The Dependence of the Representation of Precipitation Extremes in the Headwaters of the Colorado River on Regional Climate Model Grid Spacing

Andreas Prein
andreas.prein@uni-graz.at

The National Center for Atmospheric Research (NCAR)
Mesoscale and Microscale Meteorology Division
3090 Center Green Dr, Boulder CO 80301, USA

May 2012

Contents

Preface

In this report the work from my research visit at the National Center for Atmospheric Research (NCAR) is summarize. All analyses and plots within this report were done by myself.

The evaluated WRF output data were party generated by myself (in case of the WRF-12km simulation) and partly provided by the group of Roy Rasmussen from the Research Applications Laboratory (RAL) at NCAR.

Currently I am working on publishing this work in two peer reviewed publications.

Abstract

Hydrological extremes have large impacts on ecology and society and are an important factor in the water balance of a region. Precipitation extremes often depend on small processes, which makes it hard to accurately simulate them in state of the art climate simulations with typical horizontal grid spacings larger than 20 km.

In this study the impact of model grid spacing on the simulation of extreme precipitation events in the headwaters of the Colorado river is investigated. Therefore, an eight year periods is simulated with the Weather Research and Forecasting Model (WRF) on a 4 km, 12 km, and 36 km grid. Perfect boundary conditions from the North American Regional Reanalysis (NARR) were used which makes the simulation directly comparable with observations.

The 4 km simulation outperforms the coarser gridded runs especially on small scales and improves the spatial structure, and variability of extreme events. The high resolution is essential for the representation of June, July, and August (JJA) extremes because on a 4 km grid deep convection can be resolved explicitly. Thereby, not only the spatial properties of extreme precipitation events are improved, but also the wet bias in the 12 km, and 36 km runs is corrected. In December, January, and February (DJF) high resolution is beneficial to simulate the observed precipitation patterns which are strongly triggered by orography. Here the 12 km resolution is sufficient to simulate the correct amount of precipitation but only in the 4 km run the spatial correlation of extreme precipitation filds is captured.

Acknowledgments

I am very thankful to the Marshallplan-Jubiläumsstiftung who awarded me with the Marshal Plan Scholarship. Without this funding this work would not have been possible.

My gratefulness belongs to Greg J. Holland who gave me the opportunity to work with him and his Regional Climate Group at the National Center for Atmospheric Research (NCAR) in Boulder Colorado. I want to thank him and his group, especially James Done, for all the fruitful discussions and for making my visit unforgettable. In addition I want to thank Greg Holland for founding my visit at the WRF Tutorial in February 2012.

I also want to gratitude Roy Rasmussen and his group for numerous meetings discussions and for giving me access to their data. In particular I want to thank Kyoko Ikeda for her support with data preparation and for taking the time for several discussions.

My deepest thanks belongs to my girlfriend and my family. They gave me moral strength in times of trouble and supported me in every aspect from the beginning until the end of my research visit.

The computer resources for my evaluations and simulations were provided by the Computation & Information System Laboratory (CISL) of the University Cooperation for Atmospheric Research (UCAR).

The snow telemetry (SNOTEL) data set was kindly provided by the National Resource Conservation Service of the United States Department of Agriculture. Data from Parameter-elevation Regressions on Independent Slopes Model (PRISM) was made available by the Climate Group of the Oregon State University (http://prism.oregonstate.edu). I am also thankful to the Climate Prediction Center for providing their .25x.25 Daily US UNIFIED Precipitation data set.

1 Introduction

Hydrological extreme events have large impacts on society, ecosystems, and water resources. In a changing climate hydrological extremes are expected to intensify because warmer air can carry more water vapor (according to the Clausius Clapeyron relation). This means there is potentially more water available to rain out from an air parcel and more latent heat can be released when the additional water vapor is condensed. However, there are often complex feedback mechanisms in the climate system which are not regarded in this simple consideration.

Climate models are common tools to investigate feedback processes like these. This models are based on physical principals and enable simulations of the current climate and projections of future climate change. However, General Circulation Models (GCMs) currently have horizontal grid spacings of about 200 km which is often much to coarse to properly simulate hydrological extremes. Therefore, regional climate models (RCMs) are often used to refine the resolution of GCMs on an limited area of the globe. State of the art RCMs have horizontal grid spacings of about 20 km, but in complex terrain and for the simulation of precipitation extremes this resolution is still too coarse.

In this report the effect of three horizontal grid spacings in the Weather Research and Forecasting Model (WRF) on the representation of extreme precipitation events in complex terrain is analyzed. The grid spacings are 4 km, 12 km, and 36 km. In the 12 km and 36 km simulations deep convection has to be parameterized. Convection parameterizations are known to be a principal error source in the simulation of precipitation. On the 4 km grid, convection parameterizations can be switched of because deep convection can be explicitly resolved (convection permitting simulation). A second advantage of the higher grid spacing is the improved representation of orography and surface fields. This is especially important in regions with complex terrain.

In this study the region of interest are the headwaters of the Colorado river (hereafter Colorado headwaters, or headwaters) located in the western part of Colorado. This region has a very complex terrain with mountain tops reaching more than 4000 m. To accurately simulate atmospheric fields in this environment is a challenging task for regional climate models.

In Chapter 2 the used model setup, domain and time period, reference data, and statistical methods are explained. Chapter 3 focuses on the climatologically (8 year average) performance of the simulations. In Chapter 4 the average June, July, and August (JJA) and December, January, and February (DJF) precipitation extreme event is evaluated and in Chapter 5 extreme value statistic is applied to evaluate the skill of the simulations to simulate extremes. Summary and conclusions are presented in Chapter 6.

2 Data and Methods

The Weather Research and Forecasting Model (WRF) version 3.1.1 (Skamarock et al. 2005) was used to simulate the eight year period 1st January 2001 to 31st December 2008 (plus three months of spin up) with 4 km, 12 km, and 36 km horizontal grid spacing over the domain in 2.1. The most important model settings are the usage of the Noah land surface model (Chen and Dudhia (2001), Ek et al. (2003)), Mellor–Yamada–Janjić (MYJ) planetary boundary layer scheme (Skamarock et al. 2005), Community Atmosphere Model's (CAM) long-wave and shortwave schemes (Collins and Coauthors (2006), and Thompson et al. (2008) cloud micro-physics scheme. While deep convection was parameterized within the 12 km and 36 km simulations with the Betts–Miller–Janjić scheme (Betts and Miller (1986) and Janjic (1994)) no convective parameterization was used in the 4 km run since convection is assumed to be simulated explicitly at this grid spacing (Weisman et al. 1997). The initial conditions and lateral boundary forcing for the simulations were taken from the North American Regional Reanalysis (NARR; Mesinger and Coauthors (2006)) which has an update frequency of three hours over North America with 32-km horizontal grid spacing.

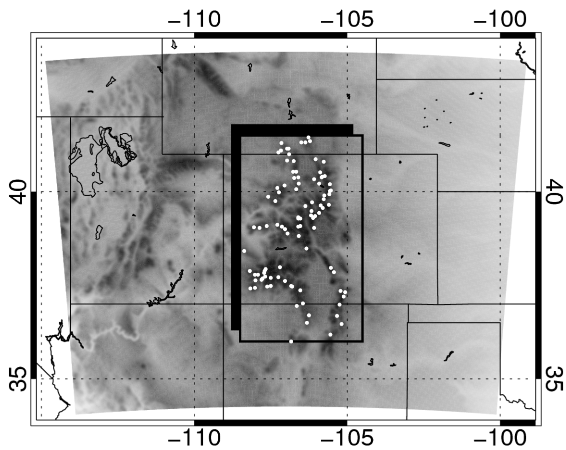

Figure 2.1: The contour shows the orography in the simulated domain. The Colorado headwaters are highlighted in the black rectangle. White dots show the location of SNOTEL stations.

2.1 Domain and Definition of Mountain Ranges

The region of interest are the headwaters of the Colorado river (hereafter Colorado headwaters or simply headwaters) which are displayed in Figure 2.1.

For the extreme value evaluations the Colorado headwaters region is split up in eight mountain ranges according to Figure 2.2. The locations of the snow telemetry (SNOTEL) stations is depicted as circles in the maps. There are different amounts of stations for the evaluation of the 8 year simulated period (99 stations, left panel) and the 30 year evaluations (39 stations, right panel) according to the availability of complete time series (no missing values).

2.2 Reference Data Sets

In this study three different reference data sets are used for the evaluation of the WRF simulations:

1. the Parameter-elevation Regressions on Independent Slopes Model (PRISM) data set,
2. the Climate Prediction Center (CPC) data set,
3. and the SNOTEL observations.

In the following subsections those data sets and their application in this study are explained.

2.2.1 SNOTEL

For the evaluation of extreme precipitation events observations from the SNOTEL station network (Serreze et al. 1999) were used. 99 of those stations are located within the Colorado headwaters region (see white dots in 2.1) and have a complete daily time series for the simulation period. The stations are located between 2400 m to 3500 m mean sea level where snowfall is highest. The measurements are performed with precipitation gauges and have a resolution of 2.5 mm.

2.2.2 PRISM

PRISM (Daly et al. 1994) is a gridded dataset for precipitation, dew point, and maximum and minimum 2 m air temperature for the entire continental USA. In this dataset a analytical tool incorporates point data, a digital elevation model, and expert knowledge of complex climatic extremes, including rain shadows, coastal effects, and temperature inversions are used to generate data on a 4 km times 4 km grid with a monthly resolution.

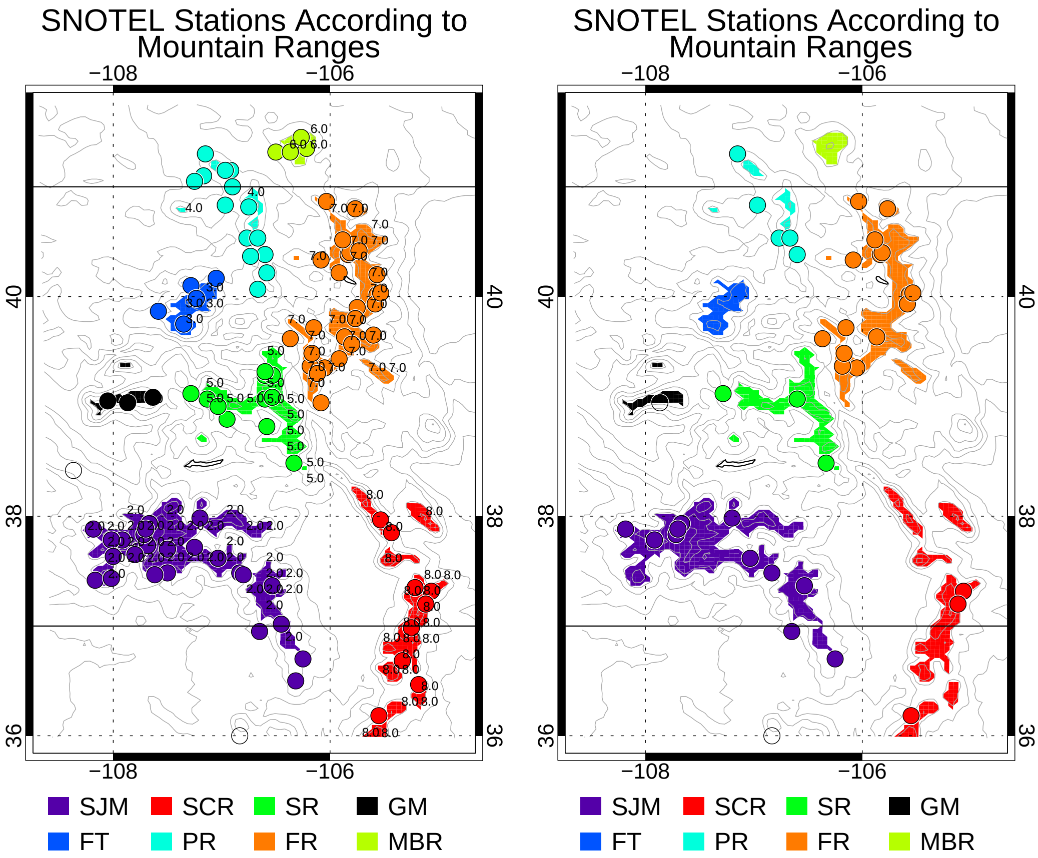

Figure 2.2: Definition of different mountain ranges within the Colorado headwaters region. The locations of the SNOTEL stations are depicted as circles. In the left panel the 99 stations for the 8 year evaluations and in the right panel the 39 stations for the 30 year evaluations are depicted. The acronyms for the mountain ranges stand for: SJM-San Juan Mountains, SCR-Sangre de Cristo Range, SR-Sawatch Range, GM-Grand Mesa, FT-Flat Tops, PR-Park Range, FR-Front Range, MBR-Medicine Bow Range.

Here the PRISM data set is only used for climatologically analyzes, because of its coarse temporal resolution.

2.2.3 CPC

The CPC data set(Higgins et al. 1996) provides precipitation data on a 0.25 degree latitude times 0.25 degree longitude grid for the entire continental USA on an daily basis. It uses station data from three sources: NOAA's National Climate Data Center (NCDC) daily co-op stations (1948-current), CPC dataset (River Forecast Centers data

+ 1st order stations - 1992-current), and daily accumulations from hourly precipitation dataset (1948-current). There are about 13000 station reports each day for 1992-current, and about 8000 reports before 1992.

In this report, the CPC data is often used as a second reference dataset (beside SNOTEL) to get a feeling for uncertainties in observed precipitation.

2.3 Calculating Station Values from Gridded Data

In many evaluations the observations from SNOTEL stations are taken as reference data set to evaluate the skill of the simulations. However, the WRF output is given on a latitude, longitude grid and so a direct comparison with station data is impossible. To make the gridded simulations comparable with the observations at single points, two methods are applied.

First, in case of the WRF-4km simulation the precipitation values of the closest four grid-points at every station are arithmetically averaged. This approach assumes that the four nearest grid-points at every station are equally contributing to the value at the point of the observation.

Second, for the coarser simulations (WRF-12km and WRF-36km) this approach is not valid any more, because the four nearest grid points can be up to 12 km/36 km apart from the observation. If the grid spacing is so coarse a better approach to derive point data is to weight the four closest grid points accordingly to their distance to the station. This method is called inverse distance weighting. Mathematically it can be written as follows:

$$u(x) = \sum_{i=1}^{N} \frac{w_i(x) u_i}{\sum_{j=0}^{N} w_j(x)} \tag{2.1}$$

where

$$w_i(x) = \frac{1}{d(x, x_i)^p} \tag{2.2}$$

In Equation 2.1 and Equation 2.2 the variables are:

$u(x)$	inverse distance weighted value at point x
N	number of grid-points for averaging (here 4)
$w_i(x)$	weights at location x
d	distance from the location of the station x to the grid points x_i
p	positive real number called the power parameter (in this study p=2)

2.4 The Discrete Cosine Transformation

The scale separation of the different atmospheric fields and the orography was performed with discrete cosine transformation (DCT). Denis et al. (2002) were the first who used the 2D DCT for limited areas. Therefore, the two dimensional field has to be mirrored at the position $i = j = -1/2$ to make it symmetric. Thereafter, the Fourier transformation can be applied, centered on $i = j = -1/2$. This is a special case of a Fourier transformation which is called DCT because the sine components of the Fourier series are zero for symmetric functions. Concerning a 2D field f_{ij} of N_i by N_j grid points, the direct and inverse DCT are defined as:

$$F(m,n) = \beta(m)\beta(n) \sum_{i=0}^{i=N_i-1} \sum_{j=0}^{j=N_j-1} f(i,j) \cos\left[\pi m \frac{(i+1/2)}{N_i}\right] \cos\left[\pi n \frac{(j+1/2)}{N_j}\right] \tag{2.3}$$

$$f(i,j) = \sum_{m=0}^{m=N_i-1} \sum_{n=0}^{n=N_j-1} \beta(m)\beta(n) F(m,n) \cos\left[\pi m \frac{(i+1/2)}{N_i}\right] \cos\left[\pi n \frac{(j+1/2)}{N_j}\right] \tag{2.4}$$

$$\beta(m) = \begin{cases} \sqrt{\frac{1}{N_i}}, & \text{for } m = 0 \quad (2.5\text{a}) \\ \sqrt{\frac{2}{N_i}}, & \text{for } m = 1, 2, \ldots, N_i - 1 \quad (2.5\text{b}) \end{cases}$$

$$\beta(n) = \begin{cases} \sqrt{\frac{1}{N_j}}, & \text{for } n = 0 \quad (2.6\text{a}) \\ \sqrt{\frac{2}{N_j}}, & \text{for } n = 1, 2, \ldots, N_j - 1 \quad (2.6\text{b}) \end{cases}$$

Thereby, f_{ij} is the value of the field at grid point number (i, j), and F_{mn} is the real spectral coefficient corresponding to the 2D-wave-number at (m, n). A more detailed derivation of the DCT can be found in the appendix of (Denis et al. 2002) and an application of this method on high resolved regional climate model (RCM) simulations can be seen in Kapper (2009).

To generate variance spectra of a 2D field, the variances have to be connected to a specific wavelength. To do so, the method of *binning* was suggested by Denis et al. (2002). It is based on dividing the wave-number field into multiple quarters of ellipses. The space between two ellipses can be connected to a specific wave-number for which the variances are summed up. For a detailed description see Denis et al. (2002).

2.5 Fitting of Extreme-Value Distribution

In Chapter 5 the Generalized Pareto distribution (GPD) and the Weibull distribution (WBL) are fitted to daily extreme precipitation events to calculate return values for several mountain ranges in the Colorado headwaters. In the following two subsections the concept of extreme value distributions and peak over threshold methods are explained.

2.5.1 Extreme-Value Distribution

Extreme values are defined statistically as values that are differing strikingly from the statistical mean. Therefore extreme data are rare and either unusually large or small. A common example for extreme values are a collection of annual maximums or block maxima (the maximum value from a block of m elements). An extreme-value dataset is derived for example by collecting the hottest day of each year out of a time series of 30 years.

In extreme-value statistics, it can be shown that the distribution of such an extreme-value dataset will fit a known distribution increasingly close with increasing m, independent of the distributions of the observation or simulation (e.g., (Leadbetter 1983), (Coles 2001)). This result is called the Extremal Types Theorem or Fisher-Tippett-Theorem (Fisher and Tippett 1928) and is the equivalent in the extreme value theory to the Central Limit Theorem which says that distributions of sums are converging to a Gaussian distribution.

The distribution derived is called general extreme value (GEV) distribution which has the following probability density function (PDF):

$$f(x) = \frac{1}{\beta}\left[1 + \frac{\kappa(x-\zeta)}{\beta}\right]^{1-1/\kappa} exp\left\{-\left[1 + \frac{\kappa(x-\zeta)}{\beta}\right]^{(} - 1/\kappa)\right\}, 1 + \kappa(x-\zeta)/\beta > 0. \tag{2.7}$$

In Equation 2.7 ζ is a location parameter, β is a scale parameter, and κ is a shape parameter. By integrating Equation 2.7 the cumulative density function (CDF) is derived:

$$F(x) = exp\left\{-\left[1 + \frac{\kappa(x-\zeta)}{\beta}\right]^{(} - 1/\kappa)\right\}. \tag{2.8}$$

By inverting Equation 2.8 a formulation for the quantile function can be formulated:

$$F^{-1}(p) = \zeta + \frac{\beta}{\kappa}\left\{[-ln(p)]^{-\kappa} - 1\right\}. \tag{2.9}$$

The parameters κ, β, and ζ can be estimated by using the method of maximum likelihood (see e.g., Wilks (2005)) or the L-moments method (Hosking (1990), Stedinger et al. (1993)) which is more often used for small data samples.

There are three special cases of the GEV distribution:

1. Gumbel, or Fisher-Tippett Type I distribution (derived when $\kappa = 0$),
2. Frechet, or Fisher-Tippett Type II distribution (for $\kappa > 0$),
3. Weibull, or Fisher-Tippett Type III distribution (for $\kappa < 0$),

which have different properties. More information about this types can be for example found in Wilks (2005).

Often derived results of extreme value analysis are quantities of large cumulative probabilities like the value of an event with an annual probability of 0.01. However, as long as the number of years n is not very large it is not possible to directly estimate those quantities. Instead a well fitted extreme-value distribution can provide values of probabilities which are larger than $1 - 1/n$.

In this report the WBL distribution is fitted to observed and simulated daily precipitation values. The advantage of using the WBL distribution instead of the GPD which is treated in the next subsection is that with the entire distribution of precipitation can be fitted with the WBL distribution. In case of the GPD only extreme values above a certain threshold are fitted.

2.5.2 Peaks over Threshold (POT) and Return Periods

Pickands (1975) showed that values of the tail of a distribution are following a GPD asymptotically if the parent distribution belongs to one of the three above mentioned extreme value distributions. The GPD is expressed as follows:

$$F(x) = 1 - \left(1 + \frac{\kappa(x-\zeta)}{\beta}\right)^{-\frac{1}{\kappa}}, \tag{2.10}$$

and is defined as the GEV distribution by three parameters (location ζ, shape κ, and scale β). After the parameters are fitted to the data, the derived probabilities can be translated into return periods $R(x)$:

$$R(x) = \frac{1}{\omega\left[1 - F(x)\right]}, \tag{2.11}$$

where ω is the average sampling frequency of the sample time frame ($\omega = 1/n$). The return period is the typical time span in which an event of the magnitude x is expected to occur once. If for example, annual maximum data are considered $\omega = 1\text{yr}^{-1}$ for an event with the cumulative probability $F(x) = 0.99$ which has a probability of $1 - F(x)$ in any given year, the value of this event will be associated with a return period of 100 years and will be called the 100 year event.

2.6 Scale Dependent Evaluation of Extremes

To find out how precipitation extremes look like at different horizontal scales and how the WRF simulations perform at those scales the following method is applied:

1. a squared search window is defined ranging in steps from zero to the twofold length of the maximum size (lat or lon) of the domain. The length of one side of this search window is denoted as horizontal scale.

2. the search window is centered above the location of each station and the daily time series of all stations within the search window are averaged. This leads to one new (spatially average) daily time series per station and can be thought of horizontal smoothing of the data.

3. all extreme events (above the 0.975 percentile) are averaged for each station to derive the value for an average extreme at this location and search window size.

4. the extreme values of all stations are averaged and their standard deviation is calculated as an estimation of spatial variability of precipitation extremes at this scale.

5. steps one to four is repeated for all considered spatial scales.

Within the above described method, the minimum search window size of zero shows the average extreme event on a station basis while the maximum window size (twofold maximum domain length) reveals the average precipitation extreme within the headwaters region.

2.7 Scale Dependent Correlation and Standard Deviation

The below described method enables to evaluate the correlation coefficient and normalized standard deviation between simulated values at the SNOTEL locations and the SNOTEL observations for different horizontal scales:

1. a squared search window is defined ranging in steps from zero to the twofold length of the maximum size (lat or lon) of the domain. The length of one side of this search window is denoted as horizontal scale.

2. the search window is centered above the location of each station and the average 18 most extreme precipitation events (according to the SNOTEL data) of all stations within the search window are averaged. This leads to one new (spationally average) mean extreme value per station and can be thought of horizontal smoothing of the data.

3. this averaging is done for all datasets and the SNOTEL observations.

4. the correlation coefficient and normalized standard deviations of the new field is calculated between the correspondent data set and the SNOTEL values.

5. repeat steps one to four for all considered spatial scales.

With this method large scales above ∼400 km should not be interpreted because the values of all stations get very similar (or identical at the largest scale). This means, the independent sample size for calculating the correlation coefficients and normalized standard deviations gets too small to be meaningful.

2.8 Correlogram and Semivariogram

Correlograms and Semivariograms are often used in Geostatistic. The basic idea of evaluating metrological fields with the variogram method is related to publications of Gebremichael and Krajewski (2004), Germann and Joss (2001), Harris et al. (2001), and Zepeda-Arce et al. (2000). The method itself was proposed by Marzban and Sandgathe (2009) and compares two fields in terms of their covariance structures. Thereby, the fields can be compared in two different ways, where one accounts for displacements and intensity errors, while the other is only sensitive to intensity errors. The analogy, which is insensitive to intensity and only accounts for displacements can be also calculated and is called *correlogram*.

The principal idea behind correlograms and semivariograms is to derive space depeneden values for the correlation and variance. The method works as follows:

1. chose a separation distance, usually referred as "lag" (h) which is the distance between two measurements. Because of the irregular distribution of observations a "lag tolerance" is needed (half of the distance between two lags to sample the entire space).

2. for each station, search for those surrounding stations which are within the search distance (lag ±lag tolerance).

3. store the data-pairs (value at each station ($z(u)$) with values within the search distance of the correspondent stations –> laged values ($z(u + h)$)).

4. calculate the correlation ($\rho(h)$) and the semivariance ($\gamma(h)$) between the values of the stations and the laged values with the following formulas for the correlation:

$$\rho(h) = \frac{C(h)}{\sqrt{\sigma_0 \cdot \sigma_{+h}}} \tag{2.12}$$

$$\sigma_0 = \frac{1}{N(h)} \sum_{\alpha=1}^{N(h)} z(u_\alpha) \tag{2.13}$$

$$\sigma_{+h} = \frac{1}{N(h)} \sum_{\alpha=1}^{N(h)} z(u_\alpha + h) \tag{2.14}$$

and for the semivariance:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [z(u_\alpha + h) - z(u_\alpha)]^2 \tag{2.15}$$

where the symbols are:

u	vector of spatial coordinates (with components for latitude and longitude)
$z(u)$	variable under consideration as a function of spatial location (e.g., precipitation at stations)
h	lag vector representing separation between two spatial locations
$z(u+h)$	lagged version of variable under consideration
$N(h)$	number of pairs separated by lag h $\pm$lag tolerance
σ_0	standard deviation of the $z(u)$ variables
σ_{+h}	standard deviation of the $z(u+h)$ variables
$\rho(h)$	correlation coefficient for lag h
$\gamma(h)$	semivariance for lag h

5. the steps 1 to 4 are repeated for different lags until the entire area of interest is sampled.

3 Climatological Evaluation

In this chapter the climatological performance of the three WRF simulations is analyzed.

First, in Section 3.1 the orography and several atmospheric fields in the Colorado Headwater region are analyzed with a DCT. Than the 8 year simulated period is compared with a 30 year climatology in Section 3.2 to find out how representative the simulated 8 years are in a climatological sense. In addition the importance of extreme events for the hydrology inn the headwaters region is treated in Section 3.3. Section 3.4 deals with simple bias maps for the entire eight year period and the splitting into seasons and reveals insights in the accuracy of precipitation sums in different areas of the study region.

In Section 3.5 the temporal performance of the WRF simulations is evaluated by looking at differences in the daily precipitation accumulation in each season compared to SNOTEL data.

The diurnal variations of the bias and the spatial correlation coefficients are analyzed in Section 3.6. If there is an elevation dependence of the bias is investigated in Section 3.7.

3.1 Scale Separation of the Orography and Atmospheric Fields

In Figure 3.2 the average absolute gradient in the orography field within the Headwaters domain is shown as a function of horizontal grid spacing. The used orography field has an original grid spacing of 1.2 km and is successively regridded to coarser grids.

The highest decrease of the average gradient of orography can be found at the left side of the plot. By using a 12 km grid instead of the original 1.2 km grid the average absolute gradient decreases by more than 50 %.

There is still a lot of fine scale orographic structure visible in sup-panel a) (cross section with 4.4 km grid spacing). This fine scale structures disappear by increasing the grid spacing to 12 km (panel b). On the 12 km grid single mountain ranges have still multiple peaks. This feature disappears by decreasing the grid spacing to 36 km (panel c). On a 100 km grid (panel d) there are no separated mountain ranges any more and the orography of the cross section appears as a single flat mountain.

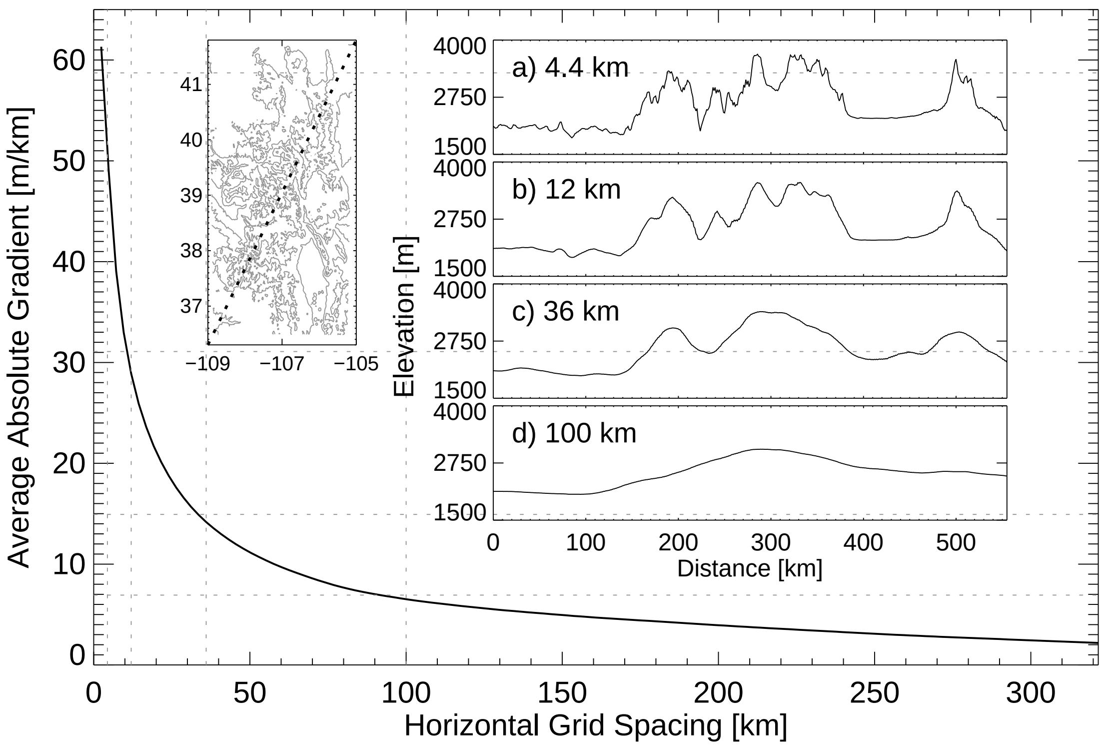

Figure 3.1: In the large panel the average absolute orography gradient in the Headwaters region is shown as a function of horizontal grid spacing. The small sub-panels show the orography along a cross section (shown in the contour plot) for 4 km (panel a), 12 km (panel b), 36 km (panel c), and 100 km (panel d).

The accumulated variance spectra of different wavelengths of the orography in the Colorado headwaters is shown in Figure 3.2 for the three simulated grid spacings.

The WRF-4km orography has higher variability for small wave lengths compared to the coarser simulations. The effective resolution of the orography in the WRF-12km simulation can be estimated from the point where the 4 km and 12 km lines start to diverge. This is approximately at 70 km which is 6 times the grid spacing of the 12 km run. This ratio is also similar for the WRF-36km orography.

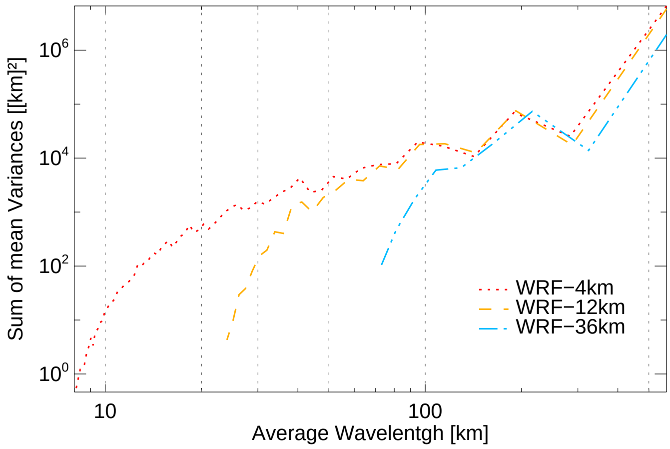

Figure 3.2: DCT of the orography in the Colorado headwaters region in the three different grids spacings: 4 km (red dotted), 12 km (yellow dashed), and 36 km (blue dotted-dashed).

In Figure 3.3 the DCT is applied to the annually and seasonally average precipitation fields. The spectra look very similar to those of orography (Figure 3.2) which shows the strong influence of the mountains on spatial precipitation patterns. The peak at ∼40 km wavelength is clearly visible in the WRF-4km and also in the WRF-12km simulation. The largest differences to the orography spectra can be found for June, July, and August (JJA) (panel d).

The effective resolution of the WRF-12km simulation varies between 4 (in March, April, and May (MAM)) to 8 (in September, October, and November (SON)) times the grid spacing. A similar ratio is found for the WRF-36km run.

The simulated spectra differ most in JJA where the WRF-12km and the WRF-36km have much higher variability at large scales compared to the WRF-4km simulation.

Also the 2 m temperature fields in Figure 3.4 are strongly correlated with the orography. This is expected because of the height dependency of temperature. The peak at ∼40 km is clearly visible in WRF-4km and WRF-12km run in all seasons.

The spectra look relatively similar in all seasons with slightly higher values in JJA (panel d) and lower values in December, January, and February (DJF).

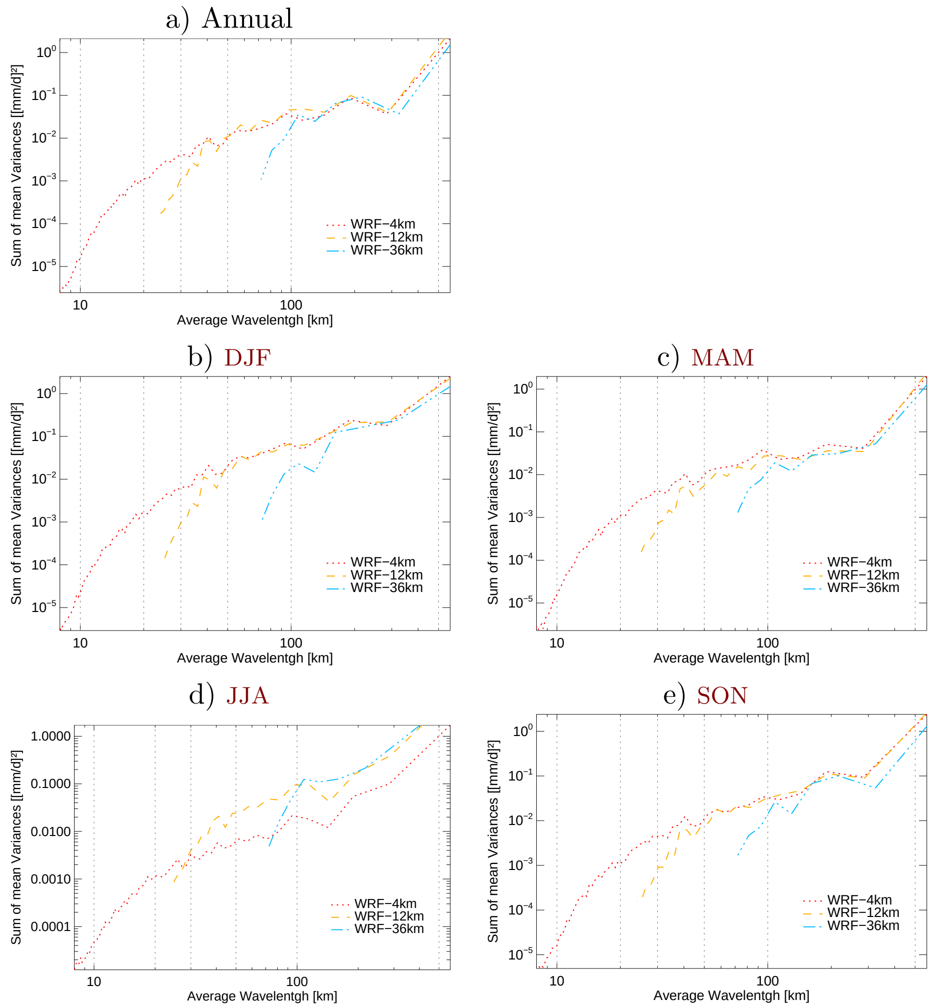

Figure 3.3: As in Figure 3.2 but for precipitation annually and in different seasons.

The effective resolution in the WRF-12km run is approximately 6 times the grid spacing. A similar ratio is found for the WRF-36km simulation.

For snow water equivalent the spectra from the DCT are displayed in Figure 3.5. Due to the height dependency of the parameter the spectra are strongly correlated with the spectra of orography. Several peaks in the orography spectra are even amplified in the spectra of snow water equivalent.

Clearly visible are higher variances at small wavelengths with increasing grid spacings. This is especially drastic in JJA (panel d) because there is much less snow in the WRF-

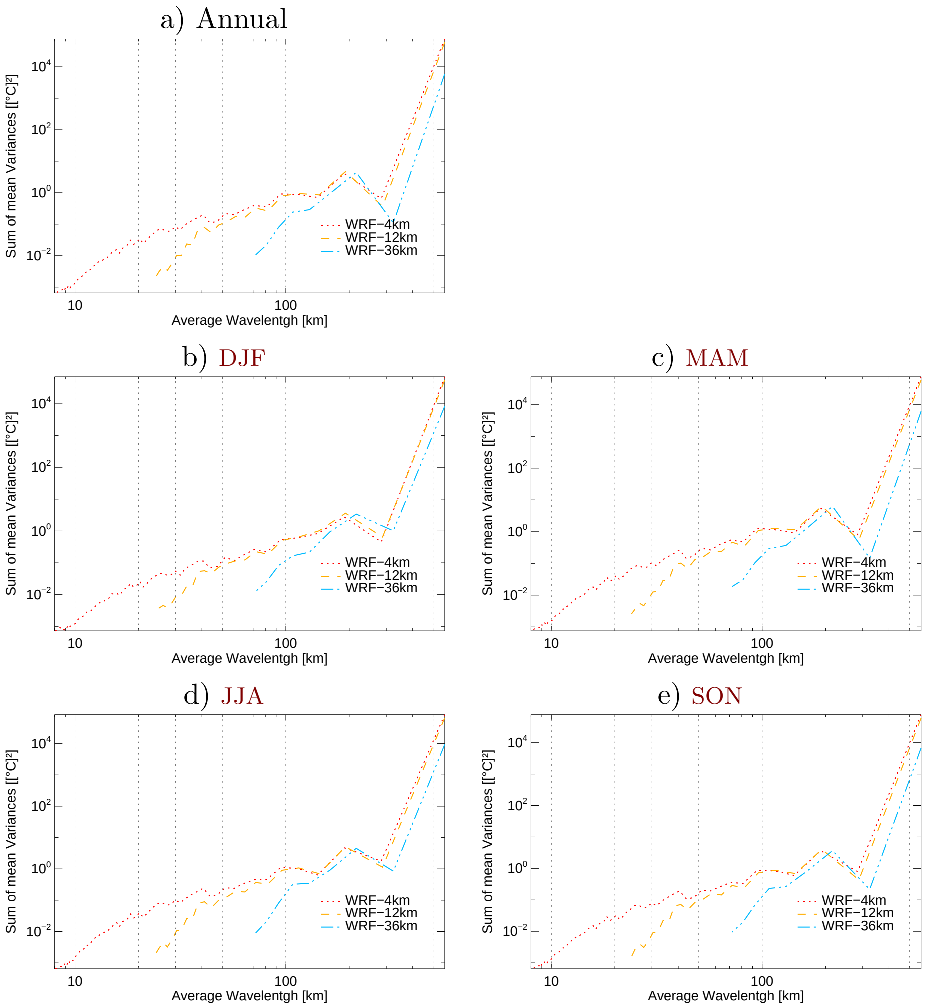

Figure 3.4: As in Figure 3.2 but for temperature two meters above ground annually and in different seasons.

12km and WRF-36km simulations due to the coarse orography (lower mountain height).

In Figure 3.6 the spectral decomposition of west wind 10 m above surface is shown. As for the previous parameters the spectra are strongly correlated wit the spectra of the underlying orography.

The spectra look very similar in different seasons. In JJA (panel d) the variance sums are lowest because of the generally lower wind speed in this season.

The effective resolution of the WRF-12km simulation is approximately 4 times the grid

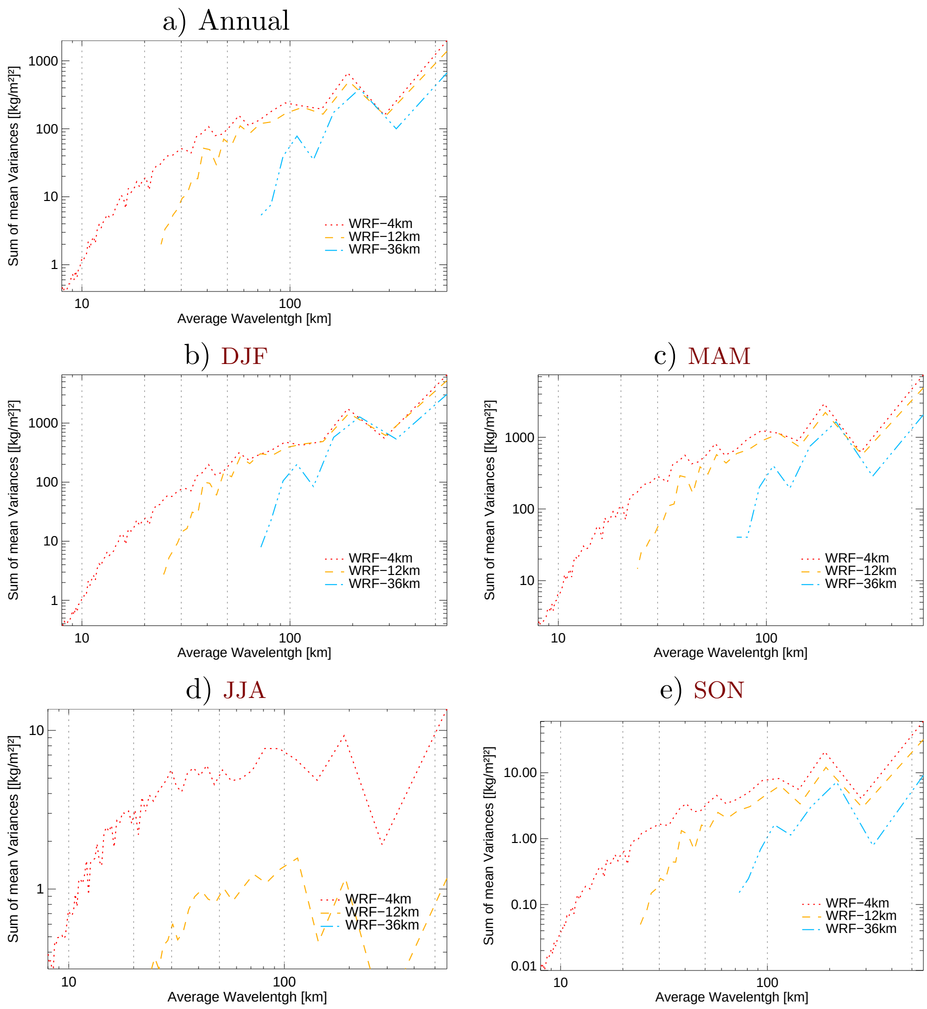

Figure 3.5: As in Figure 3.2 but for snow water equivalent annually and in different seasons.

spacing. For the WRF-36km run this factor is ∼3.

The spectra of the northward wind component are not shown because they look similar to the spectra of the westerlies.

The spectra of water vapor mixing ratio 2 m above ground is displayed in Figure 3.7. The spectra show only large scale variabilities above 30 km to 50 km. The correlation with the orography is relatively weak and strongest in DJF (panel b).

The effective resolution of the WRF-12km simulation varies with season but is ∼8 times the grid spacing. For the WRF-36km run this ratio is with ∼5 lower.

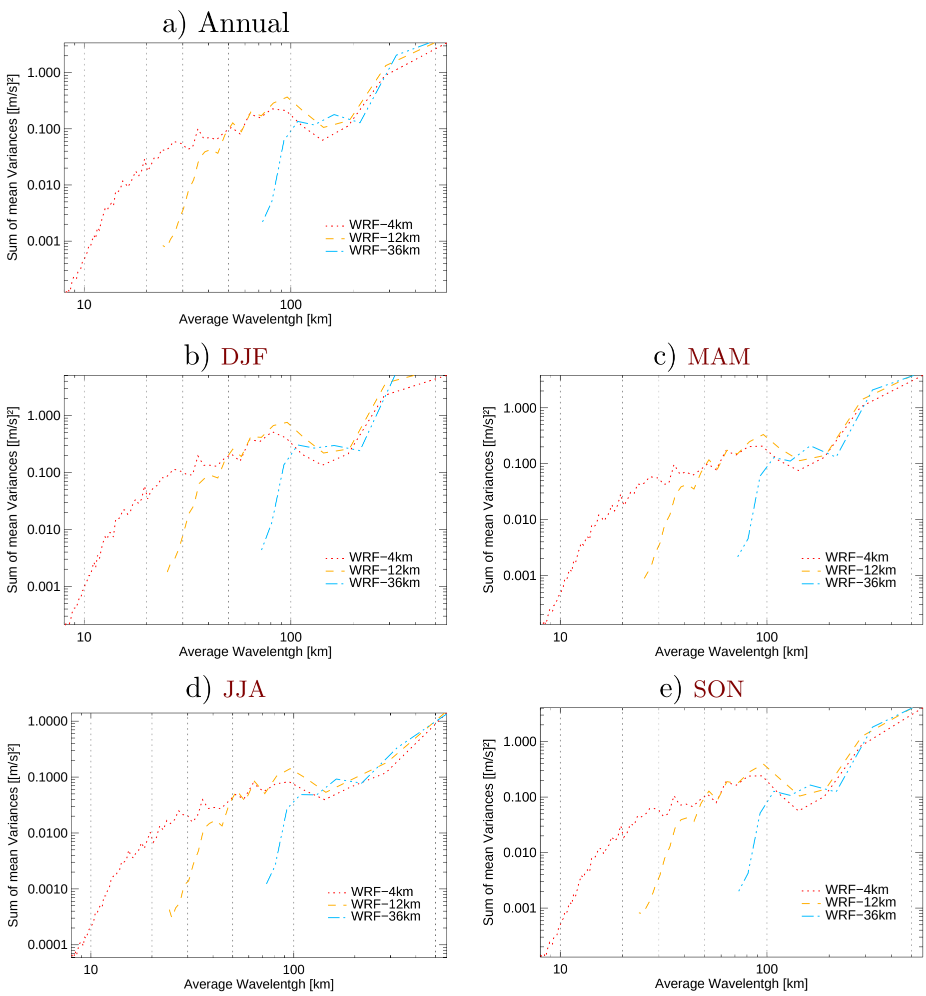

Figure 3.6: As in Figure 3.2 but for west wind 10 m above surface annually and in different seasons.

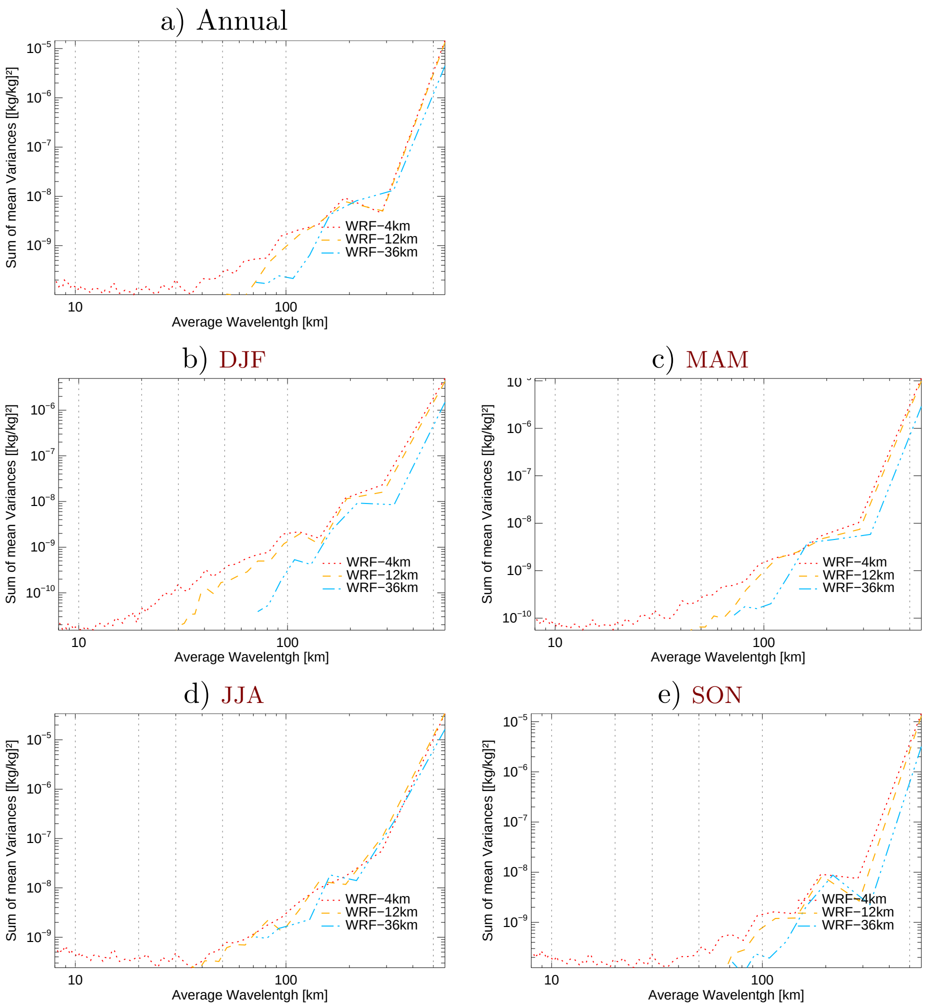

Figure 3.7: As in Figure 3.2 but for water vapor mixing ratio 2 m above ground annually and in different seasons.

3.2 Weather in the Simulated Period

To investigate how representative the simulates 8 year period is compared to a 30 year climatology, Figure 3.8 shows the average daily precipitation values for 39 SNOTEL stations within the headwaters region from October 1st 1980 to September 30th 2009. The 39 stations were chosen because they have a complete daily time-series within these period. The location of the stations can be seen in the right panel of Figure 2.2.

The left panel in Figure 3.8 shows that there is no visible trend in the total precipitation and the extremes during the considered period. The mean precipitation between October 1st 1980 to September 30th 2009 is 2.3 mm/d and therefore very similar to the average precipitation between 2001 to 2008 (2.2 mm/d). Also the day to day variability is very similar (30 year standard deviation: 2.9 mm/d, 8 year standard deviation: 2.8 mm/d). In average there were 18.7 precipitation events per year which were above the 0.95 percentile in the 30 year period while there were 17.9 in the 8 year period.

The box-whisker plot in the right panel shows that there was slightly more precipitation during DJF in the 8 year period and slightly less in MAM. This difference is more amplified if only the extreme events are considered. JJA and SON is more similar between the two periods.

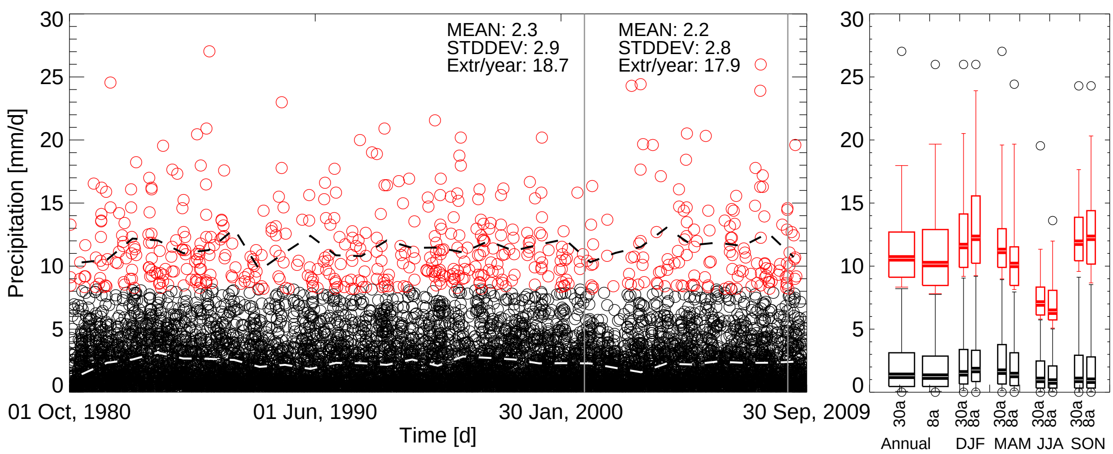

Figure 3.8: Average daily precipitation values from 39 SNOTEL observations from October 1st 1980 to September 30th 2009. Each circle in the left panel corresponds to a daily precipitation value. The red circles are values above the 0.95 percentile and depict extremes. The white and black dashed lines show the annual moving averages of the entire data and the extreme events. The two solid vertical lines frame the simulated 8 year period. In the right panel box whisker plots depict the quantiles and extremes of the entire 30 year period and the 8 year simulation period annually and in different seasons. Black boxes and whiskers correspond to the entire data while red boxes and whiskers depict the extremes.

3.3 On the Importance of Extreme Precipitation

Annually, 36 % of the precipitation within the Colorado headwaters region felt within 10 % of the days (large panel in Figure 3.9) in the 8 year simulation period. This demonstrates the importance of extreme events for the water balance in this region.

The importance of extremes is especially high in SON (small bottom right panel), where 45 % of the precipitation felt within 10 % of the days. The ratio in the other seasons is similar to the annual ratio.

The WRF simulation are overestimating the importance of extreme events. Compared to SNOTEL they have too many light precipitation events (between 0.1 mm/d and 1 mm/d in the headwaters average), underestimate the medium events (between 1 mm/d and 5 mm/d) and again overestimate the extreme events (above 5 mm/d). If only the mean precipitation is regarded those errors are canceling out and the average difference is nearly zero (e.g., see Figure 3.10).

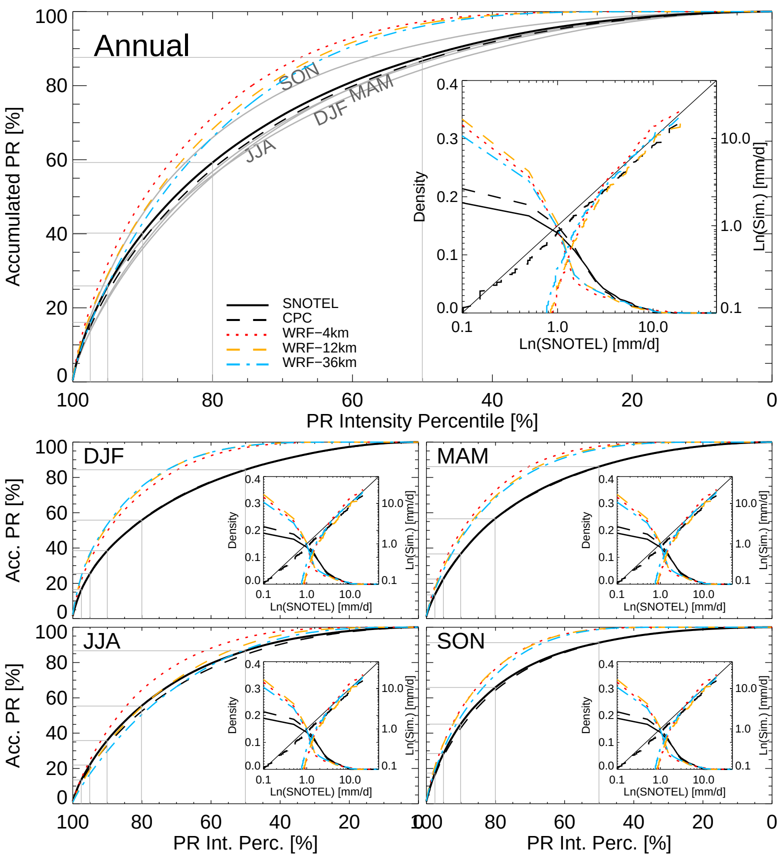

Figure 3.9: Large plots: cumulative density function of precipitation at SNOTEL locations for different seasons. Small plot: empirical density functions (left y-axis) and q-q plots (right y-axis) for precipitation at the SNOTEL locations.

3.4 Bias Maps

3.4.1 Annual

Figure 3.10 shows the precipitation fields of the WRF simulations and and the differences to the PRISM reference data for the average eight year period.

The WRF-4km simulation (panel b) looks very similar to the PRISM reference data (panel a) which is also visible in the difference field (panel B). The mean bias is 0.1 mm/d and the standard deviation in the difference field is 0.2 mm/d.

The WRF-12km simulation shows a clear underestimation of precipitation along the mountain ridges and an overestimation westward (downstream) of them. In average positive and negative differences are canceling out and the remaining average difference is 0.3 mm/d.

The WRF-36km run shows a better performance than the WRF-12km simulation. There is nearly no underestimation of precipitation along the mountain ridges but a small overestimation in the valleys. The average difference to the PRISM data-set is 0.4 mm/d with an standard deviation of 0.4 mm/d.

In Figure 3.12 the annual mean precipitation in the PRISM data-set and in the WRF simulations are compared to SNOTEL measurements.

The WRF-4km simulation (panel c and C) has an impressing good agreement with the observations and has comparable small average differences than the PRISM data-set.

As already seen in Figure 3.10 the WRF-12km run underestimates the precipitation along mountain ridges. Only in the Front Range and at lower elevated stations there is partly an overestimation visible.

The WRF-36km simulation shows an slight overestimation of precipitation especially in the south east part of the domain.

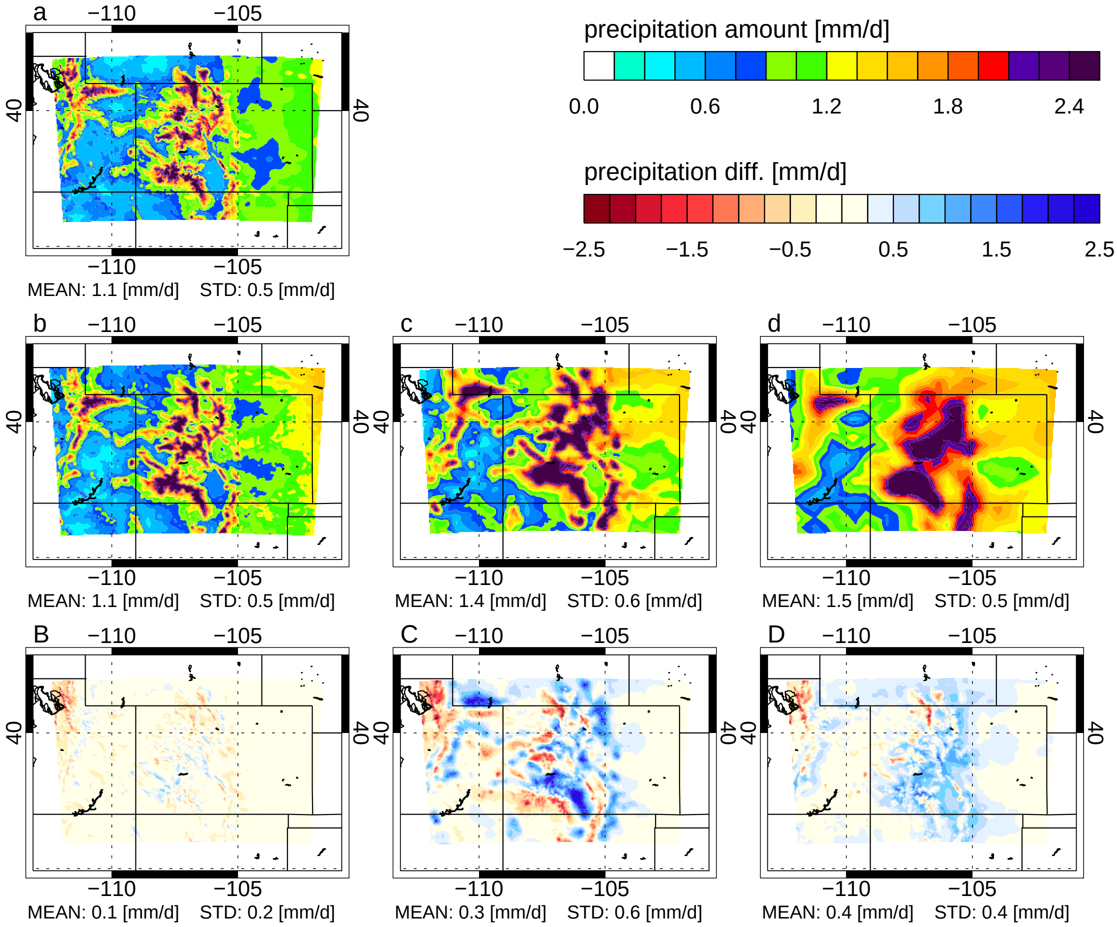

Figure 3.10: In panel a the 8 year temporal mean precipitation field of the PRISM data-set is displayed. Panel b and c show average precipitation for the WRF 4km and WRF 12km simulation while panel B and C are displaying their differences WRF minus PRISM data. The mean and the standard deviation of the correspondent field is written below each panel.

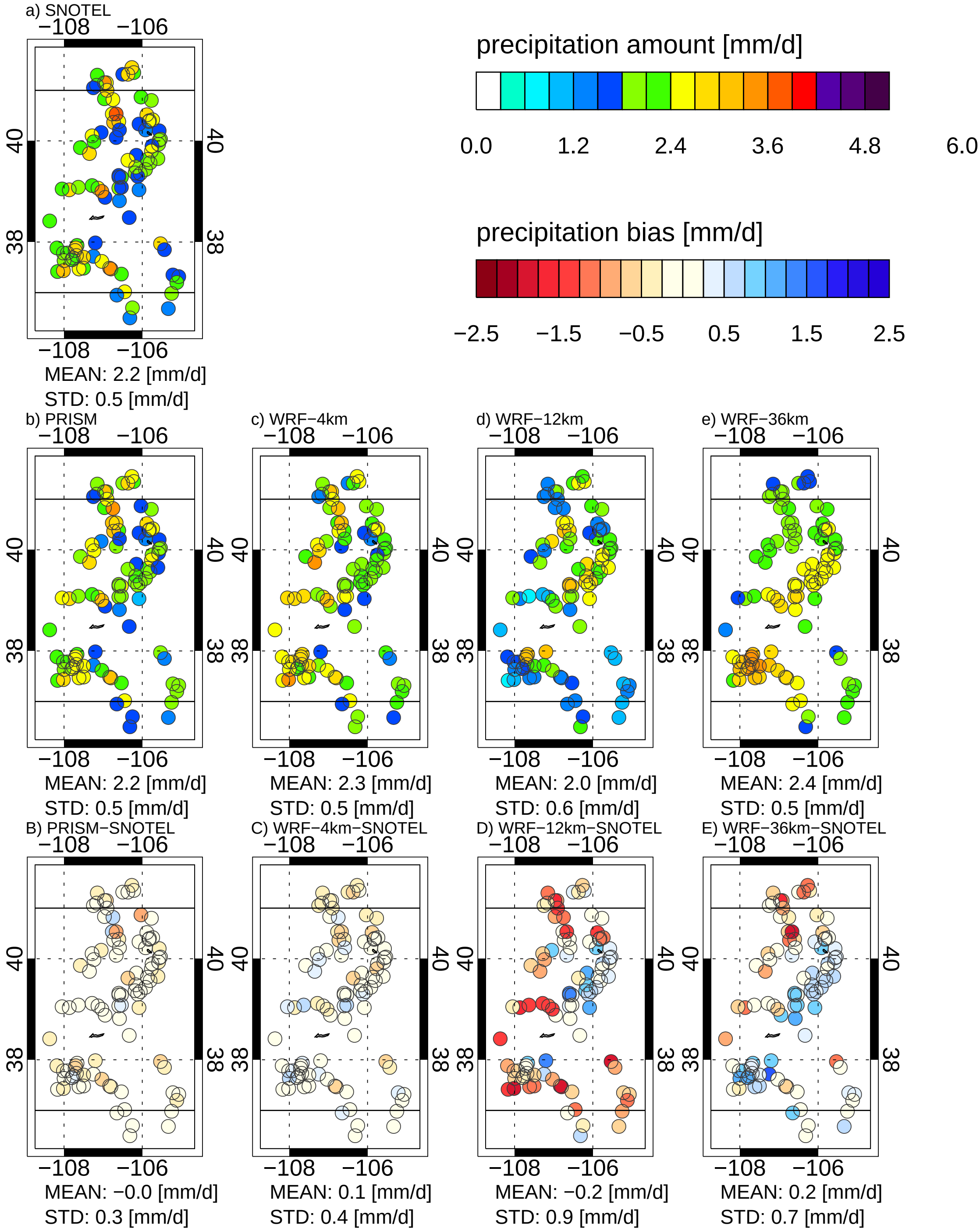

Figure 3.11: Similar as in Figure 3.10 but with SNOTEL as reference data-set.

3.4.2 DJF

The representation of precipitation in DJF is shown in Figure 3.12. Again, as annually, the WRF-4km simulation shows nearly no differences to the PRISM data-set (panel B). In DJF.

The pattern of overestimating precipitation in the west of mountain ridges and underestimating it on the ridges is strongly pronounced in the DJF WRF-12km simulation (panel c and C).

A similar dry area along mountain ridges is visible in the WRF-36km run (panel (d and D). However, there is less overestimation of precipitation in the valleys than in the WRF-12km simulation.

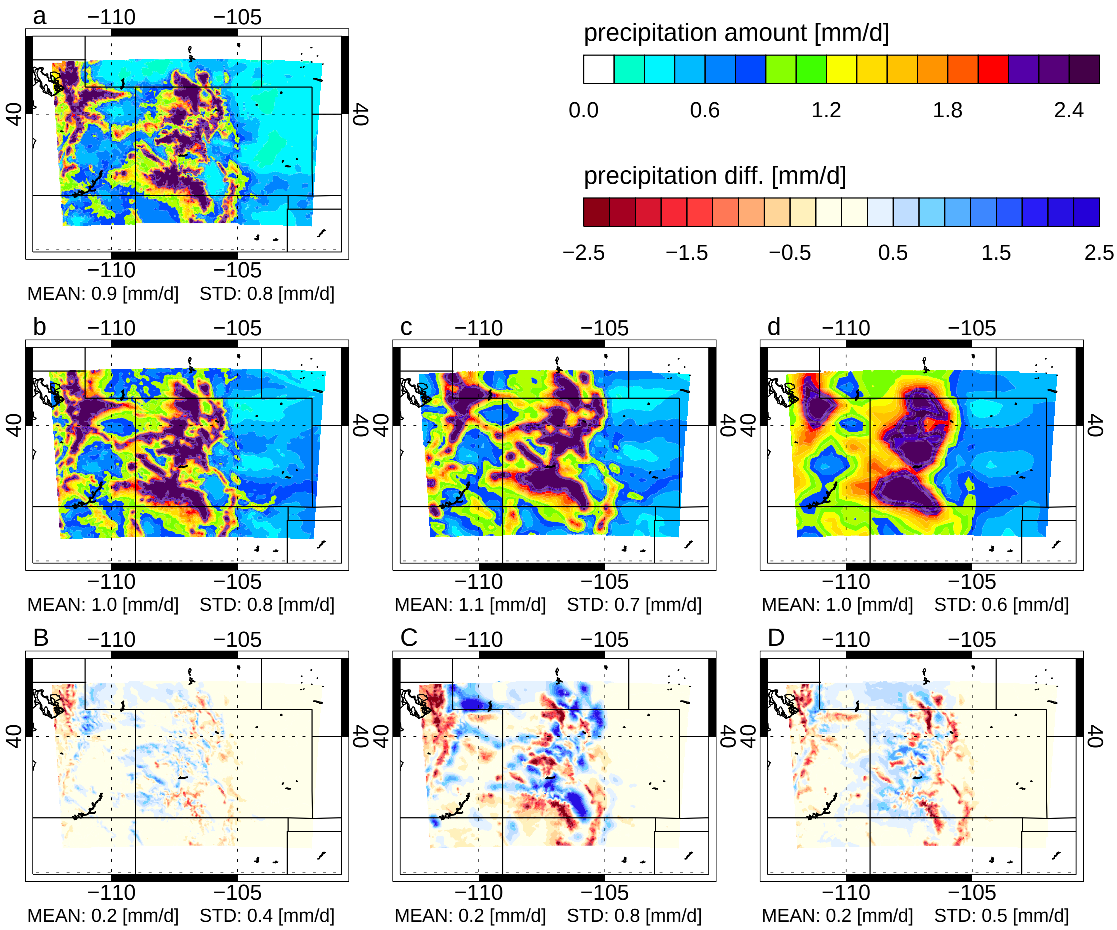

Figure 3.12: As in Figure 3.10 but for DJF.

The comparison with SNOTEL stations in Figure 3.13 shows a nearly biasfree WRF-4km run (panel b and B) which performs as good as the PRISM reference data-set.

The WRF-12km and WRF-36km simulations (panel d, D, e and E) show a relatively similar precipitation pattern with the already known underestimation of precipitation along mountain ridges. The WRF-36km has a slightly better performance in terms of standard deviation of the difference field.

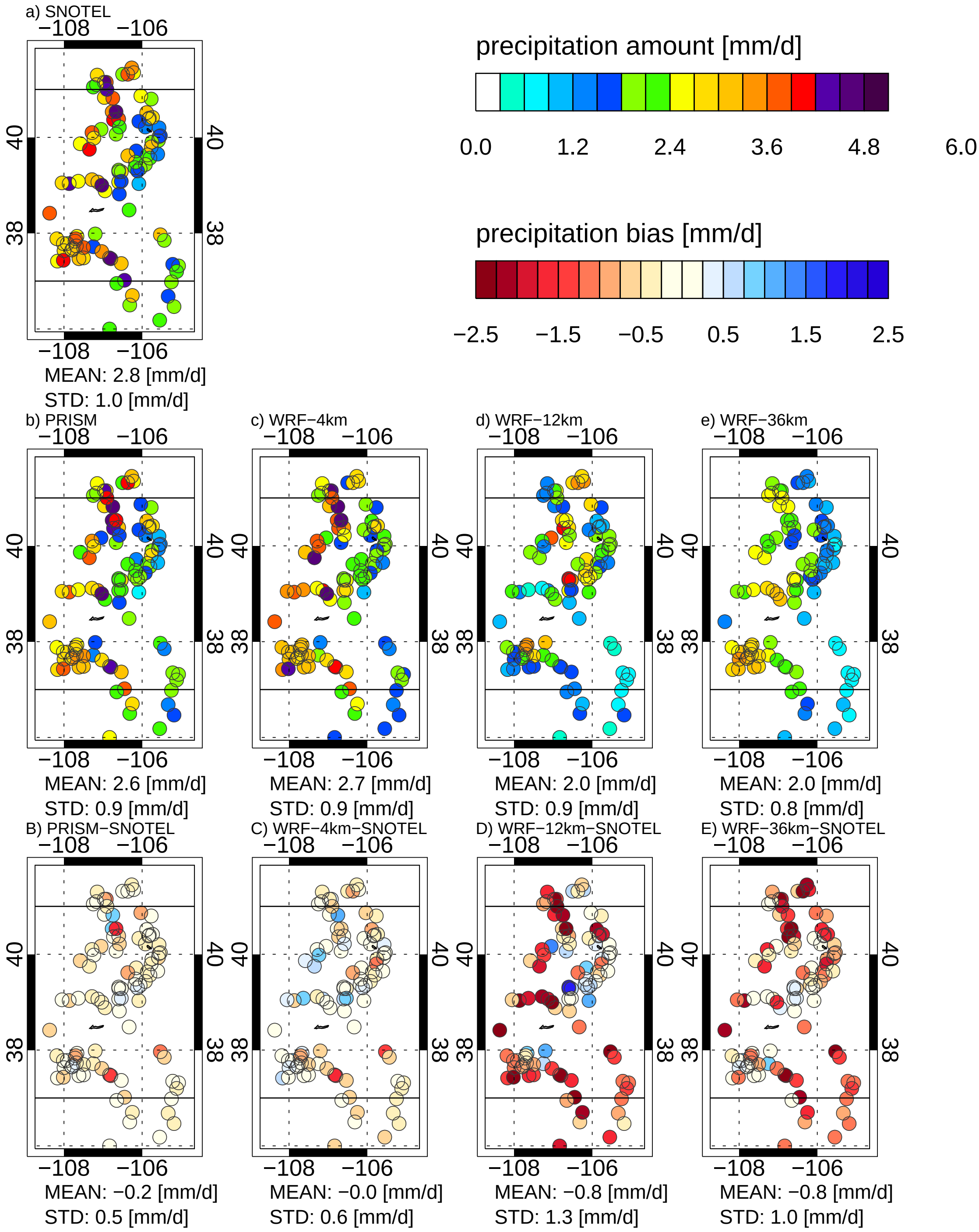

Figure 3.13: As in Figure 3.11 but for DJF.

3.4.3 JJA

In JJA there is an east-west gradient in precipitation with heavy precipitation through deep convection on the plains and less precipitation in the mountains and the great basin. The WRF-4km simulation (panel B) has again a remarkable good agreement with the PRISM data-set over large areas of the domain.

The WRF-12km simulation (panel C) overestimates the precipitation along the mountain ridges. A similar pattern is visible for the WRF-36km run with less details in the difference field compared to the WRF-12km simulation.

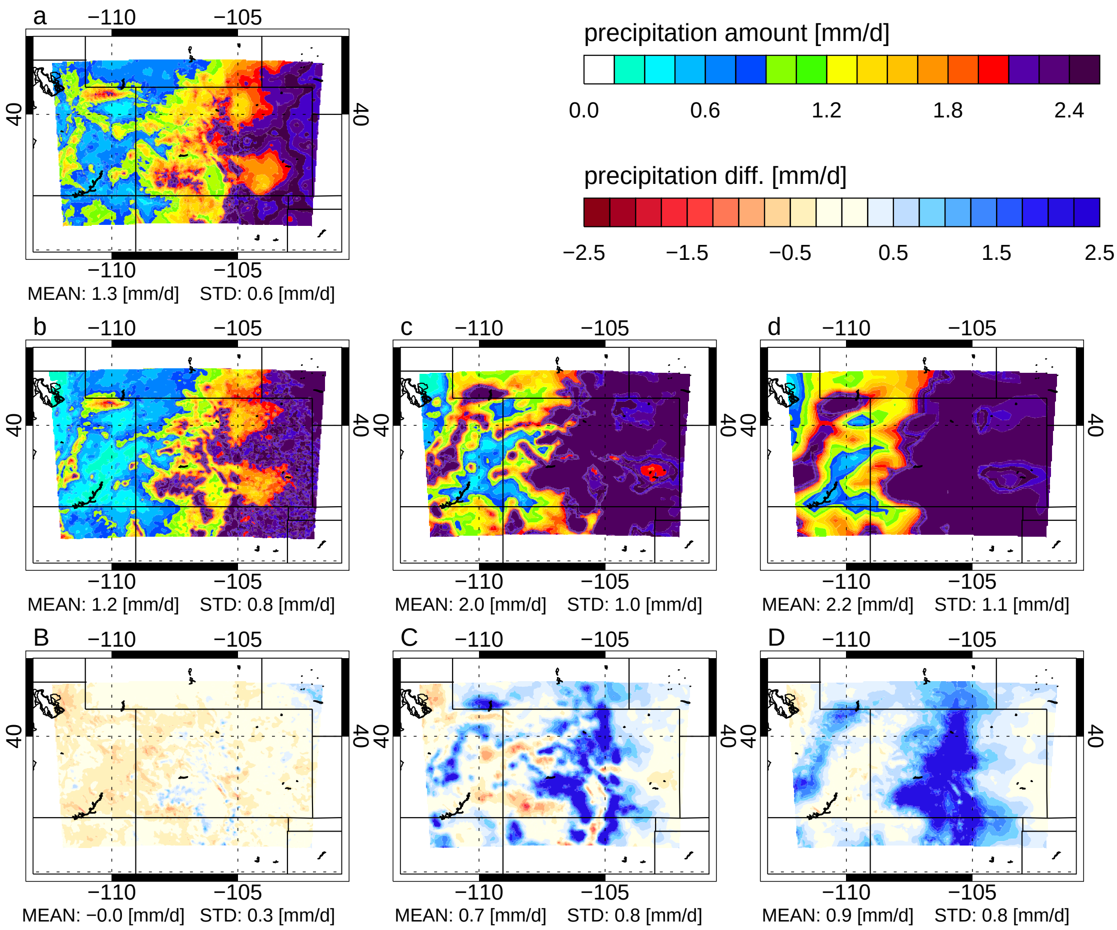

Figure 3.14: Same as in Figure 3.10 but for JJA.

The WRF-4km simulation (panel c and C in Figure 3.15) has a slight wet bias in JJA especailly in the southern part of the domain.

The WRF-12km simulation (panel d and D) overestimates precipitation especially in the Front Range while results in the San Juan Mountains and the Sangre de Cristo Range are better.

In the WRF-36km (panel e and E), precipitation is overestimated in the entire domain. The lowest overestimation can be found in the north west (Park Range, Medicine Bow Range, and Sawatch Range).

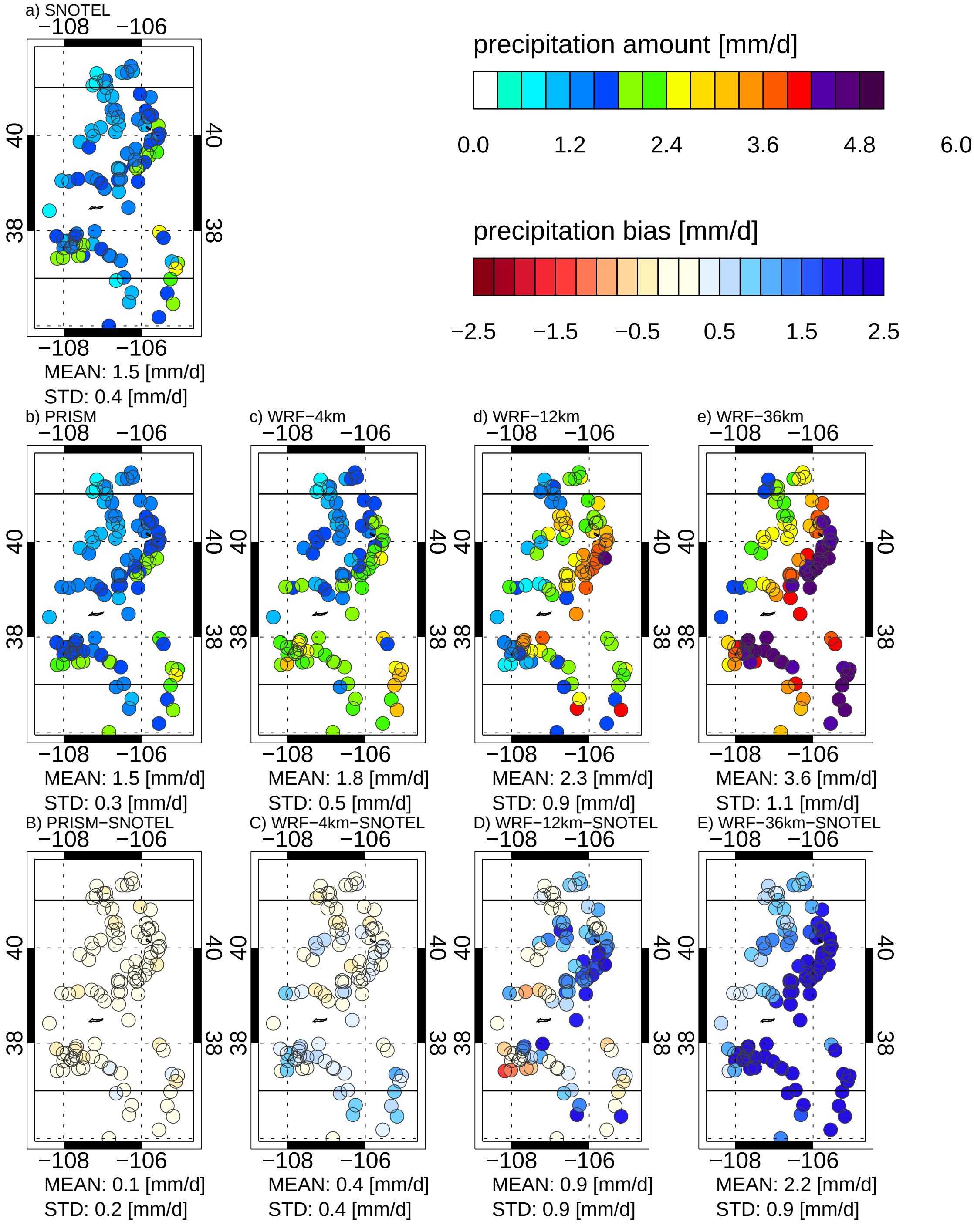

Figure 3.15: As in Figure 3.11 but for JJA.

3.5 Daily Precipitation Accumulations

Figure 3.16 displays the daily precipitation accumulations within each season of the eight year simulation period. The WRF 4km simulation (red line) shows a remarkable good agreement with the SNOTEL observations (black line) in every season.

In JJA the WRF 12km (yellow line) and WRF 36km (blue line) are overestimating precipitation in all eight simulated summer seasons. The overestimation in the WRF 12km simulation is smaller compared to the WRF 36km run. The WRF 4km simulation shows the best agreement with the observations.

In DJF and MAM precipitation amounts are typically underestimated in the WRF 12km and WRF 36km simulations. Those two simulations show very similar results during DJF and MAM. The WRF 4km simulations has only small deviations from the observations.

Differences between the simulations are smallest in SON and the precipitation sums agree fairly well with the observations.

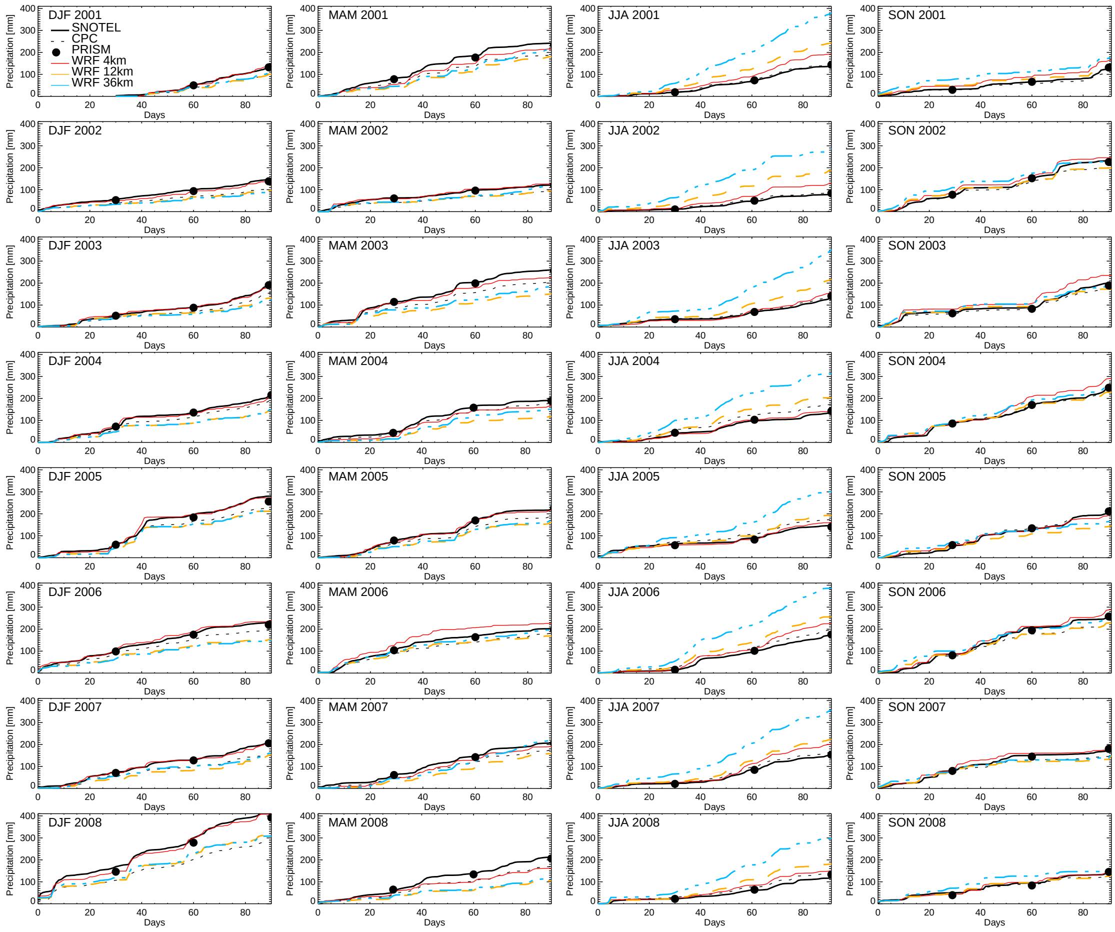

Figure 3.16: Seasonal daily precipitation accumulations over the average SNOTEL locations in the Headwaters region. Different seasons are displayed in different panels. The black line shows the results for the SNOTEL data-set. The red line corresponds to WRF 4km, the yellow line to WRF 12km, the blue line to the WRF 36km, the black dotted line to the CPC, and the blak dots to the PRISM data.

3.6 Monthly Biases and Spatial Correlations

The monthly average spatial bias and error range of the WRF 4km and WRF 12km simulation are shown in panel a and b in Figure 3.17. Here, the evaluation is done with the PRISM data-set in the headwaters. In the WRF 4km simulation no seasonal or inter-annual variability is visible in the bias or the error range. However, in the WRF 12km and WRF 36km simulation there is more precipitation and increased error ranges during the warm months of the year.

The monthly averaged spatial correlation coefficient is displayed in panel d of Figure 3.17. The WRF 4km simulation (red dotted line) has higher correlation coefficients in almost every month compared to the WRF 12km simulation. Smaller correlation coefficients are typically observed in JJAs because of the convective (non deterministic) behavior of precipitation in summertime compared to the dominant large scale, frontal precipitation in DJFs.

Second best correlation coefficients can be found for the WRF-36km simulation and third best for the WRF 12km run.

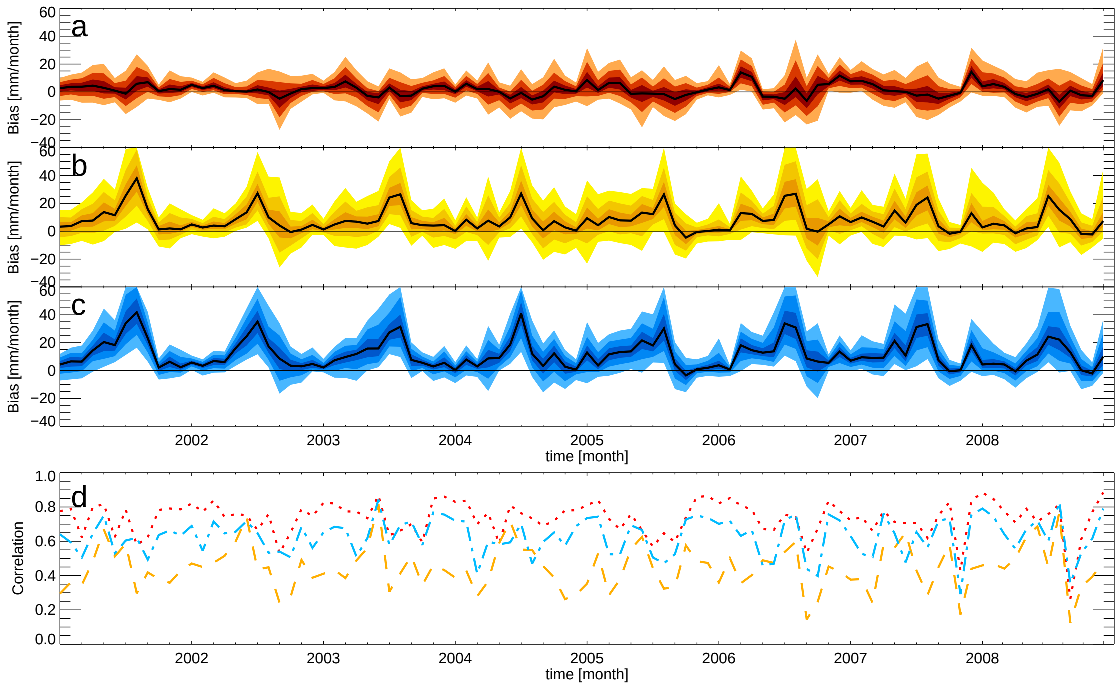

Figure 3.17: In the upper panel the relative precipitation errors between the monthly mean fields of WRF 4km (panel a), WRF 12km (panel b), and WRF 36km (panel c) minus PRISM is displayed. The black lines show the median bias, the dark contours the 40 % - 60 % quantile, the medium the 30 % - 70 % quantile and the light contour the 20 % - 80 % quantile of the difference. In panel d the spatial correlation coefficients of the monthly averaged precipitation fields (WRF compared to PRISM) are shown.

3.7 Vertical Bias Structure

The difference between the WRF-4km (Figure 3.18 panel a), WRF-12km (Figure 3.18panel b), and WRF36km (Figure 3.18panel c) simulation minus PRISM precipitation is shown in dependence of the elevation. Annually (top panel) there is nearly no height dependency visible in the bias of the WRF-4km simulation. The WRF-12km simulation underestimates precipitation below 3000 m while the WRF-36km has a constant overestimation across all height levels.

For DJF (middle left panel) the WRF-12km and the WRF-36km run show a similar characteristic with an overestimation of precipitation below approximately 2500 m and an underestimation above. The WRF-4km has much smaller differences and smaller error ranges than the coarser simulations.

In JJA (left middle panels) there is a slight underestimation of precipitation below 2600 m and an slight overestimation above that elevation in the 4 km simulation. In the 12 km run there is a general overestimation of precipitation especially at elevations below 2000 m. The WRF-36km run shows an increasing overestimation of precipitation with elevation.

In the shoulder seasons (MAM and SON. right two panels) WRF-4km has again the best performance and the smallest error ranges compared to the coarser simulations. The WRF-12km and the WRF-36km run are similar and underestimate precipitation below dim3000 m.

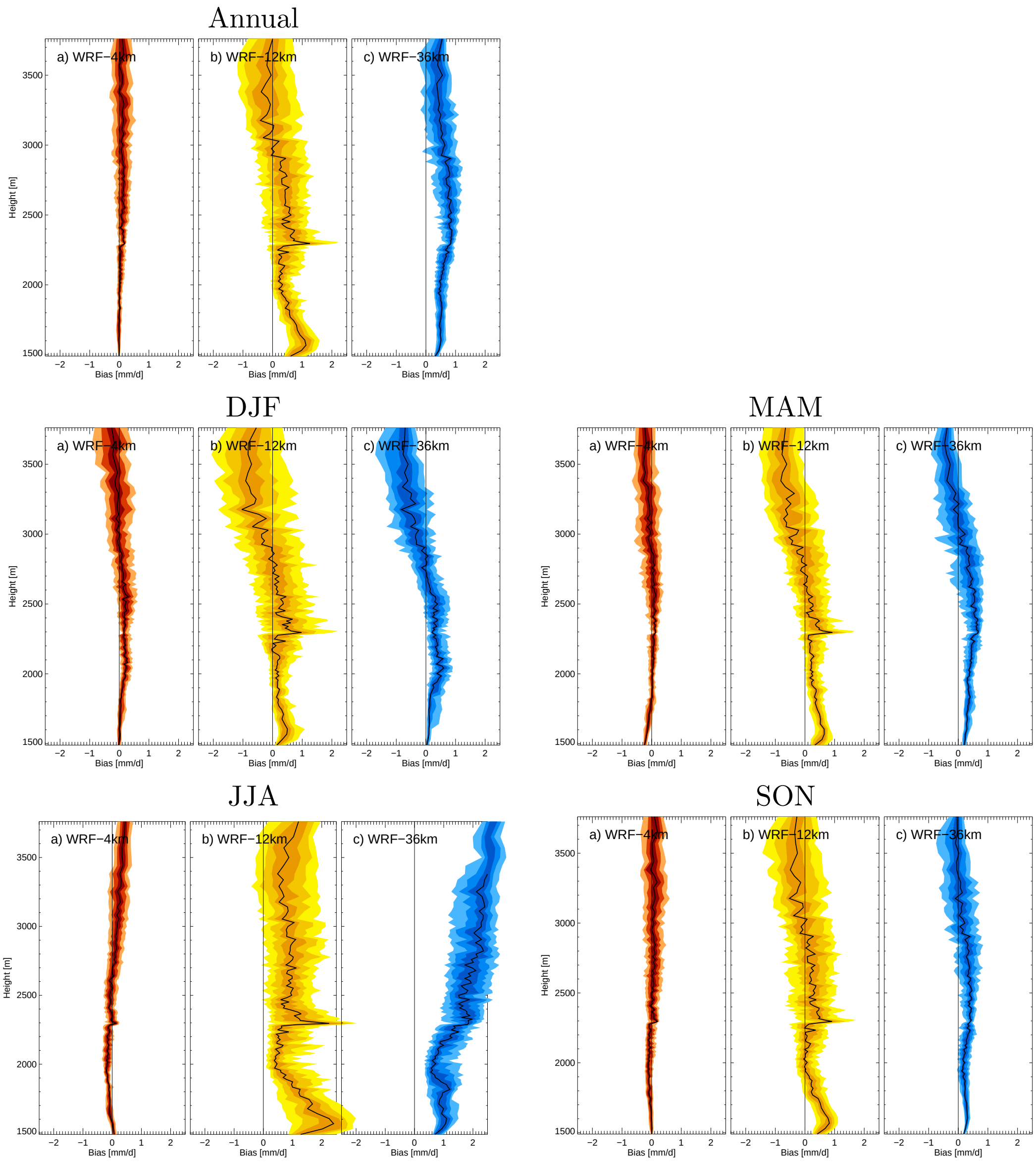

Figure 3.18: Vertical bias structures of the WRF- 4km (panels a), WRF- 12km (panels b), and WRF-36km (panels c) simulations compared to PRISM precipitation within the Headwaters region. The black lines shows the median bias, the dark contour the 40 % - 60 % quantile, the medium the 30 % - 70 % quantile and the light contour the 20 % - 80 % quantile of the difference. Different panels show different seasons according to their title. The height levels, on which the biases are calculated, are chosen to contain an equal amount of data-points.

3.8 Scaling of Extremes

In this sub-chapter extreme precipitation events are investigated concerning their spatial extension. The evaluation is based on a rather simple method which is described in Section 2.6.

Panel a in Figure 3.19 depicts precipitation extremes of SNOTEL in different seasons. The season with the highest extremes in all scales is DJF followed by SON, MAM, and JJA.

The difference between the extremes on small scales and the domain wide extremes is smallest in DJF indicating that extremes in winter are large scale and not localized. In JJA the difference is largest with more than 10 mm/d. This is related to small scale and rather localized events (convective cells).

Compared to the SNOTEL observations the WRF-4km simulation (red line) tend to overestimate the extreme precipitation events in all seasons and on all horizontal scales. The overestimation is largest in DJF (panel c) and smallest in JJA (panel e). However, measurement errors are also largest in DJF because instruments tend to underestimate precipitation due to blowing snow which is especially intense during extreme events. The 4 km simulation has a very good performance during JJA where it captures the shape of the observed curve very well above all scales.

The WRF-12km (yellow line) shows a general underestimation of extreme precipitation at small scales and a fairly good agreement at large scales.

The WRF-36km (blue line) has a similar performance for large scales than the WRF-12km run (except for JJA). For small scales it has higher values and therefore improve the representation of extremes compared to the WRF-12km run.

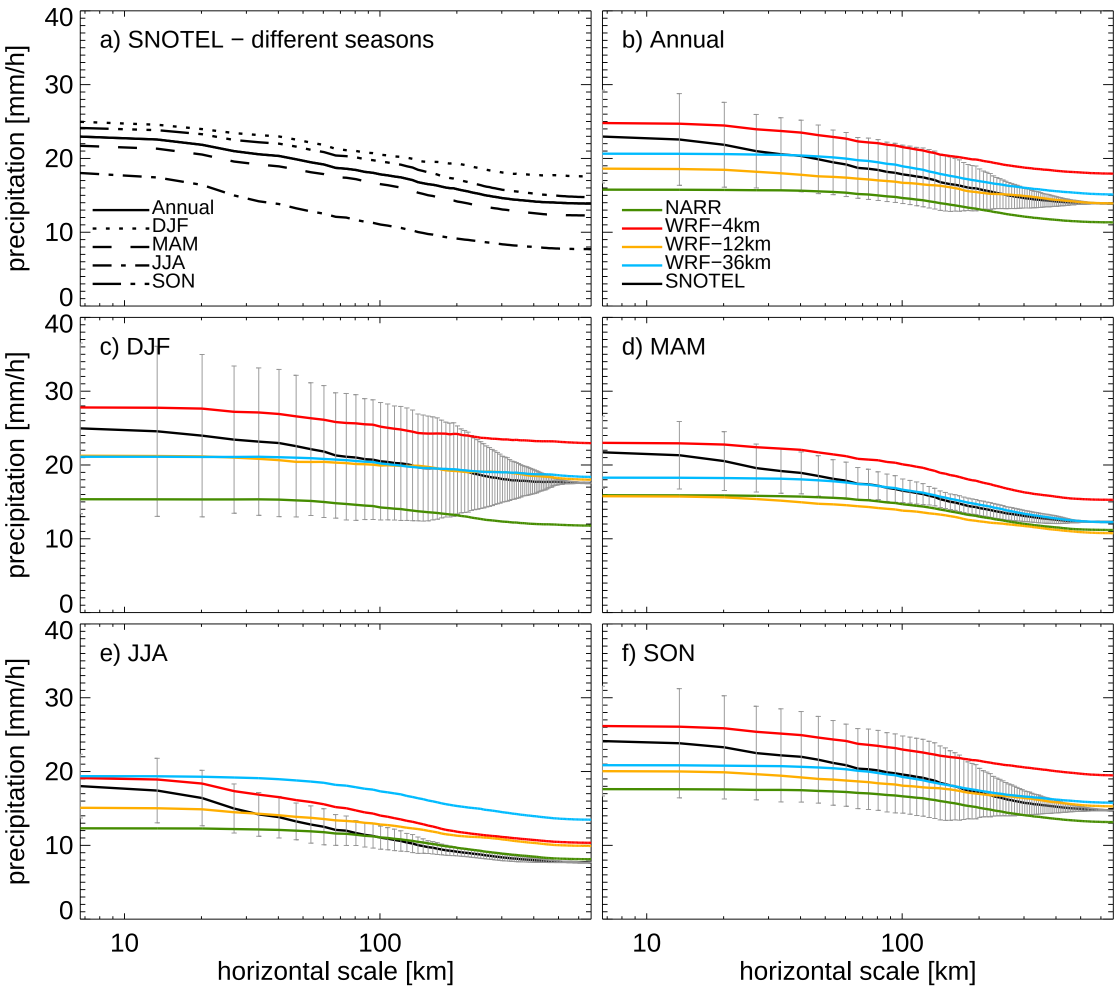

Figure 3.19: Panel a shows the extreme precipitation events of SNOTEL for different horizontal scales annually and in different seasons. The comparison between SNOTEL, WRF 4 km, WRF 12 km, WRF 36 km, and NARR is shown annually (panel b) and in DJF,MAM, JJA, and SON (panel c, d, e and f). The gray error bars show the one-fold spatial standard deviation of SNOTEL.

3.9 Selection of Extreme Events

The evaluations of extreme precipitation events in Chapter 4 focus on JJA and DJF because of the clear distinguishable predominant processes which lead to extremes in those seasons (in JJA convection and in DJF fronts). The selection is based on the domain average of the 99 SNOTEL stations within the Colorado headwaters region. All daily precipitation values above the 0.95 percentile (18 events per season) are selected and listed in Table 3.1.

Table 3.1: Dates and intensities of the twenty strongest precipitation events according to SNOTEL on domain average and grid point basis.

Extremes on average Headwaters region				Extremes on grid point basis			
DJF		JJA		DJF		JJA	
Date	Intensity [mm/d]	Date	Intensity [mm/d]	Date	Intensity [mm/d]	Date	Intensity [mm/d]
2007.12.08	29.0	2006.07.09	13.9	2007.12.08	114.3	2007.07.20	73.7
2007.12.02	27.8	2005.06.05	10.7	2007.12.08	106.7	2004.07.17	71.1
2005.01.12	20.4	2008.08.17	9.3	2007.12.08	104.1	2004.07.20	71.1
2008.01.06	19.5	2006.07.10	9.1	2005.01.12	96.5	2008.08.06	68.6
2004.01.03	18.2	2007.07.28	8.3	2008.01.06	94.0	2008.08.23	61.0
2007.12.01	18.1	2004.07.17	8.2	2007.12.02	94.0	2006.07.28	61.0
2004.12.30	18.0	2005.08.05	8.0	2007.12.02	91.4	2004.07.17	58.4
2005.01.11	16.8	2006.08.27	8.0	2005.01.05	91.4	2003.08.16	55.9
2008.12.26	16.4	2008.06.06	7.9	2005.01.12	88.9	2004.07.20	55.9
2008.01.07	15.5	2007.08.03	7.4	2008.01.06	88.9	2006.08.13	50.8
2005.01.09	15.3	2004.08.19	7.2	2007.12.08	86.4	2006.07.09	50.8
2008.12.23	15.3	2007.08.06	7.2	2007.12.08	86.4	2006.07.09	48.3
2008.01.29	15.3	2005.06.03	7.1	2007.12.02	83.8	2005.06.05	48.3
2008.12.19	14.3	2007.07.20	6.4	2008.12.26	81.3	2005.06.05	48.3
2005.12.03	14.0	2004.07.24	6.3	2005.01.12	81.3	2004.07.21	48.3
2008.02.04	13.6	2006.07.08	5.9	2007.12.08	81.3	2004.07.17	48.3
2005.01.05	13.3	2007.08.08	5.9	2007.12.02	81.3	2006.08.01	48.3
2005.01.10	12.8	2006.08.01	5.8	2007.12.01	78.7	2007.07.31	45.7

4 Ensemble Evaluation of Extreme Events

In this section an ensemble of the most extreme precipitation events (listed in Table 3.1) is analyzed for June, July, and August (JJA) and December, January, and February (DJF). The separation between JJA and DJF was done because in those two seasons different processes are responsible for extreme rainfall. In JJA convective precipitation is dominant while in DJF frontal precipitation are prevalent.

4.1 Winter Extremes

4.1.1 Mean Synoptic Situation

In Figure 4.1 the mean synoptic situation of the average 18 most extreme precipitation events in DJF is depicted. There is typically a low pressure system over the north west of the United States which brings warm and moist air from the Pacific to the Rocky Mountains in Colorado. There is a quiet strong south westerly flow over Colorado. In all 18 regarded DJF extreme events this synoptic situation is rather similar.

4.1.2 Spatial Precipitation Structure, Biases, and Error Ranges of the Mean Extreme Event

In Figure 4.2 the precipitation field of the average DJF extreme event is displayed.

Due to the predominant south westerly flow, which carries warm and moist air from the Pacific, the highest precipitation values are observed in the south west of the domain (see panel a in Figure 4.2). Downstream, the air gets drier and precipitation values get generally lower. In the spatial distribution of precipitation a clear orographically induced pattern gets visible. More precipitation is observed along mountain ridges and less in the valleys. This is because of the uplifting of air masses over mountain slopes.

The gridded reference dataset of the Climate Prediction Center (CPC) (panel b) underestimates the average precipitation extreme in DJF but captures the spatial distribution quiet well. The same is true for the North American Regional Reanalysis (NARR) driving data but here the underestimation is even larger.

The Weather Research and Forecasting Model (WRF)-4km simulation (panel d) overestimates the total amount of precipitation at the snow telemetry (SNOTEL) locations but is able to reproduce the spatial structure of precipitation. It is the only simulation which is able to capture the high spatial variability of the observation field.

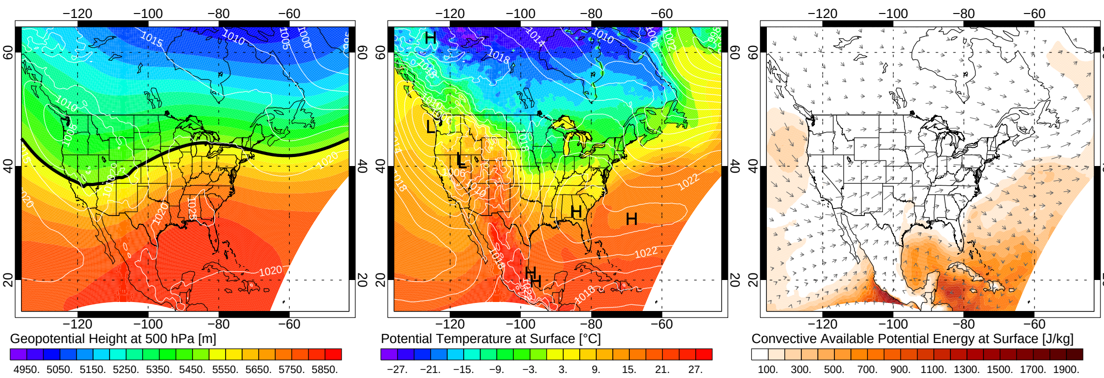

Figure 4.1: Shown is the average synoptic situation (data from the NARR dataset) of the 18 most severe DJF precipitation events. The left panel displays the Geopotential height in 500 hPa (filled contour) and the pressure at sea level in hPa (white contour). In the middle the pressure at sea level (white contours), the location of high (H) and low (L) pressure systems, and the potential air temperature at sea level (filled contour) are shown. The right panel depicts the wind speed and wind direction (arrows), and the convective available potential energy at surface (filled contours).

Compared to the WRF-4km run, the WRF-12km simulation (panel e) has a similar spatial extreme precipitation patter. There is a little less precipitation at the SNOTEL stations and less spatial variability than in the higher resolution.

The 36 km (panel f) simulations has a accurate representation of total precipitation but underestimate the spatial variability. Furthermore, the spatial structure looks more unrealistic in the 36 km simulation because of its coarse representation of the orography.

Differences between the WRF simulations minus the SNOTEL observations are displayed in Figure 4.3.

The WRF-4km simulation (panel a) shows an overestimation of extreme precipitation in the northern and western par-d of the domain. The highest overestimation is around 2800 m above sea level.

The WRF-12km simulation (panel b) has a slightly smaller precipitation bias than the WRF-4km simulation. The error pattern looks similar to the those of the 4 km simulation and also the root-mean-square error (RMSE) is approximately the same.

In the WRF-36km simulation (panel c) there is a tendency of overestimating precipitation in the valleys and underestimating precipitation in higher elevations. This simulation has the highest RMSE but the smallest bias because areas with positive and negative differences are canceling out.

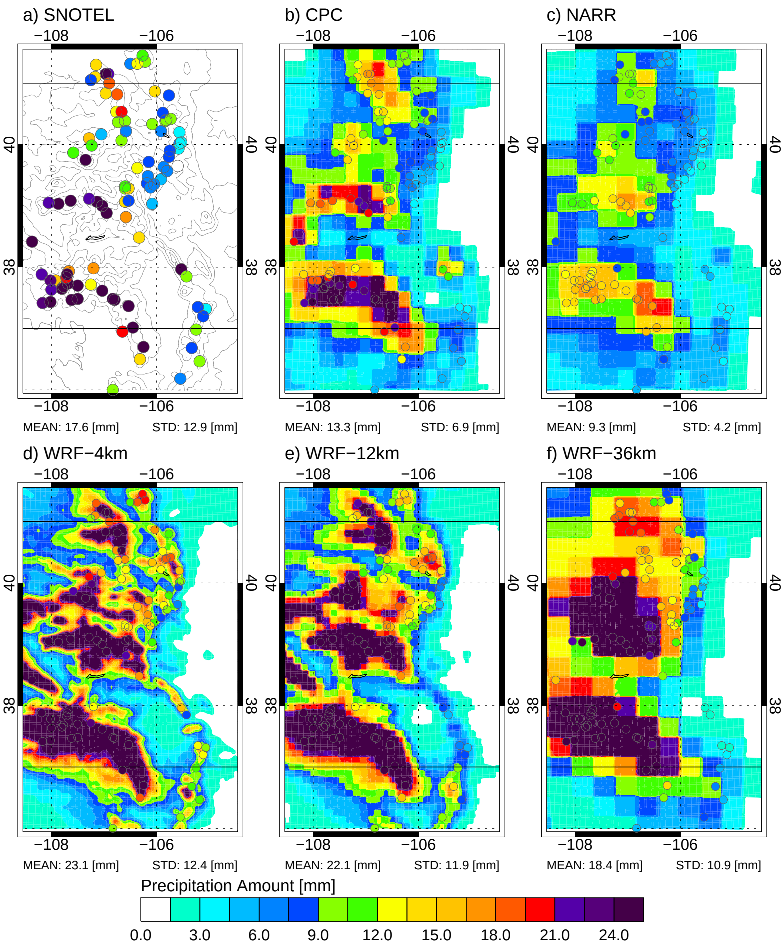

Figure 4.2: Accumulated precipitation for the average DJF extreme event from left to right for sub panel a) SNOTEL, b CPC, c) NARR, d) WRF 4km, e) WRF 12km, and f) WRF 36km. Below each map the mean accumulated precipitation and the spatial standard deviation at the SNOTEL locations (circles) is shown.

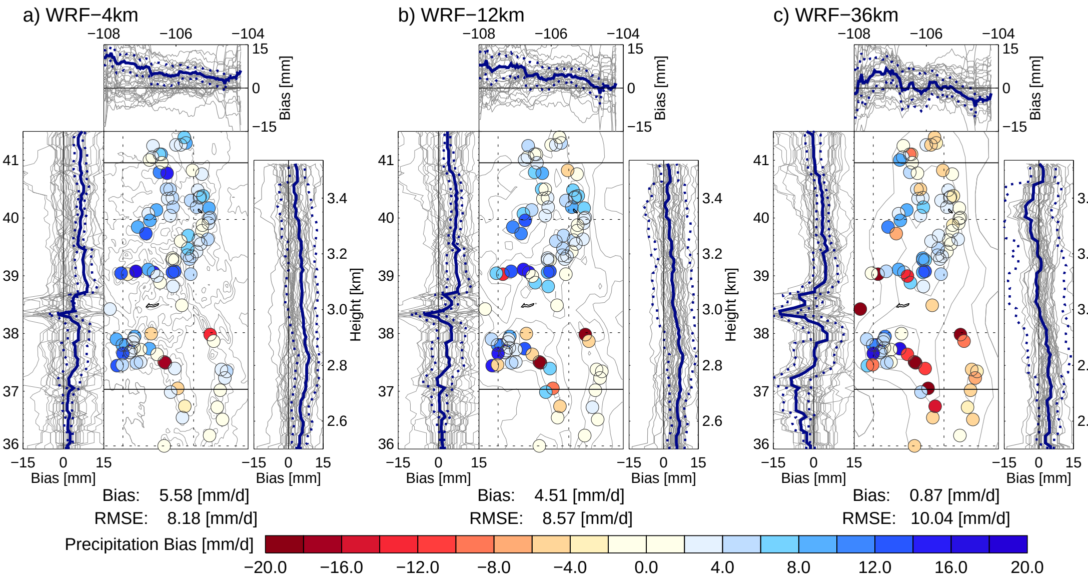

Figure 4.3: The map in the center of each sub-panel shows the differences between the correspondent WRF simulation and the observed precipitation data at the locations of the SNOTEL stations. The blue solid lines in the other sub panels shows the mean differences along the latitude (left beside maps), the longitude (above maps), and the altitude direction (right beside maps). The thick blue lines are the mean biases of the average extreme event along the latitude, longitude of elevation. The thin blue dashed lines correspond to the standard deviation of the bias and the grey solid lines show the biases of the individual extreme events.

Statistical properties of the average DJF extreme event are displayed in Figure 4.4. The box length ((25 % to 75 % distance) is smallest in the WRF 4km simulation, closely followed by WRF 12km run. In case of the whisker difference (5 % to 95 % distance) also the WRF-4km simulation performs best and is even better than the CPC data-set. Concerning the spatial error spread the WRF 4km and WRF 12km are comparable to CPC and improve the representation of their boundary data (NARR) largely. In the terms of error spread, the WRF 36km simulation has a similar performance than the NARR driving data.

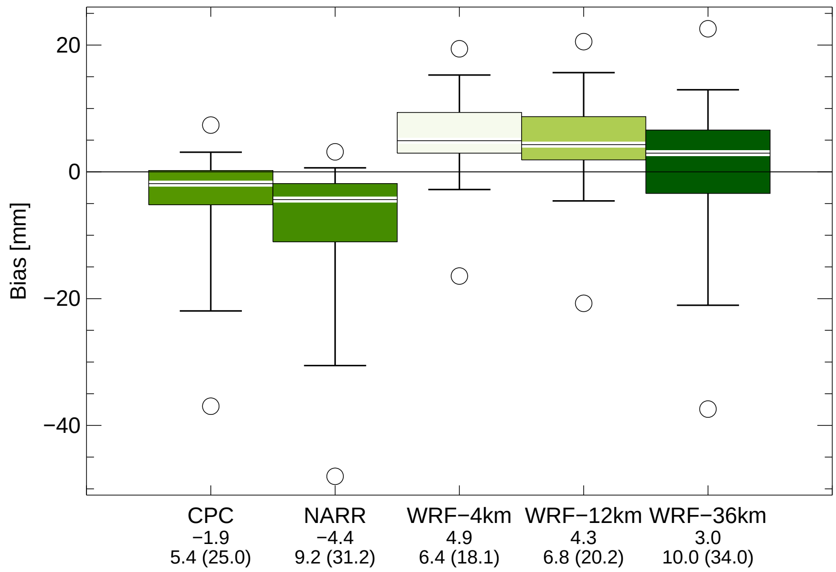

Figure 4.4: Box Whisker plots of the spatial errors of CPC, NARR, WRF 4km, WRF 12km, and WRF 36km for the mean DJF extreme event compared to the SNOTEL data set. The first number below the plot shows the median bias and the two numbers below depict the length of the box (25 % to 75 %) and the distance of the whiskers (5 % to 95 %, number in brackets).

4.1.3 Spatial Scaling of Correlation and Standard Deviation

The calculation of scale dependent correlation coefficients and normalized standard deviations is done as described in Section 2.7.

In Figure 4.5 the scale dependent correlation coefficients and normalized standard deviations are displayed for CPC, NARR, and the WRF simulations.

The WRF-4km simulation (red dotted line) has the highest correlation coefficients of all data-sets at small scales. Second best is the WRF-12km simulation (yellow dashed line) and relatively low correlation coefficients are found for WRF 36 km (blue dashed line). The correlation coefficients of the WRF-12km and NARR data-sets (green dashed-dotted line) get comparable high values to the WRF-4km simulation at scales above 100 km.

Regarding the normalized standard deviations, the WRF-4km simulation has again the best quality of all simulations for small scales closely followed by the WRF-12km run. Third best results are obtained by the WRF 36 km simulation. Above 50 km all simulations have a similar performance.

All simulations are able to improve the performance of the NARR driving data and are also better as the CPC reference data-set.

Because errors can cancel out when averaging the 18 most extreme precipitation events in DJF the evaluation in Figure 4.6 shows the same evaluation as above but for each of the 18 extreme events.

The CPC data-set (panel a) has the highest median correlation coefficients across all horizontal scales if the 18 extreme events are evaluated separately. The WRF-4km simulation (panel c) is second best followed by the 12 km simulation (panel d). The WRF-36km (panel e) simulation has similar correlation coefficients as its driving data (the NARR data-set, panel b).

Concerning the normalized standard deviations, the WRF-4km and the WRF-12km simulation (panel C and D) have the overall best performance and are close to zero across all scales. Also the WRF-36km simulation (panel E) is able to capture the spatial variability of SNOTEL nicely. All simulations are able to improve the normalized standard deviation of their driving data (panel E) remarkably.

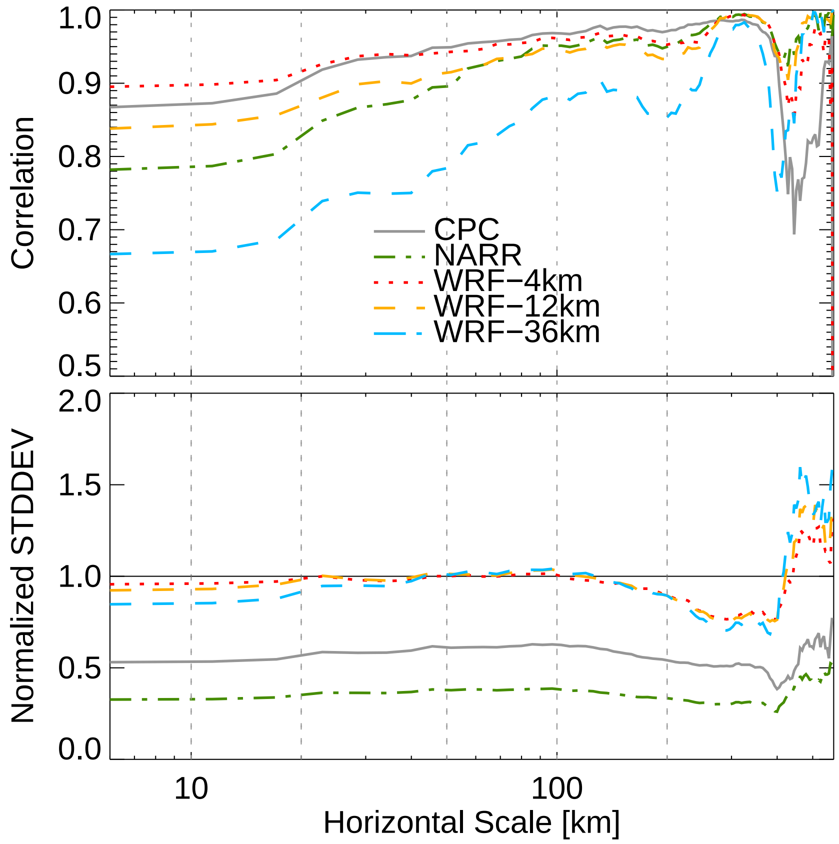

Figure 4.5: In the upper panel the correlation coefficients between the SNOTEL stations and the mean extreme precipitation data of the CPC, NARR, WRF 4km, WRF 12km, and WRF 36km data sets are shown dependent on the horizontal scale on which the data are averaged. The lower panel shows the same for the standard deviation normalized by the standard deviation of the SNOTEL data set. Horizontal scales are shown on a logarithmical x-axis.

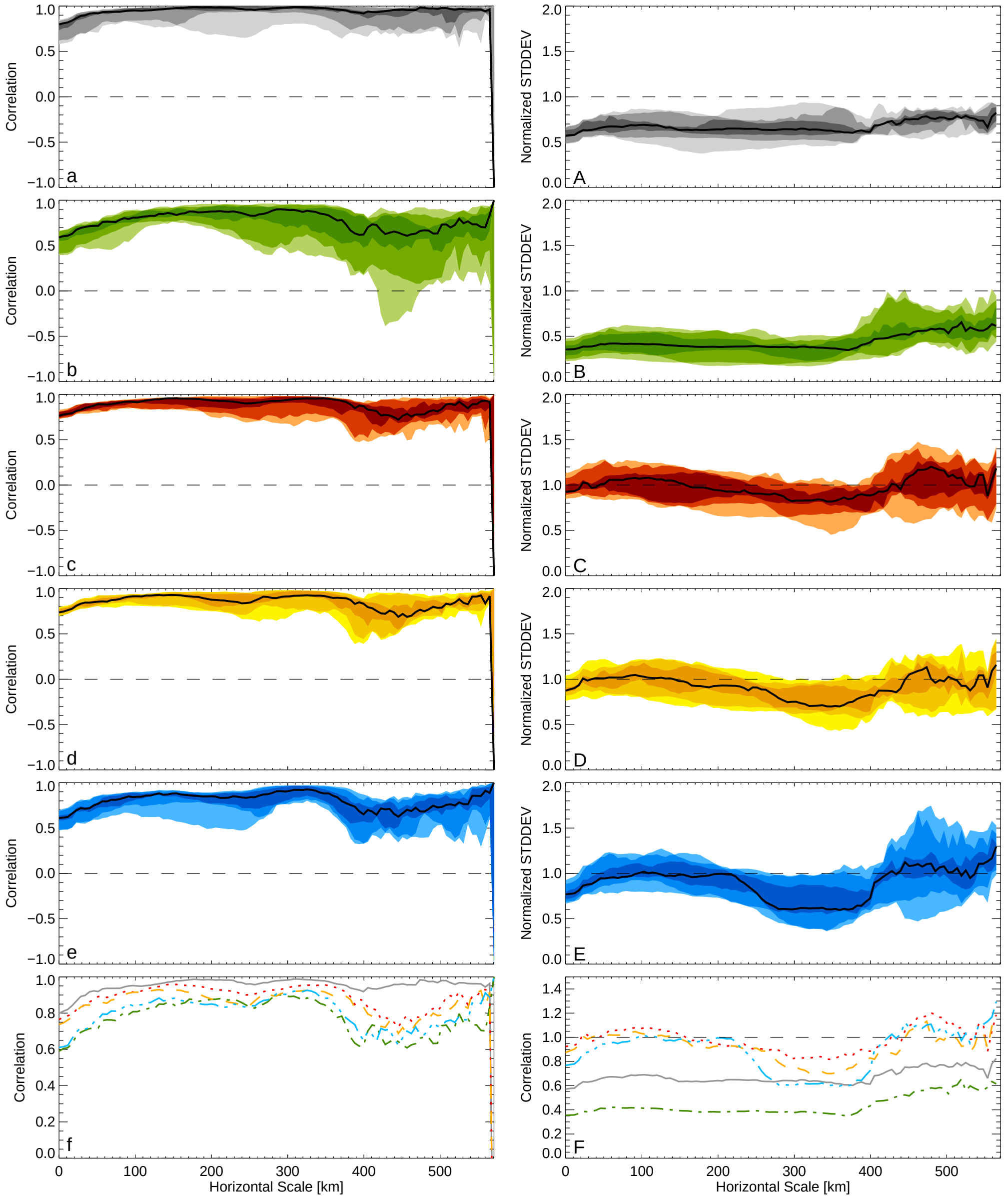

Figure 4.6: Shown are the median (black line), the 40 % to 60 % quantile (dark contour), the 30 % to 70 % quantile (medium contour), and the 20 % to 80 % quantile (light contour) of the extreme precipitation ensemble for the correlation coefficients (left panels) and the normalized standard deviation (right panels) compared to the SNOTEL data. Sub-panel a and A shows the results for CPC, b and B NARR, c and C WRF 4km, d and D WRF 12km, and e and E those of the WRF 36km data dependent on the horizontal scale on which the data are averaged. The lowest panels f and F show the medians of all data-sets in a single plot for an easier comparison of the results.

4.1.4 Similarity and Spatial Dependence of Extremes

The upper panel of Figure 4.7 displays the correlograms of the different data-sets. The closer the lines of the simulations are to the SNOTEL observations (black line) the better.

For lags below 60 km the WRF-4km simulation has the most accurate correlation values. The other simulations have a too high spatial dependency at these small scales. At about 70 km lag all data-sets are relatively close to the SNOTEL observations. No correlation exits between points which are further than 200 km apart.

The semivariogram is shown in the lower panel of Figure 4.7 and depicts the dissimilarity between two points. Also here the WRF-4km simulation is closest to the observed values in SNOTEL. A similar good performance can be found for the WRF-12km simulation at scales later than ∼70 km. The WRF-36km run has generally to low semivariance values. The simulations are able to improve the performance of their boundary data essentially in all scales and also outperform the CPC reference data-set.

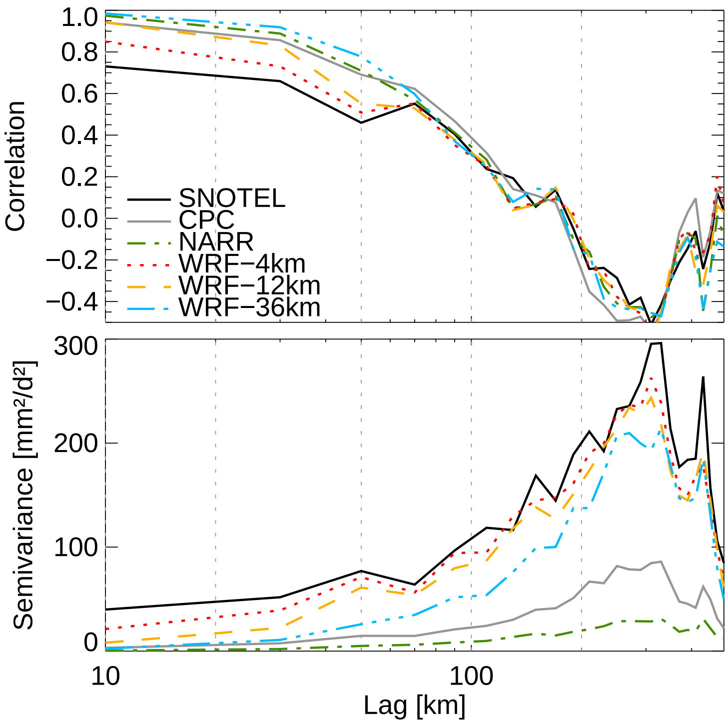

Figure 4.7: In the upper panel the correlogram of the average DJF extreme precipitation field is displayed for all data sets. The lower panel shows the semivariogram of the same data. The lag on the x-axis is displayed on a logarithmic scale.

4.2 Summer Extremes

4.2.1 Mean Synoptic Situation

In Figure 4.8 the mean synoptic situation of the JJA extreme averaged 18 most extreme precipitation events is depicted. The typical ingredients of a summertime extreme are weak pressure gradients over the entire continental USA, and very low geostrophical winds over Colorado. In addition there is some convective available potential energy at the surface which is crucial for the development of thunderstorms.

This situation is rather similar in all of the 18 JJA extreme events. Therefore, the typical summertime extreme is not triggered by large scale but by small scale processed within the simulated domain.

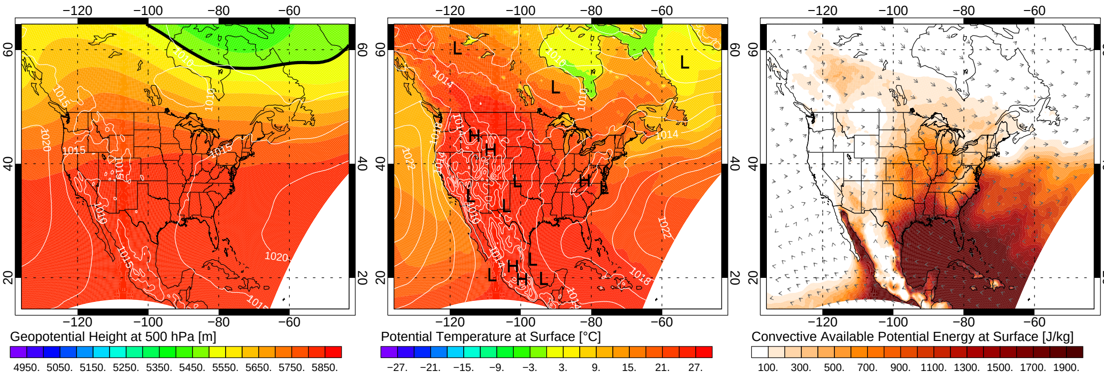

Figure 4.8: As in Figure 4.1 but for JJA.

4.2.2 Spatial Precipitation Structure, Biases, and Error Ranges of the Mean Extreme Event

Figure 4.9 displays the average extreme event in JJA for the different data-sets. Compared to DJF the domain wide precipitation is much smaller (see panel a, SNOTEL).

The WRF-4km simulation (panel a) is able to capture the domain wide precipitation and the spatial structure of the average JJA extreme. Also the spatial standard deviation is well represented.

The coarser gridded simulations overestimate the amount of precipitation by ~50 %. However, the spatial standard deviation is correctly represented.

The differences between the WRF simulations minus the SNOTEL data for the mean JJA extreme event are shown in Figure 4.10.

The WRF-4km run (panel a) shows a nearly perfect agreement with the observations and has a very small bias and RMSE.

The WRF-12km simulation (panel b) shows an overestimation of precipitation in large parts of the domain and across all height levels.

A similar result is visible in the WRF-36km simulation (panel c). The average bias and the RMSE are very closely to those in the 12 km run.

The box whisker plot in Figure 4.11 shows the spatial differences of the average JJA extreme event.

The WRF-4km simulation has a narrow box length (25 % to 75 % quantile) and a small whisker distance (5 % to 95 % quantile) which is comparable with those of the CPC reference data set.

The box and whisker distances in the WRF-12km simulation are ∼30 % higher than those of the WRF-4km simulation.

The WRF-36km simulation shows a better performance concerning the box distance (not the whisker distances) compared to the WRF-12km simulation.

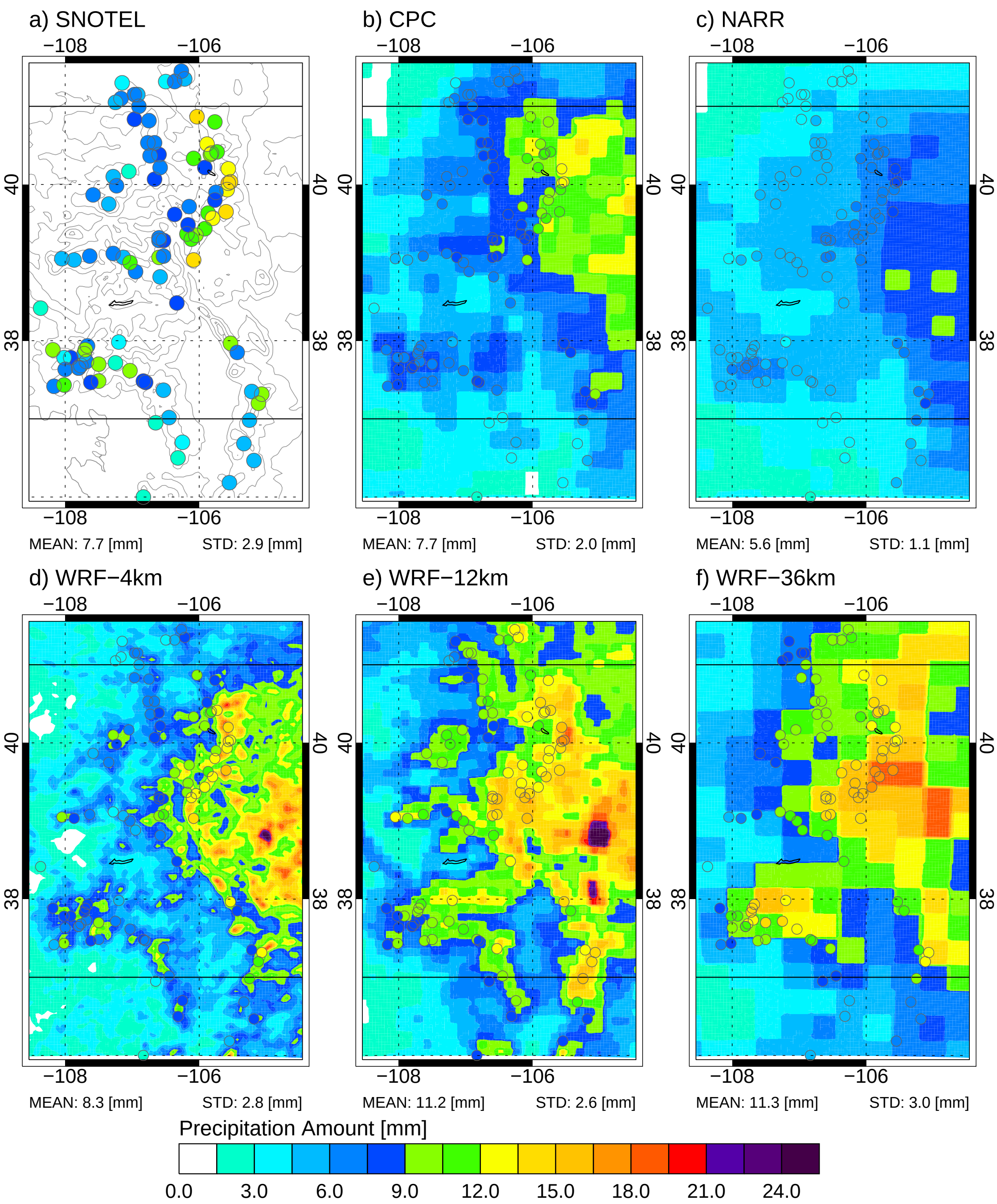

Figure 4.9: Same as in Figure 4.2 but for JJA.

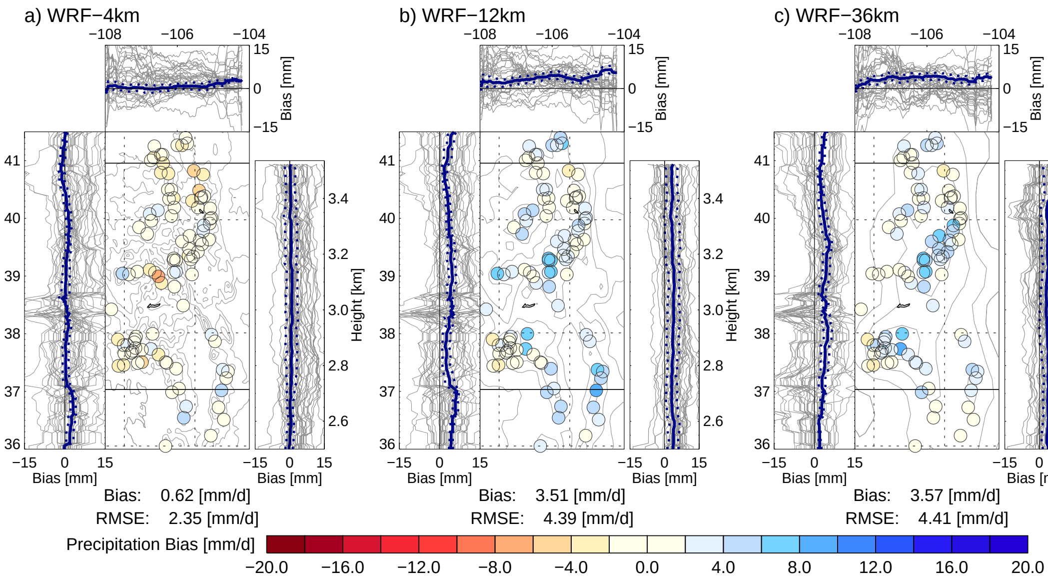

Figure 4.10: As in Figure 4.3 but for JJA.

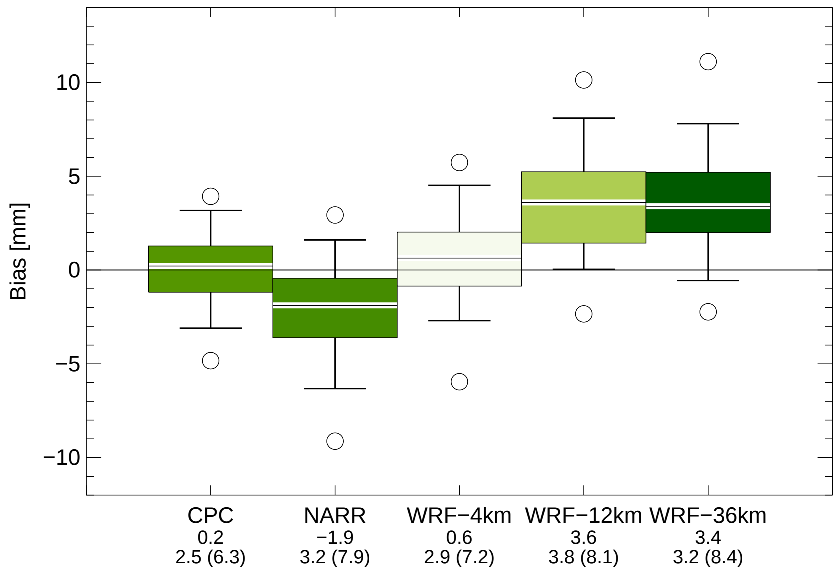

Figure 4.11: As in Figure 4.4 but for JJA.

4.2.3 Spatial Scaling of Correlation and Standard Deviation

Figure 4.12 displays the spatial correlation coefficients (upper panel) and normalized standard deviations (lower panel) for different horizontal scales in JJA (for a description of the applied method see section Subsection 4.1.3).

Compared to DJF (see Figure 4.5), the correlation coefficients are much smaller in JJA especially for small horizontal scales. This is because of the predominant convective precipitation in summer which is much less deterministic and therefore harder to simulate compared to the large scale frontal precipitation in winter.

For horizontal scales below 20 km the WRF-4km simulation has the highest correlation coefficients off all simulations. For scales above 50 km the WRF-36km simulations performs better. The WRF-12km run has generally the lowest correlation coefficients of all simulations.

In case of normalized standard deviations, the WRF-4km simulation is able to reproduce the spatial variability of the SNOTEL observations extremely well in particular for scales bellow 100 km. A similar good performance can be found in the WRF-12km run. Also the WRF-36km simulation has a accurate but slightly to high spatial variability across all horizontal scales.

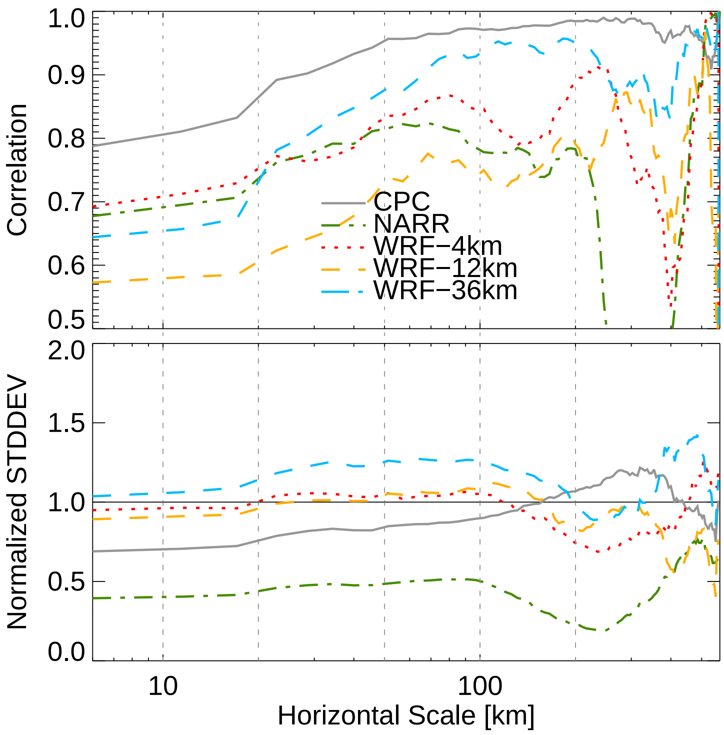

Figure 4.12: Same as in Figure 4.5 but for JJA.

The correlation coefficients and normalized standard deviations of the ensemble of the 18 most extreme precipitation events in JJA are displayed in Figure 4.13.

The spread of the correlation coefficients of the WRF-4km and WRF-12km simulation (contours in panel c and d) are smaller compared to the WRF-36km simulation(panel e) and the driving data (panel b). This means that the majority of the 18 extreme events is simulated accurately in the two higher resolved runs. The 4 km run has the highest median correlation coefficients (red dotted line in panel f) of all simulations above horizontal scales of ~70 km.

Also the spread in the normalized standard deviation (contours in panel C and D) are smaller in the WRF-4km and WRF-12km simulation than in the WRF-36km run. The median normalized standard deviation of the WRF-4km is close to one for all horizontal scales (red dotted line in panel F) and has the overall best performance of all investigated data-sets.

The normalized standard deviation of the WRF-12km simulation is slightly underestimated in the median

The WRF-36km simulation (panel e and E) has slightly to high normalized standard deviations.

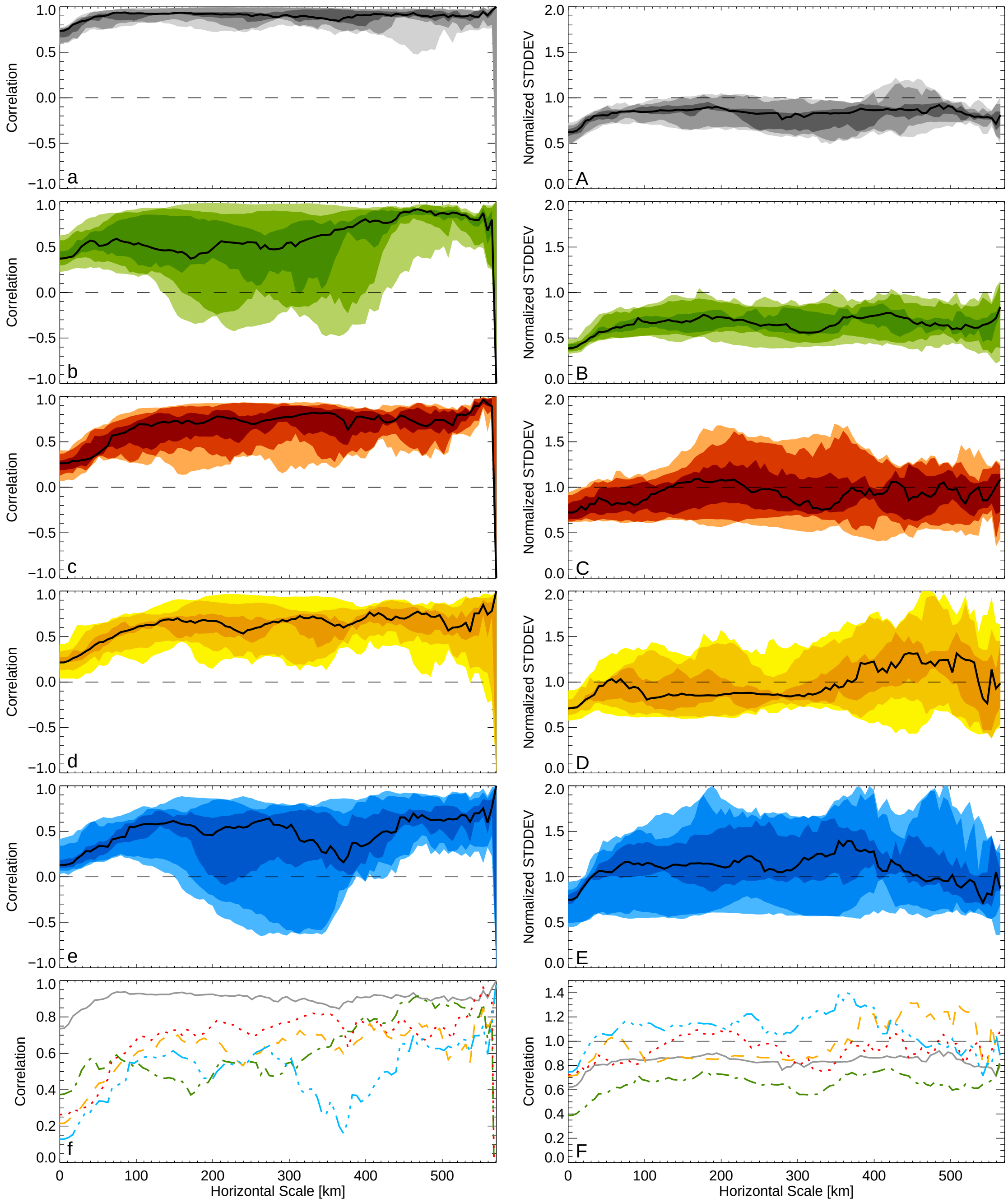

Figure 4.13: As in Figure 4.6 but for JJA.

4.2.4 Similarity and Spatial Dependence of Extremes

The upper panel of Figure 4.14 displays the correlograms of the different data-sets. Compared to DJF, the spatial similarity is much smaller and the lag where there is zero correlation between two points is with ∼100 km approximately half as small.

As in DJF the WRF-4km run (red dotted line) is able to simulate the spatial dependency best for small lags (below ∼50 km). Second best at this slam lags is the WRF-12km simulation and third best is the WRF-36km run. Above a leg of ∼50 km all data-sets perform similarly.

Also the semivariance (lower panel in Figure 4.14) is best simulated in the WRF-4km run at lags below ∼50 km. Above, the WRF-12km and WRF-36km run have a comparable performance.

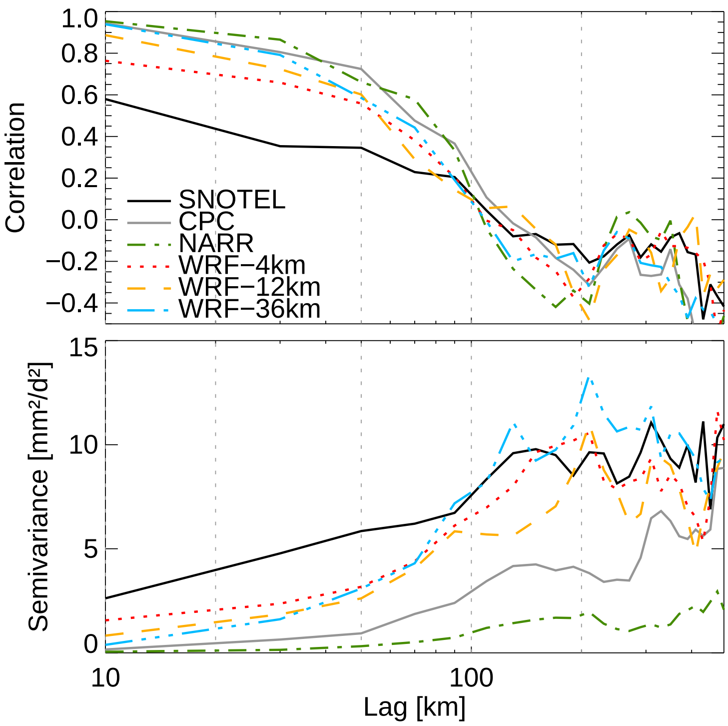

Figure 4.14: As in Figure 4.7 but for JJA.

4.3 Spectral Analysis of the Average Extreme

In this section the average DJF and JJA extreme event is spectrally decomposed wit the discrete cosine transformation (DCT).

In DJF (panel a) the spectra of the average extreme event is strongly correlated to the spectra of orography. For example, the peak at ∼40 km is clearly visible in the WRF-4km and WRF-12km simulation but is missed by the WRF-36km simulation.

The effective resolution of the WRF-12km and the WRF-36km run is ∼4 times the grid spacing.

All simulations have higher sums of mean variance for large wavelengths compared to the NARR driving data and the CPC reference data-set.

In JJA (panel b), the variance spectra of the WRF simulations are less dependent on orography as in DJF.

For large wavelengths the WRF-12km and WRF-36km simulation have higher sums of mean variances than the WRF-4km simulation. Due to this overestimation it is hard to find out the effective resolution of the WRF-12km and WRF-36km runs.

As in DJF, all WRF simulations have higher sums of mean variances than the NARR driving data and the CPC reference data set for large wavelengths.

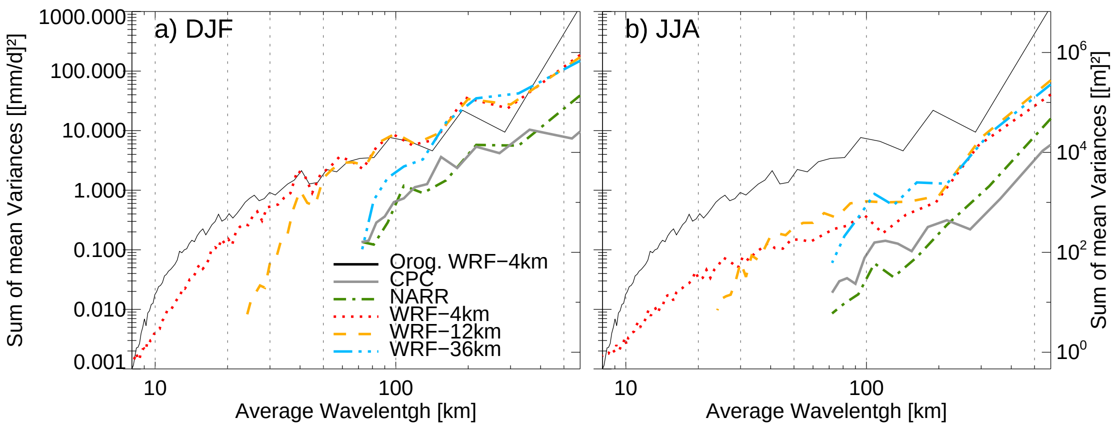

Figure 4.15: Spectral decomposition of the average DJF (panel a) and JJA (panel b) extreme events. The sum of variances for precipitation are connected to the left y-axis while the orography spectra is displayed on the right y-axis.

5 Application of Extreme Value Statistic

In this chapter methods from the extreme value theory are applied to calculate return values for extreme precipitation events in the Colorado headwater region.

5.1 Evaluation of the Fit of Pareto and Weibull Distributions

To analyze the quality of the fitted GPD and Weibull distribution (WBL) diagnosis plots were plotted for each performed fit (each data-set, season, mountain range, and elevation band). An example diagnostic plot is shown in Figure 5.1.

Displayed is the fit of the GPD and WBL distribution to the 8 year, headwater average, annually snow telemetry (SNOTEL) precipitation observations. In the upper left corner a probability plot displays fitted vs. observed probability. A perfect fit would exactly lie on the diagonal.

In the upper right corner, a quantile plot shows fitted vs. observed quantiles. Also here, a perfect fit would lie exactly on the diagonal.

The lower left panel displays a return level plot. The dots in the plots show the data from SNOTEL observations and the lines should overlay the dots as close as possible.

The last diagnostic plot in the lower right panel shows the empirical probability density function of all precipitation values above 95 % in a histogram and the two fits are displayed as red and blue lines.

In the further evaluation only the GPD was used because it lead generally better fits than the WBL distribution.

5.2 Comparison of the Simulated Period with the Climatology

In Figure 5.2 the return values of extreme precipitation events in the Colorado headwaters region are displayed. The estimations are based on 39 SNOTEL stations which provide a continues time-line from October 1980 to September 2009.

The red solid line displays the best estimate of extreme precipitation return values based on the simulated period January 2001 to December 2008 while the black line displays the same quantity estimated from a 30 year period (October 1980 to September 2009).

If only the 8 year period is considered, there is a chance of overestimating return values compared to the climatological mean. This indicates that there were more intense

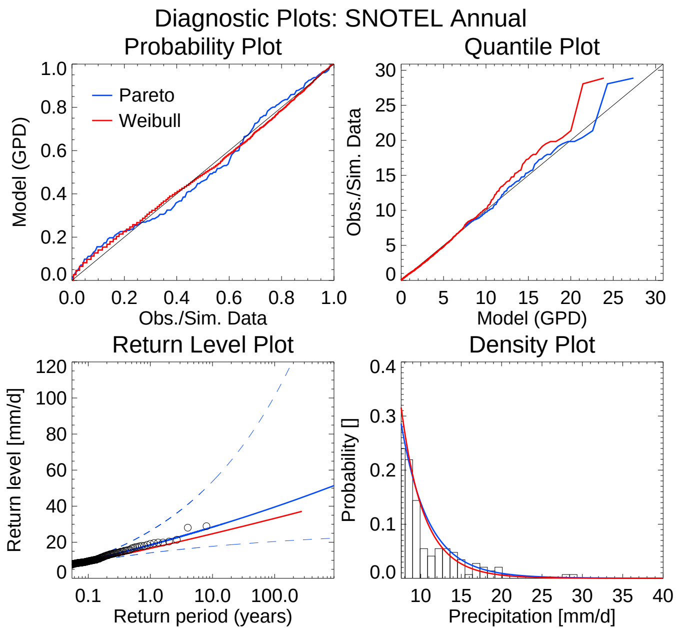

Figure 5.1: Example diagnostic plots to evaluate the fit of extreme value distributions to daily extreme value precipitation data. Displayed are the 8 year SNOTEL (99 stations) extreme values and the fitted GPD and WBL distributions.

extreme values within the 8 years than in the 30 year climatology.

The dotted lines in Figure 5.2 show the estimated uncertainty of the return values. The 8 year period has a much higher uncertainty than the 30 year period. For example the 100 year event estimated from the 8 year data set is 37 mm/d with an range from 20 mm/d to 95 mm/d while the 100 year return value of the 30 year data set is 30 mm/d with an range of 24 mm/d to 42 mm/d. This result indicates the importance of long time series for extreme value estimations.

In the following analyses the 10 year return value is chosen for the evaluations. This is a compromise between relevant return values for impact research and uncertainties in the estimation.

The 10 year best estimate return values of extreme precipitation are displayed for the 8 and the 30 year period for different seasons and different mountain ranges in Figure 5.3.

The main characteristics are similar between the two considered time periods. The 10 year return values are highest in December, January, and February (DJF) and lowest in June, July, and August (JJA) for the headwaters average. There is a south west to north east gradient in extremes with higher return values in the south eastern mountains (e.g., San Juan Mountains) and lower in the north western ones (e.g., Front Range or Park

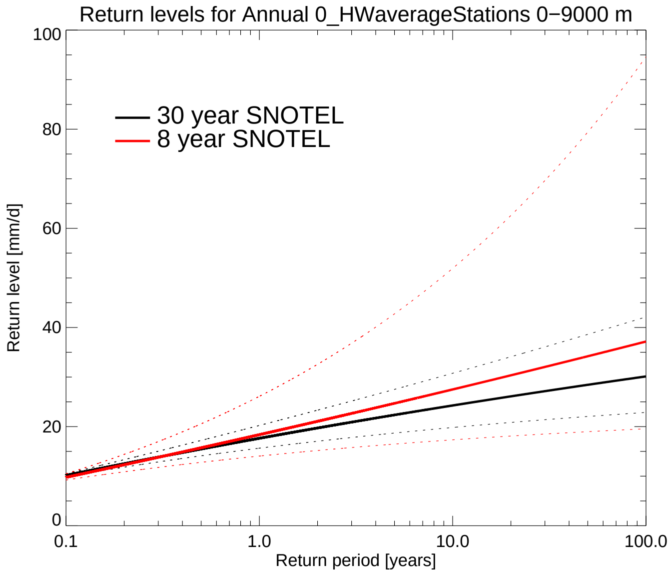

Figure 5.2: Return values of extreme precipitation events in the Colorado headwaters region. The red solid line shows the estimated return values if only SNOTEL data of the 8 simulated years are considered while the black line shows return values for a 30 year climatology. Dashed lines display the uncertainty of the estimations.

Range).

There are also differences in return values for different elevations. In general, the 10 year return value increases with increasing elevation.

The differences between the two evaluated periods are rather small (below ± 3 mm/d) on the headwater average. Larger differences occur in different mountain ranges (e.g., underestimation of return values in the Sangre de Cristo Range) or elevation bands (e.g., September, October, and November (SON) between 3400 m to 3600 m). However, there is no systematic change towards too intense or too weak extremes.

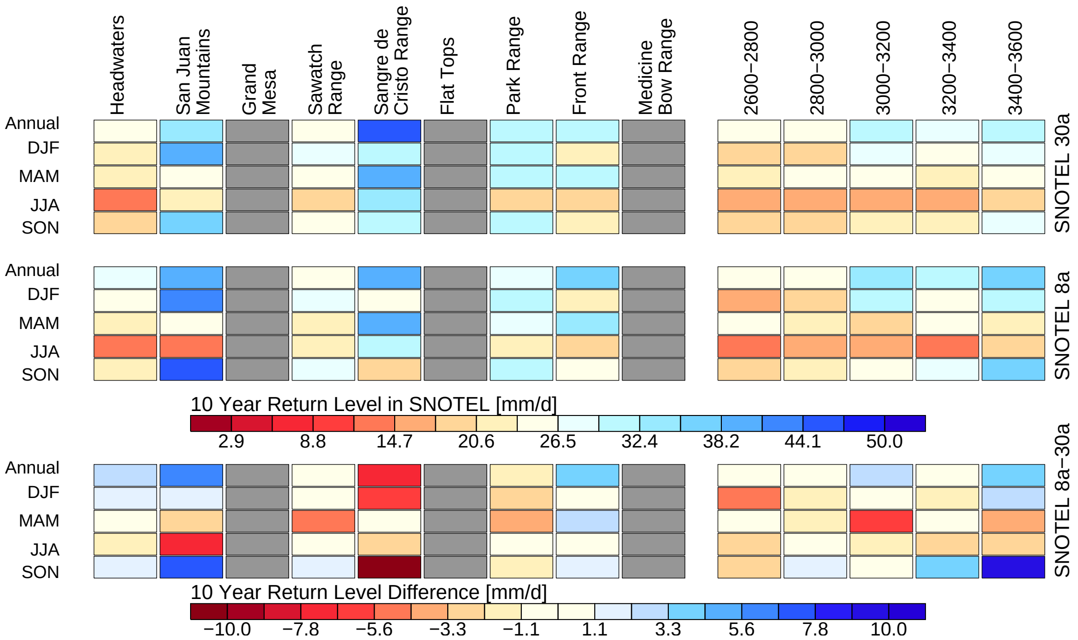

Figure 5.3: 10 year best estimate return value in the 8 year period (upper panel) and those derived from the 30 year climatology (middle panel). In the lower panel the differences between the return values of the 8 year data minus the 30 year data are displayed. Different rows in sub-panels display different seasons while different columns show different mountain ranges and different elevation bands. For gray boxes not enough SNOTEL stations were available to perform the statistics.

5.3 Evaluation of Simulated 10 Year Return Values

In Figure 5.4 the 10 year precipitation return values of the different Weather Research and Forecasting Model (WRF) simulations are compared with the SNOTEL observations.

The upper panel shows the 10 year return values derived from the SNOTEL observations. The difference to the middle panel in Figure 5.3 is that here 99 stations are used for the evaluation while in Figure 5.3 only 39 were used (because only 39 stations had continues time series for the 30 year period).

In the Headwaters average the season with the highest return values is DJF. The lowest values accrue in JJA. There is a transmission from a regime where DJF and SON are the most extreme month in the south west (San Juan Mountains and Grand Mesa) to a regime where the most extreme events occur in spring (in the north east; Park Range, Front Range, and Medicine Bow Range).

The other characteristics in SNOTEL are rather similar to those described in Section 5.2.

On an annually basis the WRF-4km simulation overestimate the 10 year return value by 4 mm/d in the average Headwaters region. The coarser simulations have smaller differences but only because over and underestimations in different mountain ranges are canceling out when they are averaged.

On an mountain range basis, the WRF-4km simulation has mostly the smallest difference of all simulations. In general, the 10 year return values are increasing when the grid spacing is decreased.

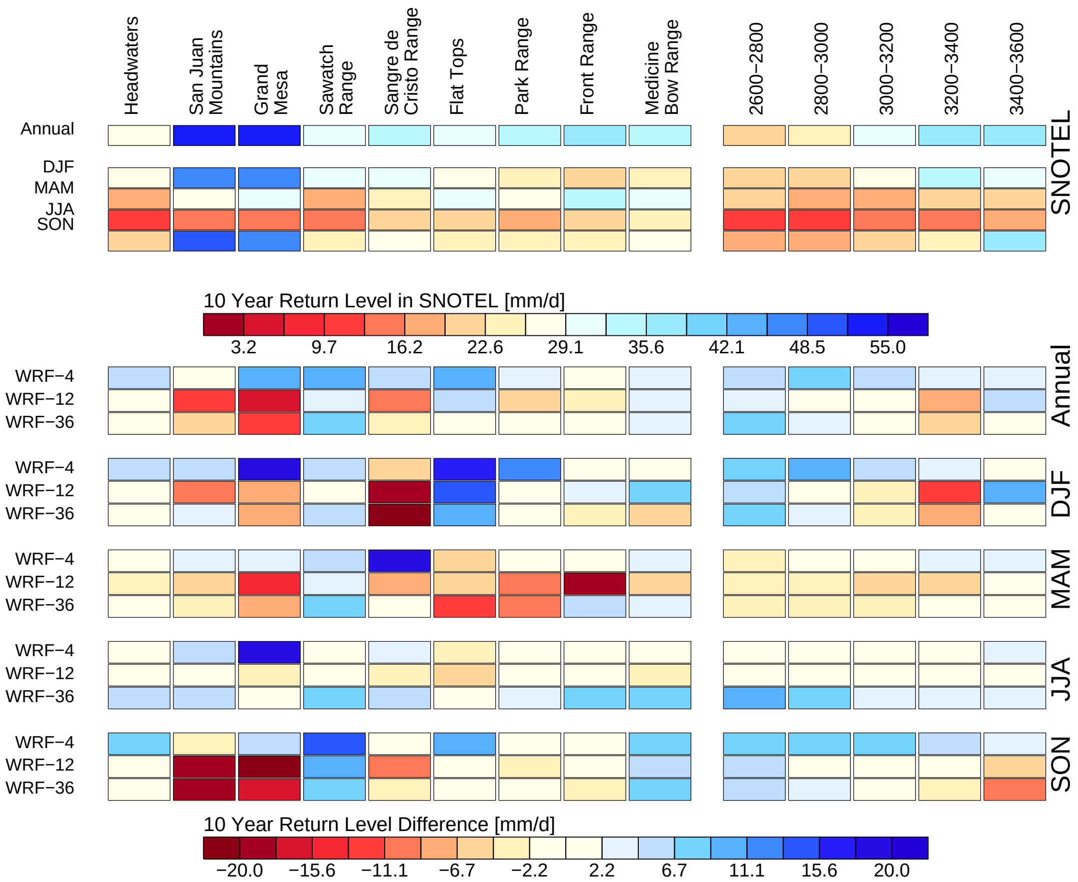

Figure 5.4: 10 year best estimate return value for the simulated 8 year period of SNOTEL (upper panel). In the lower panels the differences between the simulated return values minus the return values of the 8 year SNOTEL observations are displayed for annual, DJF, MAM, JJA, and SON (top down). Different rows in sub-panels display the different simulations while different columns show different mountain ranges and different elevation bands.

6 Summary and Conclusion

In this study, the dependence of simulating precipitation extremes over complex terrain on the horizontal grid spacing of the Weather Research and Forecasting Model (WRF) is analyzed. Therefore, the 8 year period (January 2001 to December 2008) is simulated with 4 km (WRF-4km), 12 km (WRF-12km), and 36 km (WRF-36km) grid spacing. The largest difference between the WRF-4km and the coarser simulations is that the convective parameterization was switched of in the 4 km run because deep convection can be resolved explicitly on grids ≤∼4 km. All simulations covered the same domain which was centered on the headwaters of the Colorado River. As lateral boundary data the North American Regional Reanalysis (NARR) data set was used.

The scale separation (discrete cosine transformation (DCT)) of multiple temporally averaged atmospheric fields showed the strong relationship between parameters like precipitation, 2 m temperature, wind speed, or snow water equivalent on the underlying orography. This relationship is typically strongest during the cold months of the year and gets weaker in June, July, and August (JJA).

With the DCT method also the general higher variability at small scales in the fields of the WRF-4km simulation gets visible. This higher variability is not only random noise but certainly improves the spatial patterns of precipitation.

The evaluation of precipitation in the entire 8 year period shows an astonishing good agreement between the WRF-4km simulation and the observations. Regardless of the investigated season, the simulated precipitation agrees with the observation within the typical observational error (±15 %). In the WRF-12km and especially in the WRF-36km simulations JJA precipitation is overestimated. This result shows that a correct representation of summertime precipitation is only possible if deep convection is resolved explicitly (like in the WRF-4km run). In December, January, and February (DJF) the WRF-4km simulation has a similar small bias than in JJA. In the WRF-12km and WRF-36km simulations precipitation sums are generally underestimated which is most probably because of the coarse resolution. This leads to smaller vertical wind speeds when air masses are moved upward at mountain slopes and an underestimation of orographically induced precipitation. This result was already published by Rasmussen et al. (2011).

Concerning the average precipitation extreme event in DJF (average of all events above the 97.5 percentile) the WRF-4km simulation overestimates the precipitation by 31 % and the WRF-12km by 24 %. The WRF-36km run has with 5 % difference the smallest bias. In general, evaluating precipitation biases in DJF is problematic, because observational errors are largest in this season and observed values are partly underestimating real precipitation by up to 15 % due to the under-catch of snow under normal conditions (Yang et al. 1998) (the wind at the snow telemetry (SNOTEL) stations is normally less than 2 m/s). In case of extreme events, it is very likely that wind speeds and the ratio of under-catching is higher.

By investigating the root-mean-square error (RMSE) and spatial structure of the average DJF extreme shows that the small biases in the WRF-36km simulation comes from an cancellation of positive and negative biases at different locations within the domain. The RMSE of the WRF-4km and WRF-12km run are smaller than the RMSE in the WRF-36km simulation. A clear advantage of the WRF-4km simulation is the accurate representation of spatial variability which is underestimated by 8 % in the WRF-12km and 16 % in the WRF-36km simulations.

WRF-4km has not only the most accurate standard deviation at the point scale but also for all investigated horizontal scales. Also the spatial correlation coefficients are best for all scales and especially for small scales below 50 km. For scales below 100 km the second best correlation coefficients can be found for the WRF-12km run and the worst for WRF-36km. Above 100 km the differences between the correlation coefficients in different resolutions are rather small.

The synoptic situation which lead to extremes in JJA is very different from the situation in DJF. While in DJF warm and wet air is transported with a strong south westerly flow from the Pacific to the Colorado headwaters, in JJA no large scale flow is percent. Instead a labialization of the air leads to precipitation due to deep convection. Therefore, the correct simulation of extreme precipitation in JJA is heavily dependent on the correct representation of convection within the headwaters.

In such situations the WRF-4km simulation outperforms the coarser gridded simulations significantly. The simulations with parameterized deep convection overestimate the domain-wide precipitation by ∼50 % in the WRF-36km and WRF-12km run. The WRF-4km however, stays with 8 % overestimation within the observation error range. Spatial variability is well represented in all simulations and stays within a range of ±10 %.

The spatial correlation coefficients are highest in the WRF-4km simulation for horizontal scales below 20 km. Above that value the correlation coefficients of the WRF-36km run are better than those of the WRF-4km simulation. The WRF-12km run shows generally lower correlation coefficients.

Applying the DCT to the average extreme precipitation fields of DJF shows a strong relation between the spectra of the extremes and the spectra of the orography. The effective resolution of the WRF-12km and WRF-36km simulation is roughly 4 times their grid spacing.

In JJA the extreme precipitation spectra are more independent from the spectra of the underlying orography. Clearly increased variances in long wavelengths are visible in both coarse gridded simulations but especially in the WRF-12km run.

By fitting the Generalized Pareto distribution (GPD) to the daily extreme precipitation data reveals a better representation of 10 year return values in the WRF-4km simulation. Partly the values of the WRF-12km or WRF-36km run have smaller differences to the observations at single mountain ranges or elevation bans but the overall performance is often worse than in the WRF-4km simulation. In general there is an increase in the 10 year return values when the horizontal grid spacing is increaed.

Summarizing the most important findings of this study:

- the 4 km grid spacing is especially important in JJA. Simulations with parameterized deep convection tent to overestimate extreme (but also average) precipitation during this season.
- in DJF the representation of orography is important to accurately simulate extreme precipitation. In the WRF-12km or WRF-36km runs the amplitude of the mean extreme event is accurately simulated but the correlation coefficients are worse than in the WRF-4km simulation.
- the largest improvements of the WRF-4km can be found at small spacial scales below ∼100 km. This is especially important for impact studies (e.g., hydrology studies at catchment scales) which demand for accurate high resolution climatological data.

Also remarkable is the general improvement of the WRF simulations compared to their lateral boundary condition (the NARR data-set). This is a clear prove for the ability of WRF to downscale coarser gridded weather and climate data to finer scales.

Bibliography

Betts, A. K. and M. J. Miller (1986).
'A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX, ATEX and arctic air-mass data sets'.
In: *Quart. J. Roy. Meteor. Soc.* 112, 693–709. See p. 11.

Chen, F. and J. Dudhia (2001).
'Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and ensitivity'.
In: *Mon. Wea. Rev.* 129, 569–585. See p. 11.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values.* 1st ed. Springer. See p. 16.

Collins, W. D. and Coauthors (2006).
'The Community Climate System Model version 3 (CCSM3)'. In: *J. Climate* 19, 2122–2143. See p. 11.

Daly, C. et al. (1994). 'A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountanious Terrain'. In: *J. Appl. Meteor.* 33, 140–158. See p. 12.

Denis, B. et al. (2002). 'Spectral decomposition of two-dimensional atmospheric fields on limited-area domains using the discrete cosine transform (DCT)'.
In: *Mon. Weather Rev.* 130.7 (July 2002), 1812–1829. See p. 15.

Ek, M. B. et al. (2003). 'Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta Model'.
In: *J. Geophys. Res.* 108, p. 8851. See p. 11.

Fisher, R. A. and L. H. C. Tippett (1928). 'Limiting forms of the frequency distribution of the largest or smallest member of a sample'. In: *PROCEEDINGS OF THE CAMBRIDGE PHILOSOPHICAL SOCIETY* 24 (July 1928), 180–190. ISSN: 0008-1981. See p. 16.

Gebremichael, M. and W. F. Krajewski (2004).
'Assessment of the statistical characterization of small-scale rainfall variability from radar: Analysis of TRMM ground validation datasets'.
In: *J. Appl. Meteorol.* 43.8 (Aug. 2004), 1180–1199. See p. 19.

Germann, U. and J. Joss (2001). 'Variograms of radar reflectivity to describe the spatial continuity of Alpine precipitation'. In: *J. Appl. Meteorol.* 40.6, 1042–1059. See p. 19.

Harris, D. et al. (2001). 'Multiscale statistical properties of a high-resolution precipitation forecast'. In: *J. Hydrometeorol.* 2.4, 406–418. See p. 19.

Higgins, R. W. et al. (1996). *A gridded hourly precipitation data base for the United States (1963-1993).* Tech. rep., 46 pp. See p. 13.

Hosking, J. R. M. (1990). 'L-moments: analysis and estimation of distributions using linear combinations of order statistics'. In: *J. R. Stat. Soc. Ser. B-Methodol.* 52.1, 105–124. See p. 17.

Janjic, Z. I. (1994). 'The step-mountain Eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes'. In: *Mon. Wea. Rev.* 122, 927–945. See p. 11.

Kapper, K. L. (2009). *Scale Sensitive Analysis of Regional Climate Models.* Vol. 1. Wegener Center Verlag Graz, p. 114. See p. 15.

Leadbetter, M. R. (1983). 'Extremes and local Dependence in Stationary-Sequences'. In: *Z Wahrschenilichkeit* 65.2, 291–306. See p. 16.

Marzban, C. and S. Sandgathe (2009). 'Verification with Variograms'. In: *Weather Forecast.* 24.4 (Aug. 2009), 1102–1120. See p. 19.

Mesinger, F. and Coauthors (2006). 'North American Regional Reanalysis'. In: *Bull. Amer. Meteor. Soc.* 87, 343–360. See p. 11.

Pickands, J. (1975). 'Statistical Inference Using Extreme Order Statistics'. In: *Ann. Statist.* 3.1, 119–131. See p. 17.

Rasmussen, R. et al. (2011). 'High-Resolution Coupled Climate Runoff Simulations of Seasonal Snowfall over Colorado: A Process Study of Current and Warmer Climate'. In: *J. Climate* 24, 3015–3048. See p. 75.

Serreze, M. C. et al. (1999). 'Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data'. In: *Water Resources Research,* 35.7, pp. 2145–2160. See p. 12.

Skamarock C. W., J. B. K. et al. (2005). *A Description of the Advanced Research WRF Version 2.* Tech. rep., 88 pp. See p. 11.

Stedinger, J. R. et al. (1993). *Frequency analysis of extreme events. In: D.R.Maidment, ed., Handbook of Hydrology.* 1st ed. McGraw-Hill. See p. 17.

Thompson, G. et al. (2008). 'Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization'. In: *Mon. Wea. Rev.* 136, 5095–5115. See p. 11.

Weisman, M. L. et al. (1997). 'The Resolution Dependence of Explicitly Modeled Convective Systems'. In: *Mon. Wea. Rev.* 125.4, 527–548. See p. 11.

Wilks, D. S. (2005). *Statistical Methods in the Atmospheric Sciences.* 2nd ed. Academic Press. ISBN: 0127519661. See p. 17.

Yang, D. et al. (1998). 'Accuracy of NWS 8" Standard Nonrecording Precipitation Gauge: Results and Application of WMO Intercomparison'. In: *J Atmos Sol-Terr Phy* 15, 54–68. See p. 76.

Zepeda-Arce, J. et al. (2000). 'Space-time rainfall organization and its role in validating quantitative precipitation forecasts'. In: *J. Geophys. Res.-Atmos.* 105.D8 (Apr. 2000), 10129–10146. See p. 19.