THE UNIVERSITY OF WESTERN ONTARIO
DEPARTMENT OF CIVIL AND
ENVIRONMENTAL ENGINEERING

Water Resources Research Report

Assessment of Climatic Vulnerability in the Upper Thames River Basin: Part 2

By:
Leanna King
Tarana Solaiman
and
Slobodan P. Simonovic

Report No: 066
Date: August 2010

ISSN: (print) 1913-3200; (online) 1913-3219;
ISBN: (print) 978-0-7714-2834-0; (online) 978-0-7714-2835-7;

Assessment of Climatic Vulnerability in the Upper Thames River Basin: Part 2

By

Leanna M. King

Tarana A. Solaiman

and

Slobodan P. Simonovic

Department of Civil and Environmental Engineering

The University of Western Ontario

London, Ontario, Canada

August 2010

Executive Summary

Increasing greenhouse gas emissions will cause global temperature to rise in the coming years. Understanding the effects of rising temperature on the hydrologic cycle at a local scale is important in order to assess a wide scope of climate change impacts on management of water resources. Coupled Atmosphere-Ocean General Circulation Models (AOGCMs) are state-of-the-art in climate change research, predicting the future climate based on plausible emission scenarios. Their spatial and temporal scales are quite large, so the results of their analyses must be brought to a local-scale through a downscaling process. There are several methods for downscaling AOGCM data; however each method can produce very different results. More work is necessary to develop strategies for climate change impact assessments at a local level.

In this study, statistical downscaling using a modified K-NN weather generator with perturbation and principle component analysis (WG-PCA) is employed to investigate the potential impacts of climatic change in the Upper Thames River basin. A total of 22 stations around the basin are used as inputs, each with 27 years of observed historical data. Monthly change factors are applied to the observed data from six AOGCMs, each with two to three emission scenarios. The resulting datasets are used as inputs to the WG-PCA algorithm to produce 324 years of synthetic data for two time periods, the 2020s (2011-2040) and the 2080s (2071-2100). The performance of the weather generator is evaluated by comparing a synthetic historical dataset to the observed data.

The WG-PCA algorithm is able to satisfactorily reproduce the observed monthly total precipitation values. While there is a slight overestimation in the mean of some months and an underestimation in others, the values of the observed means are well within the inter-quartile range (25^{th} and 75^{th} percentile) of the simulated data, thus performance is considered very good.

The outliers in the historical simulated values indicate the added variability by the WG-PCA weather generator. Total monthly wet-day box plots are made, and results show that there are underestimations of the mean observed data in some months. The statistical hypothesis tests show that the difference between the mean and variance of the observed and simulated precipitation are similar. Frequency distribution curves of wet-spell lengths for winter and summer months also show a very close agreement between the observed and simulated values. Overall, the performance of the WG-PCA weather generator in reproducing historical values is very good.

The AOGCM outputs are compared using box plots of total monthly precipitation values. The results show different predictions of future precipitation, however most models predict an increase in total monthly precipitation for winter for both the 2020s and the 2080s. Summer values are less conclusive as some models predict an increase in total precipitation while other predict a decrease. The general trends of wet-spell intensities show an increase in wet spell intensity for longer time spells for winter in both time periods. For summer wet-spells, both time periods predict that shorter spells will increase in intensity as long ones decrease. The results for AOGCM simulation are quite variable, thus it is important to include several models and emission scenarios in climate change impact assessments.

Table of Contents

List of Tables

List of Figures

1. Introduction

1.1 Background

In the last century, industrialization and an increased dependency on fossil fuels have caused emissions of greenhouse gases to rise significantly. It is now a scientific consensus that the increase in human generated carbon dioxide emissions is a direct contributor to climatic change. A major focus of climate change research is to predict the possible effects of rising temperatures on the natural environment. Recently, the Inter-governmental Panel on Climate Change (IPCC) published the 4th Assessment Report which predicted that increasing CO_2 emissions will cause the Earth's average temperature to rise by 3°C by the year 2080 (IPCC, 2007). The rising temperature will have a major effect on the atmospheric processes and thus impact the frequency and amount of precipitation an area receives. As such, it is crucial to perform hydrologic impact assessments of climate change in order to better understand and predict the future extreme events.

Changes in climatic variables which drive the hydrologic cycle will have a profound effect on the spatial and temporal distribution of water. Understanding these effects is critical in order to prepare the impacted population to cope with the resulting floods or droughts. The 4th Assessment Report has predicted with high confidence that in high latitude regions, annual river runoff will increase (IPCC, 2007). Future increases in the frequency or intensity of extreme precipitation events are predicted to have the largest impact on the hydrologic cycle in South-western Ontario (Prodanovic and Simonovic, 2007). Rising stream flows and runoff are a direct result of such extreme precipitation events, so it is crucial to forecast these when assessing flooding risks in a river basin (Zhang et al., 2008). Global scale predictions of the future climate are made under greenhouse gas forcing, and then scaled down to represent the local climate

which can then be analyzed in terms of hydrologic extremes (Kilsby et al., 2007). It is important to project future hydrologic extremes for effective water resource management and planning practices.

Global Climate Models are an important tool to project future climate and widely used in hydrologic impact assessment studies. Weather variables from these models are used as inputs with several plausible emissions scenarios outlined by the IPCC in the Special Report on Emissions Scenarios (CCCSN, 2010). The scenarios are based on plausible development storylines, taking into account population growth, globalization, economic growth, environmental awareness and advancements in technology (CCCSN, 2010). However, the global models have their own shortcomings: the Atmosphere Ocean integrated Global Climate Models (AOGCM) predict time-series of future meteorological variables at relatively coarser grid points around the globe, typically around 100 x 100 km (Prodanovic and Simonovic, 2007), which may not be suitable for hydrologic impact studies at small to medium basins. They are, thus, downscaled to the appropriate size for the area of interest; however, the accuracy of predictions may decrease at finer spatial scales.

In this study, statistical downscaling with the WG-PCA weather generator was used to generate 324 years of future climate data representing two time periods, 2011-2040, and 2071-2100. The Present study is a continuation of the Water Resources Research Report no. 64 (King et al, 2009) which dealt with assessment of climatic variability taking into consideration the 2050s time slice.

The WG-PCA weather generator uses a K-Nearest Neighbour algorithm with principle component analysis and perturbation. In addition to three atmospheric variables (minimum temperature, maximum temperature and precipitation) used in the previous impact assessment

study, four additional variables are included as inputs: eastward wind component, northward wind component, specific humidity and mean sea level pressure. Statistical tests are used to check the importance of the additional inputs. Synthetic datasets for six AOGCM models are created, each with up to three emissions scenarios. Monthly change factors for each model are calculated and applied to the observed data for use with the weather generator. The present report thus focuses on (a) evaluating the performance of the weather generator in simulating the observed climate, and (b) investigating the impacts of climate change on extreme precipitation events, as predicted by several AOGCM models and emissions scenarios.

1.2 Organization of the Report

The organization of the report is as follows: A review of the literature on downscaling methods and weather generators is presented in section 2. Following is the methodology, including data pre-processing and the WG-PCA algorithm used. Results and discussion can be found in section 4, with conclusions and recommendations in section 5.

2. Literature Review

Current literature outlines several methods for downscaling AOGCM data for application at a local basin scale. AOGCM outputs can be scaled down directly, however, Trigo and Palutikof (2001) found that the models failed to predict the high variability in daily precipitation, and could not accurately simulate present-day monthly precipitation amounts (Brissette et al., 2006).

Dynamic downscaling involves the use of Regional Climate Models (RCMs), which are at a much higher resolution and can simulate climatic variables more accurately in the region of

interest (Brissette et al., 2006). AOGCM output variables are used as boundary inputs to the RCM's, which provide a more accurate representation of the local climate than the coarsely gridded AOGCM data. The works of Vidal and Wade (2008), Wood et al. (2004) and Schmidli et al. (2006) compared dynamic downscaling to other methods. A limitation of the dynamic approach is the scale of RCM's (approximately 40 km x 40 km according to Brissette et al., 2006), still too coarse for application to smaller basins. The computational effort required for the dynamic approach makes it very impractical where several AOGCMs and emissions scenarios are used (Maurer, 2007). Furthermore, RCMs have only been produced for select areas and AOGCM inputs and their availability is very limited (Kay and Davies, 2008).

The second approach, statistical downscaling, is more popular in climate change impact assessments due to its computational ease and the ability to produce synthetic datasets of any desired length. Statistical downscaling is the process of developing or assuming statistical relationships between the observed basin-scale climate and the AOGCM outputs (Brissette et al., 2006). The underlying assumption in this approach is that the future climate is governed by the same relationships as the historical one.

There are several methods of statistical downscaling, the simplest of which directly reduces the scale of the gridded AOGCM data to a basin size. While the direct downscaling of minimum and maximum temperature has produced good results, precipitation values are not well reproduced directly from AOGCM data (Brissette et al., 2006). Because of the inferior quality and unavailability of daily inputs from many AOGCMs, monthly inputs should be used. Monthly change factors from any AOGCMs are applied to the observed historical data. The altered dataset is used as an input into a weather generator, which is a complex number generating algorithm that can produce synthetic datasets of any length for the time period of interest.

Weather generators have an advantage over other downscaling methods because by producing long duration rainfall series, it is possible to examine rare events and extremes in the river basin (Brissette et al., 2007; Diaz-Nieto and Wilby, 2005; Wilks and Wilby, 1999).

Three major types of weather generators: parametric, empirical or semi-parametric, and non-parametric (Brissette et al., 2007) are more common among the scientific community.

In most parametric weather generators, a Markov chain is used to determine the probability of a wet or dry day and a probability distribution is assumed to determine the amount of precipitation (Kuchar, 2004; Hanson and Johnson, 1998). Most of the parametric weather generators are extensions of Richardson's WGEN, which was developed in 1981 (Richardson, 1981). Some examples of parametric weather generators successfully employed using the Richardson approach are CLIGEN, WGENK, GEM, WXGEN, and SIMMENTO (Kuchar, 2004; Schoof et al., 2005; Hanson and Johnson, 1998; Soltani and Hoogenboom, 2003). Hanson and Johnson compared simulations using GEM to historical data using the means and standard deviations. Results from that study showed that simulated total precipitation values were significantly underestimated for some months, and annual precipitation values were considerably less than the historical record (Hanson and Johnson, 1998). A study employing the SIMMENTO weather generator found that the variability (standard deviations) of wet fractions and amounts were significantly overestimated by the synthetic historical series (Elshamy et al., 2006). A major drawback of the parametric approach is that the Markov chain does not take into account the previous days' weather, and thus rare events, such as droughts or wet spells are not well produced (Sharif and Burn, 2007; Semenov and Barrow, 1997; Dibike and Coulibaly, 2005). Another limitation of the parametric weather generators is that an assumption must be made about the probability distribution of precipitation amounts, and different distributions do not give

similar results (Sharif and Burn, 2007). Furthermore, the weather generators cannot be easily transferred to other basins as their underlying probability assumptions would change (Sharif and Burn, 2007). The computational effort is also significantly higher than other methods since many parameters must be estimated and statistically verified (Mehrotra et al., 2006). Parametric weather generators are less easily applied to multiple sites as simulations occur independently and thus spatial correlations would have to be assumed.

Semi-Parametric or Empirical weather generators include LARS-WG and the Wilks model, SDSM (Semenov and Barrow, 1997; Wilks and Wilby, 1999). LARS-WG differs from the parametric approaches described above because it employs a series-approach in which the wet and dry spells are determined by taking into account the observed values and assuming mixed-exponential distributions for dry/wet series as well as precipitation amounts (Semenov and Barrow, 1997). The wet/dry day status is first chosen, and then the amount is chosen conditional on the status. As such, the LARS-WG was able to satisfactorily reproduce wet and dry spells, unlike the parametric weather generators (Dibike and Coulibaly, 2005). Wilks (1998) improved on the parametric models of Richardson (1981) by introducing Markov-chains of higher order that have a better "memory" of the preceding weather. They further extended the Richardson (1981) model for multi-site applications by using a collection of single site models in which a conditional probability distribution is specified and thus spatially correlated random numbers can be generated (Mehrotra, 2006; Wilks, 1998). A drawback to these empirical approaches is that there is still a subjective assumption about the type of probability distribution for precipitation amounts and spell lengths, and the spatial correlation structure is empirically estimated for use with multiple sites.

Non-parametric weather generators are computationally simple and do not require any statistical assumptions to be made. They work by using a nearest-neighbour resampling procedure known as the K-NN approach (Sharif and Burn, 2007; Brandsma and Buishand, 1998; Beersma et al., 2002; Yates et al., 2003). The nearest neighbour algorithm works by searching the days in the historical record which have similar characteristics to those of the previously simulated day, and randomly selecting one of these as the simulated value for the next day (Beersma et al., 2002). This approach is easily used in multi-site studies because the values are simulated concurrently, thus spatial correlation is preserved (Mehrotra et al. 2006). The K-NN algorithm has been successfully used for multi-site hydrological impact assessments in the Rhine Basin, accurately preserving spatial correlation and climatic variability (Beersma et al., 2002; Brandsma and Buishand, 1998). Apipattanavis et al. (2007) compared a K-NN to a semi-parametric weather generator. Box plots of wet-spell lengths showed that for some months the semi-parametric model could not reproduce maximum wet spell lengths, and average spell lengths were underestimated for the traditional K-NN model. A major limitation to the K-NN approach is that the values are merely reshuffled, thus no new values are produced (Sharif and Burn, 2007). Climatic extremes are essential in predicting flooding events in response to climate change, thus Sharif and Burn (2007) modified the K-NN algorithm to produce unprecedented precipitation amounts by introducing a perturbation component in which a random component is added to the resampled data points (Sharif and Burn, 2007). Monthly total precipitation and total monthly wet day box plots were used to evaluate the performance of the Modified K-NN algorithm. The algorithm was found to satisfactorily reproduce the statistics of the original dataset while adding variability which is crucial in hydrologic impact assessments (Sharif and Burn, 2007). Prodanovic and Simonovic (2006) altered the Modified K-NN algorithm of Sharif and Burn (2007) to account for the leap year and transfer to Java programming language. Eum et

al. (2009) then revised the weather generator of Sharif and Burn (2007) by adding principal component analysis, which allows the user to include more variables for an improved selection of nearest neighbours. The updated WG-PCA model is able to more accurately define the current day's weather because several more variables can be used in the selection of the nearest neighbour (Eum et al., 2009).

3. Study Area and Data

3.1 Study Area

This study focuses on the Upper Thames River basin, which is located between the great lakes of Erie and Huron in South Western Ontario, Canada. The total population of the basin is 420,000, with 350,000 of those residing in London, the main urban centre. The basin covers three counties: Perth, Middlesex and Oxford, and has a total area of 3500km^2 (Prodanovic and Simonovic, 2008). The land use is mainly agricultural and urban, with some forested areas. The Thames river has two branches: a North branch which flows south into London from Mitchell and St. Marys, and an East branch flowing westward through Ingersoll and Woodstock. The two branches meet near London's city centre and continue westward towards Byron, ending in Lake St. Clair. The total length is 273 km, with an average discharge of 39.5 m^3/s (Prodanovic and Simonovic, 2008). Figure 1 shows a schematic location map of the Upper Thames River basin.

Seasonal flooding has historically been a major hazard for the Upper Thames River basin. Typically, flooding occurs in early March during snowmelt, and in the summer seasons as a result of extreme rainfall events. In 1937, the city of London experienced a massive flooding event which eventually sparked the creation of the Upper Thames River Conservation Authority.

Since then, three major water management reservoirs were created, namely Pittock, Wildwood, and Fanshawe (Prodanovic and Simonovic, 2007). Several weather stations are located throughout the basin to provide point measurements of climatic variables and are documented on Environment Canada's website. Stations chosen for this study are listed in Table 1.

Figure 1: Schematic location map of Upper Thames River basin

Table 1: Location of Stations in the Upper Thames River basin

Station	Latitude (deg N)	Longitude (deg W)	Elevation (m)
Blyth	43.72	81.38	350.5
Brantford MOE	43.13	80.23	196.0
Chatham	42.38	82.2	198.0
Delhi CS	42.87	80.55	255.1
Dorchester	43.00	81.03	271.3
Embro	43.25	80.93	358.1
Exeter	43.35	81.50	262.1
Fergus	43.73	80.33	410.0
Foldens	43.02	80.78	328.0
Glen Allan	43.68	80.71	404.0
Hamilton A	43.17	79.93	238.0
Ilderton	43.05	81.43	266.7
London A	43.03	81.16	278.0
Petrolia Town	42.86	82.17	201.2
Ridgetown	42.45	81.88	210.3
Sarnia	43.00	82.32	191.0
Stratford	43.37	81.00	354.0
St. Thomas WPCP	42.78	81.21	209.0
Tillsonburg	42.86	80.72	270.0
Waterloo Wellington	43.46	80.38	317.0
Woodstock	43.14	80.77	282.0
Wroxeter	43.86	81.15	355.0

3.2 Data

The inputs used in the present study are selected from several sources. It has been found that except precipitation and temperature, all stations within the area of interest do not come with other meteorological variables which can be of interest for generating precipitation. Solaiman and Simonovic (2010) have assessed the application of North American Regional Reanalysis (NARR) database for several stations in the Upper Thames River basin and concluded that the outputs from NARR can be used along with the observations in order to generate precipitation more precisely. So for the current research the following inputs are used:

- Daily weather data (precipitation, maximum temperature and minimum temperature) for the period of 1979-2005 was obtained from Environment Canada's website (http://www.climate.weatheroffice.gc.ca/climateData/canada_e.html) for each of the stations listed in Table 1. Stations were chosen based on the completeness and length of the observed data.

- North American Regional Reanalysis (NARR): The NARR is an extension of the global reanalyses which uses very high resolution Eta model (0.3° × 0.3°, 32 km grid spacing, 45 layersspatially) with the Regional Data Assimilation System (RDAS). Most of the variables are collected 8 times daily; daily and monthly means are also available at 29 pressure levels. NARR dataset has been developed by assimilating high quality and detailed precipitation observations into the atmospheric analysis, which consequently made the forcing to the land surface model component of the system more accurate and hence, a much improved analysis of land hydrology and land-atmosphere interaction has been become possible (Nigam and Ruiz-Barradas 2006). However, one significant weakness of NARR data for over Canadian regions is that the daily gauge-based data it uses for assimilation are sparse (1 degree grid), which may be insufficient for the model to perform as expected (www.emc.ncep.noaa.gov/mmb/rreanl/narr.ppt). NARR data for this study has been made available through the Data Access Integration of the Canadian Climate Change Scenarios Network of Environment Canada. For this study, the gridded data is interpolated for the stations of interest for Upper Thames River basin and used as inputs along with the observed data.

- The Canadian Climate Change Scenarios Network (CCCSN) provides access to several AOGCM models and emissions scenarios. The website allows the user to specify the range of geographical co-ordinates required, as well as the climatic variable and time period of interest.

For the purpose of this study, the time slices collected were 1960-1990 (baseline), 2011-2040 (2020's) and 2071-2100 (2080s). Seven variables were chosen: minimum temperature, maximum temperature, precipitation, specific humidity, northward wind component, southward wind component and mean sea level pressure. Six AOGCM models were collected, each with two to three emissions scenarios, as specified by the IPCC's Special Report on Emissions Scenarios (Nakicenovic et al, 2000). Full descriptions of the emissions scenarios can be found in Appendix A. Table 2 lists the AOGCM's along with the emissions scenarios available and their origin. Appendix B provides descriptions of each AOGCM.

Table 2: List of AOGCM Models and Emissions Scenarios used

GCM Models	**Sponsors, Country**	**SRES Scenarios**	**Atmospheric Resolution**	
			Lat	**Long**
CGCM3T47, 2005	Canadian Centre for Climate Modelling and Analysis, Canada	A1B, A2, B1	3.75°	3.75°
CGCM3T63, 2005		A1B, A2, B1	2.81°	2.81°
CSIROMK3.5, 2001	Commonwealth Scientific and Industrial Research Organization (CISRO) Atmospheric Research, Australia	A2, B1	1.875°	1.875°
GISSAOM, 2004	National Aeronautics and Space Administration (NASA)/ Goddard Institute for Space Studies (GISS), USA	A1B, B1	3°	4°
MIROC3.2HIRES, 2004	Centre for Climate System Research (University of Tokyo), National Institute for Environmental Studies, and Frontier Research Centre for Global Change (JAMSTEC), Japan	A1B, B1	1.125°	1.125°
MIROC3.2MEDRES, 2004		A1B, A2, B1	2.8°	2.8°

4. Methodology

4.1 Data Pre-processing

Pre-processing of AOGCM data for each time period is carried out in two steps, outlined in the following sections.

4.1.1 Spatial Interpolation of AOGCM Outputs

The AOGCM outputs are available as gridded values, so in order to produce a separate dataset for each station of interest the values are interpolated using the Inverse Distance Weighting Method (IDW). The four nearest grid points to the station of interest are chosen, and the distance, *d* from the station to each point is computed. Each gridded value is then assigned a weight, *w* using Equation 4.1. The weighted average of the variable, *p*, for the station is then computed using Equation 4.2 (where the subscript *j* represents the j^{th} grid point and the subscript *i* represents the station).

$$w_1 = \frac{1/d_1^2}{1/d_1^2 + 1/d_2^2 + 1/d_3^2 + 1/d_4^2} \tag{4.1}$$

$$p_i(t) = \sum_{j=1}^{4} w_j p_j(t) \tag{4.2}$$

4.1.2 Calculation of Change Factors from AOGCM Outputs

Using the AOGCM datasets for each station, monthly averages are computed for each variable for both the baseline (1960-1990) and the future time slice (2011-2040 or 2071-2100). For maximum temperature, minimum temperature, northward wind speed, eastward wind speed and mean sea level pressure, the monthly change factors are computed as the difference between

the baseline and the future averages. For precipitation and humidity, the change factors are taken as the percent change between the baseline and the future averages.

Next, the change factors for each AOGCM scenario are used to modify the historical daily data for each station gathered from Environment Canada. The historical daily data for humidity and precipitation are multiplied by the monthly change factors. For the rest of the variables, the change factors are added to modify the historical data. Finally, the modified datasets for each time slice and AOGCM are used as inputs into the WG-PCA weather generator in order to produce synthetic datasets for 324 years.

4.2 Weather Generator

The weather generator used in this study, WG-PCA was developed by Eum et al., (2009) and uses a K-Nearest Neighbour algorithm with perturbation and principal component analysis.

The WG-PCA algorithm with p variables and q stations has the following steps:

1) Regional means of p variables for all q stations are calculated for each day of the observed data:

$$\overline{X}_t = \lfloor \overline{x}_{1,t}, \overline{x}_{2,t}, ..., \overline{x}_{p,t} \rfloor \quad \forall t = \{1,2,..., T\} \tag{4.3}$$

$$\text{where } \overline{x}_{i,t} = \frac{1}{q}\sum_{j=1}^{q} x_{i,t}^{j} \quad \forall i = \{1,2,..., p\} \tag{4.4}$$

2) Selection of potential neighbours, L days long where $L=(w+1) \times (N-1)$ for each of p individual variables with N years of historical record, and a temporal window of size w which can be set by the user of the weather generator. The days within the given window are all potential neighbours to the feature vector. N data which correspond to the current day are deleted from the potential neighbours so the value of the current day is not

repeated (Eum et al, 2006).

3) Regional means of the potential neighbours are calculated for each day at all q stations.

4) A covariance matrix, C_t of size $L \times p$ is computed for day t.

5) The first time step value is randomly selected for each of p variables from all current day values in the historical record.

6) (a) From the covariance matrix, (C_t) the eigenvector and eigenvalue are calculated.

(b) Selection of the eigenvector corresponding to the eigenvalue which represents the highest fraction of variance in the p variables.

(c) The first principle component is calculated from Equations (4.5) and (4.6) using the eigenvector, E, found in (b). PC_t is the value of the current day and PC_k is the nearest neighbour transferred by the eigenvector in (b).

$$PC_t = \overline{X}_t E \tag{4.5}$$

$$PC_k = \overline{X}_k E \tag{4.6}$$

(d) The Mahalanobis distance is calculated with Equation (4.7) from the one dimensional matrix calculated by the above equations.

$$d_k = \sqrt{(PC_t - PC_k)^2 / Var(PC)} \quad \forall k = \{1,2,...,K\} \tag{4.7}$$

Where the variance of the first principle component is *Var(PC)* for all K nearest

neighbours (Eum et al, 2006).

7) The selection of the number of nearest neighbours, K, out of L potential values using $K = \sqrt{L}$.

8) The Mahalanobis distance d_k is put in order of smallest to largest, and the first K neighbours in the sorted list are selected (the K nearest neighbours). A discrete probability distribution is used which weights closer neighbours highest in order to resample out of the set of K neighbours. Using Equations (4.8) and (4.9), the weights, w, are calculated for each k neighbour.

$$w_k = \frac{1/k}{\sum_{i=1}^{K} 1/i} \qquad \forall k = \{1,2,...,K\} \tag{4.8}$$

Cumulative probabilities, p_j, are given by:

$$p_j = \sum_{i=1}^{j} w_i \tag{4.9}$$

9) A random number u (0,1) is generated and compared to the cumulative probability calculated above in order to select the current day's nearest neighbour. If $p_1 < u < p_k$, then day j for which u is closest to p_j is selected. However, if $p_i > u$, then the day which corresponds to $d1$ is chosen. If $u = p_K$, then the day which corresponds to day d_K is selected. Upon selecting the nearest neighbour, the K-NN algorithm chooses the weather of the selected day for all stations in order to preserve spatial correlation in the data (Eum et al, 2006).

10) In order to generate values outside the observed range, perturbation is used. A conditional standard deviation, σ, is estimated and using Equation (4.10) a bandwidth, λ is determined.

$$\lambda = 1.06\,\sigma K^{1/5} \tag{4.10}$$

Perturbation is next, using equation 4.11.

$$y_{i,t}^{j} = x_{i,t}^{j} + \lambda \sigma_{i}^{j} z_{t} \tag{4.11}$$

Where $x_{i,t}^{j}$ is the weather variable obtained in step 9, $y_{i,t}^{j}$ is the value of that variable obtained after perturbation, z_t is a random variable which is normally distributed (zero mean, unit variance) for day *t*. Negative values are prevented from being produced for precipitation by employing a largest acceptable bandwidth: $\lambda_a = x_{*,t}^{j} / 1.55\sigma_{*}^{j}$ where * refers to precipitation. If again a negative value is returned, a new value for z_t is generated (Sharif et al, 2006).

5. Results and Discussion

The WG-PCA stochastic weather generator is used to produce 324 years of synthetic data for the historical climate as well as 15 different AOGCM scenarios for two time slices (2020s and 2080s). Seven input variables are used to provide more information in selecting the nearest neighbours. A temporal window of 14 days is used along with 27 years of data, resulting in 404 days as potential nearest neighbours. The WG-PCA precipitation outputs from London are analyzed to prevent redundancy. The performance of the weather generator in reproducing the historical data is evaluated, and a comparison of AOGCM model outputs is presented in the

following sections. Box plots are used to demonstrate the variability of the datasets. The height of the box indicates the inter-quartile range (between the 25^{th} and 75^{th} percentiles) and the horizontal line across the box is the median. The whiskers show from the 5^{th} to the 95^{th} percentile with outliers marked as black dots.

5.1 WG-PCA Performance Evaluation

The WG-PCA weather generator is used to produce 324 years of synthetic historical data using the observed historical data as an input. The ability of the weather generator to reproduce precipitation amounts and frequencies is investigated. Figure 2 provides the total monthly precipitation box plots of the simulated data with the historical means plotted in red. From the plots, it can be seen that in most months there is a fairly close agreement between the simulated median and observed mean values. There is a slight overestimation of the means for total monthly precipitation in February, April, May and September. The mean values are slightly underestimated for August and October. For the remaining months, the simulated and observed values are very close. There are several values lying beyond the whiskers indicating an increased variability in the simulated data. Because of the high spatial and temporal variability of precipitation as compared to other climatic variables, the WG-PCA's performance in simulating total monthly precipitation is considered to be very good.

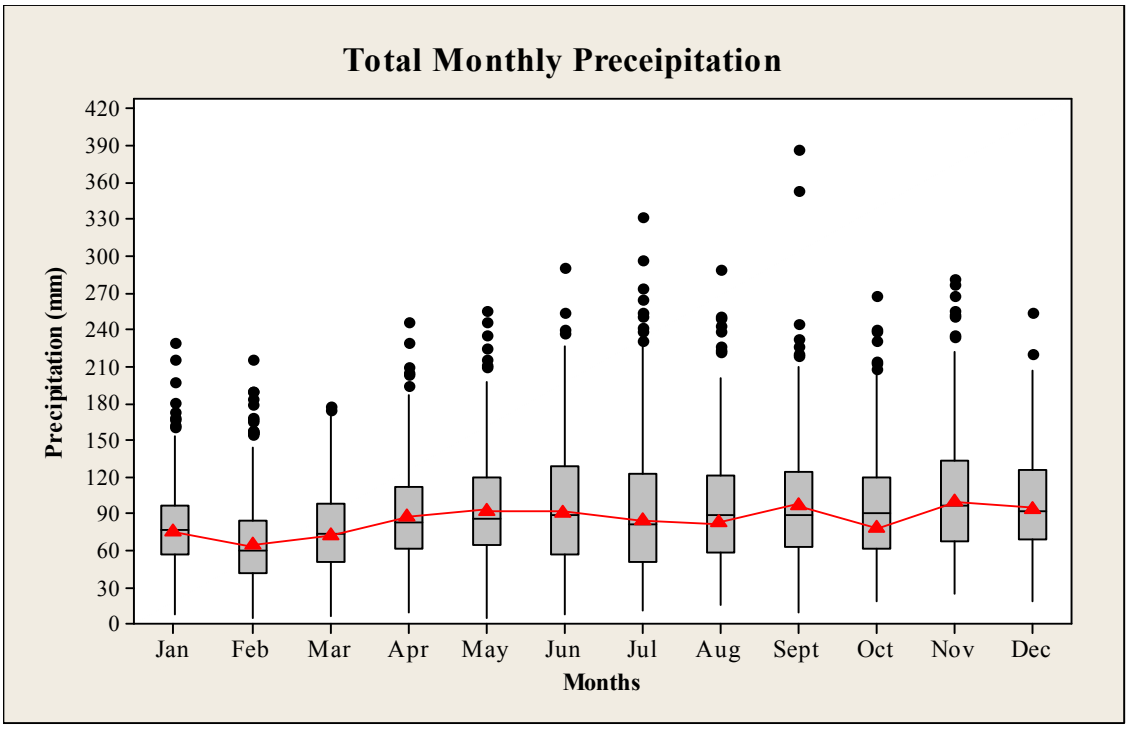


Figure 2: Total monthly precipitation box plot for historical simulated data, with the observed means shown as a line plot in red and outliers as black dots

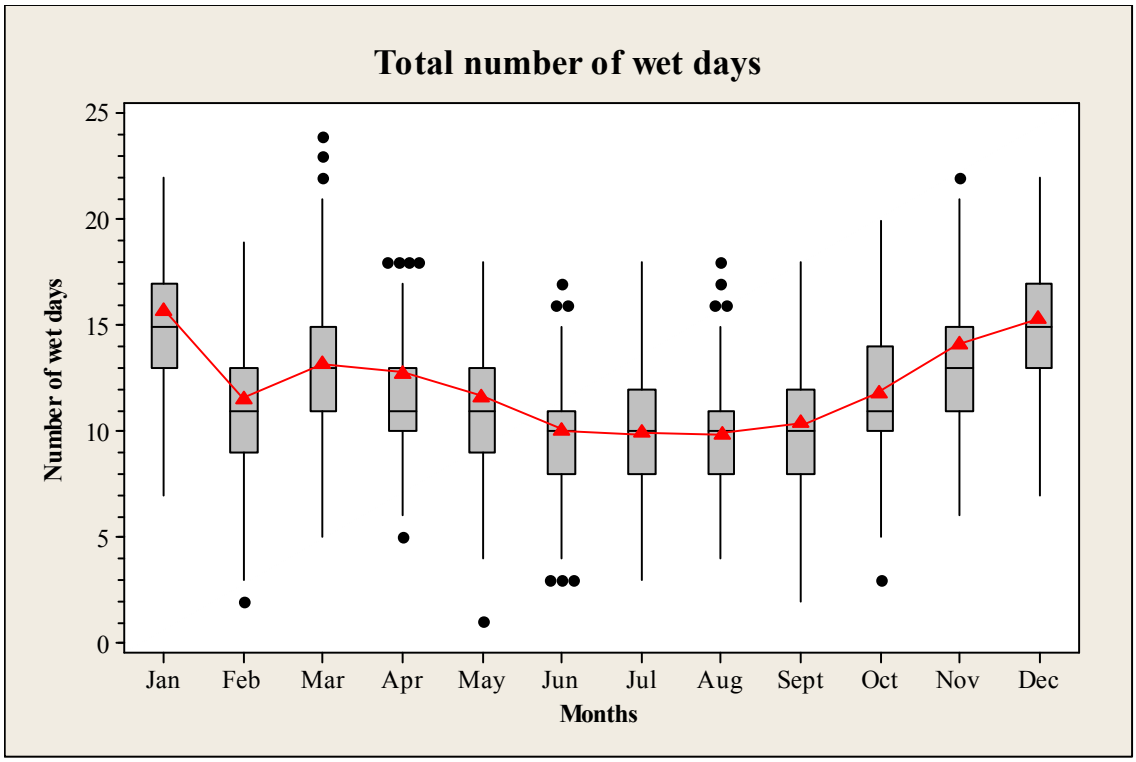


Figure 3: Total number of wet days box plot for historical simulated data, with the observed means shown as a line plot in red and outliers as black dots

Box plots of total monthly wet days in the simulated data are found in Figure 3, with the observed mean values plotted in red. Wet days are defined as days with more than 0.5 mm of precipitation. When compared with the observed means, it can be seen from the plots that the simulated median number of wet days is underestimated for winter, spring and fall seasons. The greatest underestimations are seen in April and November. There is a very close agreement in the summer values. While there is an underestimation for several months, the whiskers of the data extend at least four of five days on either side of the inter-quartile rangeindicating the high variability in the simulated data. Several outliers are seen both above and below the whiskers, with outliers of the same value plotted beside one another. In April there are four simulations with 18 monthly wet days, however the amount of rain occurring in these simulations varied from 89.5 to 228.68 mm. In March, one simulation produced 24 monthly wet days. Although many of the simulated median values significantly underestimate the observed means, the whiskers and outliers indicate that the variability in the data was very high. The performance of the weather generator was thus satisfactory in monthly wet day simulation.

Next, errors in the estimates of mean and variance of generated precipitation are evaluated using statistical hypothesis test at 95% confidence level. One of the best methods for constructing a hypothesis test *p* value for the difference of two population means is modified (Welch) *t* test. It is used to compute a confidence interval and performs a hypothesis test of the difference between two population means when distributions are unknown and samples are drawn independently from each other. This procedure is based upon *t* distribution and for small samples it works best if data are drawn from distributions that are normal or close to normal. The confidence increases with the increase of sample size. The results are shown in Table 3 for winter and summer. It is shown that in most cases the *p* value at 95% confidence level is above

the threshold 0.05 which clearly indicates that there is no evidence of different means between the observed and generated precipitations.

Table 3: Test results (*p* values) of the *t*-test (modified Welch) for the difference inmeans of observed and simulated precipitation during the period 1979-2005 at 95% confidence level

Blocks	Summer	Winter
1	0.0672	0.108
2	0.3661	0.091
3	0.7257	0.362
4	0.1636	0.029
5	0.2612	0.01
6	0.4854	0.056
7	0.0103	0.358
8	0.2696	0.024
9	0.2696	0.024
10	0.1837	0.239
11	0.1837	0.082
12	0.0269	0.527

Table 4: Test results (*p* values) of the Levene's test for the equality invariances of observed and simulated precipitation during the period 1979-2005 at 95% confidence level

Blocks	Summer	Winter
1	0.039	0.358
2	0	0.743
3	1	0.009
4	0.562	0.99
5	0	0.812
6	0.985	0.015
7	0.965	0.999
8	0.636	0.203
9	0.636	0.203
10	0.909	0.041
11	0.909	0.186
12	0.015	0.338

The second test isthe modified version of Levene's test (Levene 1980) for testing the equality of two sample population variances as proposed by Brown and Forsythe (1974). This method considers the distances of the observations from their sample median rather than their sample mean which makes the test more robust with a data following a skewed distribution. The results of the Levene's test for the equality of variances of observed and simulated precipitation at 95% confidence level are presented in Table 4. Similar to the *t* test, the p values appear above 0.05 thresholds, clearly indicating the similarity of the variances. So, the observed and the simulated precipitation in most cases can be assumed to have equal variances.

Frequency distributions of wet-spell lengths for winter and summer months are plotted in Figures 4 and 5. A comparison of observed and simulated values for wet-spell lengths shows that there is very close agreement in the frequency distributions. In all winter months, the frequency of one-day wet spells is slightly overestimated in the simulated data. For January and February, there is a very slight underestimation of two day wet-spells by the simulated data. Three day wet-spells in February and four day wet-spells in December are both underestimated by about six percent. For all other values, the frequency distributions agreed very closely. In Figure 5, it is clear that the same is true for summer wet-spell lengths. The frequency of wet-spell lengths in the simulated data for June and August are almost identical to the observed values. There is an overestimation for one day wet spells in July, and a slight underestimation for two and three day spells. In August there is a slight underestimation in the frequency of three day wet-spells by the simulated data. These slight differences could be due to the difference in data lengths, as the simulated values produce a more smoothed frequency curve since there are a greater number of data points. The performance of the weather generator in reproducing wet-spell lengths is very good.

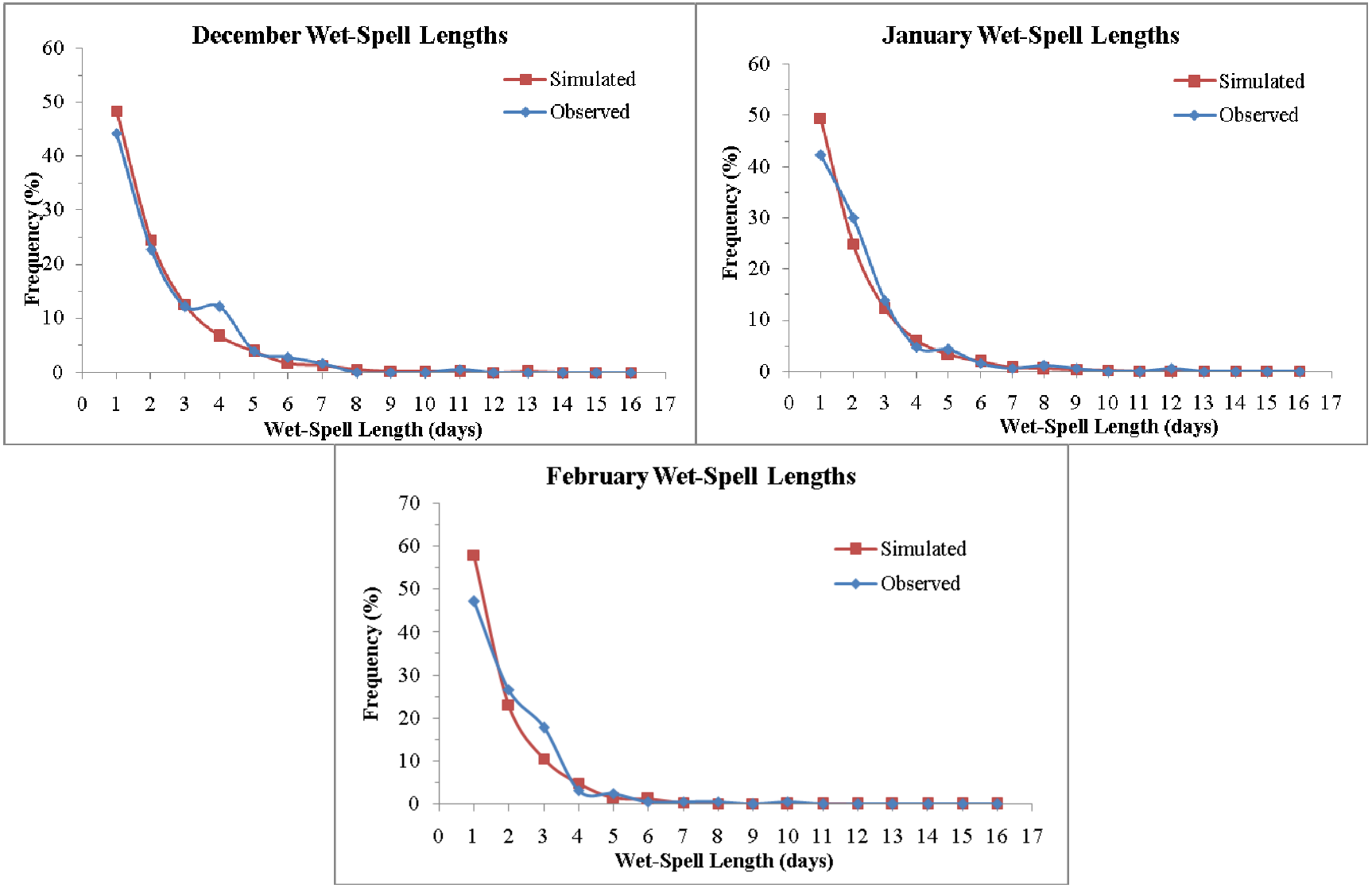


Figure 4: Frequency plots of wet spell lengths for December, January and February including historical simulated and observed data

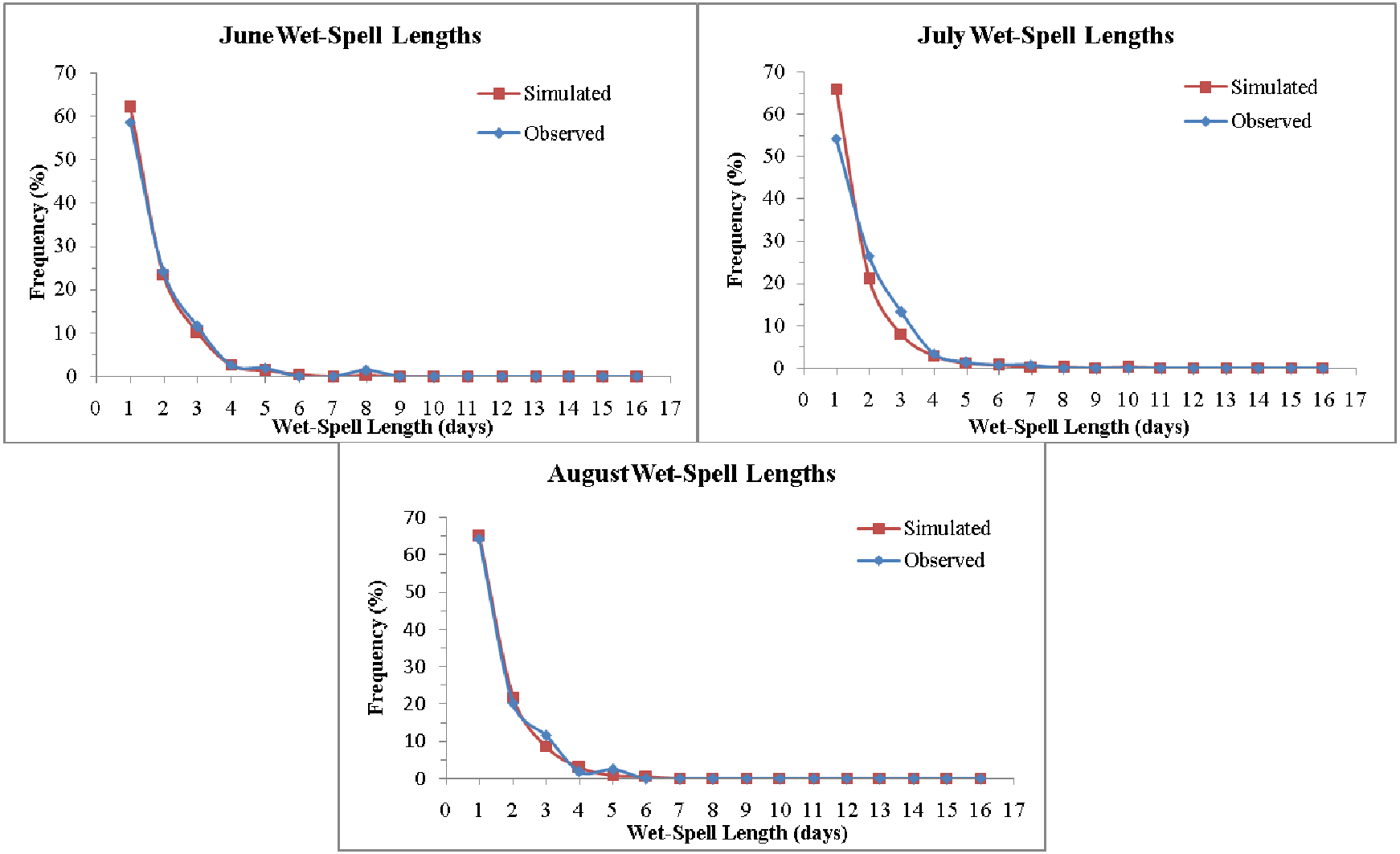


Figure 5: Frequency plots of wet spell lengths for June, July and August including historical simulated and observed data

5.2 Generation of Future Climate Scenarios

The WG-PCA weather generator is used to produce 324 years of future data for 15 AOGCM models and emission scenarios, representing the 2020s (2011-2040) and 2080s (2071-2100). Refer to Table 2 for details on the AOGCM's and emission scenarios used. Appendix C contains the Figures for 2020s which will not be included in the text to prevent redundancy.

5.2.1 Monthly Change Factors

The AOGCM predicted percent change in monthly precipitation from the baseline data to the 2020s and 2080s are shown in Figures 6 and 7. While different models project increases and decreases for the same month, many models project an increase in winter precipitation for the 2020s, and all but three models (CSIROMK3.5 B1, CGCM3T63 B1 and MIROC3.2MEDRES B1) predict an increase by the 2080s. All but three models predict an increase in November precipitation for the 2020s, and only two (from CSIROMK3.5) predict a decrease by the 2080s. While the change factors for spring precipitation in the 2020s vary, all models but three(MIROC3.2HIRES A1B, B1 and CSIROMK3.5 B1) predict an increase for the 2080s in spring. For both time periods, the predictions for summer and fall vary depending on the model. Overall, most models predict an increase in winter, spring and late fall precipitation by the 2080s. The predictions for summer vary, in both time periods, with some small increases and large decreases in precipitation by the 2080s. The varying AOGCM predictions indicate the importance of a thorough analysis employing several models and emission scenarios, as the accuracy of each model remains unknown.

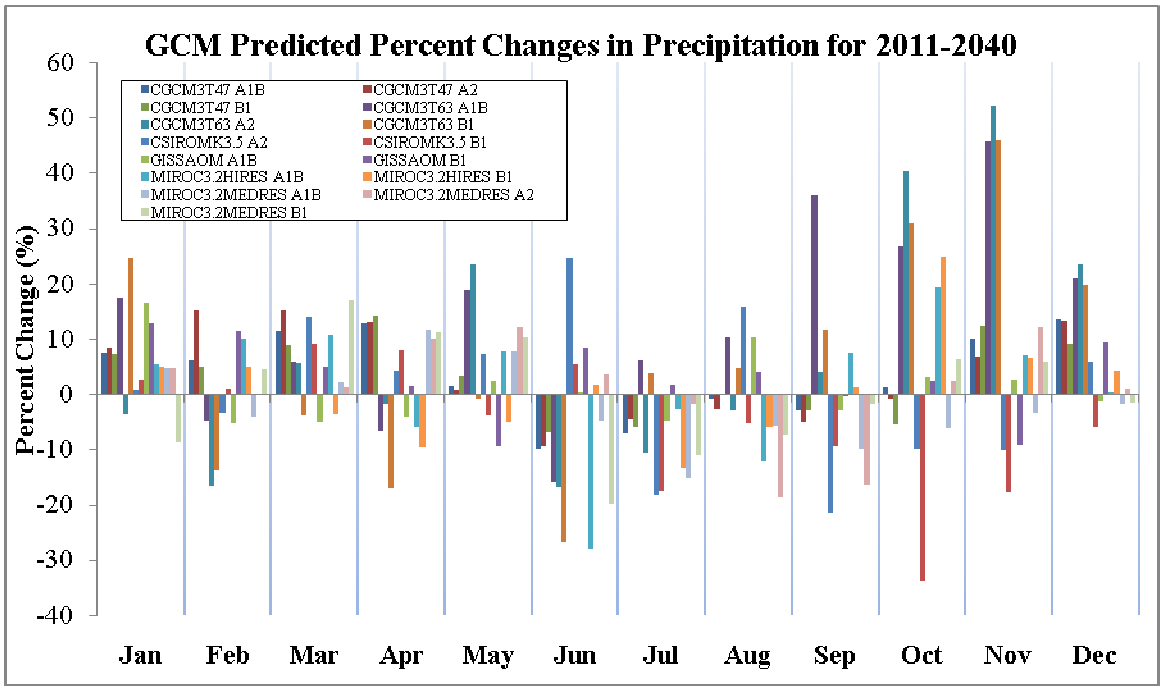


Figure 6: AOGCM predicted percent changes in precipitation for 2011-2040

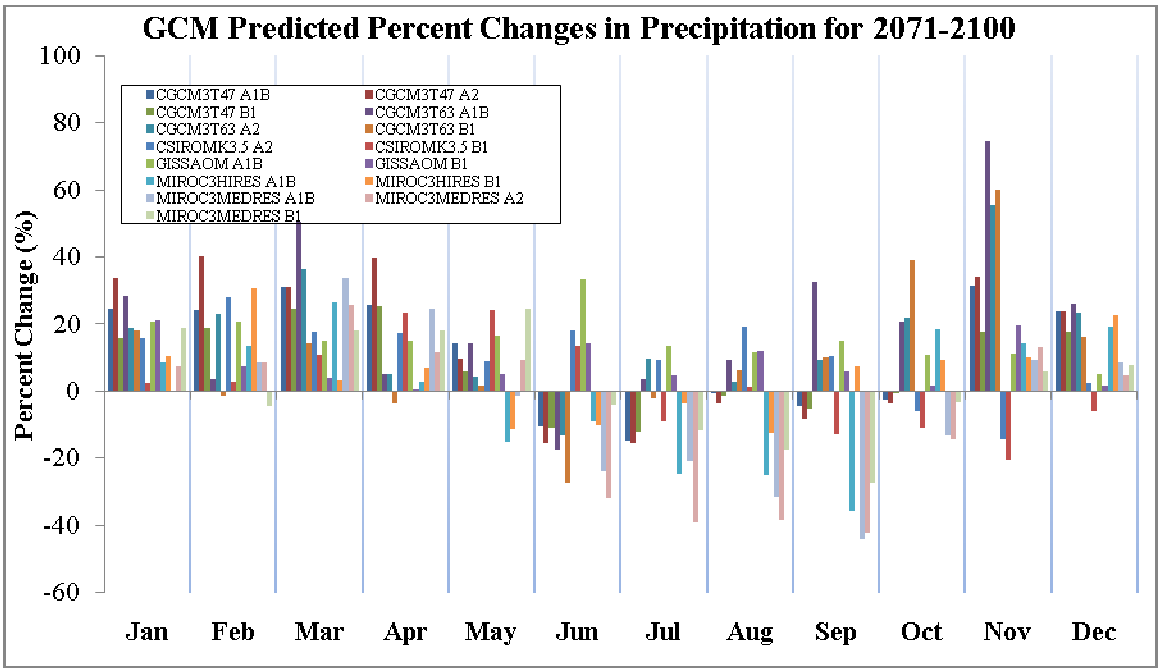


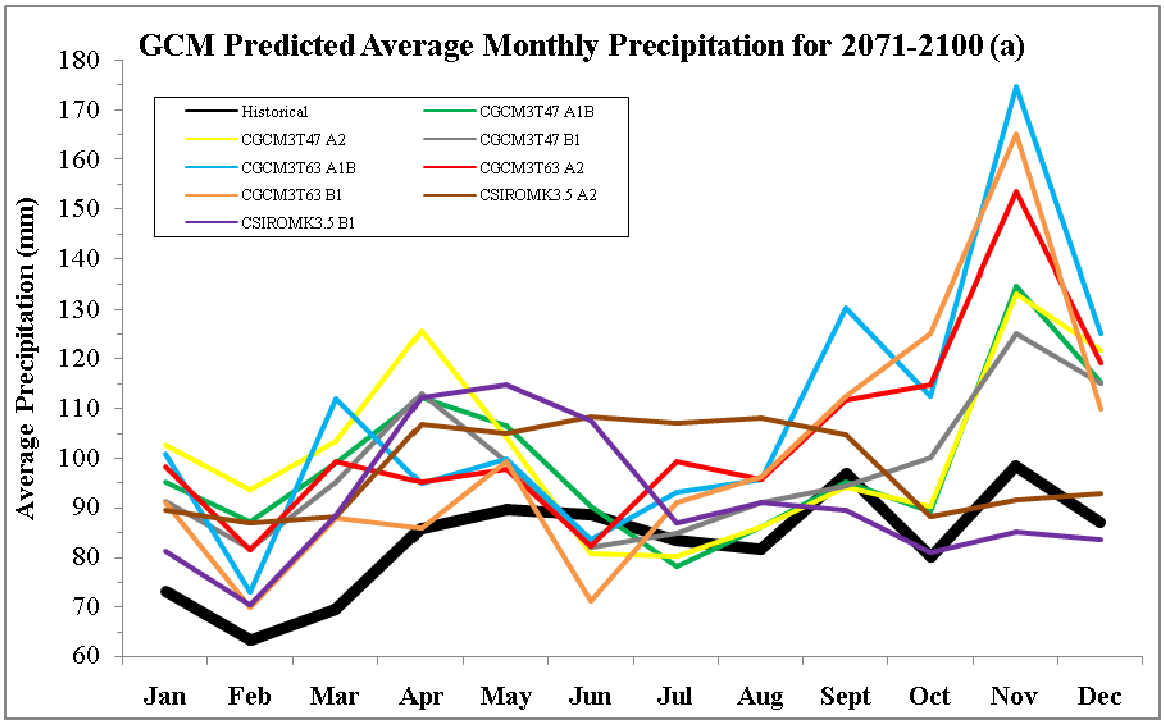
Figure 7: AOGCM predicted percent changes in precipitation for 2071-2100

5.2.2 Total Monthly Precipitation

Average values for AOGCM predicted total monthly precipitation are plotted with the observed data (in black) as line plots for both time periods. The 2080s plots are shown in Figure 8 (a, b), and the 2020s plots can be found in Figure C1, Appendix C. From the plots, it is clear that all AOGCM's but CSIROMK3.5 B1 predict an increase in total monthly precipitation for the winter months of December, January and February in the 2080s. For June to September, the predictions vary in both time periods, with some models predicting an increase and others predicting a decrease as compared to the observed values. All but two models predict an increase in November precipitation, with CGCM3T63 A1B predicting 175 mm for the 2080s, almost doubling the observed 97 mm. Predictions for September vary, with MIROC3.2MEDRES A2 and B1 averaging 57 mm compared to the observed 96 mm. Very similar results are seen in the 2020s time period, but the changes become more pronounced in the later time slice. Regardless of the time period, the variability between AOGCM projected monthly precipitation values is great, and thus several AOGCM predictions should be included in any impact assessment.

Box plots of total monthly precipitation as predicted by the AOGCM's for the 2080s are presented in Figures 9 through 14. In Appendix C, Figures C2 to C7 contain the box plots for the 2020s.

Winter and early spring precipitation is projected to increase by all emissions scenarios for CGCM3T47, with the highest increase seen in the 2080s for scenario A1B. For CGCM3T47 B1, the increase is slight in comparison with A1B and A2 for the 2080s time period. There is



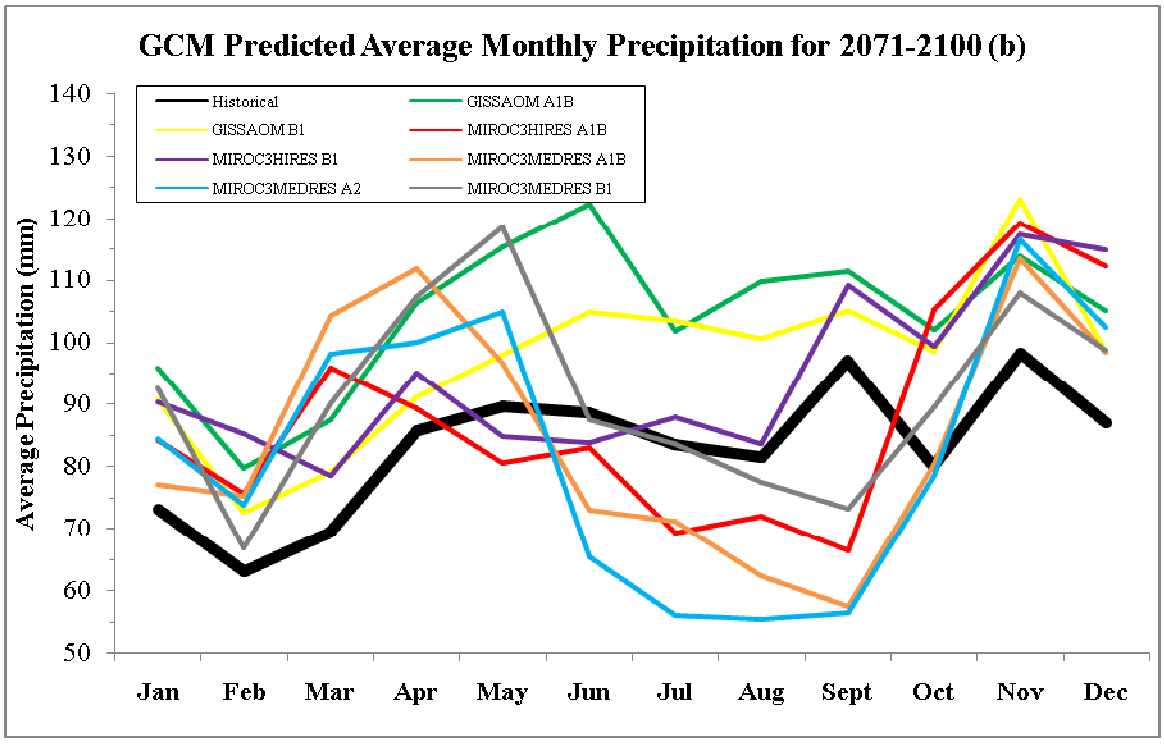


Figure 8(a, b): AOGCM predicted averages for monthly precipitation in 2071-2100, plotted against the historical observed data

a very small but less pronounced increase predicted for winter by all CGCM3T47 scenarios for the 2020s time period. For scenario A2 in the 2080s, January to March saw increases of 30-37% in total rainfall medians, and there are a large number of outliers indicating variability. In A1B, the largest increase in rainfall totals is seen in March (almost 46% for the 2080s), although December to February also predicted higher rainfall amounts, with increases of 21-36% from historical values. Results are similar for CGCM3T63 for both 2020s and 2080s time periods. Again the highest increase is seen in CGCM3T63 A1B for 2080s, projecting a March increase in median total precipitation of 43% greater than the historical mean. For CGCM3T63 B1, the increase in precipitation values for December to March is slight (8-18%), and for A2 total precipitation for these months increases by 22-36% for the 2080s. In contrast, the 2020s predictions show slight decreases in precipitation for January of CGCM3T63 A1B, B1 and in December for all three emission scenarios. CSIROMK3.5 shows different results for winter precipitation, with a decrease in December for both A2 and B1 in the 2080s. Scenario B1 projects 30% more precipitation for March in the 2080s and 15% less precipitation for December. For CSIROMK3.5 A2, up to a 23% increase in precipitation is predicted for January to February and for B1 a very small increase is seen. For CSIROMK3.5 in the 2020s, very little change from historical values is seen. In GISSAOM, winter precipitation is predicted to increase in both A1B and B1, with the most significant increase in A1B (14-28%). For GISSAOM's 2020s predictions, some months show very small increases while others showed a slight decrease. Both A1B and B1 for MIROC3.2HIRES project an increase in December precipitation by about 35% for the 2080s, and both scenarios predict slightly lower increases for January to March. In the 2020s predictions, very little change is seen however there are slight increases in February and March. For MIROC3.2MEDRES, all scenarios predict a significant increase in

March precipitation (21-45%) but B1 predicts a decrease February. Very little difference was seen for A1B and A2 in December to February precipitation, although there are some slight increases. The 2020s projections are close to historical values with some small increases and decreases.

Late spring and early summer predictions for the models vary, with some models projecting increases and others projecting decreases. In CGCM3T47, all scenarios predict an increase in precipitation for April and May of 6-29% in the 2080s and slight increases for the 2020s. For all three scenarios, there are decreases in total precipitation for June and July (6-18%), and little change in August values for the 2080s. There are very small decreases for these months in the 2020s time period. For CGCM3T63 very little changes are seen in April and May precipitation for the 2080s, and decreases in May precipitation are predicted by the 2020s time period. All scenarios for CGCM3T63 predict a decrease in June precipitation by 11-26% for both time periods. For July and August, the 2080s predictions are close to historical values except for slight increases in precipitation for the A2 scenario. The 2020s projections show increases in August values for A1B and A2 emission scenarios. For CSIROMK3.5, increases of 4-25% are seen in all emission scenarios from April to August. Scenario B1 shows the highest increase, however a slight decrease in July values is noticed. For the 2020s predictions, increases are seen for April, June and August for A2 and only in April for B1 with a decrease in May to July. Precipitation increases of 8-28% for April to August are predicted by GISSAOM for all emission scenarios in the 2080s. There is little change for scenario B1 in April and May of the 2080s. The 2020s show little change for most months, but a slight decrease in May and increase in June precipitation for B1. For MIROC3.2HIRES in the 2080s, A1B predicts no change for April but decreases in total precipitation for May to August of 11-33%. For B1 in the 2080s, the values are

close to the historical ones with decreases in precipitation for May and June. For the 2020s time period, MIROC3.2HIRES predicts a slight decrease as well for most months, with a 33% decrease in July precipitation for A1B. For all emission scenarios in the 2080s, MIROC3.2MEDRES projects a slight increase in April and May and decreases for June to August. The highest decreases in precipitation for the 2080s are seen in A2 (37-39% less than the historical average). Scenario B1 projects precipitation decreases of 6-14% for June to August, and A1B projects a 19-27% decrease for the same months. A similar trend with much more slight decreases is projected by MIROC3.2MEDRES for the 2020s time period.

For the fall months of September to November, results between models are variable. For CGCM3T47 A1B, a 25% increase is seen in November precipitation. For all emission scenarios of CGCM3T63, a 40 to 60% increase is seen for the same month as well as significant increases in October. GISSAOM predicts slight increases in fall precipitation. In contrast, decreasing fall precipitation is predicted by CSIROMK3.5.For MIROC3.2HIRES A1B, a 38% decrease in September precipitation is predicted, and both emission scenarios predict a slight increase in precipitation for October and November. A significant decrease in September precipitation of 32-47% is projected for all emissions scenarios by MIROC3.2MEDRES for September.

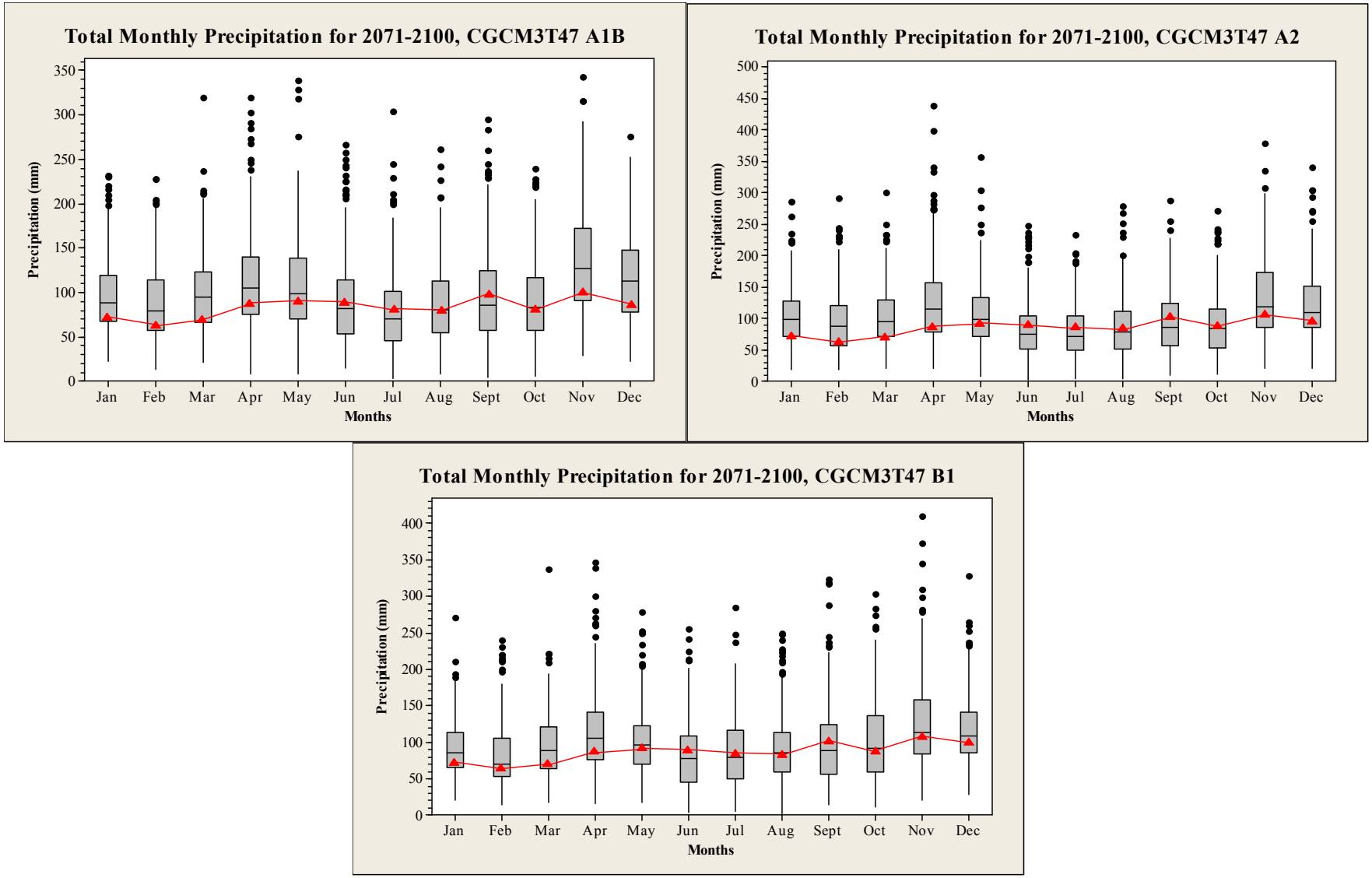


Figure 9: Total monthly precipitation box plots of CGCM3T47 A1B, A2 and B1 for the years 2071-2100 with observed historical averages marked in red.

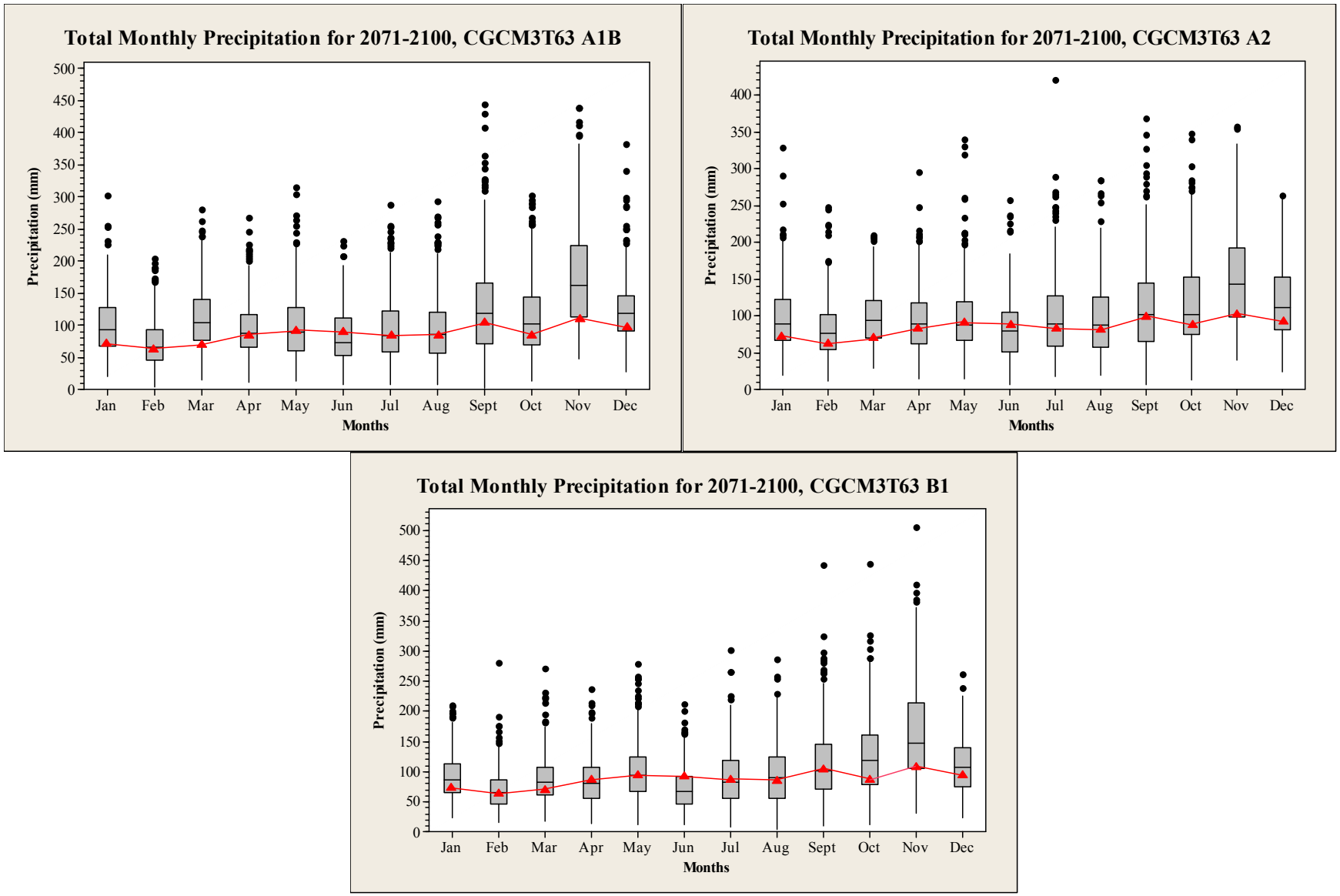


Figure 10: Total monthly precipitation box plots of CGCM3T47 A1B, A2 and B1 for the years 2071-2100 with observed historical averages marked in red.

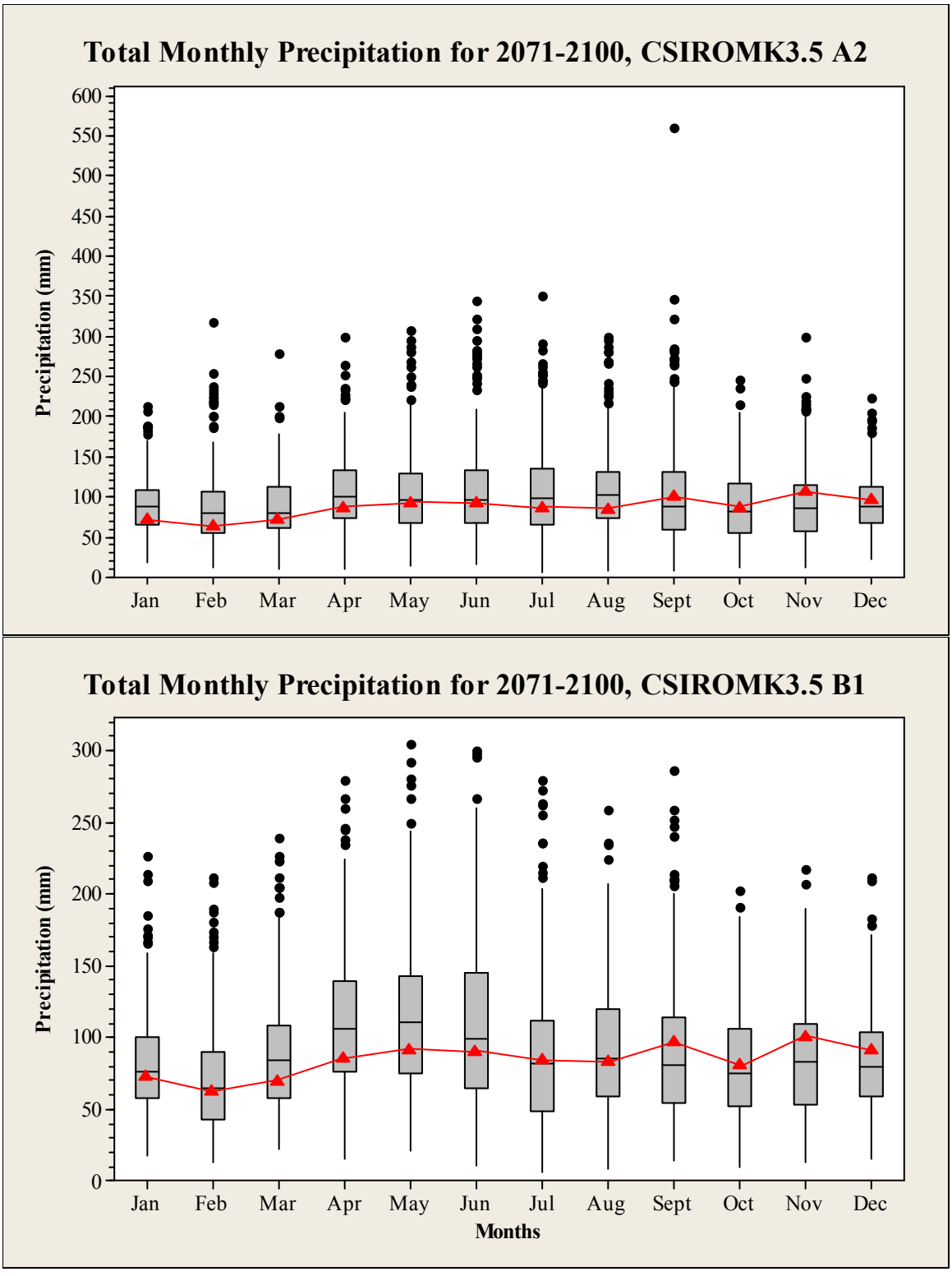


Figure 11: Total monthly precipitation box plots of CSIROMK3.5 A2 and B1 for the years 2071-2100 with observed historical averages marked in red.

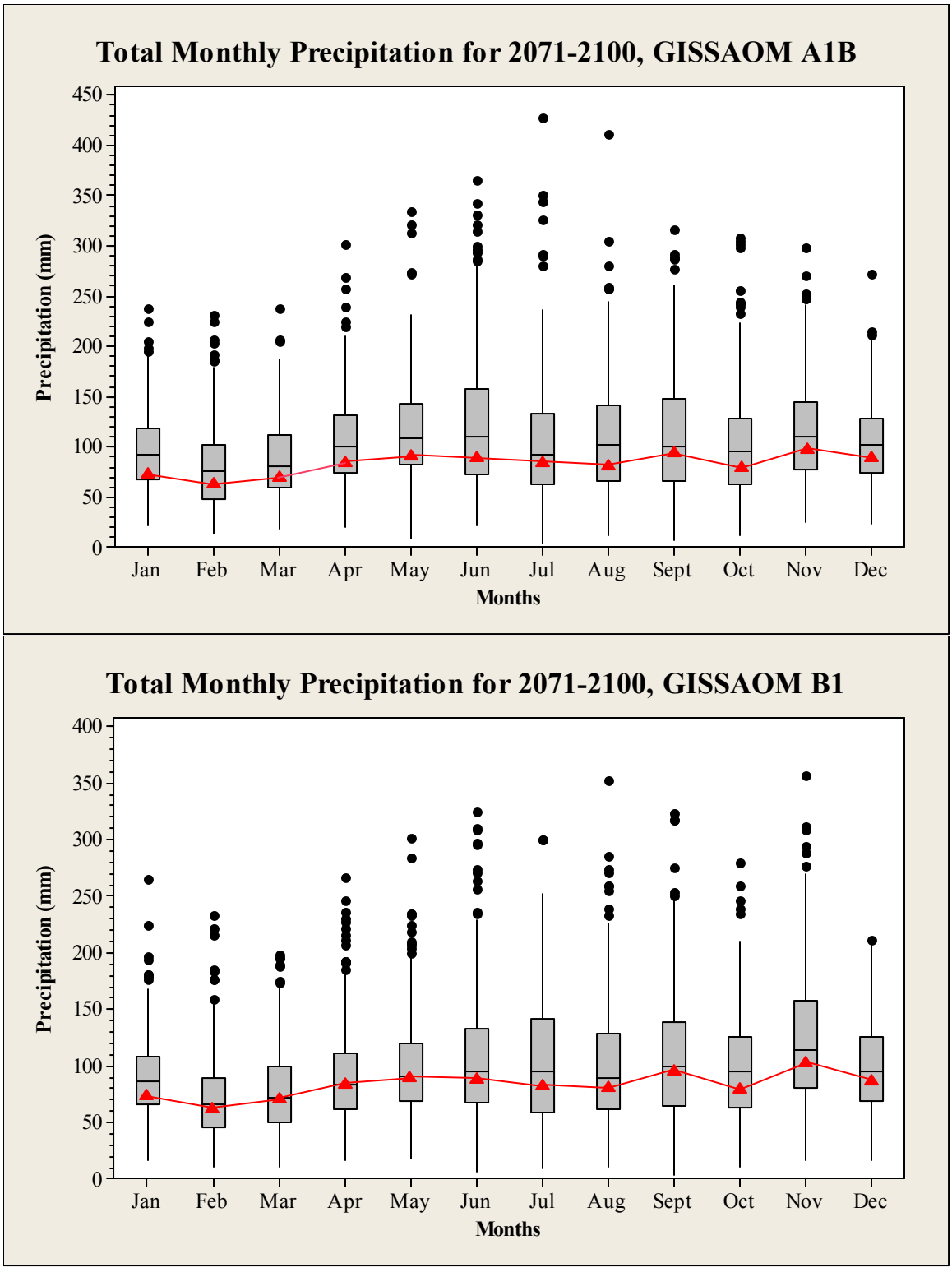


Figure 12: Total monthly precipitation box plots of GISSAOM A1B and B1 for the years 2071-2100 with observed historical averages marked in red.

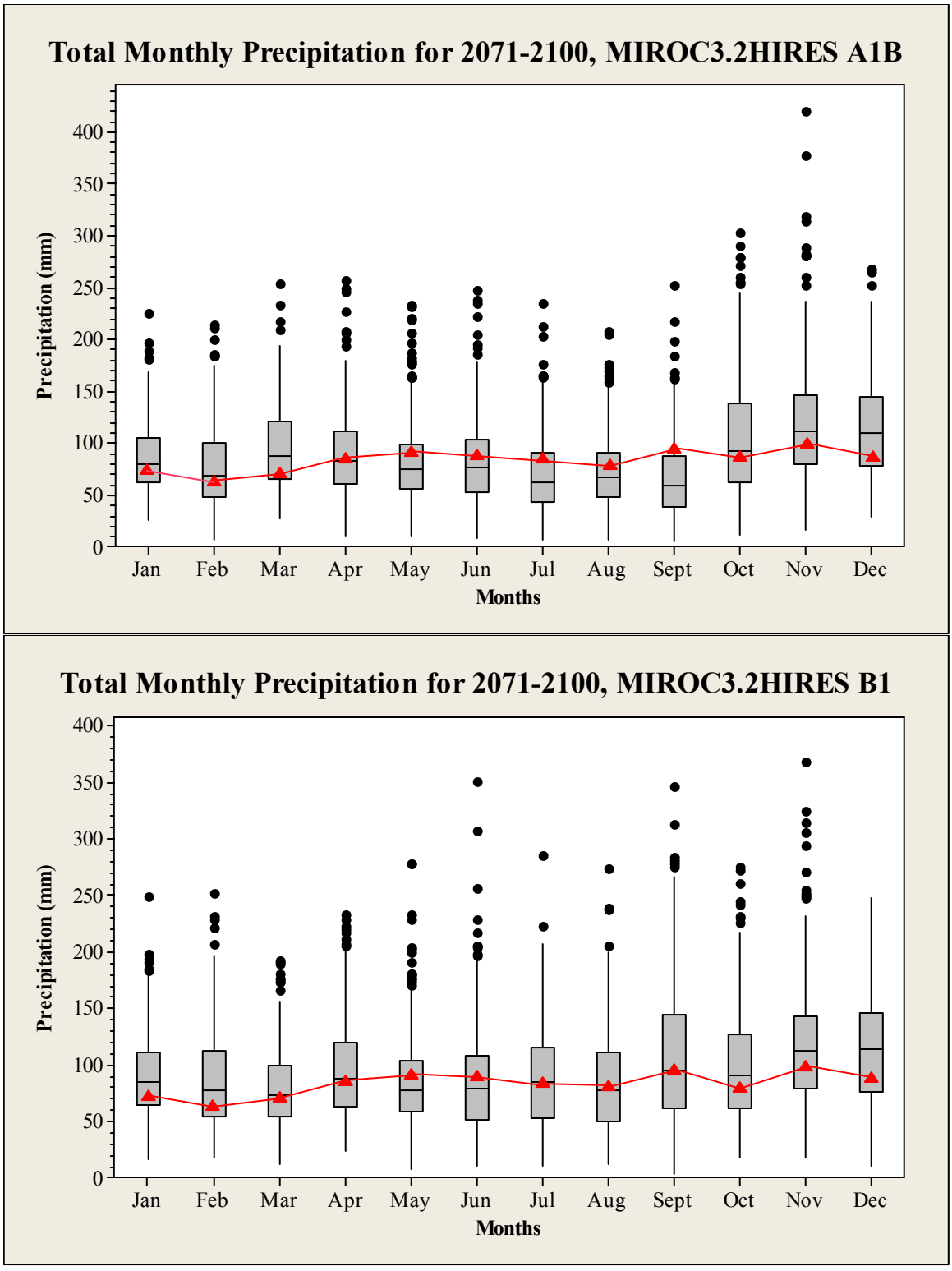


Figure 13: Total monthly precipitation box plots of MIROC3.2HIRES A1B and B1 for the years 2071-2100 with observed historical averages marked in red.

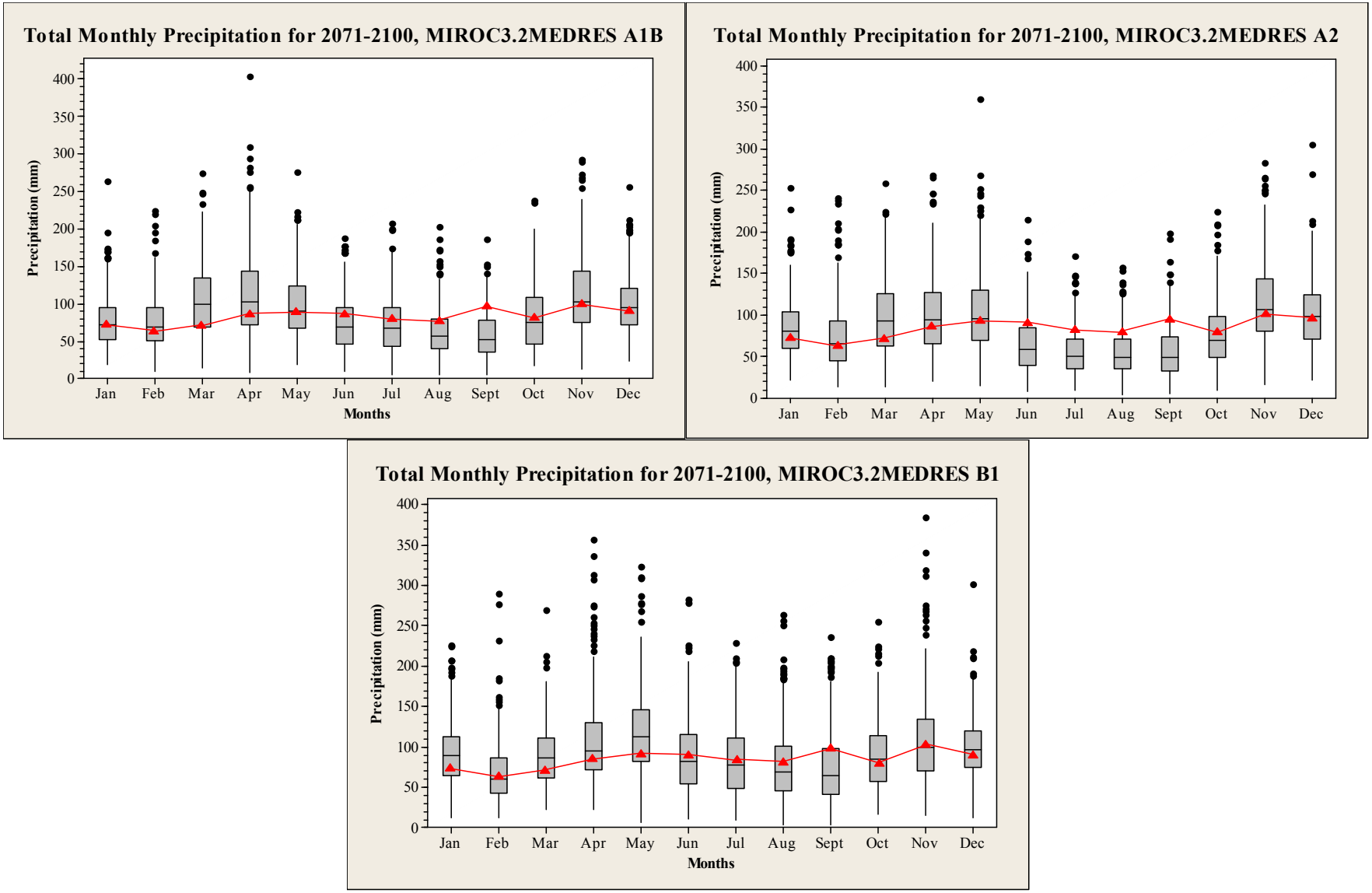


Figure 14: Total monthly precipitation box plots of MIROC3.2MEDRES A1B, A2 and B1 for the years 2071-2100 with observed historical averages marked in red.

5.2.3 Wet Spell Intensities

In order to investigate the intensity of wet spells, bar charts are made showing the percent change in wet spell intensity from the historical values to the future values. Intensities are calculated using the total amount of rain that fell during the spell over the length of the spell. The percent changes in wet spell intensities are determined for 3, 5 and 7 day wet spells. The plots are made for summer (June, July, August) and winter (December, January February) in both time periods. Figures 15 and 16 show the bar charts for summer and winter 2080s respectively, and the bar charts for the 2020s can be found in Figures C8 and C9 of Appendix C.

For summer wet spells in the 2080s, most models project an increase in 3-day intensities, except for MIROC3.2MEDRES A1B, A2 and MIROC3.2HIRES A1B. These models predict decreases of 4-17%. The most significant increase in intensity is predicted by CSIROMK3.5 A2 to be 47%. For the 2020s, all models predict an increase in 3-day intensities, ranging from 4-39%. For 5-day wet spells in the 2020s and 2080s, all models predicted an increase. Once again, CSIROMK3.5 A2 predicts the highest intensity increase of 107% for the 2080s. The smallest increase for 2080s is predicted by MIROC3.2MEDRES A2 as 16%. Most models project increases of between 35 and 79% for the 2080s. In the 2020s, increases in intensity range from 33-75%. For 7-day spells in the 2080s, all models predict a decrease in intensity except for GISSAOM A2, B1, and CGCM3T47 B1. Increases of 5-7% are predicted by these models. CSIROMK3.5 A2 projects an insignificant increase. The remaining models predict a decrease in intensity of 1 to 50% for the 2080s, with the highest being generated by MIROC3.2MEDRES A2. For the 2020s, all models project a decrease in 7-day wet spell intensity except CGCM3T47 A1B, A2 and GISSAOM B1. Most models for the 2020s project decreasing wet spell intensities of between 7 and 40%. The highest decrease for the 2020s 7-day spell intensities is predicted

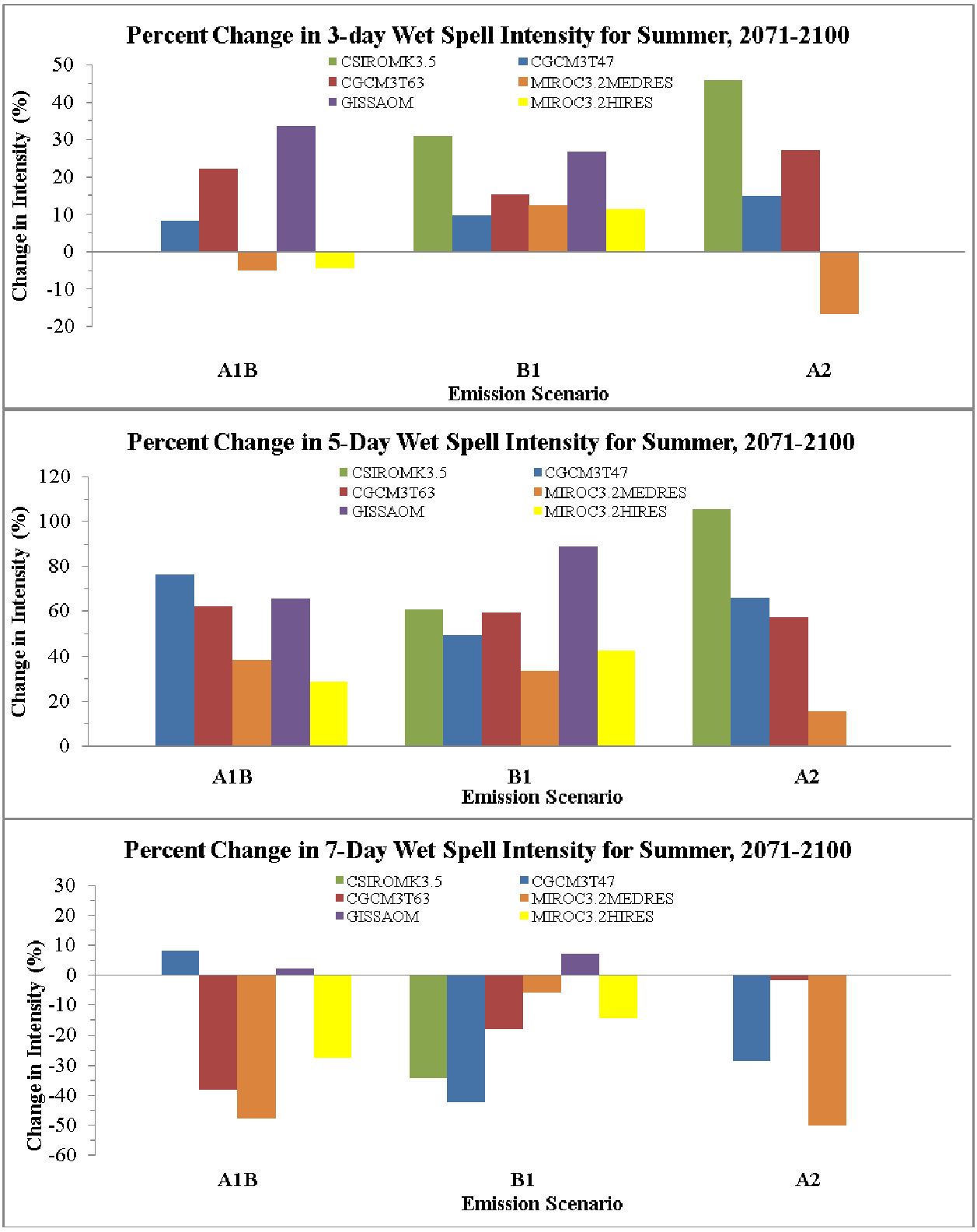


Figure 15: AOGCM predicted percent changes in wet-spell intensity for 3, 5, and 7 day spells in summer 2071-2100

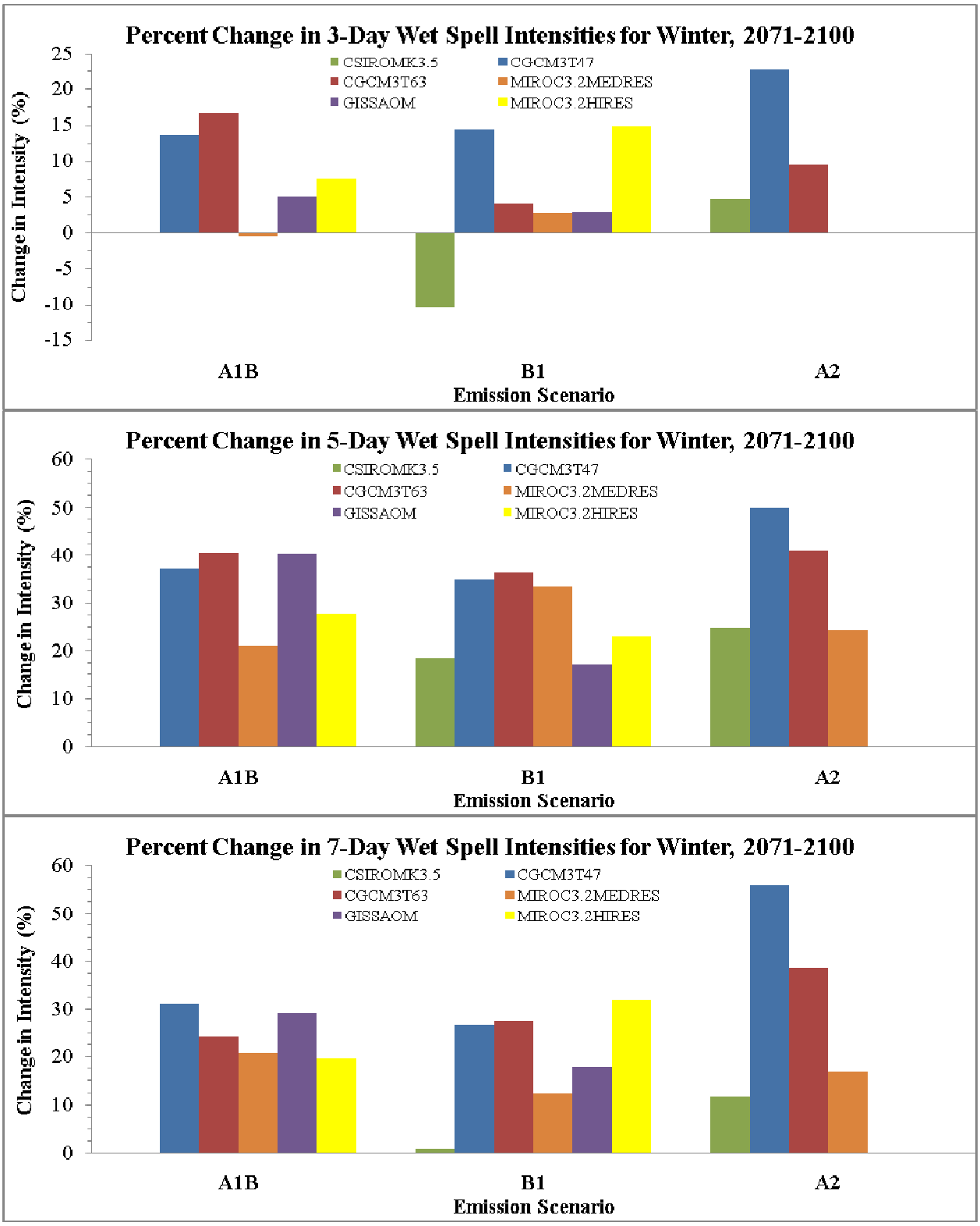


Figure 16: AOGCM predicted percent changes in wet-spell intensity for 3, 5, and 7 day spells in winter 2071-2100.

by CGCM3T47 A1B to be 49%. The general trend for summer in the future as predicted by several AOGCM's for both time periods is that shorter wet-spell intensities will increase as longer wet-spell intensities decrease.

For winter 3-day wet spells in the 2080s, all but two models (MIROC3.2MEDRES A1B and CSIROMK3.5 B1) project increasing wet-spell intensities. While CSIROMK3.5 B1 projects a 10% decrease, most other models project an increasing intensity from 2.5 to 23%. About half of the models predict a decrease in 3-day wet spell intensities for the 2020s. All models project increasing 5-day wet spell intensities for winter and both time periods. The highest increase is projected by CGCM3T47 A2 to be 50% for the 2080s. GISSAOM projects the smallest increase in 5-day wet spell intensities to be 17%. The majority of models predict that in the 2080s, intensities would increase from 23to 41%. For the 2020's period, intensity increases range from 6 to 43%. All models also predict increasing 7-day wet spell intensities for the 2080s, and all but one model, CSIROMK3.5 A2, for the 2020s. The smallest predicted increase in the 2080s is 2%, from CSIROMK3.5 B1. CGCM3T47 A2 once again predicts the highest increase of 55% for the 2080s, with most models projecting around 13-32%. For the 2020s, the models predict increases of between 5 and 35%. Overall, longer wet-spell intensities are projected to increase for winter months in the future.

5.2.4 Extreme Precipitation Events

It is important to assess the changes in extreme precipitation events because a moderate change in the precipitation can have large impact on runoff, hence the occurrence of floods. In this study three precipitation indices are tested to demonstrate the frequency, intensity and extremes of precipitation events. The highest 5 day precipitation, very wet days and the heavy precipitation days express these extreme features of precipitation (Table 5). For very wet days,

the 95th percentile reference value has been obtained from all non-zero total precipitation events for 1961-1990. It is better to use indices on percentile values rather than a fixed threshold in Canada due to large variation of precipitation intensities in various regions.

Table 5: List of extreme precipitation indices

Precipitation indices	**Definitions**	**Units**
Heavy precipitation days (>= 10 days)	no of days with precipitation>= 10 mm	days/yr
Very wet days (>= 95th percentile)	no of days with precipitation>=95th percentile of 1961-1990 observed precipitation	days/yr
Highest 5-day precipitation amount	max precipitation sum for 5 day interval	mm/5day/ yr

Table 6: Changes in extreme precipitation events for 2020s and 2080s

Models/Scenarios	**Heavy Precipitation Days**		**Very Wet Days**		**Max 5 day Precipitation**	
	2020s	2080s	2020s	2080s	2020s	2080s
CGCM3T47_A1B	5.34	12.95	6.77	14.53	3.25	7.64
CGCM3T47_A2	-0.93	0.92	-3.23	2.99	0.26	4.87
CGCM3T47_B1	0.97	-3.11	2.55	-4.87	-0.05	-7.30
CGCM3T63_A1B	5.68	8.11	11.51	17.41	17.50	20.72
CGCM3T63_A2	-2.58	-3.27	-5.24	-7.97	-5.49	-7.08
CGCM3T63_B1	-3.43	-4.00	-3.79	-3.80	-1.34	-1.53
CSIROMK3.5_A2	-5.10	-2.46	-10.73	-2.89	-9.68	-2.84
CSIROMK3.5_B1	-5.75	-6.60	-8.46	-11.81	-7.99	-9.52
GISSAOM_A1B	6.85	14.11	11.91	25.33	7.03	12.91
GISSAOM_B1	3.04	-5.82	1.87	-11.34	3.22	-2.81
MIROC3HIRES_A1B	-3.61	-9.40	-2.47	-15.42	-1.48	-15.13
MIROC3HIRES_B1	0.69	6.41	-1.88	10.26	2.32	12.86
MIROC3MEDRES_A1B	-3.25	-10.18	-4.91	-14.90	-10.71	-10.97
MIROC3MEDRES_A2	0.44	-3.23	3.23	-2.60	7.32	0.18
MIROC3MEDRES_B1	3.45	11.93	6.98	15.10	0.30	7.11

Figures 17 and C10 in Appendix C together with Table 6 show a comparison of probability plots of the number of days associated with greater than 95th percentile precipitation as predicted by AOGCMs. The parameter estimates have been displayed with Anderson-Darling (AD) goodness-of-fit statistic. The AD measures how well several distributions from several AOGCMs follow the historic observation. For comparison of several distributions with AD, the smallest AD statistic indicates the closest fit to the data. One common feature of most AOGCMs is that they are positively (rightward) skewed indicating more data points in the right tail in the upper half than expected. This clearly suggests increase in the number of very wet days. CGCM3T47 A2, CSIROMK3.5 A2, CGCM3T63 A1B scenarios predict higher occurrence of very wet days. However, scenarios, such as MIROC3.2MEDRES A1B, GISS-AOM B1 and CSIROMK3.5 A1B predict a decrease. So the differences between the AOGCMs are around 21% for 2020s with increasing uncertainty to 33% for 2080s. Same is the case with heavy precipitation days and maximum 5 day precipitation amounts. Higher AD values for MIROCHIRES B1, MIROCMEDRES A2 and MIROCMEDRES B1 indicate that they do not follow the same distribution as the observed values.

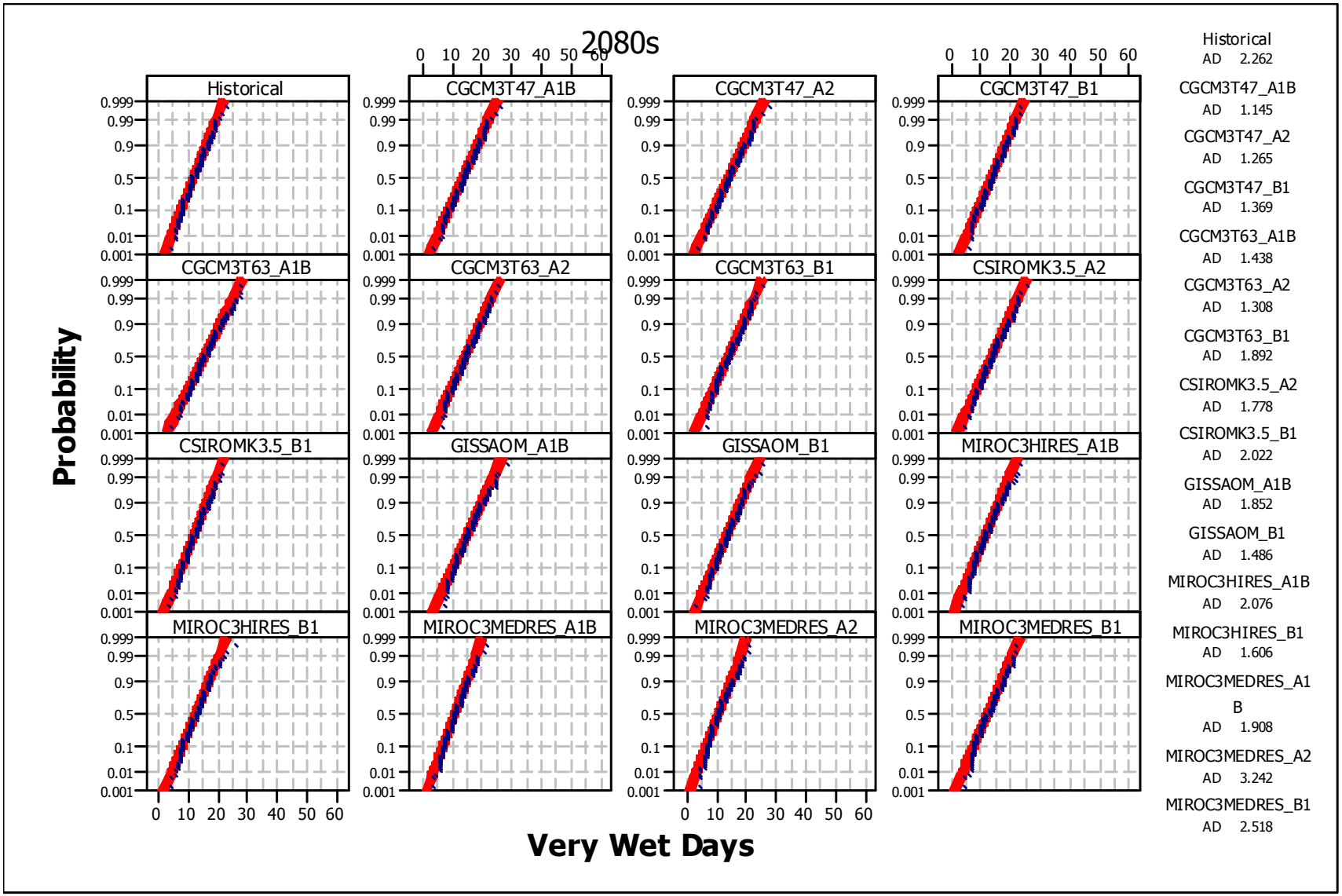


Figure 17: Probability plot of very wet days (>95th percentile of 1979-2005 precipitation) for 2080s with Anderson-Dorling's estimate

6. Conclusion

An evaluation of the WG-PCA weather generator in reproducing observed precipitation data is performed, along with an investigation of the potential impacts of climatic change on the Upper Thames River basin using six AOGCM's, each with up to three emission scenarios. Twenty-two stations around the basin are chosen, based on the length and completeness of their historical records. The weather generator is used to produce 324 years of synthetic data for the historical data as well as the AOGCM models. For the purpose of comparison, the results from the London station are analyzed.

In the reproduction of historical data, the WG-PCA performance is satisfactory. For some months the synthetic data significantly underestimates the historical mean precipitation, however the quartiles of the data extend far beyond the observed means and there are long whiskers with few outliers. Considering the higher spatial and temporal variability of precipitation, such performance is, however, considered good. Monthly wet days are also investigated using box plots. The historical data underestimates the mean values for winter, spring and fall seasons, however the observed means are well within the inter-quartile range of the synthetic data. The whiskers extend 4 or 5 days on either side of these and several outliers above those, therefore there are several extremely wet months simulated by the weather generator. The statistical hypothesis tests show insignificant differences between the means and variances of precipitations from observed and simulated datasets. The frequency distribution of wet spells is captured well by the weather generator, with very close agreements in observed and simulated values for winter and summer.

The AOGCM data is found to be highly variable. When some models predict a decrease in precipitation, others would predict an increase. However, some distinct trends are noticed during analysis. In the future, longer wet-spells in winter are projected to increase in intensity. For summer wet-spells, shorter spells will increase in intensity as longer ones decrease. Most AOGCM's predict that average winter and early spring precipitation will increase significantly. The summer results are less conclusive, as some AOGCM's predict decreasing amounts while others predict increasing amounts. The probability plot of the extreme precipitation events clearly indicates an increase in the extent of uncertainties spanning from 2020s to 2080s. Because of the high variability between AOGCM's and emission scenarios, it is crucial to consider a variety of these in climate change impact assessments. As these are predictions of future weather patterns, none can be given preference and a thorough assessment would include several models. Future work should focus on quantification of uncertainties from several AOGCM models for hydrologic modelling and flood forecasting in the Upper Thames River basin.

Acknowledgement

The research presented in this study is funded by the Canadian Foundation for Climatic and Atmospheric Sciences.

References

Apipattanavis, S., Podesta, G., Rajagopalan, P., Katz, R. W. (2007). "A semiparametric multivariate and multisite weather generator".*Water Resources Research* 43, W11401.

Beersma, J. J., Buishand, T. A., Wojcik, R. (2001). "Rainfall generator for the Rhine basin: multi-site simulation of daily weather variables by nearest-neighbour resampling".In: Generation of Hydrometeorologicalreferenceconditionsfor the assessment of flood hazard in largeriverbasins, P. Krahe and D. Herpertz (Eds.), CHR-Report No. I-20, Lelystad, p. 69-77.

Brandsma, T., Buishand, T. A., (1998). "Simulation of extreme precipitation in the Rhine basin by nearest-neighbour resampling".*Hydrology and Earth System Sciences* 2(2-3): 195-209.

Brissette, F., Khalili, M., Leconte, R. (2007). "Efficient stochastic generation of multi-site synthetic precipitation data". *Journal of Hydrology* 345: 121-133.

Brissette, F., Leconte, R., Minville, M., Roy, R. (2006). "Can we adequately quantify the increase/decrease of flooding due to climate change?". EIC Climate Change Technology Conference 2006, Ottawa, May 10-12. 1-6.

Canadian Climate Change Scenarios Network (2010). "SRES scenarios background". *Environment Canada.* Retrieved August 5, 2010 from http://ontario.cccsn.ca/?page=scen-sres.

Diaz-Neito, J., Wilby, R. L. (2005). “A comparison of statistical downscaling and climate change factor methods: Impacts on low flows in the River Thames, United Kingdom”. *Climatic Change* 69: 245-268.

Dibike, Y. B., Coulibaly, P. (2005).“Hydrologic impact of climate change in the Saguenay watershed: comparison of downscaling methods and hydrologic models”. *Journal of Hydrology* 307: 144-163.

Eum, H-I., Arunachalam V., Simonovic, S.P. (2009). “Integrated Reservoir Management System for Adaptation to Climate Change Impacts in the Upper Thames River Basin”. *Water Resources Research Report 62*, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada.

Flato, G. M. (2005)."The Third Generation Coupled Global Climate Model (2005)". *Canadian Centre for Climate Modelling and Analysis*. Environment Canada. Retrieved 23 Aug 2009 http://www.cccma.bc.ec.gc.ca/models/cAOGCM3.shtml.

Hanson, C. L., Johnson, G. L. (1998). “GEM (Generation of weather Elements for Multiple applications): its application in areas of complex terrain”. *Hydrology, Water Resources and Ecology in Headwaters IAHS* 248: 27-32.

IPCC, (2007). “Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change”. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp.

Kay, A. L., Davies, H. N. (2008). “Calculating potential evaporation from climate model data: A source of uncertainty for hydrological climate change impacts”. *Journal of Hydrology* 358: 221-239.

King, L., T.Solaiman and S. P. Simonovic (2009). “Assessment of Climatic Vulnerability in the Upper Thames River Basin”.*Water Resources Research Report no. 064*, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 62 pages. ISBN: (print) 978-0-7714-2816-6; (online) 978-0-7714-2817-3.

Kuchar, L. (2004). “Using WGENK to generate synthetic daily weather data for modelling of agricultural processes”. *Mathematics and Computers in Simulation* 65: 69–75.

Maurer, E. P. (2007). “Uncertainty in hydrologic impacts of climate changein the Sierra Nevada, California, under two emissions scenarios”. *Climatic Change* 82: 309-325.

Mehrotra, R., Srikanthan, R., Sharma, A. (2006). “A comparison of three stochastic multi-site precipitation occurrence generators”. *Journal of Hydrology* 331: 280-292.

Nakicenovic, N., Alcamo, J., Davis, G., de Vries, B., Fenhann, J., and co-authors. (2000). "IPCC Special Report on Emissions Scenarios". *UNEP/GRID-Ardenal Publications*.

Nigam, S., and A. Ruiz-Barradas.(2006). Seasonal hydroclimate variability over North America in Global and Regional Reanalyses and AMIP Simulations: Varied representation”. *Journal of Climate* 19:815-837.

PCMDI (2005). "Model Information of Potential Use to the IPCC Lead Authors and the AR4: GISS-AOM".*CMIP3 Climate Model Documentation, References, and Links*. Retrieved 23 Aug 2009 from http://www-pcmdi.llnl.gov/ipcc/model_documentation/GISS-AOM.htm.

Prodanovic, P., Simonovic, S. P. (2007). “Development of rainfall intensity duration frequency curves for the City of London under the changing climate”. *Water Resources Research Report* no. 58, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada.

Richardson, C.W. (1981). “Stochastic simulation of daily precipitation, temperature, and solar radiation”. *Water Resources Research* 17: 182–90.

Schoof, J. T., Arguez, A., Brolley, J., O’Brien, J. J. (2005).“A new weather generator based on spectral properties of surface air temperatures”. *Agricultural and Forest Meteorology* 135: 241–251.

Sharif, M., Burn, D. H., (2007). “Improved K-Nearest Neighbour weather generating model”. *Journal of Hydrologic Engineering* 12: 42-51.

Semenov, M. A., Barrow, E. M. (1997). “Use of a stochastic weather generator in the development of climate change scenarios”. *Climatic Change* 35: 397-414.

Solaiman, T.A., Simonovic, S. P., (2010). “National Centers for Environmental Predicton - National Center for Atmospheric Research (NCEP-NCAR) Reanalyses Data for Hydrologic Modelling on a Basin Scale”. *Canadian Journal of Civil Engineering* 37(4): 611-623.

Soltani, A., Hoogenboom, G. (2003). “A statistical comparison of the stochastic weather generators WGEN and SIMMETEO”. *Climate Research* 24: 215-230.

Trigo R. M., Palutikof, J. P. (2001).“Precipitation Scenarios over Iberia: A Comparison between Direct AOGCM Output and Different Downscaling Techniques”. *Journal of Climate* 14: 4422-4442.

Wilks, D.S.(1998).“Multi-site generalization of a daily stochastic precipitation model”. *Journal of Hydrology*210: 178–191.

Zhang, X., Harvey, K., Hogg, W. and Yuzyk, T. (2001). “Trends in Canadian streamflow”. *Water Resources Research* 37: 987-998.

APPENDIX A: SRES Emission Scenarios

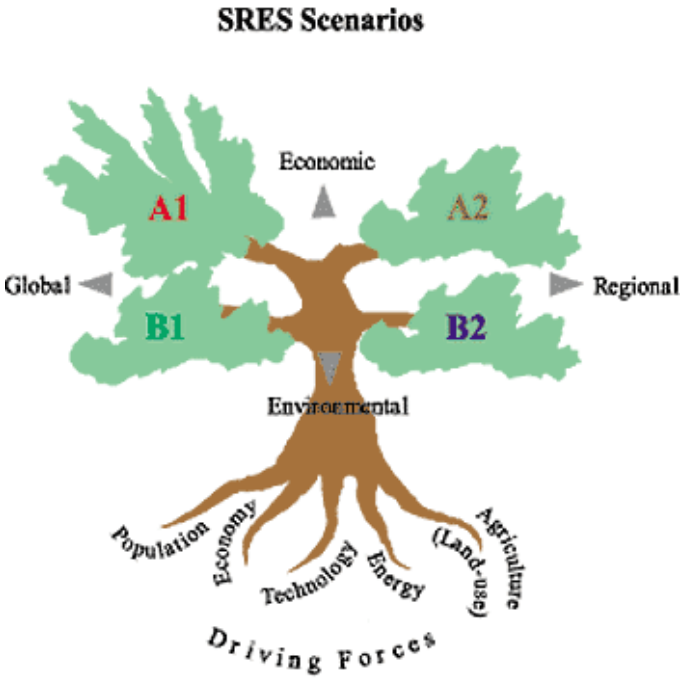


Figure A1: SRES Emission Scenarios (Nakicenovic et al, 2000)

A1B: In scenario A1B, the storyline includes rapid economic expansion and globalization, a population peaking at 9 billion in 2050, and a balanced emphasis on a wide range of energy sources (Nakicenovic et al, 2000).

B1: The storyline for the B1 scenario is much like A1B in terms of population and globalization; however there are changes toward a service and information economy with more resource efficient and clean technologies. Emphasis is put on finding global solutions for sustainability (Nakicenovic et al, 2000).

A2: For scenario A2, the storyline consists of a world of independently operating nations with a constantly increasing population and economic development on a regional level. Technological advances in this storyline occur more slowly due to the divisions between nations (Nakicenovic et al, 2000).

APPENDIX B: Atmosphere-Ocean General Circulation Model Data Description

Canadian Coupled Global Climate Model

The third generation Coupled Global Climate Model (CAOGCM3) was created in 2005 by the Canadian Centre for Climate Modelling and Analysis (CCCma) in Victoria, BC for use in the IPCC 4th assessment report to run complex mathematical equations which describe the earth's atmospheric and oceanic processes. The CAOGCM3 climate model includes four major components: an atmospheric general circulation model, an ocean general circulation model, a thermodynamic sea-ice model, and a land surface model (Hengeveld, 2000) and consists of two resolutions, T47 and T63. CAOGCM3T47 has a spatial resolution of 3.75° x 3.75° and it includes 31 vertical levels (Flato, 2005). The atmospheric resolution of CAOGCM3T63 model is 2.8° × 2.8°. The emissions scenarios A1B, A2 and B1 were used as greenhouse gas inputs in both models.

Commonwealth Scientific and Industrial Research Organization's Mk3.5 Climate Systems Model

Australia's Commonwealth Scientific and Industrial Research Organization created the AOAOGCM CSIROMK3.5, which is an improved version of the MK climate systems model. The spatial resolution of the model is 1.875 × 1.875. The SRES emissions scenarios A1B, A2, and B1 were used as inputs to the model for the IPCC 4th assessment report.

51 52

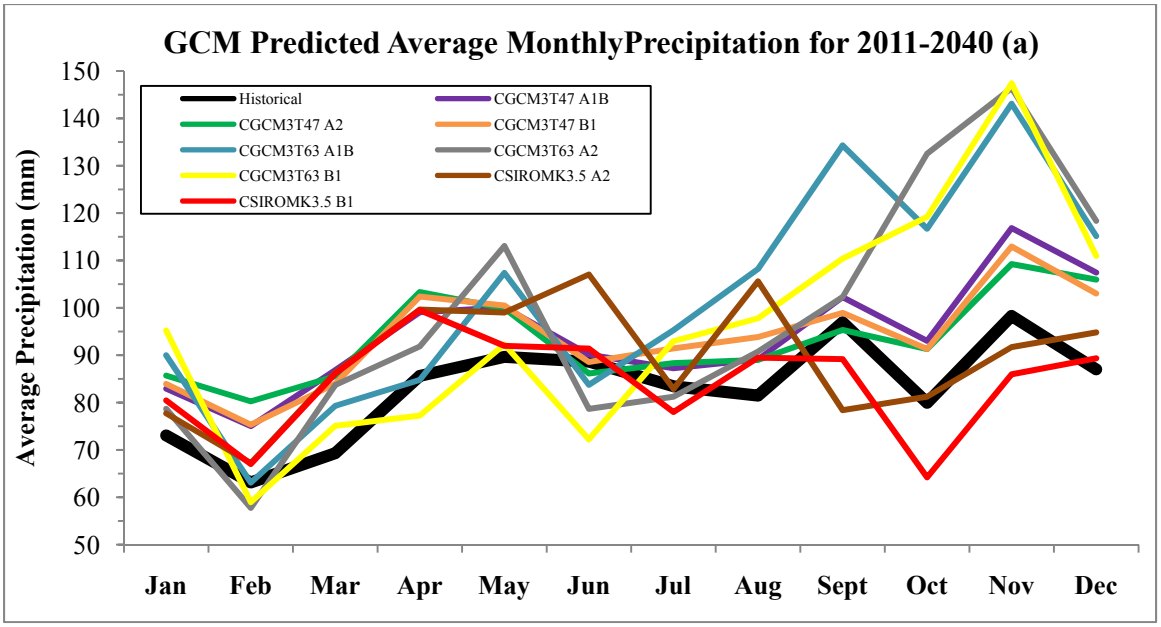
Goddard Institute for Space Studies' Atmospheric Ocean Model

The North American Space Association and the Goddard Institute for Space Studies developed the GISS-AOM climate model, first in 1995 and then a revised version was created with smaller grids in 2004 for the IPCC 4th assessment report. The resolution for the model is 4° longitude by 3° latitude (PCMDI, 2005). The atmospheric grid has 12 vertical layers (PCMDI, 2005). The emissions scenarios SRES A1B and B1 were used as greenhouse gas inputs to the model.

Model for Interdisciplinary Research on Climate version 3.2

The Japanese Model for Interdisciplinary Research on Climate version 3.2 (MIROC3.2) was developed in two resolutions: the high resolution (MIROC3.2HIRES) in 1.125° × 1.125° grid and the medium resolution (MIROC3.2MEDRES) in 2.8° × 2.8° grid. For present study, two emissions scenarios from MIROC3.2HIRES (A1B and B1) and three scenarios (A1B, A2 and B1) from MIROC3.2MEDRES were used.

APPENDIX C: AOGCM Figures for 2020s



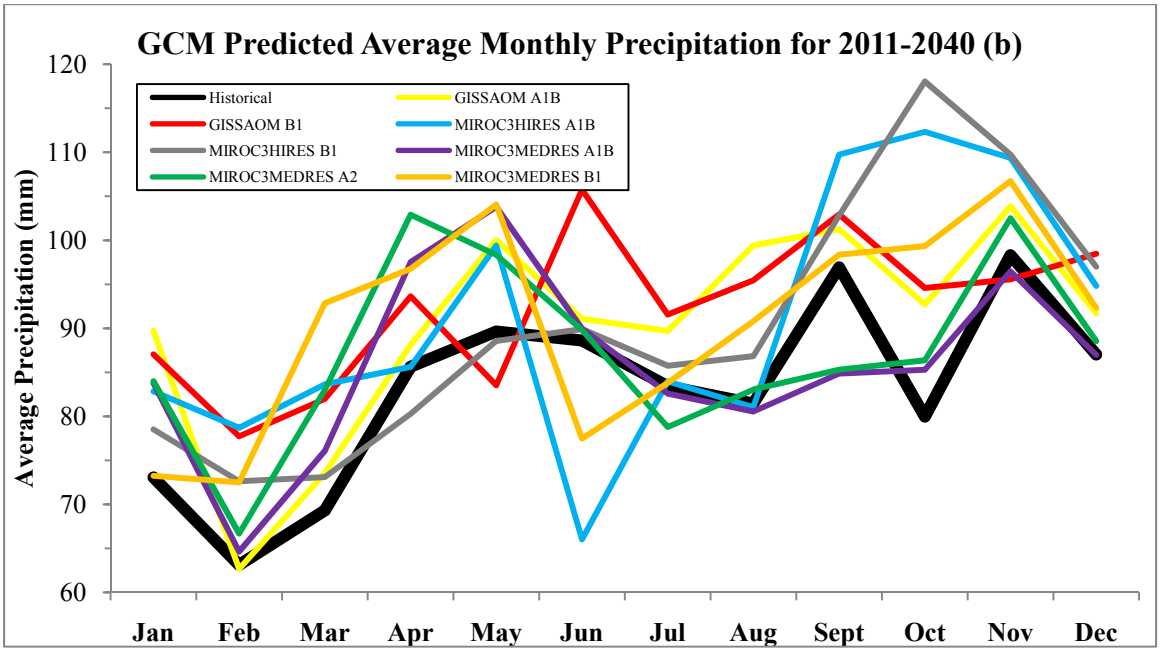


Figure C1: AOGCM predicted monthly average precipitation values for 2011-2040, plotted against the historical observed data.

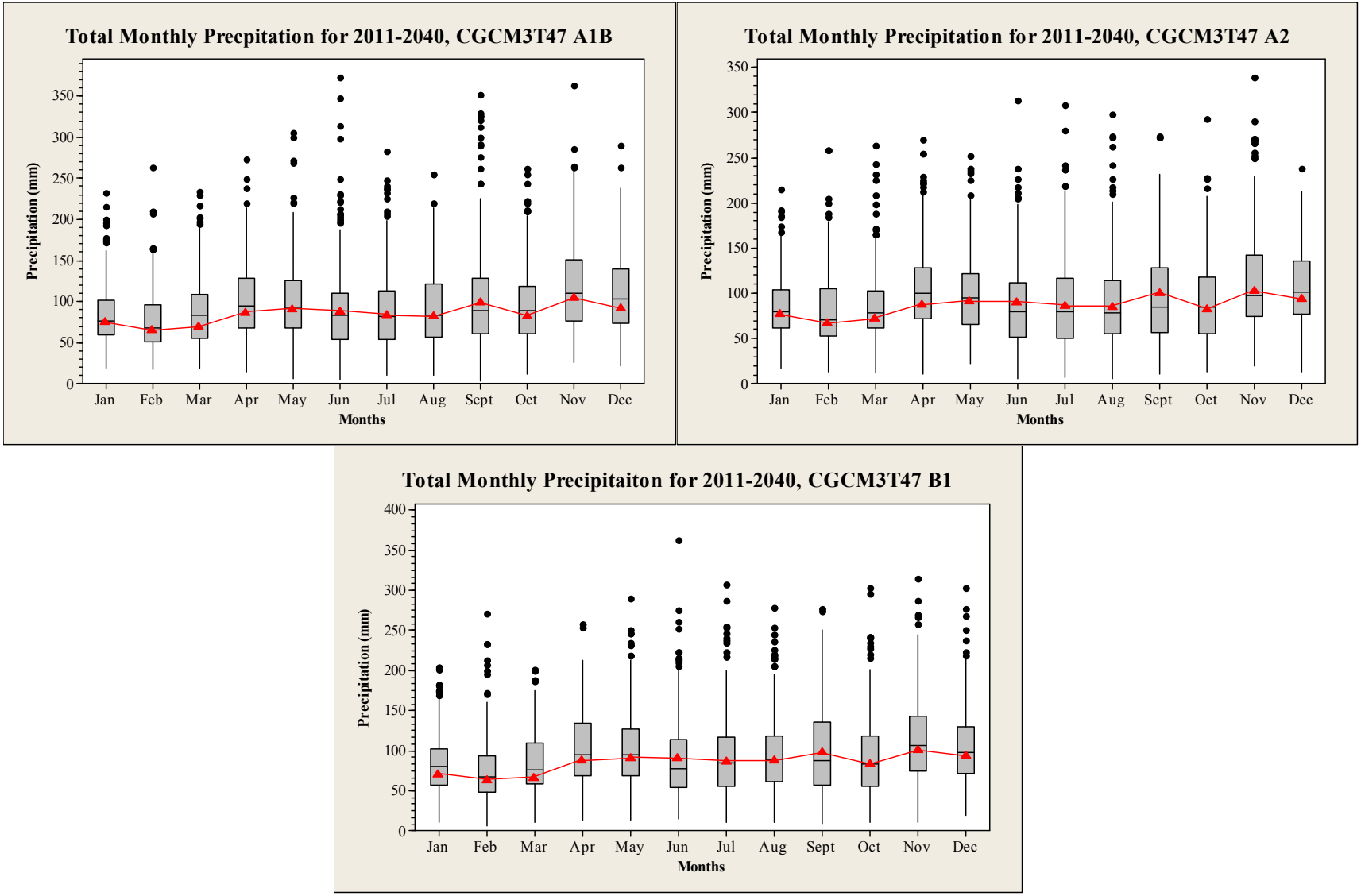


Figure C2: Total monthly precipitation box plots of CAOGCM3T47 A1B, A2 and B1 for the years 2011-2041 with observed historical averages marked in red.

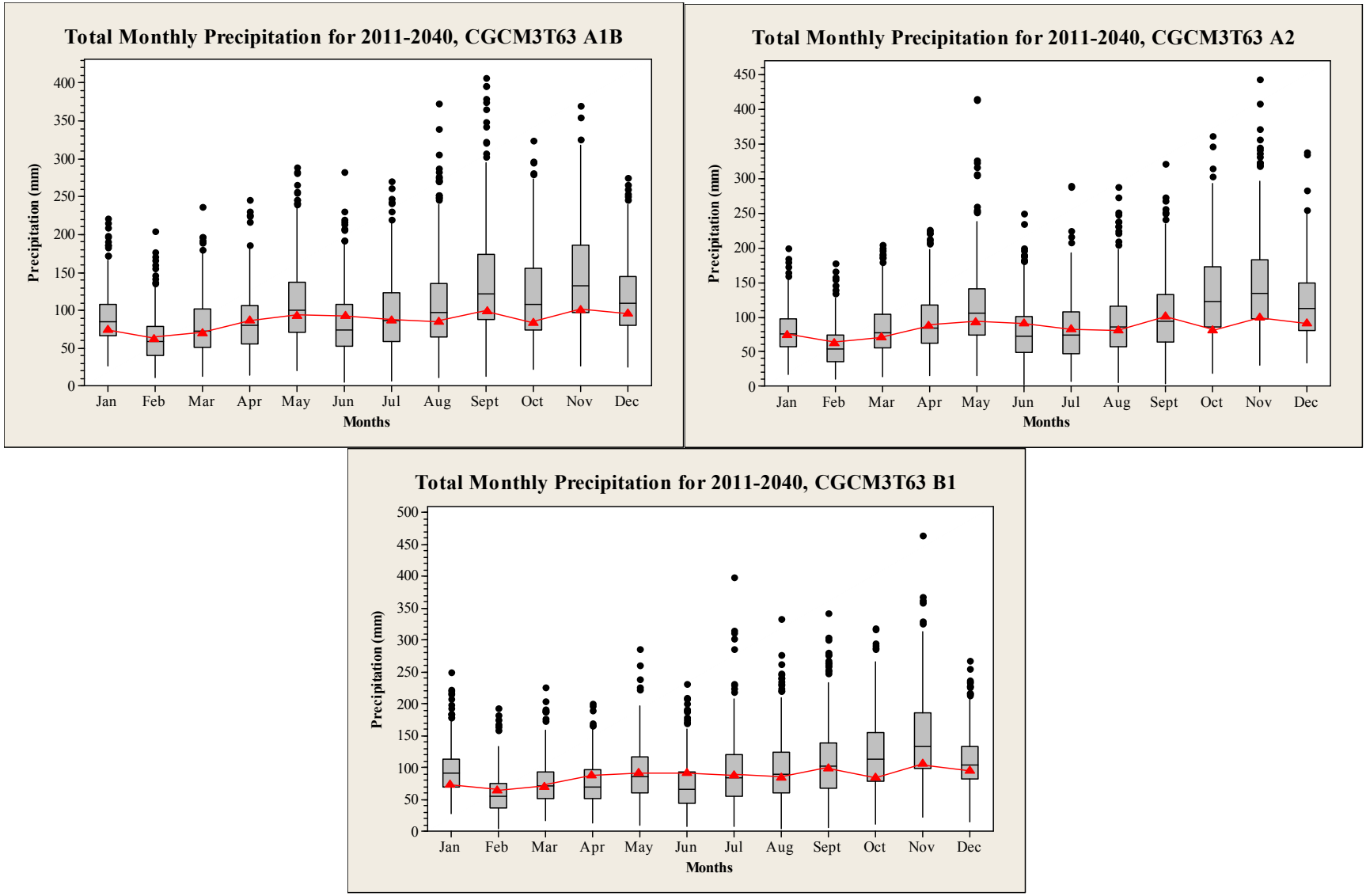


Figure C3: Total monthly precipitation box plots of CAOGCM3T63A1B, A2 and B1 for the years 2011-2041 with observed historical averages marked in red.

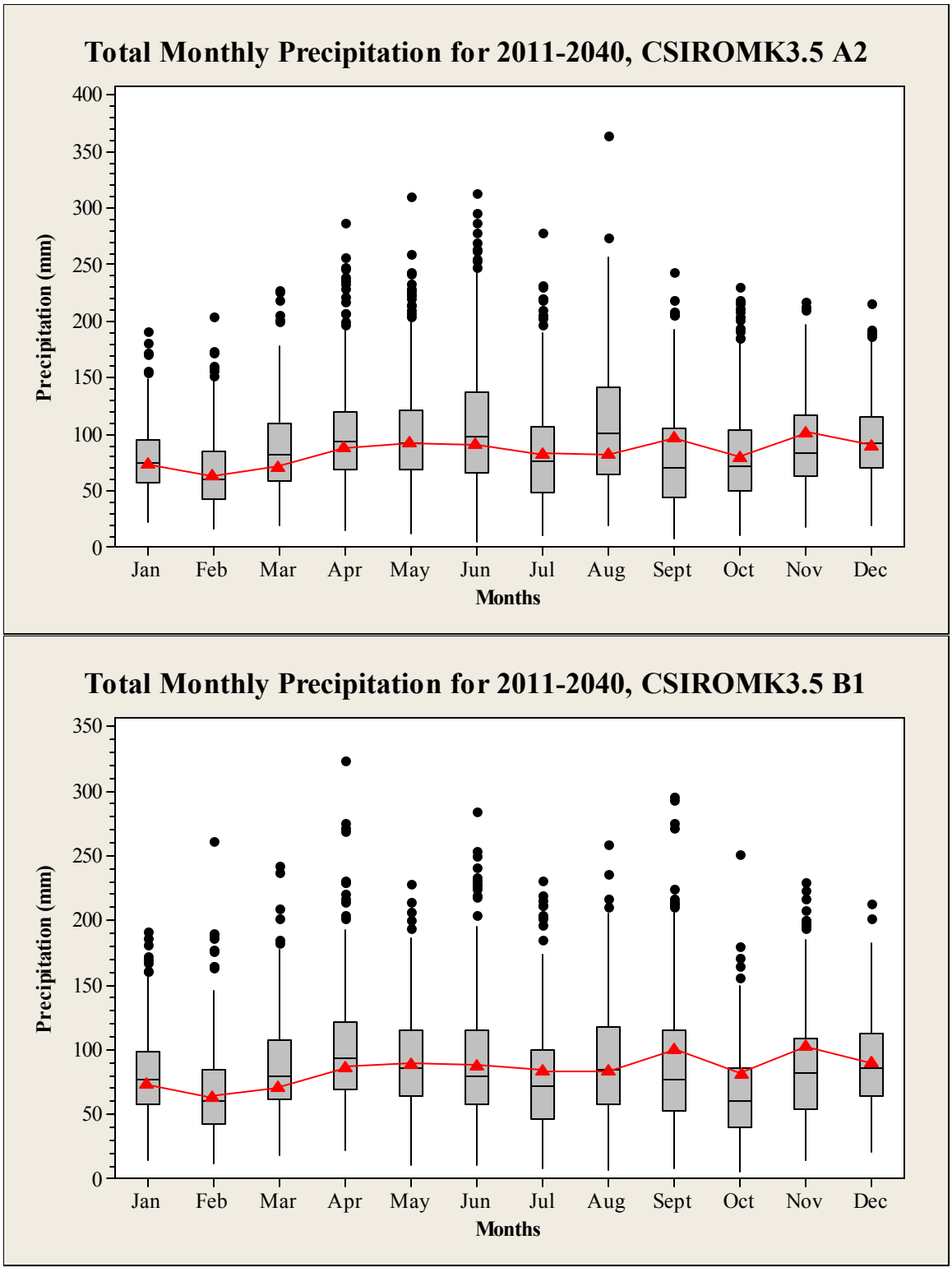


Figure C4: Total monthly precipitation box plots of CSIROMK3.5 A2 and B1 for the years 2011-2040 with observed historical averages marked in red.

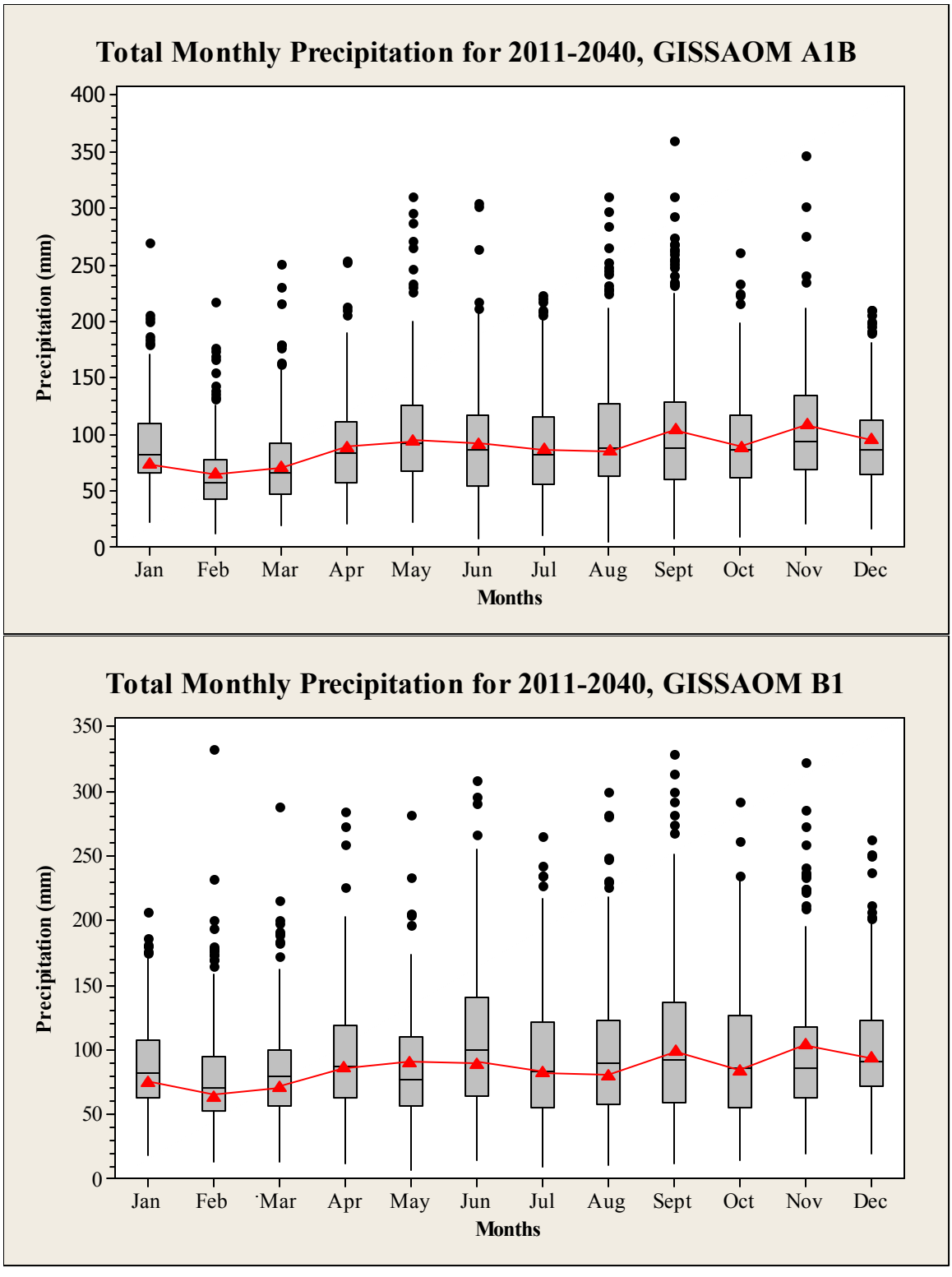


Figure C5: Total monthly precipitation box plots of GISSAOM A1B and B1 for 2011-2040 with observed historical averages marked in red.

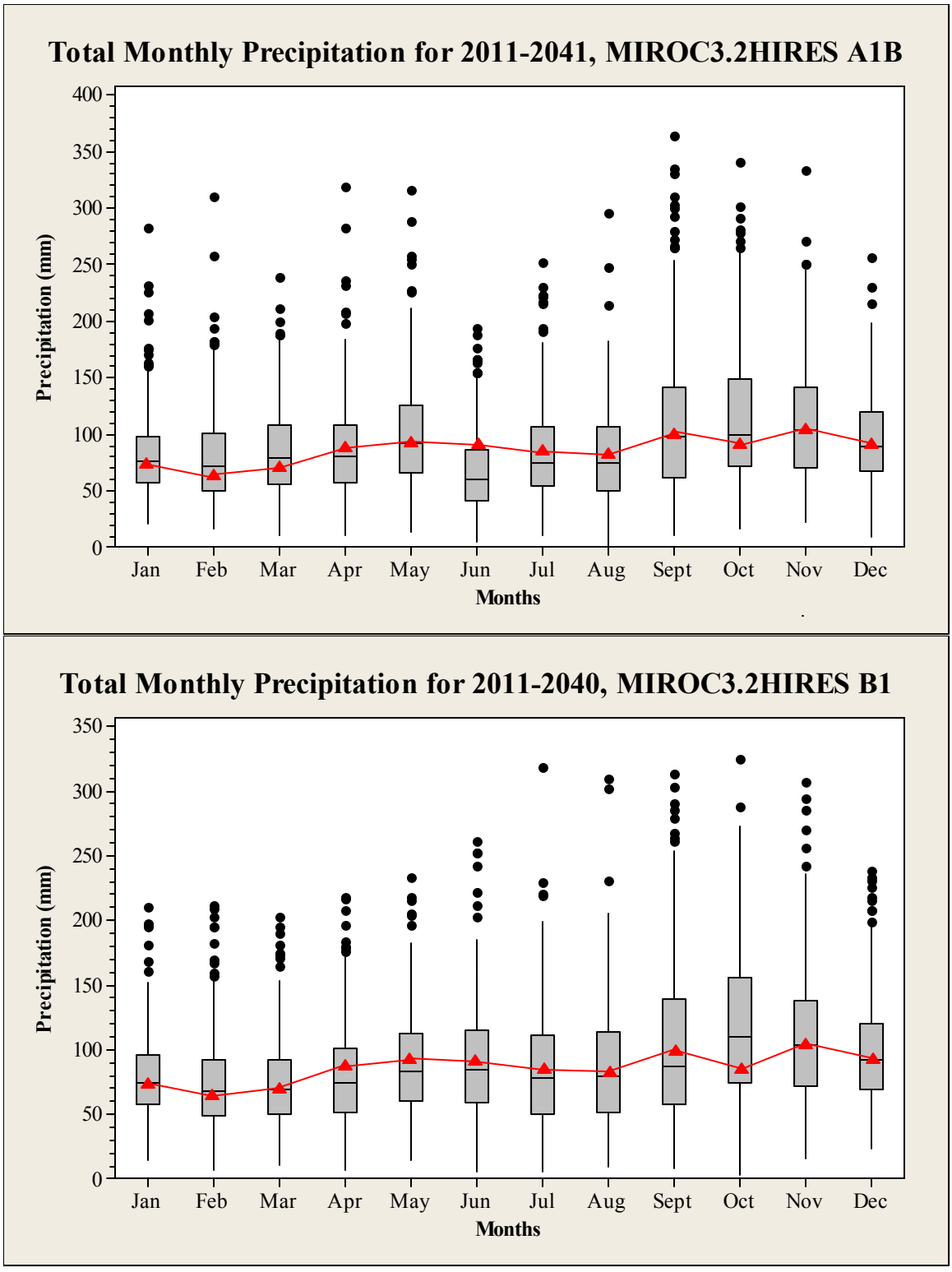


Figure C6: Total monthly precipitation box plots of MIROC3.2HIRES A1B and B1 for the years 2011-2040 with observed historical averages marked in red.

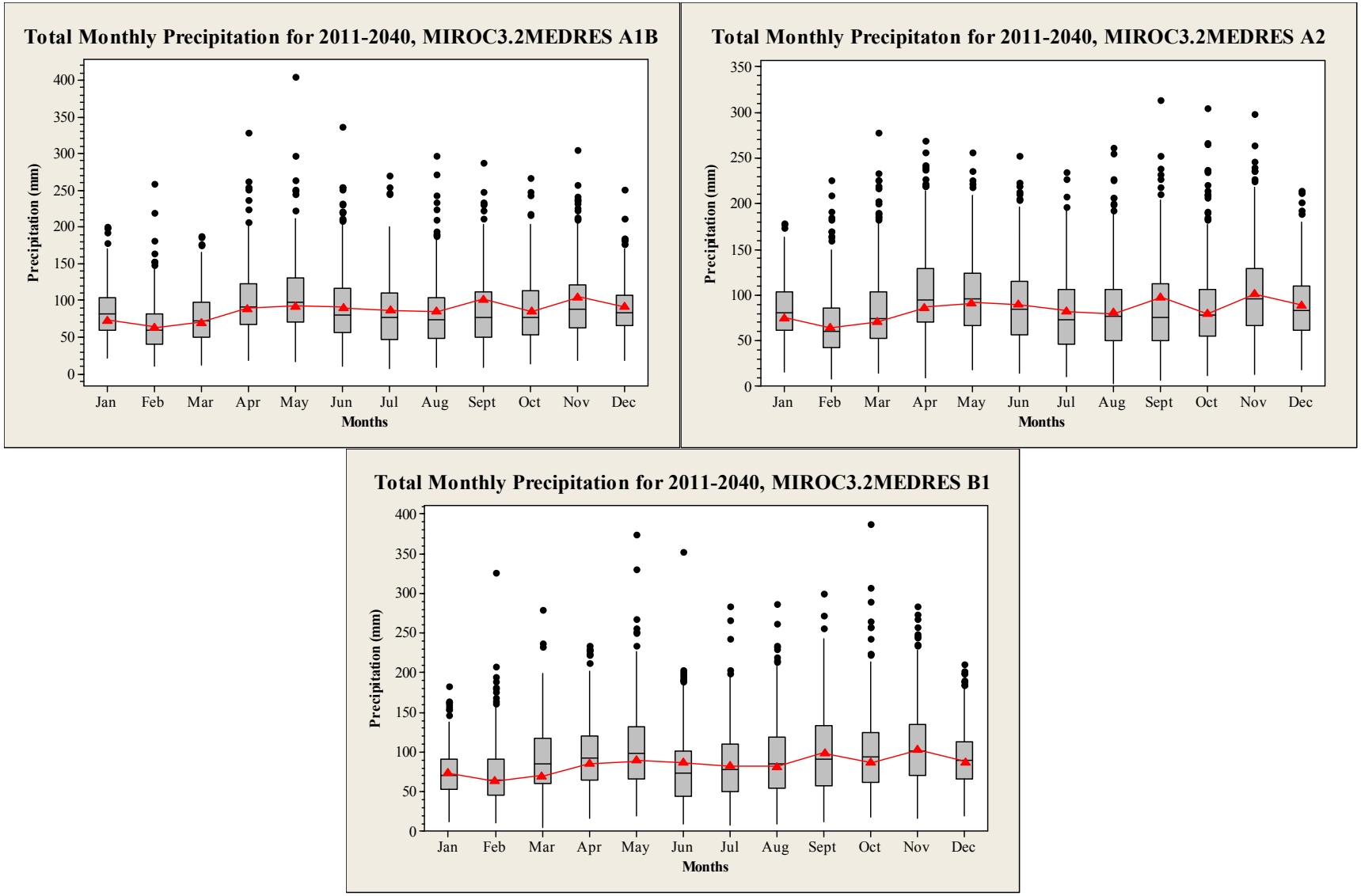


Figure C7: Total monthly precipitation box plots of MIROC3.2MEDRES A1B, A2 and B1 for the years 2011-2040 with observed historical averages marked in red

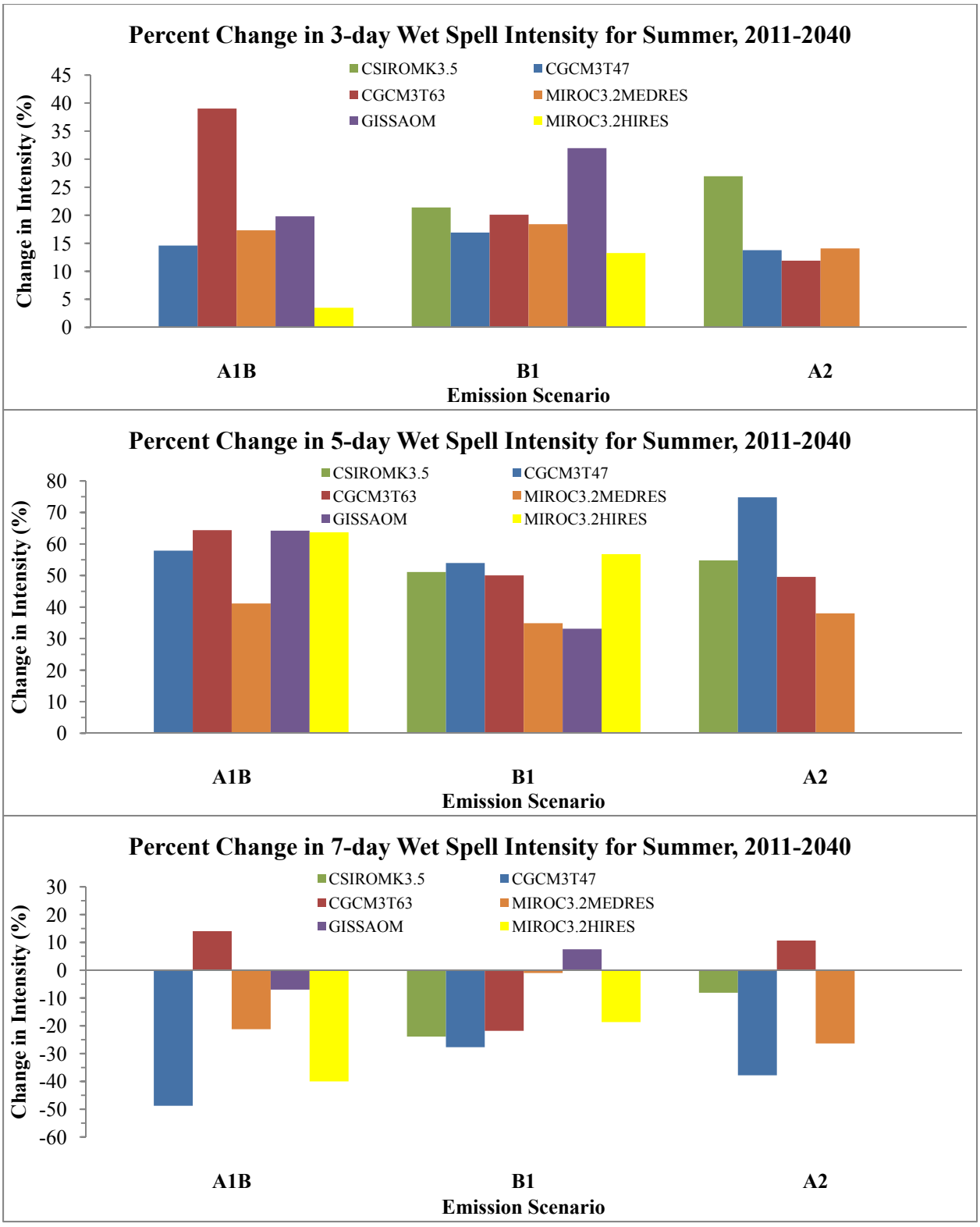


Figure C8: AOGCM predicted percent changes in wet-spell intensity for 3, 5, and 7 day spells in summer 2011-2040

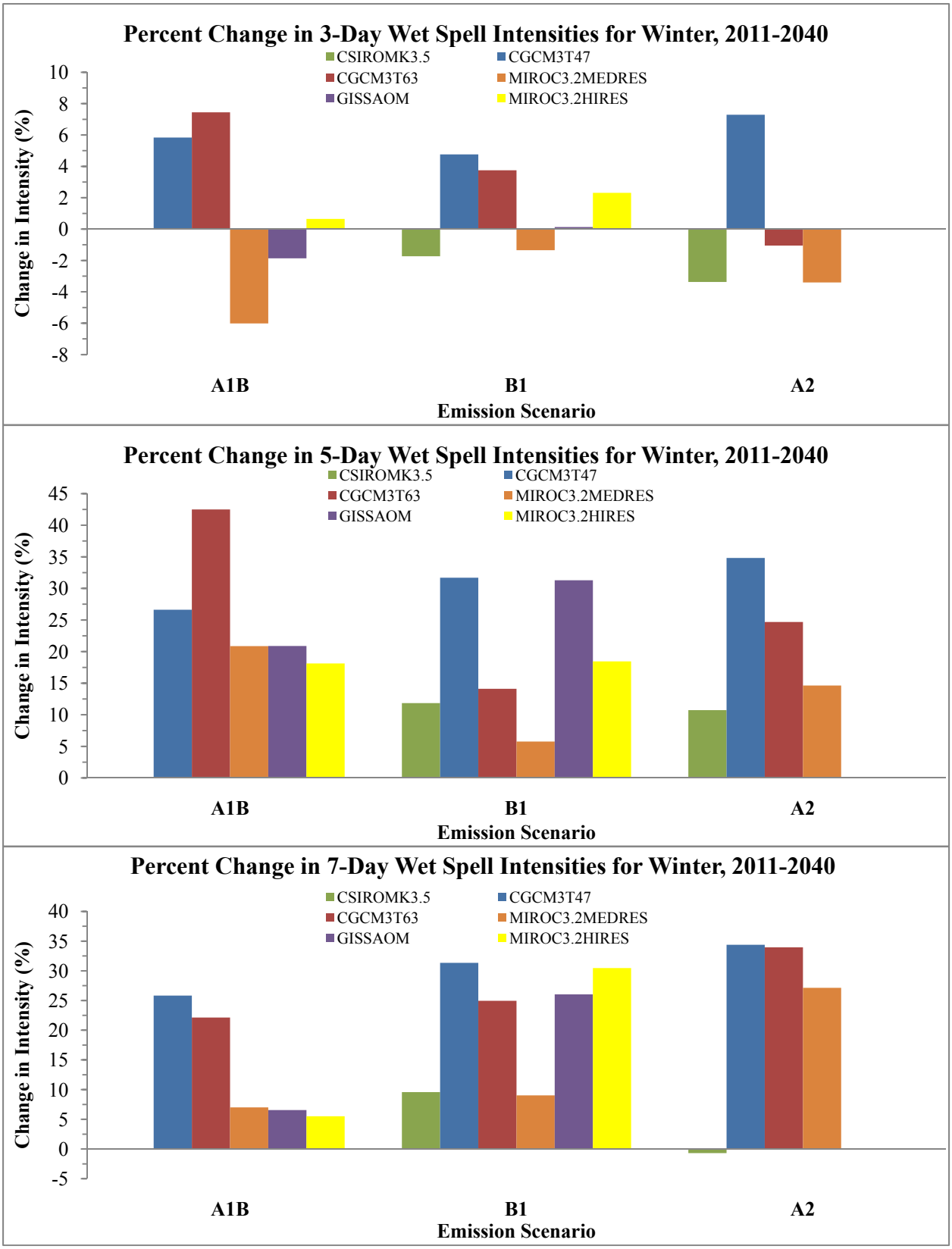


Figure C9: AOGCM predicted percent changes in wet-spell intensity for 3, 5, and 7 day spells in winter 2011-2040

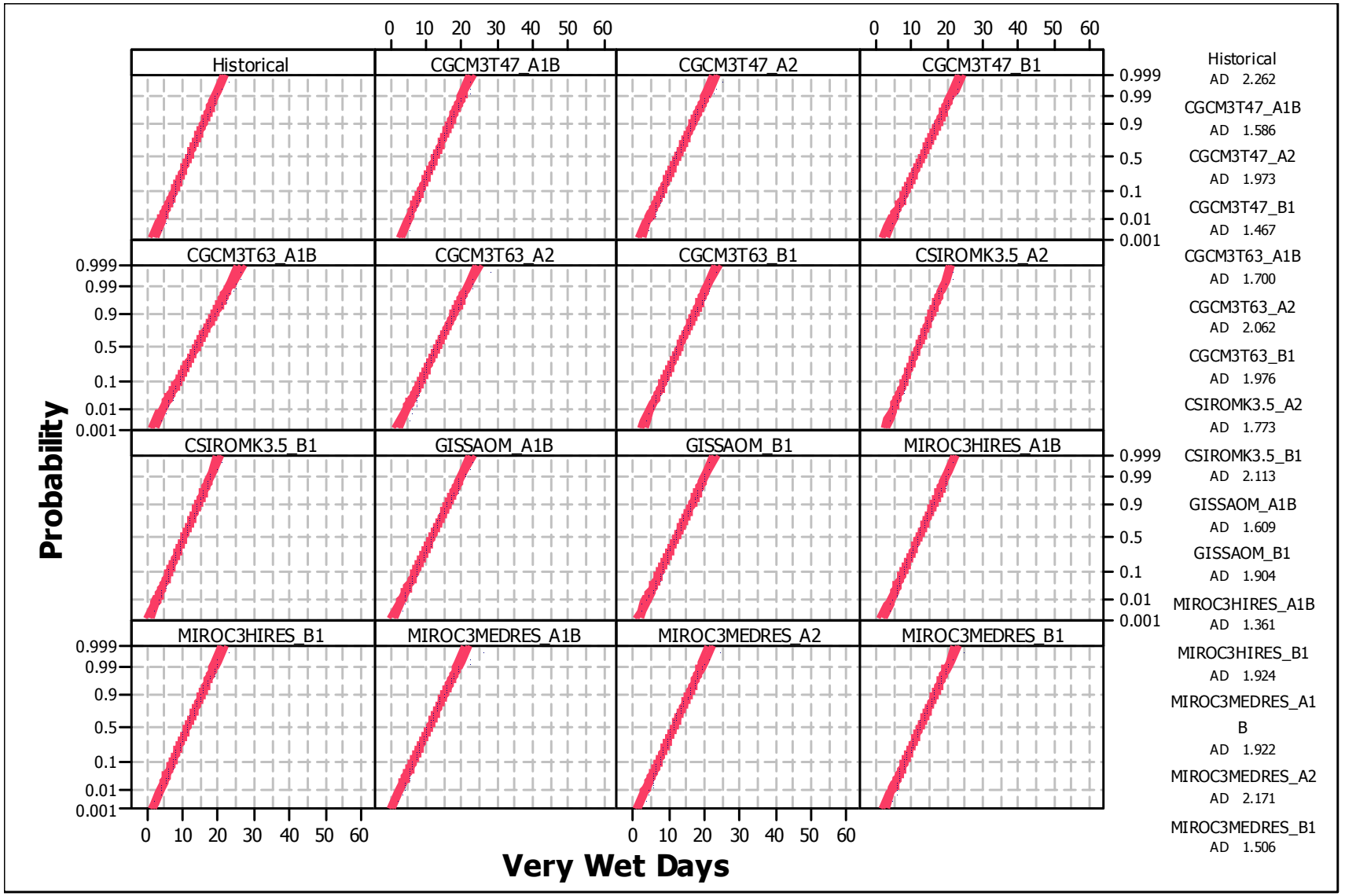


Figure C10: Probability plot of very wet days (>95th percentile of 1979-2005 precipitation) for 2020s with Anderson-Dorling's estimate

APPENDIX D: Previous Reports in the Series

ISSN: (print) 1913-3200; (online) 1913-3219

(1) Slobodan P. Simonovic (2001). Assessment of the Impact of Climate Variability and Change on the Reliability, Resiliency and Vulnerability of Complex Flood Protection Systems. Water Resources Research Report no. 038, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 91 pages. ISBN: (print) 978-0-7714-2606-3; (online) 978-0-7714-2607-0.

(2) Predrag Prodanovic (2001). Fuzzy Set Ranking Methods and Multiple Expert Decision Making. Water Resources Research Report no. 039, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 68 pages. ISBN: (print) 978-0-7714-2608-7; (online) 978-0-7714-2609-4.

(3) Nirupama and Slobodan P. Simonovic (2002). Role of Remote Sensing in Disaster Management. Water Resources Research Report no. 040, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 107 pages. ISBN: (print) 978-0-7714-2610-0; (online) 978-0-7714- 2611-7.

(4) Taslima Akter and Slobodan P. Simonovic (2002). A General Overview of Multiobjective Multiple-Participant Decision Making for Flood Management. Water Resources Research Report no. 041, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 65 pages. ISBN: (print) 978-0-7714-2612-4; (online) 978-0-7714-2613-1.

(5) Nirupama and Slobodan P. Simonovic (2002). A Spatial Fuzzy Compromise Approach for Flood Disaster Management. Water Resources Research Report no. 042, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 138 pages. ISBN: (print) 978-0-7714-2614-8; (online) 978-0-7714-2615-5.

(6) K. D. W. Nandalal and Slobodan P. Simonovic (2002). State-of-the-Art Report on Systems Analysis Methods for Resolution of Conflicts in Water Resources Management. Water Resources Research Report no. 043, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 216 pages. ISBN: (print) 978-0-7714-2616-2; (online) 978-0-7714-2617-9.

(7) K. D. W. Nandalal and Slobodan P. Simonovic (2003). Conflict Resolution Support System – A Software for the Resolution of Conflicts in Water Resource Management. Water Resources Research Report no. 044, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 144 pages. ISBN: (print) 978-0-7714-2618-6; (online) 978-0-7714-2619-3.

(8) Ibrahim El-Baroudy and Slobodan P. Simonovic (2003). New Fuzzy Performance Indices for Reliability Analysis of Water Supply Systems. Water Resources Research Report no. 045,

Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 90 pages. ISBN: (print) 978-0-7714-2620-9; (online) 978-0-7714-2621-6.

(9) Juraj Cunderlik (2003). Hydrologic Model Selection for the CFCAS Project: Assessment of Water Resources Risk and Vulnerability to Changing Climatic Conditions. Water Resources Research Report no. 046, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 40 pages. ISBN: (print) 978-0-7714-2622-3; (online) 978-0-7714- 2623-0.

(10) Juraj Cunderlik and Slobodan P. Simonovic (2004). Selection of Calibration and Verification Data for the HEC-HMS Hydrologic Model. Water Resources Research Report no. 047, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 29 pages. ISBN: (print) 978-0-7714-2624-7; (online) 978-0-7714-2625-4.

(11) Juraj Cunderlik and Slobodan P. Simonovic (2004). Calibration, Verification and Sensitivity Analysis of the HEC-HMS Hydrologic Model. Water Resources Research Report no. 048, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 113 pages. ISBN: (print) 978- 0-7714-2626-1; (online) 978-0-7714-2627-8.

(12) Predrag Prodanovic and Slobodan P. Simonovic (2004). Generation of Synthetic Design Storms for the Upper Thames River basin. Water Resources Research Report no. 049, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 20 pages. ISBN: (print) 978- 0-7714-2628-5; (online) 978-0-7714-2629-2.

(13) Ibrahim El-Baroudy and Slobodan P. Simonovic (2005). Application of the Fuzzy Performance Indices to the City of London Water Supply System. Water Resources Research Report no. 050, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 137 pages. ISBN: (print) 978-0-7714-2630-8; (online) 978-0-7714-2631-5.

(14) Ibrahim El-Baroudy and Slobodan P. Simonovic (2006). A Decision Support System for Integrated Risk Management. Water Resources Research Report no. 051, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 146 pages. ISBN: (print) 978-0-7714-2632-2; (online) 978-0-7714-2633-9.

(15) Predrag Prodanovic and Slobodan P. Simonovic (2006). Inverse Flood Risk Modelling of The Upper Thames River Basin. Water Resources Research Report no. 052, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 163 pages. ISBN: (print) 978-0-7714-2634-6; (online) 978-0-7714-2635-3.

(16) Predrag Prodanovic and Slobodan P. Simonovic (2006). Inverse Drought Risk Modelling of The Upper Thames River Basin. Water Resources Research Report no. 053, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 252 pages. ISBN: (print) 978-0-7714-2636-0; (online) 978-0-7714-2637-7.

(17) Predrag Prodanovic and Slobodan P. Simonovic (2007). Dynamic Feedback Coupling of Continuous Hydrologic and Socio-Economic Model Components of the Upper Thames River Basin. Water Resources Research Report no. 054, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 437 pages. ISBN: (print) 978-0-7714-2638-4; (online) 978-0-7714-2639-1.

(18) Subhankar Karmakar and Slobodan P. Simonovic (2007). Flood Frequency Analysis Using Copula with Mixed Marginal Distributions. Water Resources Research Report no. 055, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 144 pages. ISBN: (print) 978-0-7714-2658-2; (online) 978-0-7714-2659-9.

(19) Jordan Black, Subhankar Karmakar and Slobodan P. Simonovic (2007). A Web-Based Flood Information System. Water Resources Research Report no. 056, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 133 pages. ISBN: (print) 978-0-7714-2660-5; (online) 978-0-7714-2661-2.

(20) Angela Peck, Subhankar Karmakar and Slobodan P. Simonovic (2007). Physical, Economical, Infrastructural and Social Flood Risk – Vulnerability Analyses in GIS. Water Resources Research Report no. 057, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 80 pages. ISBN: (print) 978-0-7714-2662-9; (online) 978-0-7714-2663-6.

(21) Predrag Prodanovic and Slobodan P. Simonovic (2007). Development of Rainfall Intensity Duration Frequency Curves for the City of London Under the Changing Climate. Water Resources Research Report no. 058, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 51 pages. ISBN: (print) 978-0-7714-2667-4; (online) 978-0-7714-2668-1.

(22) Evan G. R. Davies and Slobodan P. Simonovic (2008). An integrated system dynamics model for analyzing behaviour of the social-economic-climatic system: Model description and model use guide. Water Resources Research Report no. 059, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 233 pages. ISBN: (print) 978-0-7714-2679-7; (online) 978-0-7714-2680-3.

(23) Vasan Arunachalam (2008). Optimization Using Differential Evolution. Water Resources Research Report no. 060, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 42 pages. ISBN: (print) 978-0-7714-2689-6; (online) 978-0-7714-2690-2.

(24) Rajesh Shrestha and Slobodan P. Simonovic (2009). A Fuzzy Set Theory Based Methodology for Analysis of Uncertainties in Stage-Discharge Measurements and Rating Curve. Water Resources Research Report no. 061, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 104 pages. ISBN: (print) 978-0-7714-2707-7; (online) 978-0-7714-2708-4.

(25) Hyung-Il Eum, Vasan Arunachalam and Slobodan P. Simonovic (2009). Integrated Reservoir Management System for Adaptation to Climate Change Impacts in the Upper Thames River Basin. Water Resources Research Report no. 062, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 81 pages. ISBN: (print) 978-0-7714-2710-7; (online) 978-0-7714-2711-4.

(26) Evan G. R. Davies and Slobodan P. Simonovic (2009). Energy Sector for the Integrated System Dynamics Model for Analyzing Behaviour of the Social- Economic-Climatic Model. Water Resources Research Report no. 063. Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada. 191 pages. ISBN: (print) 978-0-7714-2712-1; (online) 978-0-7714-2713-8.

(27) Leanna King, Tarana Solaiman, and Slobodan P. Simonovic (2009). Assessment of Climatic Vulnerability in the Upper Thames River Basin. Water Resources Research Report no. 064, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 61pages. ISBN: (print) 978-0-7714-2816-6; (online) 978-0-7714-2817-3.

(28) Slobodan P. Simonovic and Angela Peck (2009). Updated Rainfall Intensity Duration Frequency Curves for the City of London under Changing Climate. Water Resources Research Report no. 065, Facility for Intelligent Decision Support, Department of Civil and Environmental Engineering, London, Ontario, Canada, 64pages. ISBN: (print) 978-0-7714-2819-7; (online) 987-0-7714-2820-3.