 2 is negative ($m_{T\text{-}Sw2} < 0$). If the slope criteria are met for a particular partition, y-intercepts ($b_{T\text{-}Sw1}$ and $b_{T\text{-}Sw2}$) are determined for each group. The term "candidate" is used because there may be more than one partition at a given radar-month satisfying the two slope criteria.



(a)



(b)

Fig. 4.10. Radar-gauge pair B_G data points for Group 1 (black diamonds) and Group 2 (gray diamonds), with the T-Sw single straight lines for each group. Data are from the (a) KSJT radar domain for the 1-month period ending 31 July 2012 and the (b) KGSP radar domain for the 1-month period ending 31 March 2007.

If a partition containing a candidate B_{max} with changepoint d_p satisfies $m_{T\text{-}Sw1} > 0$ and $m_{T\text{-}Sw2} < 0$, the two T-Sw straight lines from each group are merged into a single continuous function. A restriction to the T-Sw merged maximum model is that intersection of the straight lines in a partition with a candidate B_{max} occurs at d_p and rather than the natural intersection point of the two lines. Fig. 4.10 shows an example of d_p being a poor representation (Fig. 4.10a) and d_p being a good representation (Fig. 4.10b) of the actual intersection point of two T-Sw single straight lines, given a set of radar-gauge pair B_G data points. A metric used to determine the goodness of d_p as an intersection point is the absolute difference in the two model bias values B_1 and B_2 at the intersection range

$$\left|dB_p\right| = \left|B_2\left(d_p\right) - B_2\left(d_p\right)\right| = \left|\left(m_{T-Sw2} - m_{T-Sw1}\right) d_p b_{T-Sw2} - b_{T-Sw1}\right| . \qquad (4.20)$$

Using a candidate d_p = 100 km for the San Angelo, TX (KSJT) radar domain for the 1-month period ending 31 July 2012, the two lines have an actual intersection at a range of 53.9 km and $|dB_p|$ = 1.34 (Fig. 4.10a). Fig. 4.10b shows the two lines for a candidate d_p = 40 km at the Greer, SC (KGSP) radar domain for the 1-month period ending 31 March 2007, which have an intersection at a range d_p of 41.8 km and $|dB_p|$ = 0.01.

IV.6.c. Merging in T-Sw Merged Maximum Model Form

Up to this point, the T-Sw merged maximum model form has been described by five parameters (d_p, $m_{T\text{-}Sw1}$, $m_{T\text{-}Sw2}$, $b_{T\text{-}Sw1}$, $b_{T\text{-}Sw2}$), for which there is a difference $|dB_p|$ in the two bias estimates at d_p for partitions with a candidate B_{max}. Eliminating the y-intercept of Group 2, the two straight lines can be merged into a single continuous function of range d with a model bias $B(d)$ of

$$\begin{aligned} B\left(d\right) &= m_{T-Sw1} d_p + b_{T-Sw1} + 0.5\left(m_{T-Sw1} + m_{T-Sw2}\right)\ldots \\ &\ldots + 0.5\left(m_{T-Sw2} - m_{T-Sw1}\right)\left(d - d_p\right)\operatorname{sgn}\left(d - d_p\right) \end{aligned} , \tag{4.21}$$

where

$$\operatorname{sgn}\left(d - d_p\right) = \begin{cases} -1, & \text{if } d < d_p \\ -1, & \text{if } d = d_p \\ +1, & \text{if } d > d_p \end{cases} . \tag{4.22}$$

However (4.21) results in a non-continuous first derivative at d_p characterized by an abrupt transition in the function. Replacing the sign function (4.22) in (4.21) with a hyperbolic tangent transition function

$$\operatorname{trn}\left(d - d_p\right) = \tanh\left(\frac{d - d_p}{d_{\max} \times \gamma_{trn}}\right) \tag{4.23}$$

provides a continuous first derivate at d_p. The scale parameter γ_{trn} in (4.23) is proportional to the amount of smoothing at d_p. In the T-Sw merged maximum model form, the value of the scale parameter was chosen to be $\gamma_{trn} = 0.01$, which provides a minimal amount of smoothing. Substituting the transition function (4.23) for the sign function (4.22) in (4.21) gives a model bias of

$$\begin{aligned} B\left(d\right) &= m_{T-Sw1} d_p + b_{T-Sw1} + 0.5\left(m_{T-Sw1} + m_{T-Sw2}\right)\ldots \\ &\ldots + 0.5\left(m_{T-Sw2} - m_{T-Sw1}\right)\left(d - d_p\right)\tanh\left(\frac{d - d_p}{d_{max} \times \gamma_{trn}}\right) \end{aligned} . \tag{4.24}$$

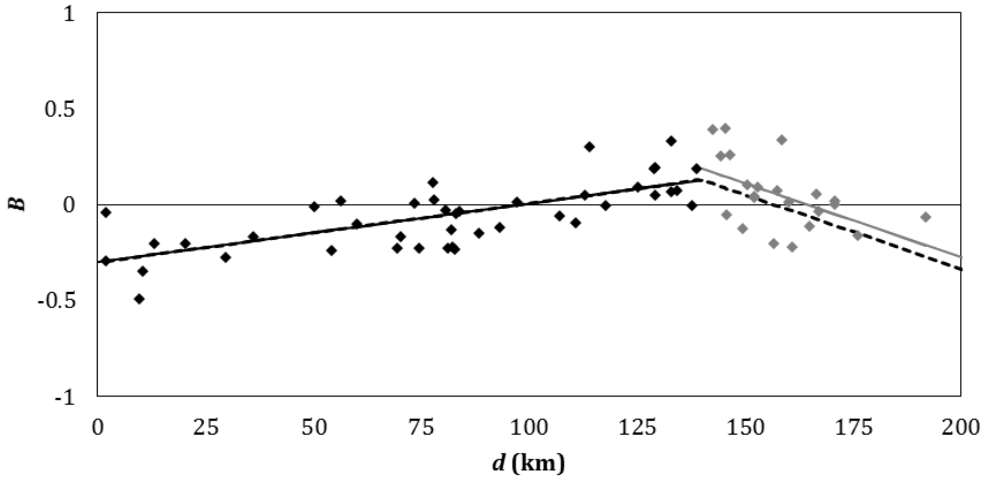


Fig. 4.11. Radar-gauge pair B_G data points for Group 1 (black diamonds) and Group 2 (gray diamonds), with the T-Sw single straight lines for each group. The dotted black line is merged model fit in Group 2 using (4.24). Data are from the KSHV radar domain for the 1-month period ending 31 January 2012.

Fig. 4.11 is an example of (4.24) using the set of radar-gauge pair bias data points at the KSHV radar domain for the 1-month period ending 31 January 2012. The merged model form in (4.24) is the single Group 1 straight line for ranges of $0 \leq d \leq d_p$ and the Group 2 straight line offset by a value $dB_p = B_2(d_p) - B_1(d_p)$ for ranges of $d_p \leq d \leq d_{max}$, with smoothing by the hyperbolic transition function near d_p. The value of dB_p for the two T-Sw single straight lines is 0.06 with a d_p value of 140 km.

The continuous 4-parameter function in (4.24) was modified to correct the systematic offset in Group 2 with the addition of a weighting term (w_D), which is

$$w_D = \begin{cases} 0.5\dfrac{d}{d_p}dB_p, \text{ if } d \le d_p \\ 0.5\left(\dfrac{d-d_p}{d_{max}-d_p}\right)dB_p, \text{ if } d \le d_p \end{cases} . \tag{4.25}$$

The inclusion of the weighting term in (4.24) gives the final T-Sw merged maximum model form

$$\begin{aligned} B(d) &= w_D + m_{T-Sw1}d_p + b_{T-Sw1} + 0.5\left(m_{T-Sw1} + m_{T-Sw2}\right)\ldots \\ &\ldots + 0.5\left(m_{T-Sw2} - m_{T-Sw1}\right)\left(d-d_p\right)\tanh\left(\frac{d-d_p}{d_{max}\times\gamma_{trn}}\right) \end{aligned} . \tag{4.26}$$

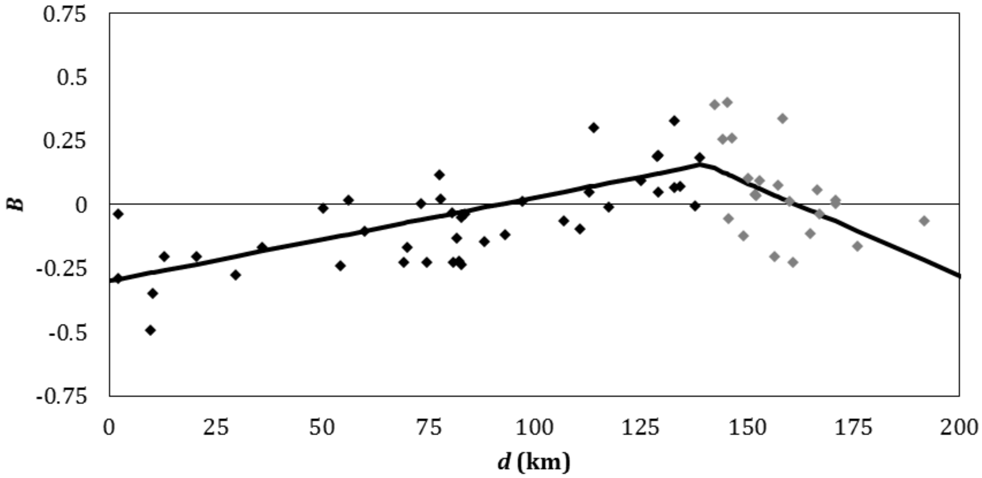


Fig. 4.12. Radar-gauge pair B_G data points for Group 1 (black diamonds) and Group 2 (gray diamonds), with the T-Sw merged model fit using (4.26) included (solid line). Data are from the KSHV radar domain for the 1-month period ending 31 January 2012.

The value of w_D at $d = 0$ is zero, the value of w_D at $d = d_p$ is $0.5dB_p$, and the value of w_D at $d = d_{max}$ is dB_p. In more general terms, the weighting term w_D allows the T-Sw merged maximum model to be a simple average of B_1 and B_2 at the intersection range. As distance increases from d_p in either direction, the T-Sw merged maximum model approaches the T-Sw single straight line from that group. Fig. 4.12 is the T-Sw merged maximum model form (4.26) using the partitioning of the same KSHV radar-gauge pair bias data points in Fig. 4.11.

IV.7. Adaptively Choosing conVPR Model Form

The development of the conVPR model uses radar-gauge pair bias data points, with choices made using the results of LOOC-V tests. A complete description will be provided for when it is appropriate to use the T-Sw single straight line model form or the T-Sw merged maximum model form. LOOC-V testing showed the T-Sw merged maximum model form had a higher overall MSE than the T-Sw single straight line model form for all partitions with a candidate B_{max} meeting the criteria of $m_{T\text{-}Sw1} > 0$ and $m_{T\text{-}Sw2} < 0$ (Table 4.1), which are the two most basic criteria for using the merged maximum model form. Given this result, there needs to be more stringent criteria for the merged maximum model form to be preferred to the single straight model form.

This section will describe the sequence of tests to determine metrics for when it is objectively desirable to use the T-Sw merged maximum model form, assuming a null hypothesis that the T-Sw single straight line model form provides a better fit. The results of each LOOC-V test will be applied to subsequent tests. The variables tested were arranged in a sequence that each test should intuitively have decreasing sensitivity to change in the variables. The general properties of the conVPR testing using radar-gauge pair bias data will be applied to Stage IV PoN precipitation bias data, described in the next section.

IV.7.a. Effect of Training Group MSE on Choice of conVPR Model Form

Testing was done for partitions with a candidate B_{max} to determine if the choice of model form should be dependent on the training group residual sum of squares. The assumption here is that with a sufficient sample size, a model fit based on the training group data should not deviate too far from a fit that includes the withheld data point. This test determines if the merged maximum model form should be used only if the RSS of the training group is lower than the RSS found using the single straight line model form. For each validation point and partition of the training data points satisfying the two slope criteria, the RSS of both the T-Sw single straight line model form (RSS_1) and the MSE of the T-Sw merged maximum model form (RSS_2) were computed.

The two possibilities tested were

1. always using the T-Sw merged maximum model form when $m_{T\text{-}Sw1} > 0$ and $m_{T\text{-}Sw2} < 0$ and,
2. when $m_{T\text{-}Sw1} > 0$ and $m_{T\text{-}Sw2} < 0$, use the T-Sw merged maximum model form if $RSS_1 > RSS_2$ and use the T-Sw single straight line model form if $RSS_s \leq RSS_c$.

The results of this LOOC-V test showed possibility 2 to have a lower overall MSE (Table 4.1), which led to a third criterion for a partition with $m_{T\text{-}Sw1} > 0$ and $m_{T\text{-}Sw2} < 0$ to be considered to have a candidate B_{max}, which is $RSS_1 > RSS_2$.

IV.7.b. Additional Attributes of T-Sw Merged Maximum Model Form Candidate Partitions

In addition to the three criteria already addressed $\{m_{T\text{-}Sw1} > 0;\ m_{T\text{-}Sw2} < 0;\ RSS_1 > RSS_2\}$, the following three metrics were used to define additional attributes for partitions with a candidate B_{max} and are

1. the absolute difference of the two model bias values B_1 and B_2 at d_p ($|dB_p|$),
2. the range of the candidate changepoint (d_p), and
3. the statistical significance of the difference between $m_{T\text{-}Sw1}$ and $m_{T\text{-}Sw2}$ (p).

Computation of $|dB_p|$ is done using (4.20) and d_p is determined by equidistant iteration (10 km) through a given set of B_G data points; the computation of the p-value for attribute 3 is a bit more complicated. The p-value is the probability of a student's t-test with a value t and df degrees of freedom and measures the statistical difference of the two partitioned slopes.

Computation of t requires knowledge of the standard error of the Group 1 (s_{m1}) slope, which has n_1 data points and slope $m_{T\text{-}Sw1}$, and the standard error of Group 2 (s_{m2}), which has n_2 data points and a slope $m_{T\text{-}Sw2}$. For a given group, the standard error of the is

$$s_m = \frac{\sqrt{\dfrac{\sum_{i=1}^{n}\left(B^i - B\left(d^i\right)\right)^2}{n-2}}}{\sqrt{\sum_{i=1}^{n}\left(d^i - \bar{d}\right)^2}} \tag{4.27}$$

and t is computed as

$$t = \frac{m_{T-Sw1} - m_{T-Sw2}}{\sqrt{s_{m1}^2 + s_{m2}^2}}. \tag{4.28}$$

The student's t-test of the difference between the two T-Sw slopes has a p-value found using a lengthy numerical solution and determined by the value of t and the number of degrees of freedom

$$df = n_1 + n_2 - 4. \tag{4.29}$$

The p-value has a range of possible values from (0, 1], with the statistical significance of the difference between $m_{T\text{-}Sw1}$ and $m_{T\text{-}Sw2}$ increasing as p-value approaches zero.

IV.7.c. LOOC-V Testing of Model Attributes

LOOC-V testing was done to determine the appropriateness of using the three attributes $|dB_p|$, d_p, and p-value as restrictive criteria for partitions containing a candidate B_{max}. To this point, there are only three restrictive criteria $\{m_{T\text{-}Sw1} > 0$; $m_{T\text{-}Sw2} < 0$; and $RSS_s > RSS_c\}$ for a candidate B_{max} in a partition. The goal of the testing was to determine the appropriate threshold values for

1. the maximum value of $|dB_p|$ (dB_{max}),
2. the minimum value of d_p (d_{pmin}),
3. the maximum value of d_p (d_{pmax}), and
4. the maximum p-value (p_{max}),

for which using the T-Sw merged maximum model form is a better fit than using the T-Sw single straight line model form when the three established criteria for a candidate B_{max} were met. For each attribute, several different values were tested (Table 4.3) and each LOOC-V test was done independently of the other attributes and in the order listed. The attribute threshold value j with the lowest overall MSE (MSE^j) was considered to be superior to all other threshold values. The tests were conducted in the order listed and the results of each test were applied to subsequent tests.

Table 4.3. LOOC-V test values for different thresholds of T-Sw merged maximum model attributes and the one chosen having the smallest MSE.

Parameter	**Minimum**	**Interval**	**Maximum**	**Smallest MSE**
dB_{max}	$dB_{max} \leq 0.05$	0.05	$dB_{max} \leq 5.00$	$dB_{max} \leq 1.70$
d_{pmin}	$d_{pmin} > 0$ km	10 km	$d_{pmin} > 340$ km	$d_{pmin} > 20$ km
d_{pmax}	$d_{pmax} \leq 10$ km	10 km	$d_{pmax} \leq 350$ km	$d_{pmax} \leq 310$ km
p_{max}	$p_{max} \leq 0.05$	0.05	$p_{max} \leq 1.00$	$p_{max} \leq 1.00$

IV.7.c.1) Finding an Ideal Maximum Threshold Value for $|dB_p|$

The first attribute tests seeks to find the maximum value of $|dB_p|$ a partition may have to consider the conVPR model form preferable to the T-Sw single straight line model form. For a given radar-month and a partition with a candidate B_{max}, found using the training data for validation point i, the value of $|dB_p|$ is computed. For each attribute threshold value j with a dB_{max} less than $|dB_p|$, the squared error of the T-Sw single straight line model (SE_s) at point i was added to SSE$_j$. For each attribute threshold value j with a dB_{max} greater than or equal to $|dB_p|$, the squared error of the T-Sw merged maximum model (SE_c) at point i

was added to SSE_j. The overall LOOC-V testing indicated that $dB_{max} = 1.70$ was an ideal threshold value and use of the T-Sw merged maximum model should be limited to partitions with $|dB_p| \leq 1.70$.

Table 4.4. Summary of criteria necessary for a partition of data to have a candidate B_{max}.

Criteria	**Restriction**
1	$m_{T\text{-}Sw1} > 0$
2	$m_{T\text{-}Sw2} < 0$
3	$RSS_1 > RSS_2$
4	$dB_{max} \leq 1.70$
5	$d_{pmin} > 20$ km
6	$d_{pmax} \leq 310$ km
7	$p_{max} \leq 1.00$

IV.7.c.2) Finding Ideal Values for d_p and for p-value

The testing for minimum and maximum values of d_p and for a maximum p-value was done for partitions restricted to $m_{T\text{-}Sw1} > 0$, $m_{T\text{-}Sw2} < 0$, and $RSS_s > RSS_c$, and $|dB_p| \leq dB_{max}$. The LOOC-V testing for the minimum threshold value of d_p indicated $d_{pmin} = 20$ km was ideal when looking at partitions every 10 km and the maximum threshold value for d_p was found to be $d_{pmax} = 310$ km. Therefore, the use of the T-Sw merged maximum model should be limited to the interval 20 km $\leq d_p \leq$ 310 km in addition to the four aforementioned restrictions. The testing for a maximum threshold p-value found there should be no restriction as $p_{max} = 1.0$, given that the range of possible p-values is (0, 1]. A summary of the criteria for a partition at range d_p to be considered to have a candidate B_{max} is found in

Table 4.4; if these all these criteria are not met, a T-Sw single straight line model form will be used as a fit to the bias data points in the conVPR model.

IV.7.d. Choosing Single Bias Maximum in Merged Maximum Model Form

Since the VPR maximum in the MK2011 model form only has a single range d_{bb}, LOOC-V testing determined an appropriate method for choosing which partition among a set of candidate should be chosen. Given a set of radar-gauge pair bias grid points, the values of $B(d_p)$, $|dB_p|$, MSE_2, and the p-value were computed for each partition considered to have a candidate B_{max} based on the criteria of Table 4.5. LOOC-V testing for the choice of partition with range $d_\cap$ and chosen bias maximum $B_{\max}$ looked at four different possibilities, each one based on the aforementioned metrics. The possible choices were

1. the partition with the largest candidate B_{max},
2. the partition with the smallest value of $|dB_p|$,
3. the partition with the smallest RSS_2, and
4. the partition with the smallest p-value.

The results of this LOOC-V testing procedure showed that choice 3, using the candidate partition with the RSS_2 value equal RSS_{min} had the lowest overall MSE (Table 4.5).

Table 4.5. LOOC-V test MSE for choosing a single B_{max} among a set of candidates meeting all the required criteria for consideration.

Largest B_{max}	Smallest $\lvert dB_p\rvert$	**Smallest RSS_2*	Smallest p-value
0.2049	0.2048	0.2043	0.2051

IV.8. The conVPR Model Algorithm

The LOOC-V testing in the previous section provided objective indicators for determining which conVPR model choice was appropriate given a set of radar-gauge pair bias data points. The focus in this section shifts to the actual algorithm for determining the conVPR model form and the parameters of the model form using all the available data for a given radar domain, ending date *t*, and accumulation period *a*.

IV.8.a. Radar-Gauge Pair Bias Data Points

For a set of radar-gauge pair bias data, criteria have been established to objectively determine which conVPR model forms provides the best fit for the given data points; the T-SW single straight line or the T-Sw merged maximum model form. Fig. 4.13 is a flowchart of the algorithm for objectively determining which conVPR model form should be chosen. The key to the decision is whether or not at least one partition exists that meets the all the criteria (Table 4.4) necessary to have a candidate B_{max}.

The model fitting algorithm is based largely on the results of the previous LOOC-V test of narrowing down a set of B_{max} candidates to a single choice. The final choice of a T-Sw merged maximum model form means the selected partition must have an RSS_2 value (RSS_{min}) both less than the all the other candidate partitions and the T-SW single straight line model fit RSS_1. Therefore, before iteration through all the possible partitions, the initial value of $RSS_{min} = RSS_1$. When iterating through each possible value of d_p for the grid points for a given radar-month, the T-Sw merged maximum parameter $d_\cap$ is stored only for partitions when the value of RSS_2 is less than that of RSS_{min} and the criteria listed in Table 4.5 are met (Fig. 4.13). If there is not a candidate partition meeting the criteria, the conVPR model is represented by T-Sw single straight line with slope $m_{T\text{-}Sw}$.

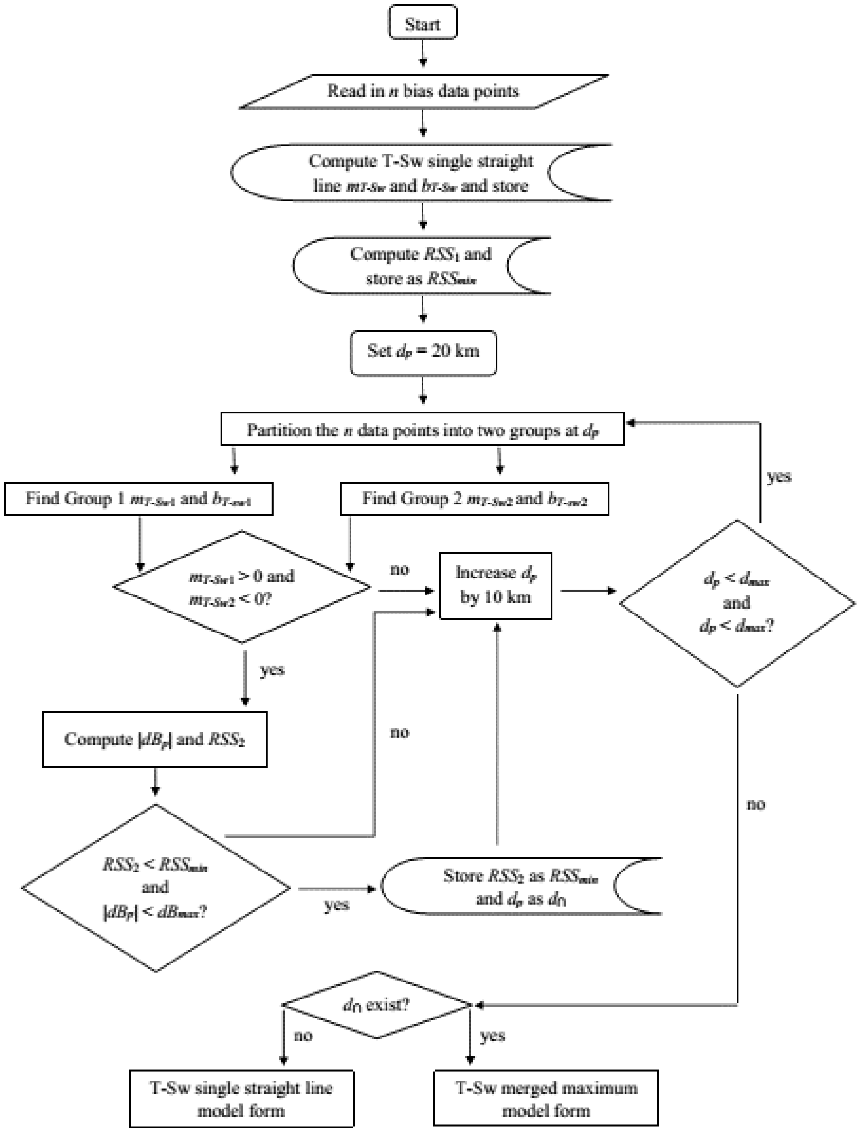


FIG. 4.13. Flowchart of the conVPR algorithm to determine the appropriate model form for a set of B_G data points.

IV.8.b. Stage IV PoN Precipitation Bias Data Points

The Stage IV PoN precipitation bias data conVPR model fitting algorithm is more straightforward than the radar-gauge pair bias model fitting algorithm. As opposed to the random and irregular spatial coverage of gauges, the Stage IV data have complete spatial coverage in a given radar domain. Consequently, a collection of Stage IV PoN precipitation bias data points for a given radar domain can be approximated to provide a continuous estimate of the bias as a function of range for all azimuths. A caveat to this approximation is that azimuthal coverage is not complete for all ranges because all radar domains are polygons with sharp edges.

Despite this caveat, we will move forward with the idea that the Stage IV PoN precipitation bias data points approximates a true representation of the range-dependent bias for a given radar-month. Therefore, the only restrictions for using the T-Sw merged maximum model form is that $m_{T\text{-}Sw1} > 0$ and $m_{T\text{-}Sw1} < 0$ and $RSS_1 > RSS_2$. If there is at least one partition with a candidate B_{max}, the partition having the smallest RSS_2 value is chosen for the T-Sw merged maximum model form with changepoint $d_{\cap}$. As with the radar-gauge pair bias data points, if $RSS_1 = RSS_{min}$ after examining all candidate partitions, the conVPR model will use the T-SW single straight line model best-fit with slope $m_{T\text{-}Sw}$.

The Stage IV PoN bias data points for the one-month period ending 31 January 2012 at the KSHV radar domain shows a clear B_{max}, which is captured by the T-Sw merged maximum model form (Fig. 4.14). Though not nearly as evident, the radar-gauge pair biases for the KSHV December 2012 radar-month have a nearly identical conVPR parameters for the best model fit. Agreement of the conVPR model parameters between the two bias data types indicates a high degree of confidence in the range dependence of the bias. The KSHV is a radar domain known to suffer from systematic range-dependent biases, which appear even to the naked eye when looking at the 36-month PoN precipitation ending 31 December 2012 (Fig. 4.15), so the high degree of confidence in the model fits (Fig 4.14) is visually justified.

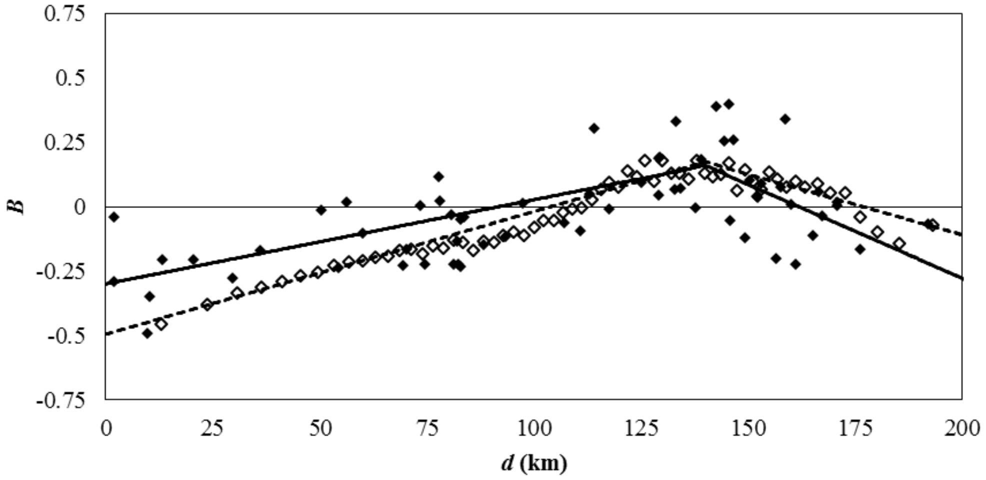


Fig. 4.14. Radar-gauge pair (solid diamonds) and Stage IV PoN precipitation (open diamonds) bias data points. The T-Sw merged model fit is included for the B_G data (solid line) and the B_N data (dashed line) from the KSHV radar domain for the 1-month period ending 31 January 2012.

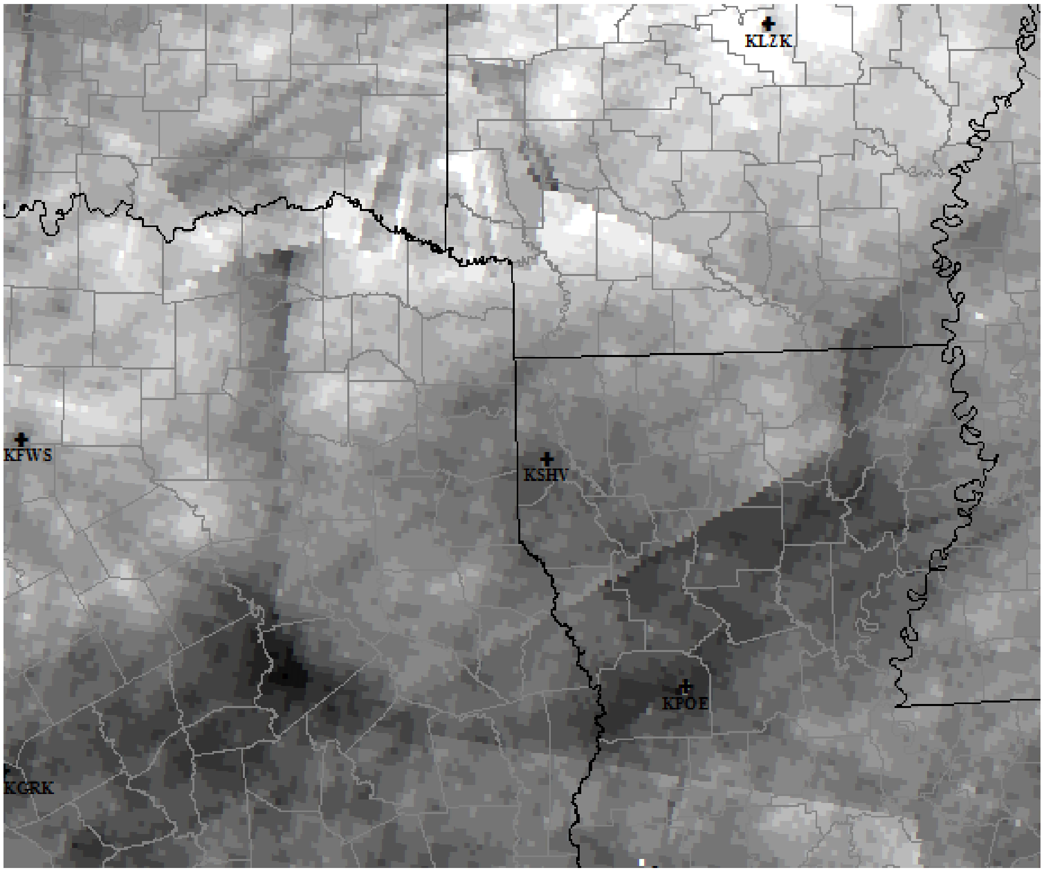


Fig. 4.15. Stage IV PoN precipitation for the 36-month ending 31 January 2012, with radar locations denoted by the white crosshairs. The minimum value (black) is *PoN* = 50% and the maximum value (white) is *PoN* = 125%.

IV.9. Combining Bias Data Point Types into Single conVPR Model

IV.9.a. Overview

In this section, the conVPR models found individually for each bias data type are combined into a single conVPR model (combi-conVPR model). At each HRAP grid cell, the resultant model bias value will be used to correct the Stage IV P_1 precipitation value. Like the conVPR model, the combi-conVPR model can either have a T-Sw single straight line or a T-Sw merged maximum model form. Rather than combining the radar-gauge pair and Stage IV PoN bias data conVPR models, the combi-conVPR model initially fuses only a single T-Sw estimated parameter from each of the best model fits. For the T-Sw single straight line model form, the radar-gauge pair bias $m_{T\text{-}Sw}$ (m_G) and the Stage IV PoN precipitation bias $m_{T\text{-}Sw}$ (m_N) are fused into a single $m_{T\text{-}Sw}$ (m). If the choice is the T-Sw merged maximum model form, the $d_\cap$ from the radar-gauge pair bias (d_G) and the Stage IV PoN precipitation (d_N) model fits are fused into a single intersection parameter.

For a given radar domain, ending date t, and accumulation period a, the first step of the combi-conVPR model is determining the conVPR model form that best represents each available bias data type. If there is agreement in the best-fit radar-gauge pair and Stage IV PoN precipitation model forms, the combi-conVPR model will take the same form; the relative weight of the parameters from each conVPR model is proportional to the statistical confidence of the parameter values. If the best-fit model form differs between the two bias data types, the choice of the combi-conVPR model is the one with more confidence in the parameter estimation.

For both the T-Sw single straight line form and the T-Sw merged maximum forms of the conVPR model, the level of confidence in the model fit increases as uncertainty in the parameter estimation decreases. The confidence for a given conVPR model fit will be assessed solely by the level of uncertainty (σ^2) in the T-Sw slope parameters. This is because the T-Sw intercept parameters for the two bias data types are not independent. Each of the Stage IV PoN precipitation bias data points were adjusted by amplitude A so that the median value of B_N is equal to the median value of B_G. Because of this, the relative confidence of the T-Sw intercept values for radar-gauge pair and Stage IV PoN precipitation conVPR model are assumed to be equal.

IV.9.b. Radar-Gauge Pair Bias T-Sw Slope Parameter Uncertainty

For a given set of radar-gauge pair bias data points, the T-Sw slope parameter is the median of n_m slopes, each computed using a unique pair of (d, B) points. A traditional measure of uncertainty for the T-Sw slope estimate $m_{T\text{-}Sw}$ is variance (var_m), which is formulated as

$$var_m = \frac{1}{n_m} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(m^{ij} - m_{T-Sw} \right)^2 , \tag{4.30}$$

where range $d^i < d^j$ for each computed slope m^{ij}. However, (4.30) fails to account for the weighting term w^{ij} used in the Sievers (1978) T-Sw method of determining the slope parameter. A more appropriate measure of uncertainty in the T-Sw slope parameter estimation for a set of radar-gauge pair biases ($\sigma_G{}^2$) is

$$\sigma_G{}^2 = \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(m^{ij} - m_{T-Sw} \right)^2}{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} w^{ij}} , \tag{4.31}$$

and is used to compute the uncertainty in the conVPR T-Sw single straight line model. Determining $\sigma_G{}^2$ for the T-Sw merged maximum form merges the uncertainty values from the Group 1 ($\sigma_{G2}{}^2$) and Group 2 ($\sigma_{G1}{}^2$) slope estimation parameters, each computed separately using (4.31), with

$$\sigma_G{}^2 = \frac{\sigma_{m1}{}^2 d_{\cap} + \sigma_{m2}{}^2 \left(d_{max} - d_{\cap}\right)}{d_{max}}. \tag{4.32}$$

IV.9.c. Stage IV PoN Precipitation Bias T-Sw Slope Parameter Uncertainty

Unfortunately, determining the uncertainty of the Stage IV PoN precipitation T-Sw slope parameter estimation ($\sigma_N{}^2$) is not as straight-forward as the computation of $\sigma_G{}^2$. As with the radar-gauge pair bias conVPR model, the uncertainty is determined independently for the two slopes if the model has the T-Sw merged maximum form.

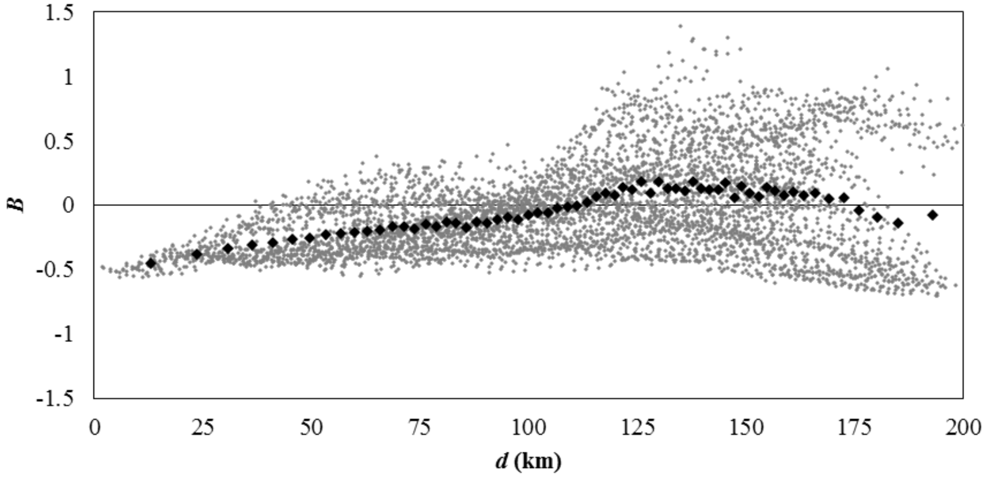


Fig. 4.16. B_N values (black diamonds) and Stage IV PoN precipitation values for all HRAP grid cells (gray dots), both adjusted by amplitude factor A. Data are from the KSHV radar domain for the 1-month period ending 31 January 2012.

Each of the Stage IV PoN precipitation bias data points is average of several data points, so each B_N value has a its own variance σ_N. Fig. 4.16 shows the all the Stage IV PoN precipitation values, each adjusted by amplitude A, and the Stage IV PoN precipitation bias data points for the KSHV December 2012 radar-month. The computation of σ_N^2 would be straight-forward if we used all the HRAP grid cells to compute (4.31); however the number of potential T-Sw slopes n_m is

$$n_m = \sum_{i=1}^{n_N - 1} i \, , \tag{4.33}$$

where n_N is the number of Stage IV PoN precipitation bias grid points. In Fig. 4.16, the number of data points used in the Group1 T-Sw slope estimate $m_{T\text{-}Sw1}$ was manageable (n_{N1} = 45), which leads to n_m = 990. However, if the Group 1 T-Sw slope was computed using all 2,617 grid cells in the KSHV domain with a d value less than the changepoint range $d_\cap$ = 140 km, there would be $n_m = 5.79 \times 10^6$ possible slopes.

Because of the computational expense is too large, the uncertainty σ_N^2 in the T-Sw slope parameter for a given set of B_N data will be estimated using the two slope estimates and the uncertainty in the gauge data. The estimation begins by assuming that both the radar-gauge pair slope (m_G) and the Stage IV PoN slope (m_N) are estimates of the expected slope value <m>, each with an error term ε, such that

$$\langle m \rangle = m_G - \varepsilon_G = m_N - \varepsilon_N \, . \tag{4.34}$$

Rearranging the terms in (4.34) to solve for ε_N and squaring each side gives you

$$(m_N - m_G)^2 = (\varepsilon_N - \varepsilon_G)^2 . \tag{4.35}$$

The expected value of $(m_N - m_G)^2$ is the variance of the quantity $(m_N - m_G)$, which will be denoted as σ_d^2. Similarly, $<\varepsilon_N^2> = \sigma_N^2$, $<\varepsilon_G^2> = \sigma_G^2$, and the expected value of $\varepsilon_N\varepsilon_G$ is zero assuming that the estimates of m_G and m_N are independent. Using the expected values of the quantities in (4.35) and rearranging the terms gives an estimate for the uncertainty σ_N^2 that is

$$\sigma_N^2 = \sigma_d^2 - \sigma_G^2 . \tag{4.36}$$

The quantity $\sigma_d^2 - \sigma_G^2$ for a given radar-month should not be used as a direct estimate of σ_N^2. Instead, (4.36) should be limited to a general approximation, such as estimating $<\sigma_N^2>$ by using $<\sigma_d^2 - \sigma_G^2>$.

IV.9.d. Function for Uncertainty of Stage IV PoN Precipitation Slope Estimate

Given the approximation in (4.36), it was determined that σ_N^2 for each given radar-month should be estimated as a parametric function of the Stage IV PoN precipitation bias data points. The two parameters of the σ_N^2 estimation function are

1. the RSS of the of the B_N data points (a_σ) and
2. the magnitude of the slope estimate (b_σ),

so that the estimate of σ_N^2 is

$$\sigma_N^2 = f(a_\sigma) + f(b_\sigma) . \tag{4.37}$$

The determination of the form and parameters for the function $f(a_\sigma)$ was accomplished by plotting expected values of $\sigma_d^2 - \sigma_G^2$, used as an approximation of σ_N^2 in (4.36), as a function a_σ for all radar-months (Fig. 4.17). The values of $\sigma_d^2 - \sigma_G^2$ for all radar-months were ordered by the independent variable a_σ and split into 20 groups. Fig. 4.17 shows the plot of the expected values of $\sigma_d^2 - \sigma_G^2$ versus the expected values of a_σ from each group. For both quantities, the expected value for each group was represented by the median value. A power function fit to the data points with the form

$$f(a_\sigma) = Aa_\sigma^{\ B}. \tag{4.38}$$

and constants $A = 3.87$ and $B = 0.57$ with an $R^2 = 0.9293$.

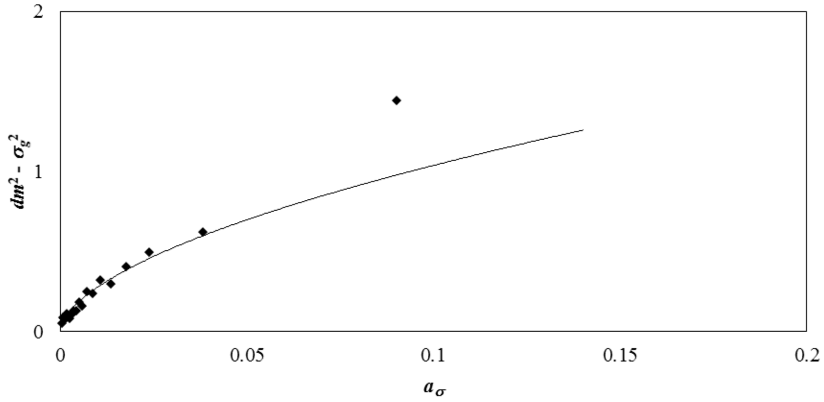


Fig. 4.17. The approximation for σ_N^2 ($\sigma_d^2 - \sigma_G^2$, black diamonds) as a function of a_σ for all radar-months in the testing period. The data were ordered by a_σ and split into 20 groups, with each point representing the median from each group for both a_σ and $\sigma_d^2 - \sigma_G^2$. The solid line is a power function that is fit to the data points.

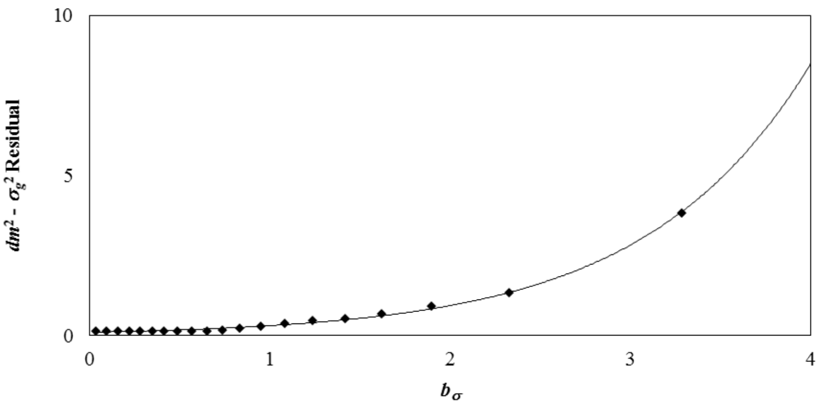


Fig. 4.18. The residuals of $\sigma_d^2 - \sigma_G^2$ from the power function in Fig. 4.17 as a function of b_σ for all radar-months in the testing period. The data were ordered by b_σ and split into 20 groups, with each point representing the median from each group for both b_σ and the residuals. The solid line is an exponential function that is fit to the data points.

However, the data points in Fig. 4.17 show heteroscedasticity, so the residuals of $\sigma_d^2 - \sigma_G^2$ to the model fit $f(a_\sigma)$ were plotted as a function of the magnitude of the slope estimate (b_σ). Fig. 4.18 shows the residuals of $\sigma_d^2 - \sigma_G^2$ for all radar-months ordered by b_σ and split into 20 groups to reduce the noise in the plot. An exponential function fit to the data points with the form

$$f\left(b_\sigma\right) = C\exp\left(Db_\sigma\right) \tag{4.39}$$

and constants $C = 0.10$ and $D = 1.10$ with an $R^2 = 0.9586$. Combining (4.38) and (4.39), σ_N^2 can be represented as a combination of the functions $f(a_\sigma)$ and $f(b_\sigma)$, such that

$$\sigma_N{}^2 = Aa_\sigma^B + C\exp\left(Db_\sigma\right). \tag{4.40}$$

Given that the constant C is the value the exponential function when b_σ is zero (Fig. 4.18), the final form of the σ_N^2 equation (4.41) subtracts constant C and is

$$\sigma_N^{\;2} = Aa_\sigma^{\;B} + C\exp\left(Db_\sigma\right) - C\,. \tag{4.41}$$

If the conVPR model for a set of Stage IV PoN precipitation bias data is the merged maximum model form, the two independent estimates of uncertainty (σ_{N1}^2 and σ_{N2}^2) are combined into a single σ_N^2 estimate by using the values of $d_\cap$ and d_{max} in (4.32). Now that there is a satisfactory method for determining σ_N^2, we can now turn our attention to fusing the two parameters from each individual conVPR model into a single combi-conVPR model parameter.

IV.9.e. Combi-conVPR for T-Sw Single Straight Line Model Form

Fusion of the two T-Sw single straight line slope parameters (m_G and m_N) into a single combi-conVPR slope parameter (m) is given by

$$m = \frac{m_G\sigma_N^{\;2} + m_N\sigma_G^{\;2}}{\sigma_N^{\;2} + \sigma_G^{\;2}}\,, \tag{4.42}$$

regardless of the appropriate fit for each bias data type. After finding m using (4.42), the intercept b^i is computed for each of the n radar-gauge pair bias data points and the n Stage IV PoN precipitation bias data points (4.15). The combi-conVPR model intercept is the median of all the $2n$ computed b^i values.

IV.9.f. Combi-conVPR for T-Sw Merged Maximum Model Form

Rather than fusing the Group 1 and Group 2 slope parameters in the conVPR T-Sw merged maximum model form, the choice was made to fuse the changepoints $d_\cap$ from the models created using each bias data type. No fusion is required if one of the bias data types does not have a single changepoint $d_\cap$ with a B_{max}, and the changepoint $d_\cap$ of the other bias data type will be chosen. In the case that the best-fit conVPR model forms are both the T-Sw merged maximum form, the fusion of the $d_\cap$ parameters (d_G and d_N) is given by

$$d_\cap = \frac{d_G \sigma_N{}^2 + d_N \sigma_G{}^2}{\sigma_N{}^2 + \sigma_G{}^2}. \qquad (4.43)$$

Once the combi-conVPR $d_\cap$ range is determined, both the radar-gauge pair and *the* B_N data points are partitioned into two groups. For each group,

1. the Sievers (1978) T-Sw method is used to independently estimate m_G and m_N,
2. determine the measure of uncertainty for m_G (σ_G) and m_N (σ_N),
3. fuse m_G and m_N into a single slope parameter value m using (39), and
4. use the fused slope value m to find the combi-conVPR intercept b.

Within each group the intercept is determined by fitting the slope value m to each and every radar-gauge pair bias and Stage IV PoN bias data point. Using the two slope parameters (m_1 and m_2) and two intercept parameters (b_1 and b_2), the difference $|dB_\cap|$ between the two combi-conVPR model estimates at $d_\cap$ is computed (4.20). The final combi-conVPR T-Sw maximum merged model form merges the two individual single straight lines into a single continuous function at $d_\cap$ with the hyperbolic transition function

(4.23). The combi-conVPR model fit at any range *d* is given by (4.26), and the fit for the KSHV January 2012 radar-month is shown in Fig. 4.19.

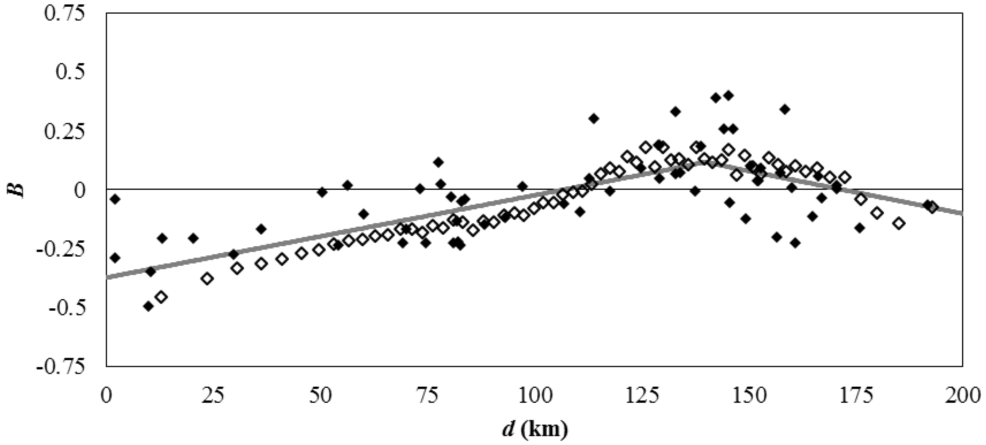


FIG. 4.19. Same as Fig. 4.14, but with the combi-conVPR model fit (gray line) included instead of the two independent model fits.

IV.9.g. Application of Combi-conVPR Model to Multi-Month Periods

The combi-conVPR model of range-dependent and mean field biases for a given radar domain are used to correct the all Stage IV PoN precipitation values for the specific time period of the data from which the model was built. For example, correction for range-dependent and mean field biases for 12-month PoN precipitation will be based on the model fit using 12-month B_G and B_N data. This is different than the beam blockage corrections of 1-month precipitation totals that were aggregated to multi-month precipitation totals. Beam blockage is temporally invariant, but the mechanisms causing range-dependent biases differ in magnitude and variability on different time scales.

The combi-conVPR model for an accumulation period *a* of length *n* days is computed using (4.5), and at each gauge *i*, using only days with non-missing gauge data. As the length *a* of accumulation periods increases, the percentage of available gauges (defined here as have at least a single non-missing gauge value within the period) with missing data increases. Fig. 4.20 shows the percentage of available COOP and WBAN gauges surpassing different thresholds of data availability, and even for smaller values of *a*, there are not many gauges with a complete time series of daily values. Therefore, it would be naïve to use only temporally complete stations to determine the combi-conVPR model fit for a radar, accumulation period, and ending date, particularly if the value of *a* large.

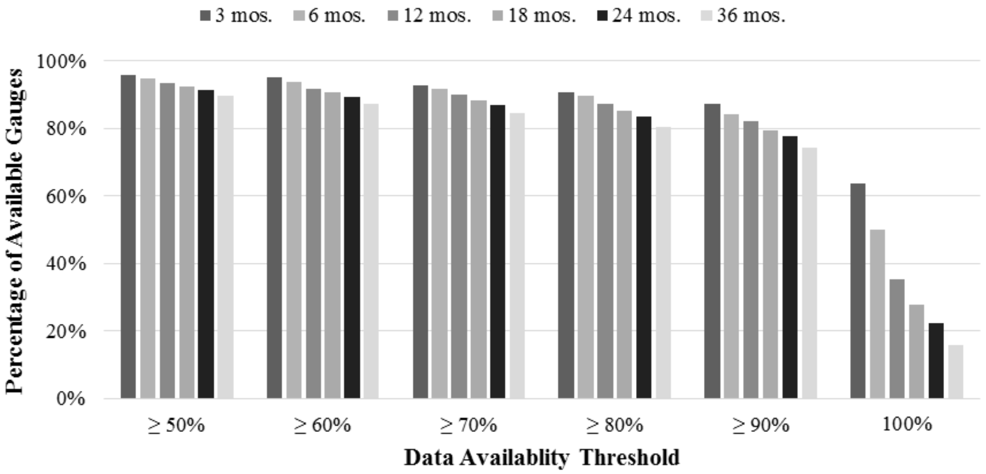


FIG. 4.20. The percentage of available gauges meeting different data availability threshold criteria for different values of *a*; each value is a summary over all available gauges and possible ending dates.

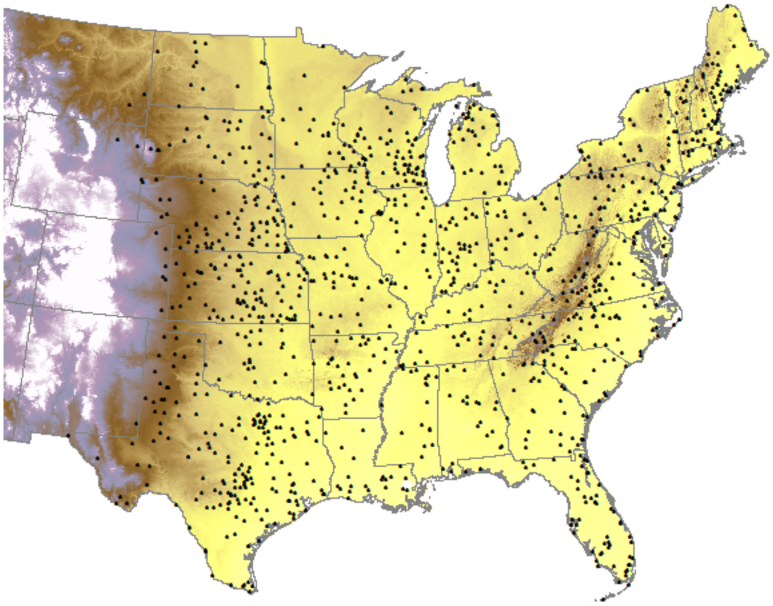
A LOOC-V testing procedure looked at the effects that different data availability thresholds had on the robustness of the combi-conVPR fit for several different accumulation periods (Table 4.6). The MSE values from the LOOC-V test indicate that at least 90% data availability for radar-gauge pair biases in combi-conVPR models for all time scales. The network of gauges with 100% data completeness is sparse for 36 months (see Fig. 4.21a), whereas the number of gauges increases significantly for lower thresholds (see Fig. 4.21b).

Table 4.6. LOOC-V test MSE for choosing an optimal data availability threshold in determining the combi-conVPR model fit, for different accumulation periods (*a*).

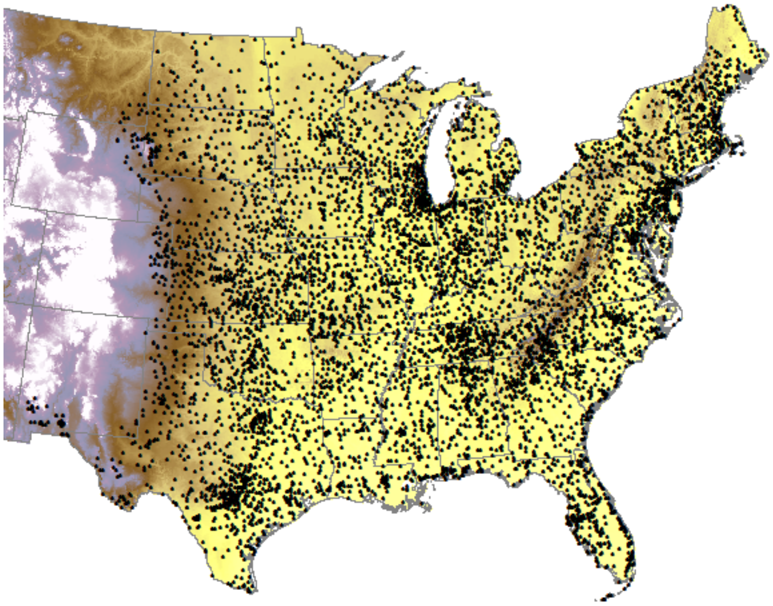
a	**Data Availability Thresholds**					
	≥ 50%	**≥ 60%**	**≥ 70%**	**≥ 80%**	**≥ 90%**	**100%**
1	0.22088 (6)	0.22084 (5)	0.22070 (4)	0.22055 (3)	***0.22043 (1)***	0.22046 (2)
6	0.07252 (4)	0.07253 (5)	0.07250 (2)	0.07252 (3)	***0.07246 (1)***	0.07307 (6)
12	0.01710 (3)	0.01710 (2)	0.01712 (5)	0.01710 (4)	***0.01709 (1)***	0.01809 (6)
18	0.01489 (4)	0.01489 (3)	0.01489 (5)	***0.01486 (1)***	***0.01486 (1)***	0.01623 (6)
24	0.01367 (4)	0.01367 (3)	0.01369 (5)	***0.01364 (1)***	***0.01364 (1)***	0.01519 (6)
36	0.01287 (4)	0.01286 (3)	0.01287 (5)	0.01280 (2)	***0.01278 (1)***	0.01506 (6)

Once the combi-conVPR model bias (B_3) is determined, computation of a Stage IV precipitation value P_3 corrected for beam blockage, mean-field bias, and range-dependent bias at an HRAP grid cell is straightforward and is

$$P_3 = \frac{P_1}{B_3\left(d,\theta,r\right)+1}. \tag{4.44}$$



(a)



(b)

Fig. 4.21. The available gauges meeting the data availability thresholds of (a) 100% availability and (b) ≥ 90% availability for the period ending 31 December 2012.

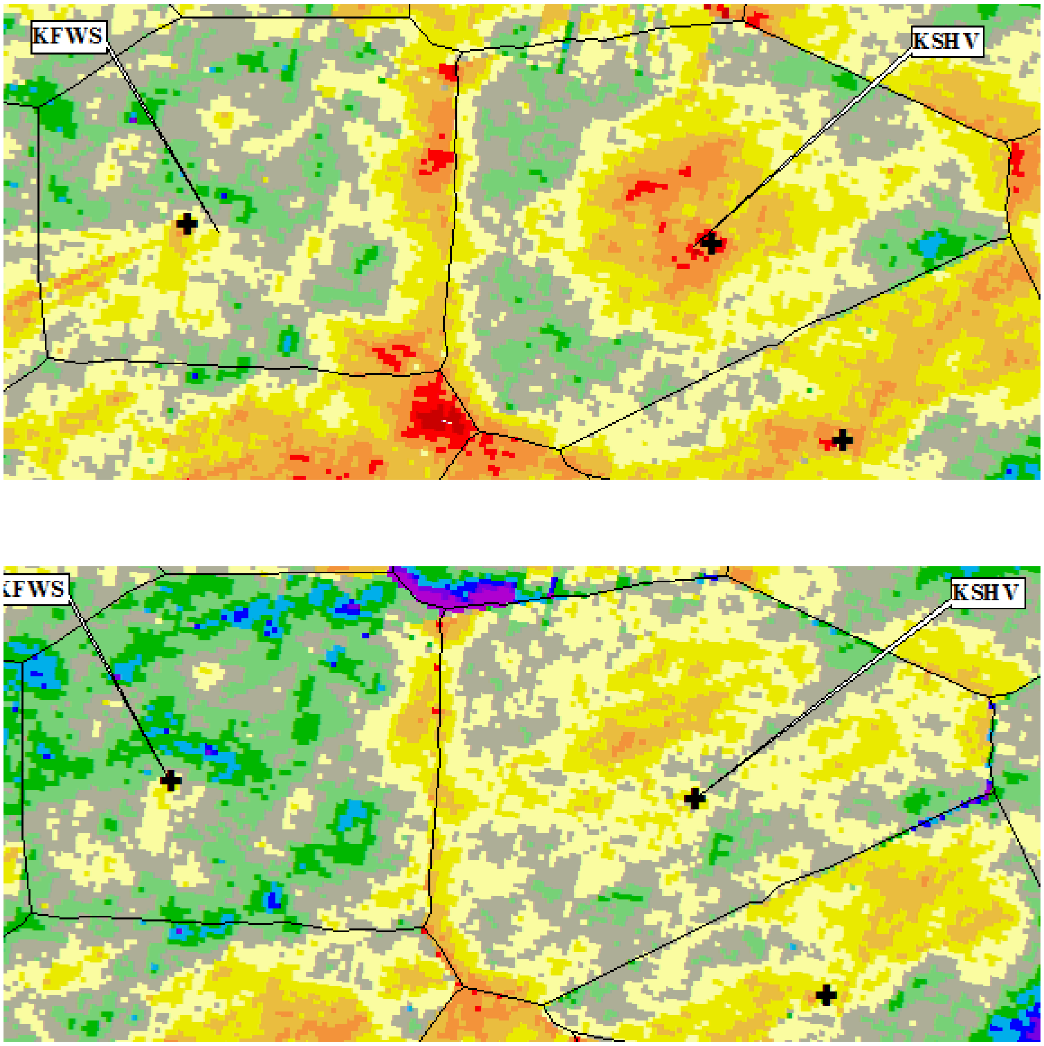
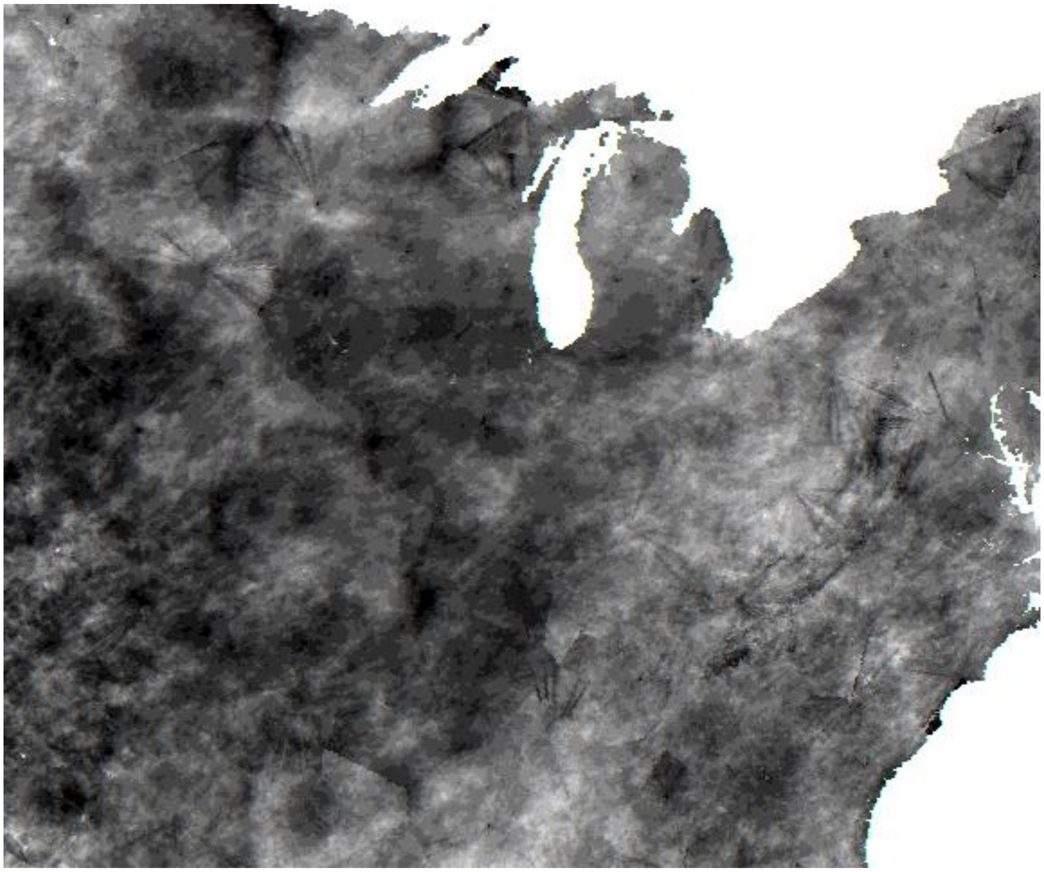


Fig. 4.22. The Stage IV 36-month PoN precipitation for the period ending on 31 December 2012, with a focus on the KFWS and KSHV radar domains; data shown are (top) uncorrected and (bottom) corrected for beam blockage, mean-field biases, and range-dependent biases.

The changes in the Stage IV estimates due to the corrections for biases are shown in Fig. 4.22 for the 36-month period winding on 21 December 2012, highlighting the Dallas/Fort Worth, TX (KFWS) and KSHV radar domains. The primary improvements to the KFWS radar domain are related to the removal of beam blockage to the south of the radar location, whereas the KSHV improvements are related to corrections for range-dependent biases, i.e., no more donut hole near the radar location.



(a)

Fig. 4.23. The Stage IV 12-month PoN precipitation for the period ending on 31 December 2012 that is (a) uncorrected and (b) corrected for beam blockage, mean-field biases, and range-dependent biases. The minimum value (black) is *PoN* = 50% and the maximum value (white) is *PoN* = 150% in both maps.

(b)

Fig. 4.23. Continued.

A broader spatial view is provided by Fig. 4.23, which looks at the difference between the initial Stage IV 12-month precipitation field and the bias-corrected field for same ending date as Fig. 4.22. The improvements due to corrections for beam blockage are more obvious than those for the bias correction procedures in this chapter. However, closer inspection shows a reduction of the range-dependent bias artifacts in the Stage IV estimates for several of the radar domains in the northwest quadrant of Fig. 4.23b.

IV.9.h. Combi-conVPR Model Verification

A straight-forward verification exercise was designed to ensure that the correction procedures for beam blockage, mean-field biases, and range-dependent biases are providing improved estimates relative to the original Stage IV precipitation data. The basic methodology of the verification procedure is to withhold 20% of the available radar-gauge pairs in a validation group and use the remaining 80% of the pairs as a training group for developing the Combi-ConVPR model at each radar. This verification procedure is similar to LOOC-V, except that 20% of the radar-gauge pairs are used for verification rather than a single pair. The bias-adjusted Stage IV precipitation estimate (P_3) is compared to the gauge value (G) at each of the withheld radar-gauge pairs and at each pair, and an RMSE value will be computed using (3.25). At each withheld pair, the RMSE values for the uncorrected Stage IV precipitation (P_0) and beam blockage-corrected Stage IV precipitation (P_1) estimates are also computed for comparison.

Evaluation of the errors in the Stage IV P_0, P_1, and P_3 data was done for four different four accumulation periods (a) and ending times (t), with $\mathbf{a}$ = {1 month; 4 months; 12 months; 36 months} and $\mathbf{t}$ = {January 2008; July 2009; April 2011; October 2012}. The procedure was run 10 different times for each possible permutation of (a, t). Each run has a unique set of randomly selected radar-gauge pairs in both the training and validation groups, with each pair belonging exclusively to either the training or validation group. The average RMSE over the 10 runs for each permutation of (a, t) for each of the Stage IV precipitation datasets is summarized in Table 4.7, which also includes the overall reduction in error of the P_1 and P_3 datasets relative to the P_0 dataset.

In all but three of the 16 (a, t) permutations, the overall RMSE was reduced by more than a third after corrections for beam blockage and mean field and range-dependent biases. Given that the errors in this procedure were evaluated at gauges not used in any of the bias corrections, it is reasonable that any randomly selected grid cell should have a similarly improved estimation of the true surface precipitation when using the P_3 value instead of the value of P_0. The evaluation procedure summarized in Table 4.7 provides empirical evidence in addition to the visual evidence (i.e., Fig. 4.23) that the three types of bias corrections provide improved Stage IV precipitation estimates at all time scales.

Table 4.7. Validation test RMSE values for the P_0, P_1, P_3 datasets for different accumulation periods (a) and ending dates (t).

		RMSE (mm)			**P_0 RMSE Reduction**	
t	**a**	**P_0**	**P_1**	**P_3**	**P_1**	**P_3**
January 2008	**1**	63.0	62.6	36.6	0.71%	41.91%
	4	189.7	182.8	110.4	3.64%	41.78%
	12	509.9	488.2	298.3	4.25%	41.49%
	36	1328.5	1267.2	849.0	4.61%	36.09%
July 2009	**1**	101.5	100.6	53.0	0.93%	47.78%
	4	257.4	253.5	152.3	1.53%	40.84%
	12	1026.6	1002.5	798.7	2.35%	22.20%
	36	1695.0	1602.9	1187.3	5.43%	29.95%
April 2011	**1**	89.3	88.4	48.3	0.98%	45.89%
	4	192.1	191.2	110.4	0.50%	42.53%
	12	467.6	456.7	232.2	2.32%	50.33%
	36	1568.1	1478.2	1105.3	5.74%	29.51%
October 2012	**1**	76.3	75.9	32.2	0.50%	57.72%
	4	211.5	209.8	111.6	0.78%	47.24%
	12	436.4	423.7	260.0	2.91%	40.42%
	36	1133.7	1087.7	671.3	4.06%	40.79%

CHAPTER V

TWO-DIMENSIONAL BIAS CORRECTIONS

V.1. Introduction

This chapter will build upon the prior bias correction procedures for beam blockage, mean-fields biases, and range-dependent biases and look to identify any remaining two-dimensional biases in the P_3 dataset. Of particular importance are spatial anisotropies in the P_3 data that lead to biases which are a function of the RFC and radar domain that a particular HRAP grid cell is located in. The original Stage IV P_0 dataset contains anisotropies related to the individual RFCs at which the analyses were produced, since each RFC produces an independent daily analysis and these anisotropies lead to residual biases in the P_3 product. Anisotropies attributed to the WSR-88Ds are due to the prior bias corrections in the P_3 data being determined independently at each radar domain. The most visible manifestation of these spatial anisotropy sources are discontinuities in analyzed P_3 Stage IV precipitation fields at boundaries of both radar domains and RFC areas of responsibility.

Additionally, the bias corrections of the previous chapters were done for a single dimension. In Chapter III, the beam blockage corrections used grid cells with approximately the same d value. The mean-field biases were corrected using a single constant that is a function only of the WSR-88D and the range-dependent bias corrections are solely a function of range at each radar.

The remaining anisotropies in the P_3 dataset, the corrections being done independently at each radar, and the one-dimensional nature of the bias corrections up to this point will be addressed in a final bias correction procedure, which will minimize any two-dimensional biases still existing in the two-dimensional Stage IV precipitation fields. Given the imitations of the previous bias corrections, is important that the bias corrections in this chapter are done

1. in two dimensions and
2. with no knowledge of the radar domains or RFC areas of responsibility.

From our conceptual equation of bias sources for radar precipitation estimates (1.1), the equation for Stage IV data biases in this chapter is

$$\frac{P_0}{P_T} = B(\Phi, \Lambda). \tag{5.1}$$

The Stage IV 1-month PoN precipitation field for December 2012 (Fig. 5.1) is an example of both sources of spatial anisotropies showing up prominently. It is possible to determine many of the radar domain and RFC boundaries (given in Fig. 3.1a) simply using the Stage IV PoN precipitation field without any prior knowledge.

In addition to the already discussed limitations, the bias corrections of the previous two chapters may have unintentionally accounted for the anisotropies related to RFC boundaries in the corrections for range-dependent biases. In Fig. 5.2, all of the biases in the North Central RFC (NCRFC) are biased low, whereas most of the biases in the

Missouri Basin RFC (MBRFC) are biased high. These regional bias differences can have a large impact on the individual radar scale (Fig. 5.2). Given these issues, even an optimal procedure for minimization of beam blockage, mean-field biases, and range-dependent biases at the individual radars will leave significant residual biases related to the RFC areas of responsibilities. In the following section, a two-dimensional bias adjustment procedure will adjust the bias data from the combi-conVPR model to account for two-dimensional anisotropies.

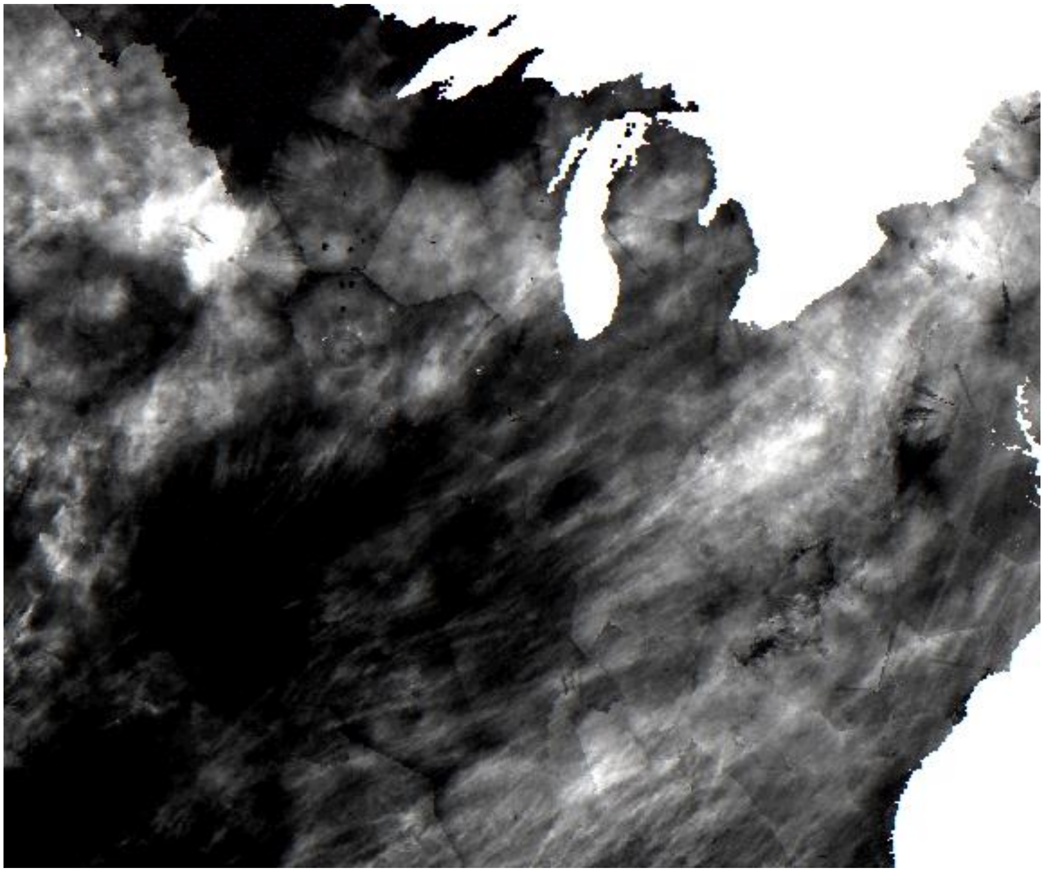


Fig. 5.1. Same as Fig. 3.1, but without radar locations or boundaries for radar domains.

(a)

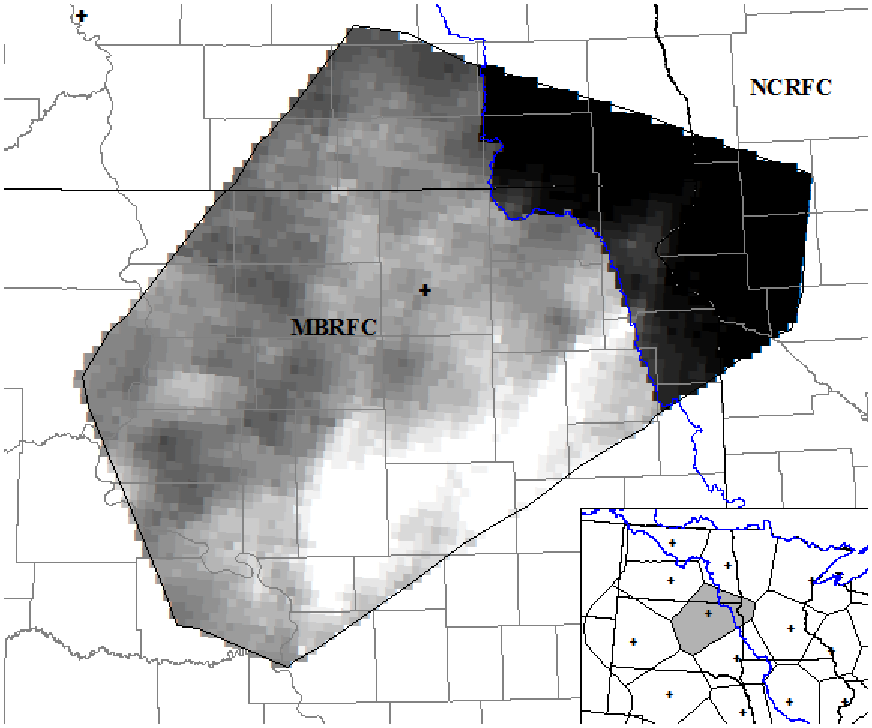
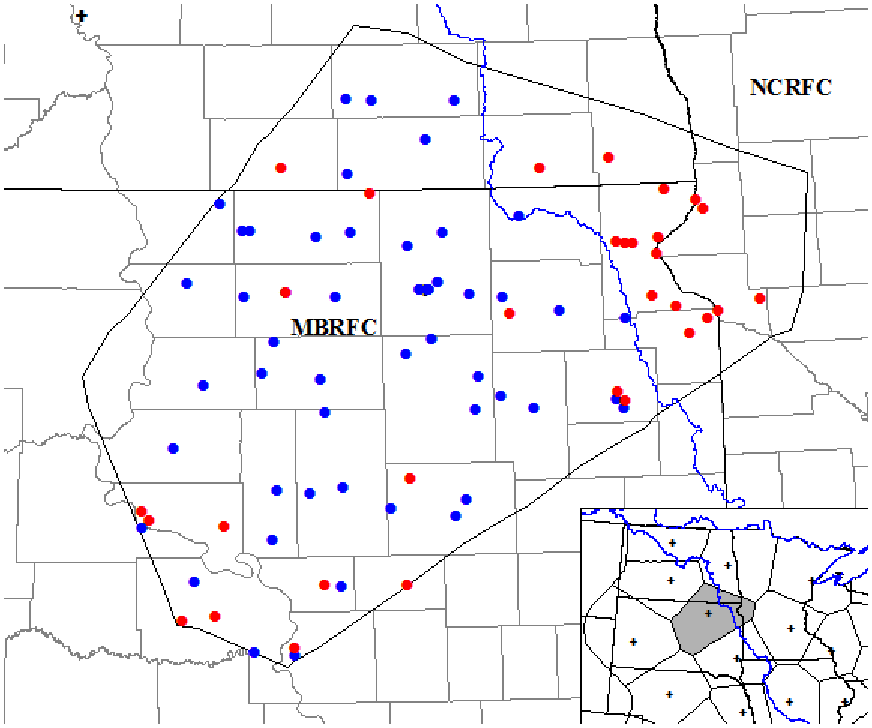
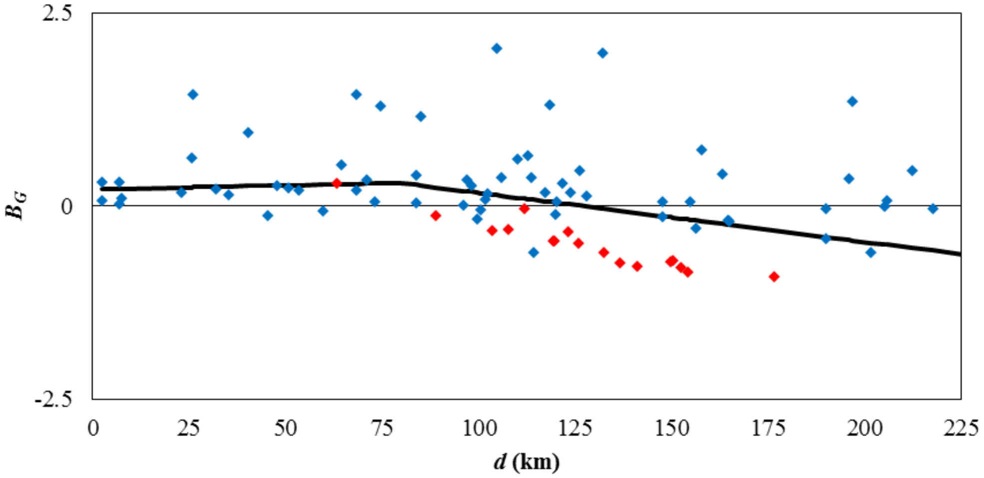


Fig. 5.2. (a) Stage IV 1-month PoN precipitation from December 2012 in the KABR radar domain with RFC boundary included; minimum value (black) is *PoN* = 50% and the maximum value (white) is *PoN* = 150%. (b) Spatial distribution of B_G data with values > 0 (blue) and values < 0 (red). (c) The same data points in (b) plotted as a function of range, divided into the NCRFC (red diamonds) and the MBRFC (blue diamonds).



(b)



(c)

Fig. 5.2. Continued.

V.2. Two-Dimensional Bias Adjustment Procedure

V.2.a. Overview

The two-dimensional bias adjustment procedure will utilize the combi-conVPR bias data, which are used to correct for beam blockage $B(d, \theta)$, mean field-biases $B_{M\text{-}F}$, and range-dependent biases $B(d)$ in the P_3 dataset. At each HRAP grid cell, the combi-conVPR model for the given radar domain determined a model bias value. The corrected Stage IV P_3 dataset, which used modeled biases to correct Stage IV P_1 precipitation estimates, was an improvement over the original Stage IV P_0 dataset uncorrected for any biases (results in Table 4.7). In the two-dimensional bias adjustment procedure, an assumption is that the combi-conVPR model bias value provides the best initial estimate accounting for the three known types of one-dimensional biases that appear in the Stage IV radar precipitation estimates. The combi-conVPR model has the desired spatial properties of

1. reducing the noise in the radar-gauge pair biases associated with the random errors in the gauge data and
2. estimating biases in the Stage IV precipitation dataset where gauge information is not available.

Even though combi-conVPR bias model reduces the overall noise of the radar-gauge pair biases, at any given HRAP grid cell containing a gauge, the computed radar-gauge pair bias (using 4.5) is the best estimate of the true bias in the Stage IV precipitation estimate. However, the combi-conVPR model bias values don't maintain information about the radar-gauge pair biases beyond the model construction. There may be two-dimensional patterns in the radar-gauge pair bias data points that go unnoticed in the one-dimensional bias correction steps. In other words, the bias values do not account for the possible two-dimensional spatial anisotropies since the combi-conVPR model is a function of only the distance *d* from the radar location in a given domain.

A desirable two-dimensional bias adjustment procedure should combine the information provided by the one-dimensional bias correction (combi-conVPR model data) and the two-dimensional spatial properties of the radar-gauge pair biases. For example, the grid cells in the union of the KABR radar domain and the North Central RFC (Fig. 5.2b) show a clear two-dimensional consistency of negative biases that is not apparent in the one-dimensional combi-conVPR bias model (Fig. 5.2c).

V.2.b. Data Assimilation Procedure

The approach for adjusting the biases resulting from the combi-conVPR model for two-dimensional anisotropies will be data assimilation, a technique in which "observations" are combined with a "first guess" model field (Kalnay 2003). At each time step, or "increment," a gridded first guess field is merged with information from irregularly spaced observations, which have been interpolated to the regularly spaced grid. Differences between the first guess field and the interpolated observation field are known as "observational increments" and are used to correct the first guess field.

For a given time *t* and accumulation period *a*, the two-dimensional bias adjustment procedure is done for a single increment, using the spatially complete combi-conVPR model bias field as a first guess field. The radar-gauge pair biases for the same values of *a* and *t* are computed using (4.5) are the "observations". At each available radar-gauge pair bias data point *i*, the observational increment (OI^i) is defined as the difference between the observation B_G and the first guess (B_3) and is computed as

$$OI^i = B_G^{\ i} - B_3^{\ i} . \tag{5.2}$$

The subscript "3" in B_3 refers to the combi-conVPR model bias value that accounts for three different types of biases (beam blockage, mean-field, and range-dependent). After the observational increments are computed at available each radar-gauge pair *i*, the OI data are interpolated onto the HRAP grid using ordinary Kriging analysis. The final step is to merge an analysis of the observational increments with the first guess field. The merging of the Kriged observational increment value $OI(\Phi_H, \Lambda_H)$ to the first guess bias field $B_3(\Phi_H, \Lambda_H)$ at each HRAP grid cell is

$$B_f(\Phi_H, \Lambda_H) = B_3(\Phi_H, \Lambda_H) + OI(\Phi_H, \Lambda_H). \tag{5.3}$$

This merged bias value B_f resulting from (5.3) will be used to correct the P_I value to produce a bias-corrected Stage IV precipitation estimate P_f.

V.2.c. Optimization of the Two-Dimensional Bias Adjustment Procedure

The optimization of the two-dimensional bias adjustment procedure is based solely on the choices for interpolating the irregularly spaced OI data to the HRAP grid. Ordinary Kriging was chosen as the interpolation method, with a detailed discussion of its principles included in section III.9. Discussions in this section will be limited to the choices made in this specific interpolation scheme.

For given values of *a* and *t*, interpolation of the observation increments to a continuous two-dimensional field make use of each available radar-gauge pair. Additionally, all the available *OI* data will be used to construct an empirical semivariogram constructed, with only a single semivariogram constructed for each set of observational increment data. The optimal procedure for the ordinary Kriging of analysis increments from irregularly spaced observations to HRAP grid cells has two primary considerations, which are

1. the type of model semivariogram used to fit the semivariance data and
2. the number of neighbors used in the Kriging interpolation.

The consideration of model type was narrowed to two choices after investigating numerous empirical plots of semivariogram data. Fig. 5.3 shows the empirical semivariograms from April 2012, July 2012, and December 2012, each using the 1-month OI data. As one would anticipate, the autocorrelation is greater (semivariance lower) for cold season precipitation than warm season precipitation.

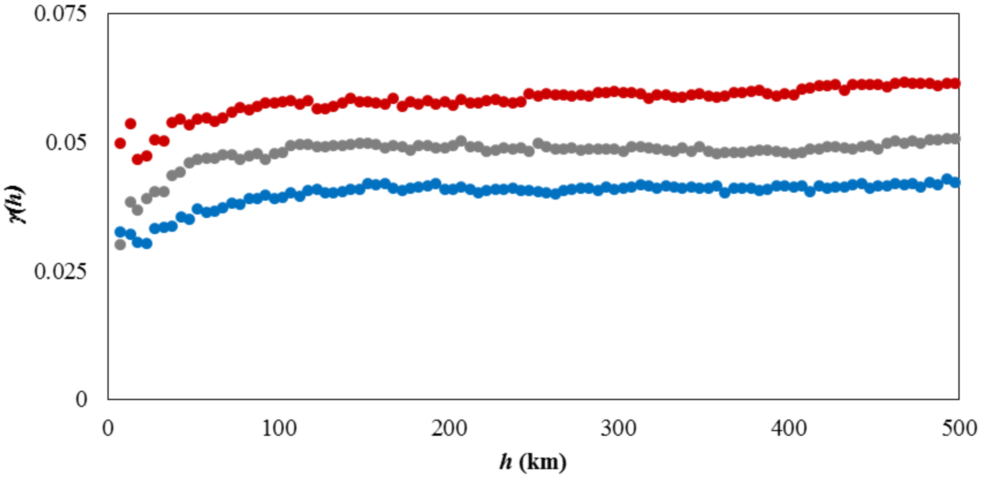


Fig. 5.3. Empirical semivariogram constructed using pairs of 1-month OI data points from April 2012 (gray circles), July 2012 (red circles), and December 2012 (blue circles).

The two semivariogram model types investigated were a single straight line model fit and a spherical model fit. For both types of fits, a sill value c is assumed for values of h greater than lag a (the value $h = a$ is where the sill begins) and a nugget value b when $h = 0$.

The spherical model function $g_S(h)$ can be written, using Bohling (2005a), as

$$g_S(h) = \begin{cases} b_S + (c - b_S)\left[\frac{3}{2}\left(\frac{h}{a}\right) - \frac{1}{2}\left(\frac{h}{a}\right)^3\right] & \text{if } h < a \\ c & \text{if } h \geq a \end{cases} . \tag{5.4}$$

The value of the sill c in this study is found by averaging the semivariance $\gamma(h)$ data between 250 km and 500 km, values at which the data is assumed to asymptote with increasing lag. The lag a at which the sill begins in the spherical model is determined using the slope m_s of the $\gamma(h)$ data of the 12 nominal lags with the lowest values of h. The slope

m_S and the intercept (or "nugget" using Kriging terminology) b_S of the 12 $\gamma(h)$ data points is found using the SLR method. The lag a at which the sill begins in the spherical model is

$$a = \frac{3}{2}\left(\frac{c - b_S}{m_S}\right), \tag{5.5}$$

with the constant 3/2 in (5.5) to account for the curvature of the sill. The single straight line model function $g_L(h)$ can be written as

$$g_L(h) = \begin{cases} b_L + (c - b_L)\left(\dfrac{3h}{2a}\right) & \text{if } h < \dfrac{2}{3}a \\ c & \text{if } h \geq \dfrac{2}{3}a \end{cases}, \tag{5.6}$$

using the same values of a and c as the spherical model. Since there is no curvature in the single straight line model $g_L(h)$, the straight line in this model is assumed to intercept the sill at $2/3a$. The nugget value b_L used in the $g_L(h)$ model is found by regressing a straight line through all the $\gamma(h)$ values at lags $h \leq 2/3\text{a}$.

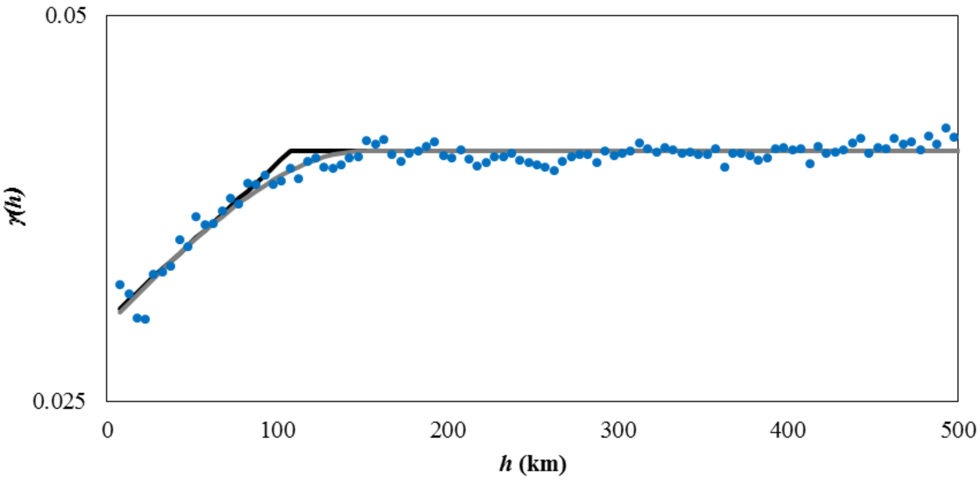


Fig. 5.4. The single straight line (black line) and spherical (gray line) semivariogram model fits for the 1-month semivariance data from January 2012 (blue circles).

Fig. 5.4 shows an example of the two different model fits to semivariances from the 1-month OI data from January 2012. A procedure to determine the ideal choices of the three considerations listed for optimization of the objective analysis procedure, including the semivariogram model type, are discussed in the following section.

V.2.d. Optimization of the Interpolation Procedure

The procedure for optimizing the interpolation of radar-gauge pair OI values to the HRAP grid cells used an 80/20 cross-validation scheme similar to that used for verification of the combi-conVPR model. For each permutation of (a, t), 20% of the OI values were withheld for verification and the remaining 80% of the values were used in the interpolation scheme. The two ordinary Kriging interpolation considerations tested were (with values) were

1. the model type: $g = \{g_L, g_S\}$ and

2. number of neighboring data points used in the interpolation: $n_I = \{4, 8, 12, \ldots, 32, 36, 40\}$.

The optimization procedure used OI data with $a = \{1 \text{ mo.}, 4 \text{ mos.}, 12 \text{ mos.}, 36 \text{ mos.}\}$ and all possible values of t in the period January 2010 – December 2012. The optimization procedure used pairs with data availability $A \geq 90\%$, a criterion chosen based on the results of the combi-conVPR model optimization testing. The radar-gauge pairs withheld for validation were randomly selected from the subset of pairs with complete data availability ($A = 100\%$). For each permutation, the number of withheld pairs was exactly 20% of the total number of pairs with $A = 100\%$ used in the testing procedure.

The results of the testing procedure indicated that the straight line model $g_L(h)$ outperformed the spherical model $g_S(h)$ for all accumulation periods (Table 5.1). In Fig. 5.4, the nominal OI data seem to approach the sill suddenly rather than gradually asymptote the sill. Additionally, the single straight line model is a more robust estimator of the slope of the nominal semivariance points in the first 100 km given it uses a higher number of data point. The vast majority of neighbors used for the OI interpolation to target HRAP grid cells have lags less than 100 km, which would tend to favor use of the $g_L(h)$ semivariance model.

Table 5.1. The MSE for different values of a using different permutations of g and n_I from the procedure to determine optimal values for ordinary Kriging interpolation of bias analysis increments.

n_I	**a = 1 mo.**		**a = 4 mos.**		**a = 12 mos.**		**a = 36 mos.**	
	$g_L(h)$	**$g_S(h)$**	**$g_L(h)$**	**$g_S(h)$**	**$g_L(h)$**	**$g_S(h)$**	**$g_L(h)$**	**$g_S(h)$**
4	0.0627	0.0627	0.0270	0.0270	0.0143	0.0143	0.0107	0.0107
8	0.0585	0.0588	0.0252	0.0253	0.0133	0.0133	0.0100	0.0101
12	0.0568	0.0573	0.0244	0.0246	0.0128	0.0130	0.0097	0.0098
16	0.0559	0.0567	0.0239	0.0243	0.0126	0.0128	0.0095	0.0097
20	0.0554	0.0564	0.0237	0.0242	0.0124	0.0127	0.0094	0.0096
24	0.0550	0.0561	0.0235	0.0240	0.0123	0.0126	0.0094	0.0096
28	0.0548	0.0560	0.0234	0.0240	0.0123	0.0126	***0.0093***	0.0096
32	0.0547	0.0558	0.0234	0.0239	***0.0123***	0.0125	0.0094	0.0095
36	***0.0546***	0.0556	0.0233	0.0238	0.0123	0.0125	0.0094	0.0095
40	0.0546	0.0555	***0.0233***	0.0237	0.0123	0.0124	0.0094	0.0095
44	0.0547	0.0554	0.0233	0.0237	0.0123	0.0124	0.0094	0.0095
48	0.0549	0.0552	0.0234	0.0236	0.0124	0.0124	0.0095	0.0094
52	0.0550	0.0551	0.0234	0.0236	0.0124	0.0123	0.0095	0.0094
56	0.0551	0.0550	0.0235	0.0235	0.0124	0.0123	0.0095	0.0094
60	0.0552	0.0549	0.0235	0.0235	0.0124	0.0123	0.0095	0.0094

The number of neighbors n_I need for the ordinary Kriging interpolation of OI values decreases with an increasing length of accumulation period a. A reasonable extension of these results would be that $n_I = 32$ for $a = 18$ months and $n_I = 28$ for $a = 24$ months given the results in Table 5.1.

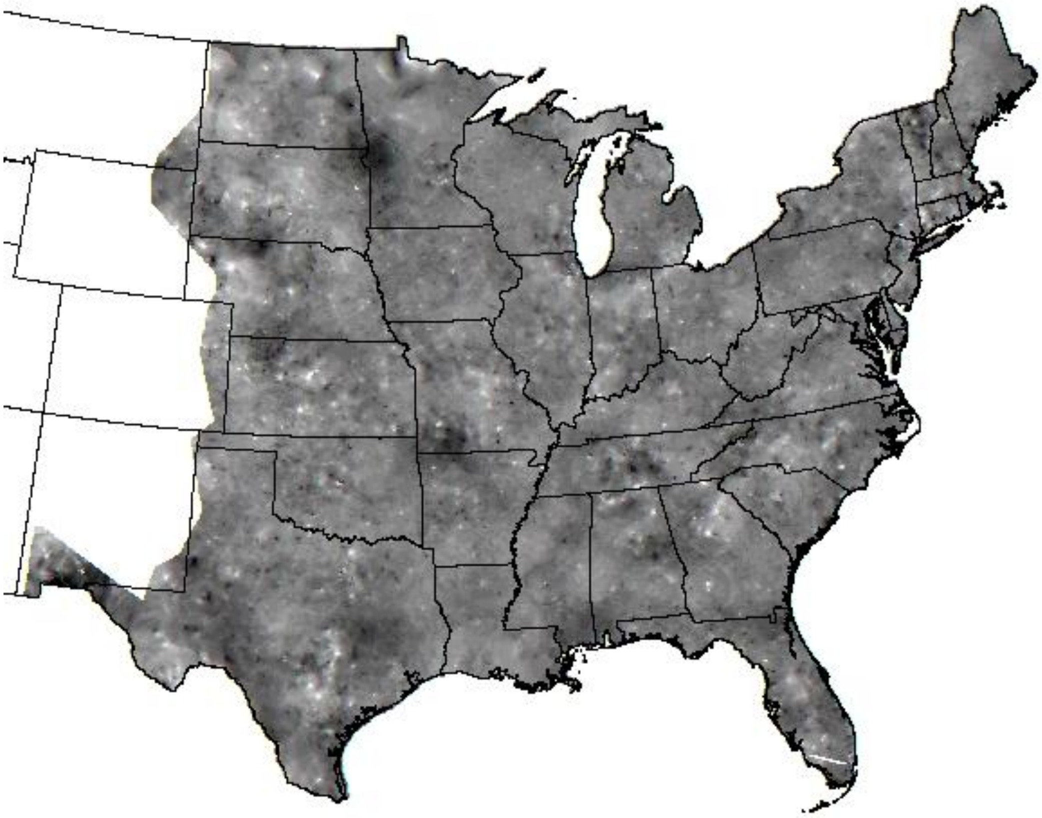


Fig. 5.5. Field of observational increments for the 12-month period ending 31 December 2012. The minimum value (black) is *OI* = -0.25 and the maximum value (white) is *OI* = 0.25.

One would expect that the spatial autocorrelation between OI values in a given region would increase with increasing length of accumulation period *a*. In other words, the bias field should get less random with increasing *a* values, which should reduce the number of neighbors needed to provide accurate estimates of the OI where gauges are not available. Fig. 5.5 is an example of an analyzed OI field for all the HRAP grid cells in the central and eastern United States for the 12-month period ending 31 December 2012. Fig. 5.6 is the first guess analysis of the 12-month bias field for the same values of *a*, *t* in Fig. 5.5.

The radar domain boundaries are clearly distinguishable in Fig. 5.6 since the combi-conVPR bias models were developed independently in each radar domain.

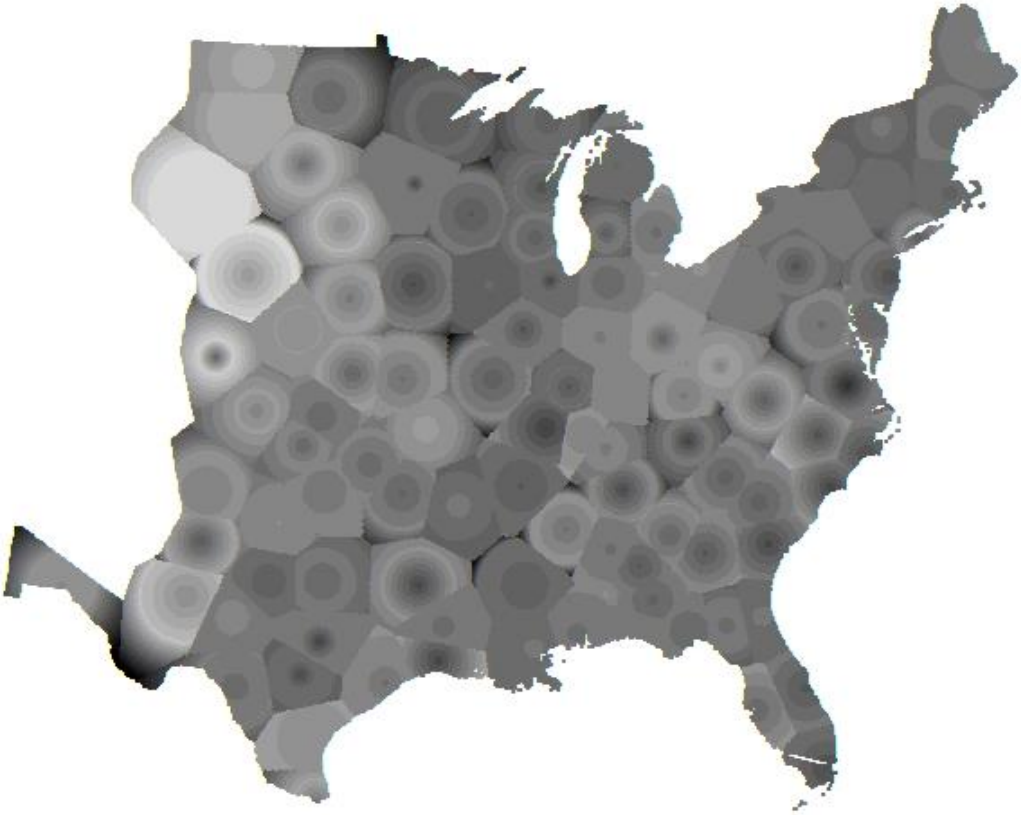


Fig. 5.6. The first guess bias field (B_3) of the Stage IV P_1 precipitation for the 12-month period ending 31 December 2012. The minimum value (black) is B_3 = -0.5 and the maximum value (white) is B_3 = 0.5.

The final bias value B_f is computed at each HRAP grid cell in using (5.4) and visually is the addition of the *OI* spatial field from Fig. 5.5 to the first guess bias field (B_3) in Fig. 5.6. The 12-month B_f field for the period ending 31 December 2012 (Fig. 5.7) shows the more complex two-dimensional variation that cannot obviously be captured by the first guess field using the combi-conVPR model. The most visibly apparent contribution of the first guess field in the B_f spatial field (Fig. 5.7) is seen at the edges of some of the radar domains, but the two-dimensional variations in B_f have more in common with Fig. 5.5.

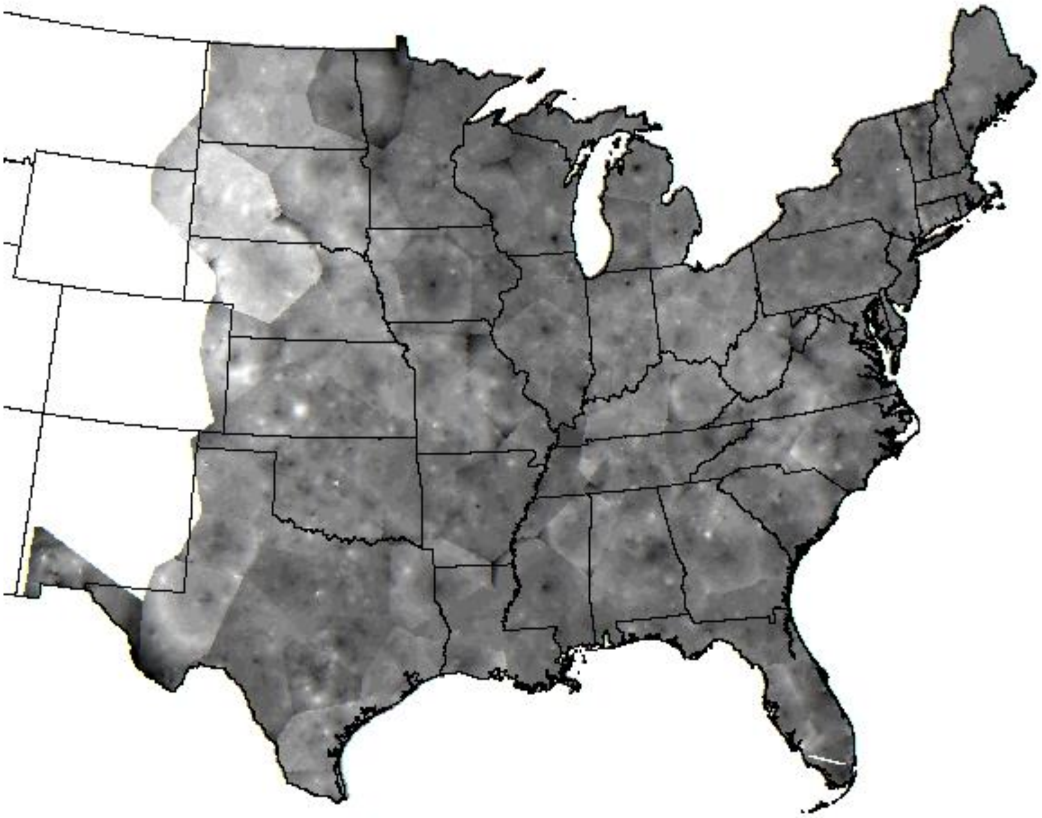


Fig. 5.7. The final bias field (B_f) of the Stage IV P_1 precipitation for the 12-month period ending 31 December 2012. The minimum value (black) is $B_\mathrm{f} = -0.5$ and the maximum value (white) is $B_\mathrm{f} = 0.5$.

The final step for producing a precipitation field (P_f) corrected for all the biases in (1.1) is to apply the bias field B_f to the Stage IV P_1 precipitation data, which has only been corrected for beam blockage. The B_f field accounts for range-dependent, mean-field, and two-dimensional anisotropies, so at each grid cell the value of P_f is

$$P_f\left(\Phi_H,\Lambda_H\right)=\frac{P_1\left(\Phi_H,\Lambda_H\right)}{B_f\left(\Phi_H,\Lambda_H\right)+1}. \tag{5.7}$$

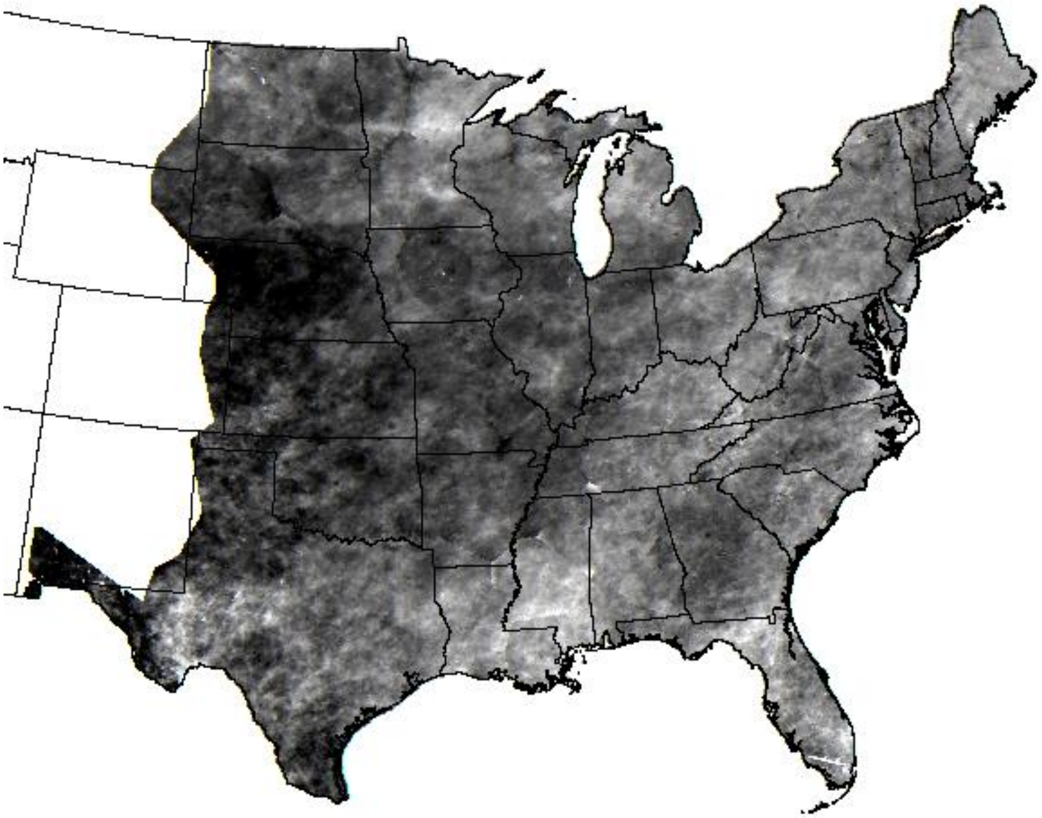


Fig. 5.8. The final Stage IV PoN precipitation field, computed using P_f for the 12-month period ending 31 December 2012. The minimum value (black) is PoN = 50% and the maximum value (white) is PoN = 150%.

The analysis for the 12-month Stage IV PoN field for the period ending 31 December 2012 is shown in Fig. 5.8 and at each HRAP grid cell is computed as P_f / P_n. The final precipitation field P_f (Fig. 5.9) uses the bias field from Fig. 5.7 to correct the Stage IV P_1 field.

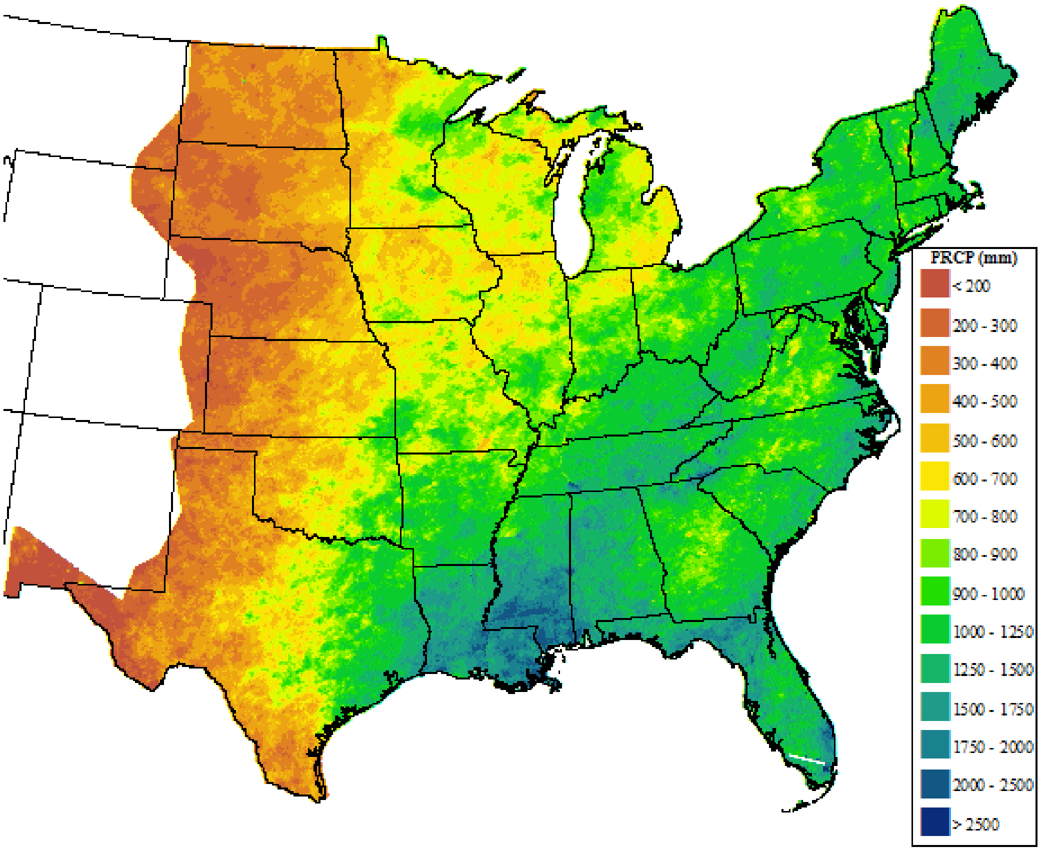


Fig. 5.9. The final Stage IV precipitation field P_f for the 12-month period ending 31 December 2012.

V.2.e. Two-Dimensional Bias Adjustment Procedure Verification

The verification procedure described in this section was carried out to confirm that B_f, which accounts for the four types of known biases in (1.1), provides an improved Stage IV precipitation estimate (P_f) compared to P_3, which accounts for only beam blockage, mean-field biases, and range-dependent biases. The basic methodology of the verification procedure is to withhold 20% of the available radar-gauge pairs from accumulation period *a* and ending date *t*, with each withheld pair having a combi-conVPR model B_3 value already determined. The remaining 80% of the pairs are used to determine the model semivariogram and to act as neighboring data points for an ordinary Kriging analysis. Observation increments are interpolated to each of the withheld pairs and a $B_f(\Phi_H, \Lambda_H)$ value is determined. The final Stage IV precipitation estimate (P_f) is computed using (5.8) and compared to the P_3 value computed from (4.44) without the OI adjustment. The two precipitation estimates at each withheld pair are compared to the gauge value (G) and RMS errors are computed.

As in the evaluation procedure of the combi-conVPR model from the previous chapter, evaluation of the errors in this procedure used $\mathbf{a}$ = {1 month; 4 months; 12 months; 36 months} and $\mathbf{t}$ = {January 2008; July 2009; April 2011; October 2012} and 10 different randomizations of withheld pairs for each permutation of *a* and *t*. The average RMSE value over the 10 runs for each permutation of (*a*, *t*) for each of the Stage IV precipitation datasets is summarized in Table 5.2.

For all but one of the 16 permutations of *a* and *t*, the withheld P_f had a smaller RMSE value than the P_3 data, and the overall improvement of the P_f estimates relative to the P_3 estimates was about 4.6%. In fact, the overall RMSE value was smaller for 141 out of the 160 runs. Seven out of the 19 runs where B_f was a worse overall fit came from the 10 randomizations where a = 36 months and t = January 2008. The verification procedure validated B_f as a superior assessment of the overall bias compared to the B_3 assessment.

Table 5.2. Validation test RMSE values for the P_3 and P_f datasets for different accumulation periods (a) and ending dates (t).

t	***a***	**RMSE (mm)**	
		P_3	***P_f***
January 2008	**1**	14.3	13.9
	4	40.8	39.3
	12	117.7	112.5
	36	274.9	278.0
July 2009	**1**	26.5	25.5
	4	60.2	56.3
	12	297.0	284.4
	36	364.2	344.9
April 2011	**1**	23.8	21.5
	4	47.8	46.6
	12	114.3	107.3
	36	370.3	345.4
October 2012	**1**	18.6	17.8
	4	55.8	54.1
	12	107.6	104.8
	36	315.8	299.5

CHAPTER VI

SUMMARY AND FUTURE WORK

VI.1. Summary

This dissertation addresses the need for a reduction in the biases that lead to erroneous Stage IV radar precipitation estimates. The general approach for solving this problem was to 1) determine the mechanisms of the radar measurements responsible for the biases, 2) decide in what manner the radar estimates were affected by these biases, 3) develop quantitative methods for evaluating the spatial properties of the biases, and 4) use the evaluation of the bias field properties to correct for biases in the Stage IV estimates. The bias evaluations and corrections were done primarily with gauge data, which were used to compute radar-gauge pair biases, and percent of normal (PoN) precipitation which was used as a proxy for biases in the Stage IV estimates.

The bias minimization procedures were developed to deal specifically with biases due to beam blockage, range-dependent biases related to the vertical profile of reflectivity (VPR), and mean-field biases related to issues with radar calibration and were done independently at each radar domain. An ordinary Kriging interpolation procedure was designed to minimize any remaining two-dimensional anisotropies following the minimization of the one-dimensional biases. The resulting bias correction procedures will be used to generate an improved dataset of Stage IV precipitation estimates on the spatially continuous HRAP grid that is continuously updated in real-time.

The first bias correction procedure dealt with beam blockage in the Stage IV estimates, which visually appear as sudden discontinuities in a long-term precipitation field that can be traced along an azimuth back to the radar location. The longest accumulation period in the HIRDTT suite of products is 36 months, and PoN precipitation at this time scale was utilized in the beam blockage detection procedure. The guiding principle of the procedure was that, at a fixed range, the Stage IV PoN precipitation data should be a smooth and continuous function for a complete azimuthal sweep. Beam blockage was flagged at HRAP grid cells with PoN precipitation values that had extreme minimum residuals relative to a Fourier series fit. For the detection algorithm to consider a region of a given radar domain "blocked", there must be a spatiotemporal consistency in the flagging of grid cells within the region. Ordinary Kriging interpolation of data from neighboring HRAP grid cells estimated the PoN precipitation at HRAP grid cells considered blocked.

The approach to the second bias correction procedure resulted in the creation of a combined mean-field and range-dependent bias model to quantifying VPR and other range-dependent effects of radar precipitation estimates as a smooth and continuous function of range. The inspiration for this bias quantification model was the Krajewski et al. (2011) model of the climatological VPR as a function of height, which was adapted to model bias in the Stage IV estimates as a function of range. The bias model data points were radar-gauge pair biases computed at gauges and collocated HRAP grid cells and Stage IV PoN precipitation data, the latter of which were averaged over range intervals and adjusted for magnitude by the radar-gauge pair biases. The resulting combi-conVPR model identifies a VPR maximum if signaled by the data points (four parameter model) or uses a single regression line (two parameter model) if the existence of a VPR maximum is doubtful. Extensive LOOC-V testing determined objective measures for choosing which of the two model forms were appropriate, both of which use weighted Theil-Sen (Sievers 1978) for parameter estimation.

The third and final step in the bias minimization procedure was to remove any residual two-dimensional anisotropies from the prior two bias corrections procedures. For a given Stage IV precipitation field already corrected for beam blockage, mean-field bias, and range-dependent biases at the individual radar domains, radar-gauge pair biases were computed at each available gauge using the corrected values. A two-dimensional bias field was computed from the bias data at the individual gauges using ordinary Kriging interpolation, and the corrected precipitation value at each HRAP grid cell was adjusted based on that cell's Kriged bias value.

For a given time *t* and accumulation period *a*, the bias-correction algorithm is designed to give an accurate spatial assessment of precipitation for that specific time period. Because the bias-correction adjustments are done independently for each precipitation field, an additional filter would need to be applied for the bias-adjusted precipitation totals across different time periods and accumulation periods to be consistent. For instance, it is unlikely that the bias-adjusted precipitation total for a given radar and 12-month period can be computed using an aggregation of totals from sub-periods within that timeframe. The stated goal of this work is to provide an improved quantification of drought, not the best estimate of total precipitation at each grid cell; potential users of this dataset should keep this caveat in mind.

In addition to providing improved Stage IV precipitation mosaics for use in drought monitoring, there are other useful applications of the techniques developed throughout the course of this project, particularly in the analysis of precipitation at individual radars. The beam blockage detection algorithm offers an improved understanding of the apparent obstacles within each radar domain. Precipitation estimates at most radars use very recent gauge information, dating back only several hours, to correct for spatial and mean-field biases. The aggregation of precipitation totals allows for an improved understanding of the spatial properties of existing biases in each radar domain. Additionally, the aggregation of the Stage IV precipitation totals provides a better identification of each radar's effective coverage area. If radar-gauge pair bias data is limited to time scales of hours or days, use of the bias assessment algorithms can still be used to better assess spatial biases.

VI.2. Future Work

The precipitation field resulting from the three bias correction procedures is considered to have biases minimized and be an improvement on the Stage IV precipitation estimate as an assessment of the true surface precipitation. Based on the preliminary bias-corrected PoN precipitation fields (i.e., Figs. 5.8 and 5.9), some additional work is needed to produce an optimal field that can be directly interpreted by non-experts. A great deal of time and effort went into building appropriate conceptual models for correcting biases; however analysis of the precipitation fields resulting from these corrections is in its infancy.

Thorough visual inspection of the remaining and new errors in the bias-corrected dataset is needed for tweaking of the algorithms to produced less-biased estimates. Further refinements of the correction procedures will be explored in the coming months. The author is confident that the basic bias correction procedures are fundamentally sound, so any improvements will have a marginal impact on the methodology laid out in this dissertation. Some details of the correction methodology that need to be addressed are as follows.

1. There appears to be a systematic overcorrection of range-dependent biases toward the edge of radar domains, where bias data are sparse.
2. Extreme minima of Stage IV estimates in the HRAP grid cells near the radar location seems to be having an impact on choice of partition for the radar-gauge pair bias conVPR model.
3. There needs to be an additional step in the two-dimensional anisotropy correction procedure to specifically address the discontinuities at the boundaries of radar domains and the RFCs.

Once there is confidence in the resulting bias-corrected precipitation fields, the adjusted Stage IV estimates will be used as the real-time, high-resolution precipitation input for the HIRDTT products. At this point, the broader stated goal of this study will be accomplished with the existence of high-resolution drought monitoring products that do not require user expertise.

REFERENCES

Anagnostou, E. N., W. F. Krajewski, D. J. Seo, and E. R. Johnson, 1998: Mean-field rainfall bias studies for WSR-88D. *J. Hydrol. Eng.*, **3**, 149-159.

Andrieu, H., and J. D. Creutin, 1995: Identification of vertical profiles of radar reflectivity for hydrological applications using an inverse method. Part I: Formulation. *J. Appl. Meteorol.*, **34**, 225-239.

Andrieu, H., Creutin, J. D., Delrieu G., and D. Faure, 1997: Use of a weather radar for the hydrology of a mountainous area. Part I: Radar measurement interpretation. *J. Hydrol*., **193**, 1-25.

Austin, P., 1987: Relation between measured radar reflectivity and surface rainfall. *Mon Wea. Rev.*, **115**, 1053-1070.

Baeck, M. L., and J. A. Smith, 1998: Rainfall estimation by the WSR-88D for heavy rainfall events. *Wea. Forecasting*, **13**, 416-436.

Battan, L. J., 1973: *Radar Observation of the Atmosphere*. The University of Chicago Press, 324 pp.

Bellon, A., GW Lee, and I. Zawadzki, 2005: Error statistics of VPR correction in stratiform precipitation. *J. Appl. Meteor.*, **44**, 998-1015.

Bohling, G., 2005a: Introduction to geostatistics and variogram analysis. C&PE 940. Kansas Geological Survey, 20p.

Bohling, G., 2005b: Kriging. C&PE 940. Kansas Geological Survey, 20p.

Breidenbach, J. P., D. J. Seo, and R. Fulton, 1998. Stage II and III post-processing of NEXRAD precipitation estimates in the modernized Weather Service. Preprints, *14th Intl. Conf. on Interactive Information and Processing Systems (IIPS) for Meteor., Oceanography, and Hydrology*, Phoenix, AZ, Amer. Meteor. Soc., pp. 263-266.

Chen, S., and Coauthors, 2013: Evaluation and Uncertainty Estimation of NOAA/NSSL Next-Generation National Mosaic Quantitative Precipitation Estimation Product (Q2) over the Continental United States. *J. Hydrometeor*, **14**, 1308-1322.

Chrisman, J., D. Rinderknecht, and R. Hamilton, 1995: WSR-88D clutter suppression and its impact on meteorological data interpretation. Preprints, *First WSR-88D User's Conference*, Norman, OK, WSR-88D Operational Support Facility, 9-20.

Ciach, C. J., and W. F. Krajewski, 1999: On the estimation of radar rainfall error variance. *Adv. Water Resour.*, **22**, 585-595.

Ciach, C. J., W. F. Krajewski, and G. Villarini, 2007: Product-Error-Driven Uncertainty Model for Probabilistic Quantitative Precipitation Estimation with NEXRAD Data. *J. Hydrometeor..*, **22**, 1325-1347.

Cifelli, R., N. Doesken, P. Kennedy, L. D. Carey, S. A. Rutledge, C. Gimmestad, and T. Depue, 2005: The Community Collaborative Rain, Hail, and Snow Network: Informal education for scientists and citizens. *Bull. Amer. Meteor. Soc.*, **86**, 1069-1077.

Crum, T. D., and R. L. Alberty, 1993: The WSR-88D and the WSR-88D operational support facility. *Bull. Amer. Meteor. Soc.*, **74**, 1669-1687.

Crum, T. and Coauthors, 2013: 2013 Update on Access to Real-Time and Archive NOAA Weather Radar Data: 2013. *29th Conf. on Environmental Information Processing Technologies*, Austin, TX, Amer. Meteor. Soc., J12.1.

Cunha, L. K., J. A. Smith, M. L. Baeck, and W. F. Krajewski, 2013: An Early Performance Evaluation of the NEXRAD Dual-Polarization Radar Rainfall Estimates for Urban Flood Applications. *Wea. Forecasting*, **28**, 1478–1497.

Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor.*, **33**, 140-158.

Donaldson, R., 1964: A demonstration of antenna beam errors in radar reflectivity patterns. *J. Appl. Meteor.*, **3**, 611–623.

Doviak, R. J., V. Bringi, A. Ryzhkov, A. Zahrai, D. Zrnić, 2000: Considerations for Polarimetric Upgrades to Operational WSR-88D Radars. *J. Atmos. Oceanic Technol.*, **17**, 257–278.

Dow, K., R. L. Murphy, and G. J. Carbone, 2009: Consideration of user needs and spatial accuracy in drought mapping. *J. Amer. Water Res.*, **45**, 187-197.

Duchon, C. E., and G. R. Essenberg, 2001: Comparative rainfall observations from pit and aboveground rain gauges with and without wind shields, *Water Resour. Res.*, **37**, 3253-3263.

Elliott, J., and Coauthors, 2013: Predicting agricultural impacts of large-scale drought: 2012 and the case for better modeling. No. 111. Grantham Research Institute on Climate Change and the Environment, 5 pp.

Fabry, F., G. L. Austin, and D. Tees, 1992: The accuracy of rainfall estimates by radar as a function of range. *Quart. J. Royal Meteorol. Soc.*, **118**, 435-453.

Fulton, R. A., 2002: Activities to improve WSR-88D radar rainfall estimation in the National Weather Service. *Proc. Second Federal Interagency Hydrologic Modeling Conf.*, Las Vegas, NV, Subcommittee on Hydrology, Advisory Committee on Water Data.

Fulton, R. A., 2005: Multisensor Precipitation Estimator (MPE) Workshop. *Advanced Hydrol. Applications Course*, Kansas City, MO, National Weather Service.

Fulton, R. A., J. P. Breidenbach, D. Seo, D. A. Miller, and T. O'Bannon, 1998: The WSR-88D rainfall algorithm. *Wea. Forecasting*, **13**, 380-391.

Glaudemans, M., P. Tilles, and B. Lawrence, 2008: Interactive quality control and operational product generation of hourly multi-sensor precipitation estimates in the NWS. Preprints, *23rd Conference on Hydrol.*, New Orleans, LA, Amer. Meteor. Soc., 5B.3.

Groisman, P. Y., and D. R. Legates, 1994: The accuracy of United States precipitation data. *Bull. Amer. Meteor. Soc.*, **75**, 215-227.

Habib, E., and L. Qin, 2013: Application of a radar-rainfall uncertainty model to the NWS multi-sensor precipitation estimator products. *Meteor. Appl.*, **20**, 276-286.

Habib, E., W. F. Krajewski, and A. Kruger, 2001: Sampling errors of tipping-bucket rain gauge measurements. *J. Hydrol. Eng.*, **6**, 159–166.

Habib, E., B. F. Larson, and J. Graschel, 2009: Validation of NEXRAD multisensor precipitation estimates using an experimental dense rain gauge network in south Louisiana. *J. Hydrol.*, **15**, 463-478.

Harrison, D. L., S. J. Driscoll, and M. Kitchen, 2000: Improving precipitation estimates from weather radar using quality control and correction techniques. *Meteorol. Appl.*, **6**, 135-144.

Hastie, T., R. Tibshirani, and J. Friedman (2005): *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. 2nd ed. Springer, 745 pp.

Henkel, A., and C. Peterson, 1996: Can deterministic quantitative precipitation forecasts in mountainous regions be specified in a rapid, climatologically consistent manner with Mountain Mapper functioning as the tool for mechanical specification, quality control, and verification? *Extended Abstracts, Fifth National Heavy Precipitation Workshop*, State College, PA, NWS/NOAA, 31 pp.

Holleman, I., 2006. Bias adjustment of radar-based 3-hour precipitation accumulations. KNMI Technical Report TR-290, 56 pp.

Hosking, J. R. M., and J. R. Wallis, 1997: *Regional Frequency Analysis: An Approach Based on L-Moments*. Cambridge University Press, 224 pp.

Jameson, A. R., 1991: A comparison of microwave techniques for measuring rainfall. *J. Appl. Meteor.*, **30**, 32-54.

Jayakrishnan, R., R. Srinivasan, and J. G. Arnold, 2004: Comparison of raingage and WSR-88D Stage III precipitation data over the Texas-Gulf basin. *J. Hydrol.*, **10**, 135-152.

Joss, J., and A. Waldvogel, 1990: Precipitation measurement and hydrology. *Radar in Meteorology*, D. Atlas, Ed., Amer. Meteor. Soc., 577-600.

Kalinga, O. A., and T. Y. Gan, 2012: Merging WSR-88D stage III radar rainfall data with rain gauge measurements using wavelet analysis. *Intl. J. Remote Sens.*, **33**, 1078-1105.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation, and Predictability*. Cambridge University Press, Cambridge, UK, 341 pp.

Karl, J. W., and B. A. Maurer, 2010: Spatial dependence of predictions from image segmentation: A variogram-based method to determine appropriate scales for producing land-management information. *Ecolo. Inform.*, **5**, 194-202.

Kim, D., E. I. Tollerud, S. V. Vasiloff, and J. Caldwell, 2009: Comparison of manual and automated quality control of operational hourly precipitation data of the National Weather Service. Preprints, *23rd Conf. on Hydrology*, Phoenix, AZ, Amer. Meteor. Soc., J6.3.

Klazura, G. E., and D. A. Imy, 1993: A description of the initial set of analysis products available from the NEXRAD WSR-88D system. *Bull. Amer. Meteor. Soc.*, **74**, 1293-1311.

Krajewski, W. F., and J. A. Smith, 2002: Radar hydrology: rainfall estimation. *Adv. Water Res.*, **25,** 1387-1394.

Krajewski, W. F., and B. Vignal, 2001: Evaluation of anomalous propagation echo detection in WSR-88D Data: A large sample case study. *J. Atmos. Oceanic Technol.*, **18**, 807-814.

Krajewski, W. F., B. Vignal, B.-C. Seo, and G. Villarini, 2011: Statistical model of the range-dependent error in radar-rainfall estimates due to the vertical profile of reflectivity. *J. Hydrol.*, **402**, 306-316.

Lawrence, B., M. Shebsovich, M. Glaudemans, and P. Tilles, 2003: Enhanced precipitation estimation capabilities at National Weather Service field offices using Multi-Sensor Precipitation data mosaics. *19th Conference on IIPS*, Long Beach, CA, Amer. Meteor. Soc., 15.1.

Legates, D. R., 2000: Real-time calibration of radar precipitation estimates. *Prof. Geogr.*, **52**, 235-258.

Lin, Y., and K. E. Mitchell, 2005: The NCEP stage II/IV hourly precipitation analyses: Development and applications. Preprints, *19th Conf. on Hydrology,* San Diego, CA, Amer. Meteor. Soc., 1.2.

Marshall, J. S., and W. McK. Palmer, 1948: The distribution of raindrops with size, *J. Atmos. Sci.*, **5**, 165-166.

Marzen, J. and H. E. Fuelberg, 2005: Developing a high resolution precipitation dataset for Florida hydrologic studies. Preprints, *19th Conference on Hydrology*, San Diego, CA, Amer. Meteor. Soc., J9.2.

McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration to time scales. Preprints, *Eighth Conf. on Applied Climatology*. Anaheim, CA, Amer. Meteor. Soc., 179–184.

McRoberts, D. B., John W. Nielsen-Gammon, 2012: The use of a high-resolution Standardized Precipitation Index for drought monitoring and assessment. *J. Appl. Meteor. Climatol.*, **51**, 68–83.

Neff, E. L., 1977: How much rain does a rain gage gage? *J. Hydrol.*, **35**, 231-220.

Nespor, V., and B. Sevruk, 1999: Estimation of wind-induced error of rainfall gauge measurements using a numerical simulation. *J. Atmos. Oceanic Technol.*, **16**, 450-464.

New, M., M. Hulme, and P. Jones, 2000: Representing twentieth-century space-time climate variability. Part II: development of 1901-96 monthly grids of terrestrial surface climate. *J. Climate*, **13**, 2217-2238.

Overeem, A., I. Holleman, and A. Buishand, 2009: Derivation of a 10-year radar-based climatology of rainfall. *J. Appl. Meteorol. Climatol.*, **48**, 1448-1463.

Peterson, T. C., and D. R. Easterling, 1994: Creation of homogeneous composite climatological reference series. *Intl. J. Climatol.*, **14**, 671-679.

Quiring, S. M., 2009: Developing Objective Operational Definitions for Monitoring Drought. *J. Climate Appl. Meteor.*, **48**, 1217–1229.

Rinehart, R. E., 2004: *Radar for Meteorologists*. 4th ed., Rinehart Publications, Columbia, MO, 482 pp.

Rogalus III, M. J., and F. L. Ogden, 2007: Comparison of GCIP and Stage III weather radar rainfall estimates over the Mississippi River Basin. *J. Hydrol.*, **341**, 177-185.

Rogalus III, M. J., and F. L. Ogden, 2011: Assessment of estimation bias around rainfall gage sites for five years of WSR-88D data over the Mississippi River Basin. *Proc.*

World Environmental and Water Resources Congress 2011, Palm Springs, CA, Environmental and Water Resources Institute.

Rousseeuw P. J., and A. M. Leroy, 2003: *Robust Regression and Outlier Detection*. Wiley, 360 pp.

Ryzhkov, A., and D. S. Zrnić, 1995: Comparison of dual-polarization radar estimators of rain, *J. Atmos.. Oceanic Technol.*, **12**, 249-256.

Ryzhkov, A., M. Diederich, P. Zhang, and C. Simmer, 2014: Potential Utilization of Specific Attenuation for Rainfall Estimation, Mitigation of Partial Beam Blockage, and Radar Networking, *J. Atmos. Oceanic Technol.*, **31**, 599-619.

Schaake, J., A. Henkel, and S. Cong, 2004: Application of PRISM climatologies for hydrologic modeling and forecasting in the western U.S. Preprints, *18th Conf. on Hydrology*, Seattle, WA, Amer. Meteor. Soc., 5.3.

Schmidt, J., B. Lawrence, and B. Olsen, 2000: A comparison of National Weather Service River Forecast Center operational precipitation processing methodologies. Tech. Memo. NWS SR-205, NOAA, Washington, DC, 4 pp.

Scofield, R. A., and R. J. Kuligowski, 2003: Status and outlook of operational satellite precipitation algorithms for extreme-precipitation events. *Wea. Forecasting*, **18**, 1037-1051.

Seo, B.-C., 2010: Towards a better representation of radar-rainfall: filling gaps in understanding uncertainties. Ph.D. dissertation, University of Iowa, 169 pp.

Seo, D. J., 1998: Real-time estimation of rainfall fields using radar rainfall and rain gage data. *J. Hydrol.*, **208**, 37-52.

Seo, D. J., 2002: Multisensor Precipitation Estimator (MPE). *NWS Flash Flood Workshop*, Boulder, CO, National Weather Service.

Seo, D. J., and J. P. Breidenbach, 2002: Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements. *J. Hydrometeor.*, **3**, 93-111.

Sevruk, B. 1991: International workshop on precipitation measurement: Preface. *Hydrol. Proc.*, **5**, 229-232.

Sieck, L. C., S. J. Burges, and M. Steiner, 2007: Challenges in obtaining reliable measurements of point rainfall. *Water Resour. Res.*, **43**, W01420.

Sievers, G. L., 1978: Weighted rank statistics for simple linear regression. *J. Amer. Statis. Assoc.*, **73**, 628-631.

Smith, A. B., and R. W. Katz, 2013: US billion-dollar weather and climate disasters: data sources, trends, accuracy and biases. *Nat. Hazards*, **67**, 387-410.

Smith, J. A., and W. F. Krajewski, 1991: Estimation of the mean field bias of radar rainfall estimates. *J. Appl. Meteor.*, **30**, 397-412.

Smith, J. A., D. J. Seo, M. L. Baeck, and M. D. Hudlow, 1996: An intercomparison study of NEXRAD precipitation estimates. *Water Resour. Res.*, **32**, 2035-2045.

Smith, P. L., 1998: On the minimum useful elevation angle for weather surveillance radar scans. *J. Atmos. Oceanic Technol.*, **15**, 841-843.

Steiner, M., J. A. Smith, S. J. Burges, C. V. Alonso, and R. W. Darden, 1999: Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation. *Water Resour. Res.*, **35**, 2487-2503.

Story, G., 2011: Recent Changes Made in the Hydrometeorological Analysis and Support Function at the NWS West Gulf RFC. *2011 NHWC Training Conference & Exposition*, San Diego, CA, National Hydrologic Warning Council.

Straka, J. M., D. S. Zrnić, and A. V. Ryzhkov, 2000: Bulk hydrometeor classification and quantification using polarimetric radar data: Synthesis of relations. *J. Appl. Meteor.*, **39**, 1341–1372.

Svoboda, M., and Coauthors, 2002: The Drought Monitor. *Bull. Amer. Meteor. Soc.*, **83**, 1181-1190.

Tapiador, F. J., and Coauthors, 2011: Global precipitation measurement: Methods, datasets and applications. *Atmos. Res.*, **104**, 70-97.

Uijlenhoet, R., M. Steiner, and J. A. Smith, 2003: Variability of raindrop size in a squall line and implications for radar rainfall estimation. *J. Hydrometeor.*, **4**, 43-61.

Vasiloff, S. V., 2012. Evaluation of dual-polarization QPE: Initial results for spring and summer 2012. Final Report MOU Task 1.1., 48 pp.

Vasiloff, S. V., and Coauthors, 2007: Improving QPE and very short term QPF: An initiative for a community-wide approach. *Bull. Amer. Meteor. Soc.*, **88**, 1899-1911.

Vignal, B., and W. F. Krajewski, 2001: Large-sample evaluation of two methods to correct range-dependent error for WSR-88D rainfall estimates. *J. Hydrometeor.*, **2**, 490-504.

Vignal, B., G. Galli, J. Joss, and U. Germann, 2000: Three methods to determine profiles of reflectivity from volumetric radar data to correct precipitation estimate. *J. Appl. Meteor.*, 39, 1715-1726.

Villarini, G., and W. F. Krajewski, 2010a: Sensitivity studies of the models of radar-rainfall uncertainties. *J. Appl. Meteor. Climatol.*, **49**, 288-309.

Villarini, G., and W. F. Krajewski, 2010b: Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall. *Surv. Geophys.*, **31**, 107-129.

Villarini, G., W. F. Krajewski, and J. A. Smith, 2009: New paradigm for statistical validation of satellite precipitation estimates: Application to a large sample of the TMPA 0.25 3-hourly estimates over Oklahoma. *J. Geophys. Res.*, **114**, D12106.

Westrick K. J., C. F. Mass, and B. A. Colle, 1999: The limitations of the WSR-88D radar network for quantitative precipitation measurement over the coastal western United States. *Bull. Amer. Meteorol. Soc.*, **80**, 2289-2298.

Yang, D, B. E. Goodison, J. R. Metcalfe, V. S. Golubev, R. Bates, T. Pangburn, and C. L. Hanson, 1998: Accuracy of NWS 8" Standard nonrecording Precipitation Gauge: Results and application of WMO intercomparison. *J. Atmos. Oceanic Technol.*, **15**, 54-68.

Yilmaz, K., T. Hogue, K. L. Hsu, S. Sorooshian, H. Gupta, and T. Wagener, 2005: Intercomparison of rain gauge, radar, and satellite-based precipitation estimates with emphasis on hydrologic forecasting. *J. Hydrometeorol.*, **6**, 497- 517.

Young, C. B., B. R. Nelson, A. A. Bradley, J. A. Smith, C. D. Peters-Lidard, A. Kruger, and M. L. Baeck, 1999: An evaluation of NEXRAD precipitation estimates in complex terrain. *J. Geophys. Res.*, **104**, 19691-19703.

Young, C. B., A. A. Bradley, W. F. Krajewski, A. Kruger, and M. L. Morrissey, 2000: Evaluating NEXRAD Multisensor Precipitation Estimates for operational hydrologic forecasting. *J. Hydrometeor.*, **1**, 241-254.

Zhang, J., C. Langston, and K. Howard, 2008: Brightband identification based on vertical profiles of reflectivity from the WSR-88D. *J. Atmos. Oceanic Technol.*, **25**, 1859-1872.

Zhang, J., K. Howard, S. Vasiloff, C. Langston, B. Kaney, A. Arthur, S. V. Cooten, and K. Kelleher, 2009: National Mosaic and QPE (NMQ) System – Description, results and future plans. *34th Conf. on Radar Meteorology*, Williamsburg, VA, Amer. Meteor. Soc., 7A.1.

Zhang, J., and Coauthors, 2011: National Mosaic and Multi-Sensor QPE (NMQ) System: Description, Results, and Future Plans. *Bull. Amer. Meteor. Soc.*, **92**, 1321-1338.