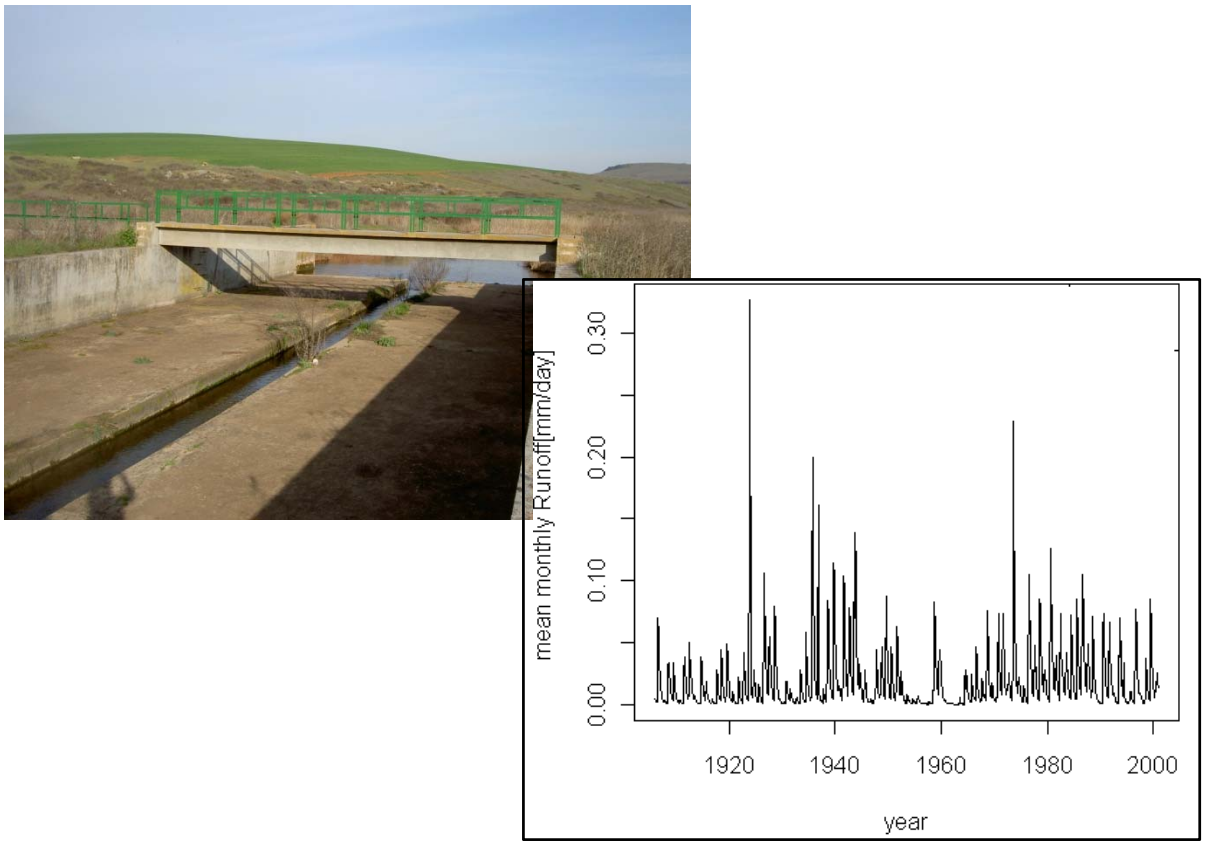
Technical Report No. 40

MULTI-MODEL ANALYSIS OF DROUGHT AT THE GLOBAL SCALE: DIFFERENCES IN HYDROLOGICAL DROUGHT BETWEEN THE FIRST AND THE SECOND PART OF THE 20TH CENTURY



Sihine T. Estifanos, Marjolein H.J. van Huijgevoort, Pieter Hazenberg,

Graham P. Weedon, Nathalie Bertrand, Sonja Folwell, Sandra Gomes,

Frank Voβ & Henny A.J. van Lanen

August 2011

WATCH is an Integrated Project Funded by the European Commission under the Sixth Framework Programme, Global Shift and Ecosystems Thematic Priority Area (contract number: 036946).

The WATCH project started 01/02/2007 and will continue for 4 years.

Title:	Multi-model analysis of drought at the global scale: differences in hydrological drought between the first and the second part of the 20th century
Authors:	Sihine T. Estifanos, Marjolein H.J. van Huijgevoort, Pieter Hazenberg, Graham P. Weedon, Nathalie Bertrand, Sonja Folwell, Sandra Gomes, Frank Voβ & Henny A.J. van Lanen
Organisations:	- Wageningen University - Hydrology and Quantitative Water Management Group (WUR) - UK Met Office · Laboratoire de Météorologie Dynamique (LMD), France · Centre for Ecology and Hydrology, UK · Centro de Geofisica da Universidade de Lisboa, Portugal · University of Kassel, Germany
Submission date:	25 August 2011
Function:	This report is an output from Work Block 1; Task 1.3.4 Methodology for evaluating uncertainty of means and extremes from climate model data, and Work Block 4; Task 4.1.1 Investigate processes controlling the propagation of drought.
Deliverable	WATCH deliverables D 1.3.4 Report on the uncertainty of the global water cycle of the 20th Century, D 4.1.4 Report on the increased understanding of the propagation of drought in different hydro-climatological regions, physical catchment structures and different scales, D 4.1.3a Methodologies to quantify the space-time development of droughts at regional and continental scales, D 4.1.7 Report on the trends/shifts in hydrological extremes in the 20th century and the attribution of climate and human influences to possible trends,.and it contributes to M4.1-3 Space time development and propagation of drought: regions using selected LSHMs, M4.1.6a Overview of major historical events; part: drought.

Photo cover:

Left:	Low flow in the Upper-Guadiana Basin (2008)
Right:	Simulated hydrograph with WaterGAP of a grid cell in Mexico for the whole 20th Century

Table of Contents

List of tables

List of figures

Abstract

Understanding the occurrence and distribution of global hydrological drought has paramount importance in assessing drought monitoring schemes, seasonal drought forecasting and drought prediction for future climate. We assessed global hydrological drought for the twentieth century based on runoff simulations of two global hydrological models (WaterGAP and GWAVA), two land surface models (HTESSEL and Orchidee), and an ensemble mean derived from six global models used within the EU-WATCH project. Global hydrological drought trend was analysed by splitting the 20^{th} century in to the first part (1906-1957) and the second part (1958-2000) based on the different approaches used to generate the WATCH meteorological forcing data. Global gridded (0.5^0 resolution) time series of simulated total runoff (sum of the surface and sub-surface runoff) obtained from each model is used to characterize drought.

The influence of the different approaches used to generate the WATCH Forcing Data (for the two periods of the century) on simulated runoff, is assessed using two-tailed t-test. A combination of variable threshold level and consecutive dry period approach is used to characterize drought. For both methods a threshold value of 20 percentile is used to define drought. Trends in spatial extent of drought for the two periods of the century are assessed at global and continental scale using the Mann-Kendall non parametric test.

Statistically significant trend shifts in simulated runoff between the two periods of the century are observed in nearly half of the land surface cells. Sliding time windows are used to determine thresholds to smooth the effect of these trend shifts and also to capture inter-decadal climatic variability. Percentage of the globe in drought is computed for the whole simulation period at a monthly time step for each model. Most of the world wide droughts, that are often associated with the ENSO phenomena, are identified by the four models as well as the multi-model ensemble mean though there exist differences in percentage of area in drought among models. Inter-comparing the two periods of the century, consistent statistically significant decreasing trends are observed in area of the globe in drought for the first part (1906 -1958) of the century, whereas increasing trends are observed for the second part(1958-2000) of the century among all models. The difference can be attributed to the different approaches used to drive the meteorological forcing data for the first and second part of the 20^{th} century. Most of the models capture the documented major historical droughts at continent scale. Similar analysis for continental drought trend shows that there is a consistent increase in percentage of area in drought in the second part of the century for Africa, Asia, whereas there is a decreasing trend for North America, and Australia.

Key words: drought; hydrological drought, global hydrological models; land surface models; multi-model analysis; trend analysis; twenty century

1 Introduction

1.1 Background

Drought is defined as "a sustained and regional occurrence of below average natural water availability"(Tallaksen and Van Lanen, 2004) which implies that drought can occur in any hydro-climatological region and at any time of the year (Fleig *et al.*, 2006). Droughts are triggered by precipitation deficiency that is often associated with an increase in evapotranspiration rates (Corzo Perez *et al.*, 2011). Droughts can be classified as meteorological drought (deficit in precipitation), agricultural drought (deficit in soil moisture) or hydrological drought (low discharge of streams and low ground water levels) (Tallaksen and Van Lanen, 2004; Fleig *et al.*, 2006). A better understanding about the occurrence of droughts (initiation, persistence and recovery) is important in order to improve drought management including drought monitoring, seasonal forecasting of droughts and prediction of its occurrence for the future climate (Sheffield and Wood; 2007, 2008).

Within the EU-WATCH project, the global water cycle is modelled with different land surface and global hydrological models to be able to analyse, quantify and predict the current and future global water cycle (www.eu-watch.org). The outputs of these models are later analysed for characterization of hydrological extremes including drought at a global scale. One of the inputs required to model the global water cycle is the meteorological data of the whole globe for the simulation period (Sheffield *et al.*, 2006; Weedon *et al.*, 2011). These meteorological data could be obtained from Global Climate Models (GCMs) that integrate atmosphere-land interactions (Corzo Perez *et al.*, 2011). However, such models have coarse resolutions and large biases to reliably simulate the land surface processes (Weiland *et al.*, 2010; Corzo Perez *et al.*, 2011). Hence, meteorological data from GCMs have to be downscaled to finer resolutions and biases have to be corrected to get reliable simulations (Corzo Perez *et al.*, 2011; Weedon *et al.*, 2011).

WATCH Forcing Data (WFD) are derived separately for the two periods of the 20th century: the first part (1901-1957) and the second part (1958-2001) (Weedon *et al.*, 2011). Different techniques of reanalysis are used for producing the forcing data for these two periods. This is mainly because the ERA-40 reanalysis product is not available prior to 1958 and also due to limited availability of data especially for the first half of the 20th century. The WFD for the second part of the 20th century is derived from ERA-40 re-analysis product of the same period by downscaling and bias correcting the ERA-40 reanalysis product (Weedon *et al.*; 2010,2011), whereas the WFD for the first part of the century are derived by reordering the ERA-40 reanalysis product (1958-2000). Previous studies have shown that the results of land surface and global hydrological models are sensitive to the atmospheric forcing datasets that drive them (Sheffield, 2008). Hence, this study will focus on analysing the differences in hydrological drought between the first and the second half of the 20th century using these two subsets of the WATCH forcing data. All the land surface and hydrological models in the EU-WATCH project used these WATCH Forcing Data as an input (see Figure 1-1). Therefore,

the discrepancies caused by the difference in atmospheric forcing are reduced and comparisons among model results are possible. The main focus of this study is to analyse 20th century global drought using multi-model approach.

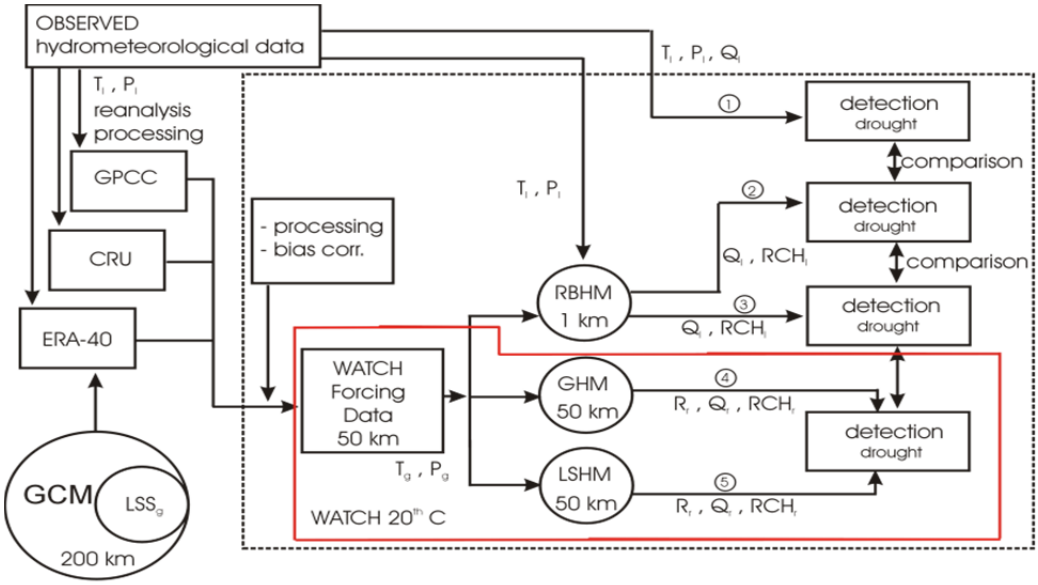


Figure 1-1. Flow chart showing the context of the multi-model intercomparison of global drought for the whole 20th century within WATCH

1.2 Objective

The specific objectives of this study are:

- To analyse the differences in global drought between the first and the second part of the 20th century using the outputs of two land surface, two global hydrological models and ensemble mean of six models used within the EU-WATCH project
- To assess whether there exists a discrepancy or consistency in the droughts characterized by the different models and model ensemble mean.

1.3 Research approach

The total simulated runoff (sum of surface and sub-surface runoff) obtained from two Global hydrological models (WaterGAP and GWAVA), two land surface models (HTESSEL and Orchidee), and an ensemble mean of six models used within the EU-WATCH project are used to assess the occurrence of drought at global scale

1.4 Outline

Chapter 2 presents the methodologies employed with in this research. This includes a brief description of the different approaches used to derive the WFD for the first and the second part of the century. A general overview of the models, the statistical tests used to assess the trend shifts between the two periods of the century and the three approaches used to characterize droughts are also discussed in Chapter 2. Chapter 3 describes the trend shifts in simulated runoff between the first and the second part of the century using different approaches. Chapter 4 presents global drought characteristics for the first and the second part of the twenty century using different large scale models and the model ensemble mean. The conclusions and recommendation of the study are summarized in Chapter 5.

2 Methodology

In this chapter, the research methodologies implemented within this study are presented. The different approaches used to generate the WFD for the first (1906-1957) and the second (1958-2000) part of the 20th century are discussed briefly. A general overview of the WaterGAP model and a summary of the other models are also provided. The statistical test used for assessing the presence of a trend shift in the simulated total runoff between the two periods of the century is also described. At last, different drought characterization approaches are explained.

2.1 The WATCH Forcing Data

The Watch Forcing Data (WFD) are meteorological forcing data for the whole 20th century that are derived within the European Union WATCH project using the ERA-40 reanalysis product as a basis (Uppala *et al*., 2005; Weedon *et al*., 2011). The ERA-40 dataset is reanalysis of meteorological observations that is based on information provided by the data-assimilation system (model and analysis) and on external information from observations and boundary conditions (Uppala *et al*., 2005; ERA-40 Atlas, 2011). The ERA-40 dataset is available for the period 1958-2002 (Uppala *et al*., 2005). In order to be able to derive the WFD for the whole 20th century from the ERA-40 dataset, different approaches were used for the period 1901-1957 and 1958-2000 (Weedon *et al*., 2011).

For the second part of the 20th century(1958-2000), the WFD are derived by downscaling the ERA-40 reanalysis product from one degree grid to half degree grid using bilinear interpolation (Weedon *et al*.; 2010,2011). The half degree gridded values of temperature, pressure and specific humidity are then elevation and bias corrected using CRU and GPCC observations. In addition, the precipitation of ERA-40 reanalysis product is also bilinear interpolated to a half degree grid and a number of adjustments are made in order to improve most of its limitations (details are discussed in Uppala *et al*., 2005; Weedon *et al*., 2011). The resulting WATCH forcing dataset is then validated using gap filled FLUXNET data (Weedon et al.; 2010, 2011).

For the first part of the 20th century (1901-1957), the WATCH forcing data are computed from re-ordered ERA-40 data since the ERA-40 data are not available prior to 1958. The ERA-40 data (1958-2001) were extracted and assigned randomly to the years 1901-1957. The remaining 13 years are again randomly assigned. The assignment is done in such a way that leap years are assigned to leap years and non-leap years to non-leap years. Once the assignment is accomplished, the same procedures that are applied to the second part of the century are performed (Weedon *et al*., 2011). However, for the period prior to 1950, the spatial and temporal coverage of CRU and GPCCv4 observations is limited for precipitation coverage and cloud cover observations. Hence, for specific months, those cells that have too few observations, the values for the local month are obtained from CRU 1961-1990 climatologically average. In addition, the reordering of years lead to end year discontinuity in the meteorological variables and these discontinuities are also adjusted (details are discussed in Weedon *et al*. (2011)).

2.2 Large-scale models

To assess the characteristics of global drought in the first and the second half of the century, different models are used in this study. Two Global Hydrological Models (WaterGAP and GWAVA), two Land Surface Models (HTESSEL and Orchidee) and an ensemble mean of six models are used to assess the differences in drought characteristics between the first and the second part of the century. Models that are classified as Global Hydrological Models (GHMs) are models that solve the water balance at the land surface and run in a daily time steps whereas Land Surface Models (LSMs) are those that solve both the water and the energy balance and they run in sub-daily time steps (Haddeland *et al*., 2011; Weedon *et al*., 2011). In the following section, a general overview of the WaterGAP model is discussed as an example and a summary of the other models is presented.

The Water Global Assessment and Prognosis (WaterGAP) model is one of the global hydrological modelling tools designed for regional and global assessment of fresh water resources. The model comprises two components. A global hydrological model that takes into account physical and climatic factors in order to simulate surface and subsurface runoff, groundwater recharge and river discharge, Whereas a water use model that takes into account socio-economic factors to estimate the domestic, industrial and agricultural water use (Alcamo *et al*., 2003). The schematic representation of the WaterGAP model is given Figure 2-1.

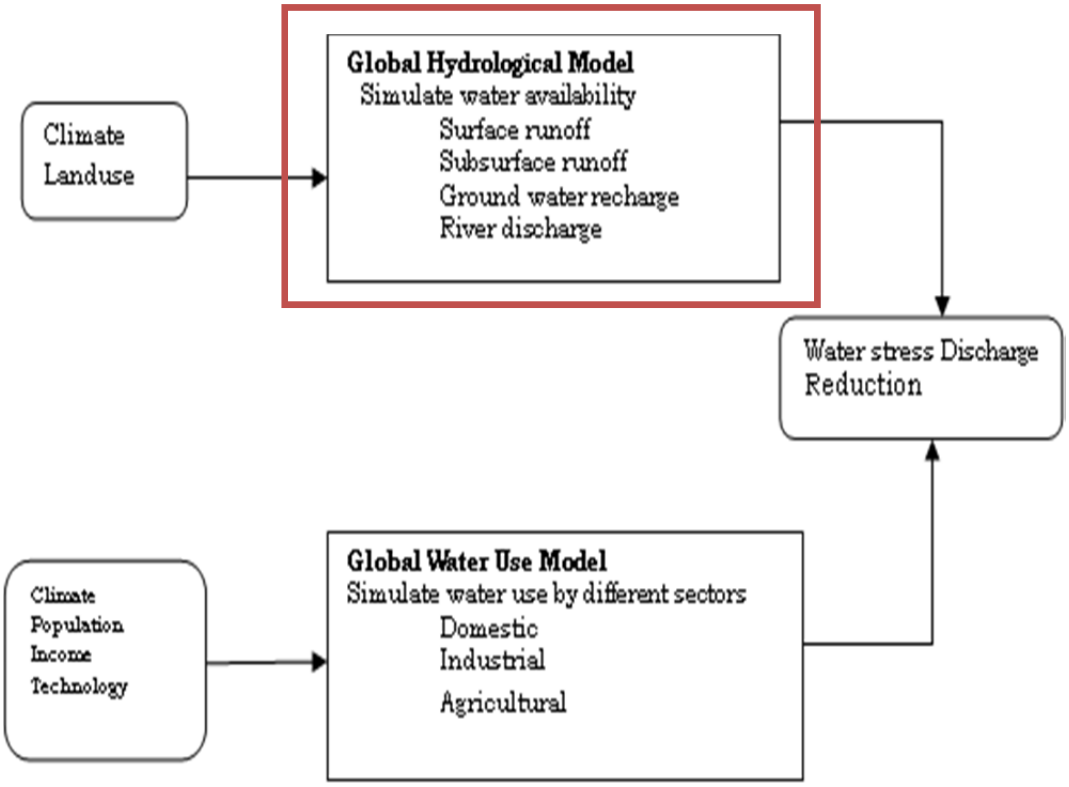


Figure 2-1. Schematic representation of the WaterGAP model (Alcamo et al., 2003). The red box shows part of the model this study focuses

In depth discussion about how each model works is beyond the scope of this study. Therefore, the summary of the models used in this study is given in P: Precipitation (rain or snow distinguished in the model), T: Mean daily air temperature, Tmax: Maximum daily air temperature, Tmin: Minimum daily air temperature, W: Wind speed, Q: Specific humidity, LW: Longwave radiation flux (downward), LWn: Longwave radiation flux (net), SW: Shortwave radiation flux (downward), SP: Surface pressure

As WaterGAP, all the other models used in this study have a spatial resolution of 0.5o covering the whole globe. IN this study, global drought analysis is carried out only for land point cells defined according to the CRU mask. Theses land points are distributed over 67421 cells. Simulated total runoff (*Qtot*)(the sum of the surface runoff(*Qs*) and the subsurface runoff(*Qsb*)) that is obtained from each model, is used to characterize global drought. Apart from the four models given in Table 2-1, an ensemble mean total runoff is also calculated from six models in order to investigate the combined effect of the models on drought characterization. The models that participate in the ensemble mean calculation is shown in Error**! Reference source not found..** (Haddeland et al., 2011). 5o covering the whole globe. IN this study, global drought analysis is carried out only for land point cells defined according to the CRU mask. Theses land points are distributed over 67421 cells.

Table 2-1. Summary of the Characteristics of models used in this study (adapted from Haddeland et al., 2011)

Model Type	Model Name	Model time step	Meteorological forcing variables	Solve energy balance	Evapotranspiration scheme	Runoff scheme	Snow scheme
Global Hydro-logical models (GHMs)	WaterGAP	daily	P, T, LWn, SW	No	Priestley-Taylor	Beta function	Degree Day
	GWAVA	daily	P, T, W, Q, LWn, SW, SP	No	Penman-Monteith	Saturation excess/Beta function	Degree Day
Land Surface Models (LSMs)	HTESSEL	1hr	R, S, T, W, Q, LW, SW, SP	yes	Penman-Monteith	Infiltration excess/ Darcy	Energy Balance
	Orchidee	15 min	R, S, T, W, Q, SW, LW, SP	Yes	Bulk formula	Saturation excess	Energy Balance

P: Precipitation (rain or snow distinguished in the model), T: Mean daily air temperature, Tmax: Maximum daily air temperature, Tmin: Minimum daily air temperature, W: Wind speed, Q: Specific humidity, LW: Longwave radiation flux (downward), LWn: Longwave radiation flux (net), SW: Shortwave radiation flux (downward), SP: Surface pressure

As WaterGAP, all the other models used in this study have a spatial resolution of 0.5^{o} covering the whole globe. IN this study, global drought analysis is carried out only for land point cells defined according to the CRU mask. Theses land points are distributed over 67421 cells. Simulated total runoff (Q_{tot})(the sum of the surface runoff(Q_s) and the subsurface runoff(Q_{sb})) that is obtained from each model, is used to characterize global drought. Apart from the four models given in Table 2-1, an ensemble mean total runoff is also calculated from six models in order to investigate the combined effect of the models on drought characterization. The models that participate in the ensemble mean calculation is shown in **Error! Reference source not found.**.

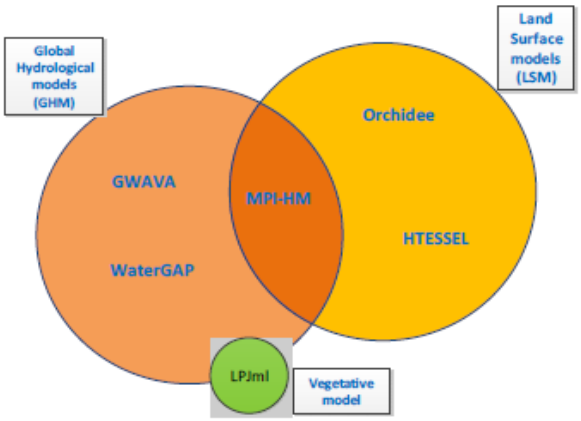


Figure 2-2. Schematic representation of the the models that participate in the model ensemble mean

2.3 Statistical tests for trend shift of simulated total runoff

Prior to drought analysis, this study also assesses whether there exist trend shift in simulated total runoff between the first (1906-1957) and the second (1958-2000) part of the century. This is because different reanalysis approaches were used to generate the WFD for the two periods of the century that are used to drive the GHMs and LSMs. The assessment of trend shift is necessary since it has implications on the time windows to be used for threshold determination and consequently on drought characterization. Among the different statistical tests that can be used to detect trend shifts, two-tailed t-test (Rodionov, 2005; Xiong and Guo, 2004) is chosen.

The two-tailed t-test is used since the expected shift of point is known in advance, which is supposed to be the point where different approaches are used to derive the WFD i.e. 1957/58. Two-tailed t-test is performed to detect if there exist statistically significant shift in mean runoff for land point cells of the globe at 5% significance level assuming the same variance for the two periods of the century. In addition, the magnitude of the shift in mean between the two periods of the century is quantified as the difference in mean total run off of the first part and the second part of the century and is given as:

$$\Delta Q = mean(Q_{fp}) - mean(Q_{sp}), \tag{2.1}$$

where ΔQ is the magnitude of shift in mean, Q_{fp} is a dataset of runoff for the first part (1906-1957) of the century and Q_{sp} is a dataset of runoff for the second part (1958-2000) of the century. However, due to the spatial climatic differences all over the world and higher magnitude of shift in areas with relatively higher runoff, it is not possible to clearly see the spatial variability of the magnitude of shift. Therefore, the magnitude of shift given in equation 2.1 is scaled by the mean runoff of the whole century for a given cell

$$\Delta Q_{\,scaled} = \frac{\Delta Q}{mean(Q_c)}, \tag{2.2}$$

Where, ΔQ_{scaled} is scaled magnitude of shift and Q_c is a dataset of runoff for the whole century (1906-2000).

2.4 Drought characterization methods

One approach to define drought is measuring the degree of departure from normal for a given variable of interest such as precipitation, soil moisture, groundwater level or discharge (e.g, Hisdal *et al*., 2004; Andreadis *et al*., 2005; Corzo Perez *et al*., 2011). In this study, anomalies of simulated total runoff (sum of surface and sub-surface runoff) are used to determine hydrological drought. Two different methods were implemented to derive hydrological drought from the simulated time series total runoff, the threshold level method (TLM) (Yevjevich, 1967; Hisdal *et al*., 2004) and the consecutive dry period (CDP) approach.

For situations where runoff values are larger than zero, the TLM (Yevjevich, 1967; Hisdal *et al*., 2004) is used. With this method, drought occurs when the variable of interest (e.g. stream flow, precipitation, recharge) is below a predefined threshold level. In this study, a threshold the beginning and ending of drought is defined based on a threshold level of total runoff percentile value of 20%. This value corresponds to the value below which 20 percent of the simulated total runoff for a given time series data of total runoff. The 20 percentile value was selected in order to be consistent with other global and large scale studies (e.g. Sheffeld et al. (2009); Andreadis et al. (2005)). This value is an indication of low flows rather than only focusing on severe drought events. The start of a drought event is indicated by the point in time when the variable falls below this threshold and the event continues until the threshold is exceeded again. Drought characteristics commonly derived with this method are beginning, end, duration, deficit volume, and minimum flow during an event (Hisdal et al., 2004; Fleig et al., 2006). Both a fixed and variable (seasonal, monthly, or daily) threshold can be used. However, in the current study it was decided to apply the variable TLM for the monthly runoff data (Figure 2-3a). The threshold approach has difficulties to make a distinction between the impact of a drought for regions where zero runoff is observed during part of the year, either because precipitation occurs in the form of snow (winter period in higher latitudes) or is not observed at all (arid and semi-arid regions).

Therefore, for cells which have a zero runoff for at least 5% of their time series, a Consecutive Dry Period (CDP) approach was developed. This technique is related to the consecutive dry day approach originally develop to study meteorological drought (Vincent & Mekis, 2006; Groisman & Knight, 2008; Deni & Jemain, 2009). However, in its current implementation, this method is combined with the TLM applied for positive runoff data. An implementation of the CDP approach is presented in Figure 2-3b. The runoff data series plotted in Figure 2-3a contains multiple periods with zero runoff. The first step is to apply the TLM method for a given threshold, to identify months that are in drought but still contain positive runoff values (Figure 2-3a). Based on this information, two different series are identified. The first contains the consecutive number of months with zero runoff (Figure 2-3b, red line), while the second contains the number of consecutive months for a pixel either being in a drought (based on the TLM) or containing zero runoff month. Based on all data in

the former series and the threshold of interest, threshold value CDP is calculated using the same threshold level of 20 percentile (Figure 2-3b, dashed line). This consecutive number threshold value is compared to the latter series, which contains both the TLM and zero runoff information. In case the monthly value of that series contains is larger as compared to the consecutive number threshold value for a month with zero runoff data, the cell experiences drought. Combining both the TLM and CDP approach, one is able to identify whether the simulated runoff for a given month within the cell corresponds to a drought event or not (Figure 2-3c)

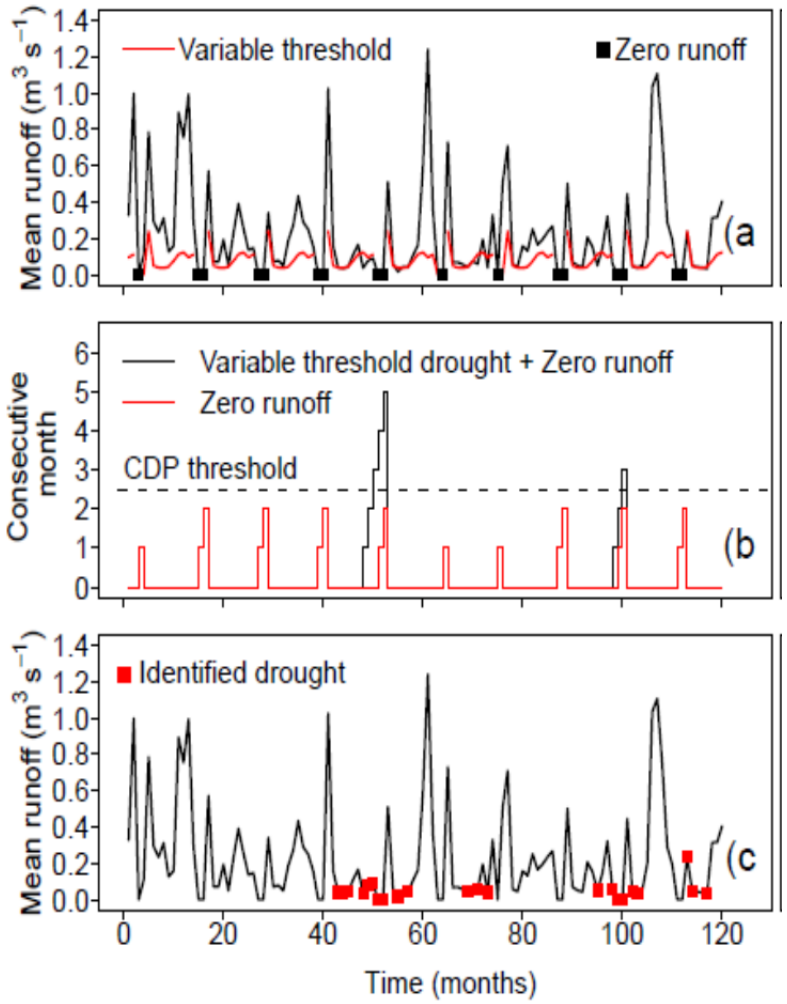


Figure 2-3. Schematic to demonstrate drought characterization methodologies used in this study: A) variable threshold Method, B) consecutive dry period approach, and C) identified drought combing the two methods

Once the spatial extent of drought is computed for the whole simulation period, Mann-Kendall non parameteric test is used to asses the trends in global drought in the 20th century and separetly for the two periods of the century using the four models and the multi-model ensemble mean. Similar trend test in spatial extent of drought is also carried out at contint scale.

3 Assessment of trend shift in simulated runoff between the two periods of the 20th century

This chapter provides a comprehensive analysis of trend shift in simulated runoff between the first (1906-1957) and the second part (1958-2000) of the 20th century for the WaterGAP model using two tailed t-test. This is important to assess the influence of the different approaches used to drive the WFD on simulated runoff. Based on this comprehensive analysis for the WaterGAP model, a common procedure that could be applicable to the other models is identified. Using the identified common procedure the trend shifts for the other models are assessed.

3.1 Comprehensive analysis for WaterGAP

Though eventually four models and a multi-model ensemble mean will be used for analysis of drought occurrence, a comprehensive analysis of trend shift in simulated runoff between the two periods of the century is carried out only for the WaterGAP model. In this study, the two-tailed t-test discussed in Section 2.3 is used to identify the existence of trend shift in simulated runoff between the two periods of the century for each for each land surface grid cell.

3.1.1 T-test using daily runoff data

The WaterGAP model as well as the other models used within the EU-WATCH project simulates total runoff on a daily basis. The trend shift test for WaterGAP is performed using these daily simulation results and those cells that show statistically significance differences in mean daily runoff for the two periods of the century at 5 percent significant level are identified (Figure 3-1(left)). In addition, the magnitude of the shift in the mean of total daily runoff between the two periods is computed according to Equation 2.1 for those cells that have statistically significant shifts (Figure 3-1(right)).

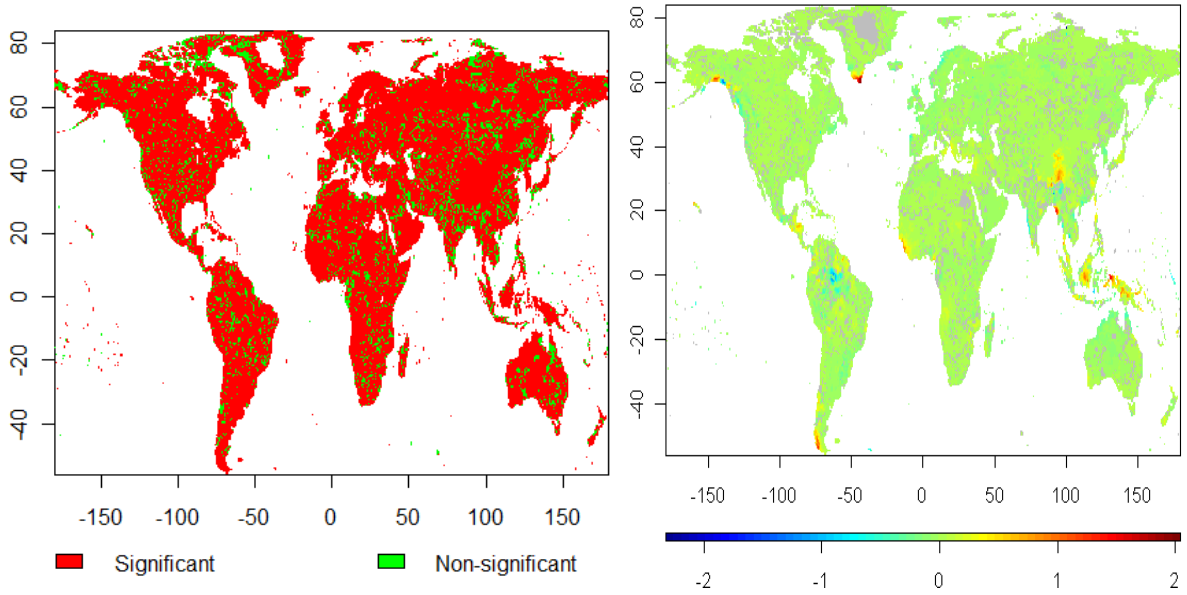


Figure 3-1. Left: Spatial location of cells that have statistically significant shift (red) and non-significant shift (green) in mean daily runoff. Right: The absolute magnitude of shift in mean daily runoff of cells that show statistically significant shift.

Though Figure 3-1(right) shows the absolute value of the magnitude of the shift in the mean daily runoff between the two periods of the century, it is difficult to identify the spatial variation due to the influence of spatial climatic differences and their effect on the associated runoff. Hence, in order to improve the identification of spatial variation differences with respect to the magnitude of shift, there values were scaled using Equation 2.2. The results from this procedure are shown in Figure 3-2.

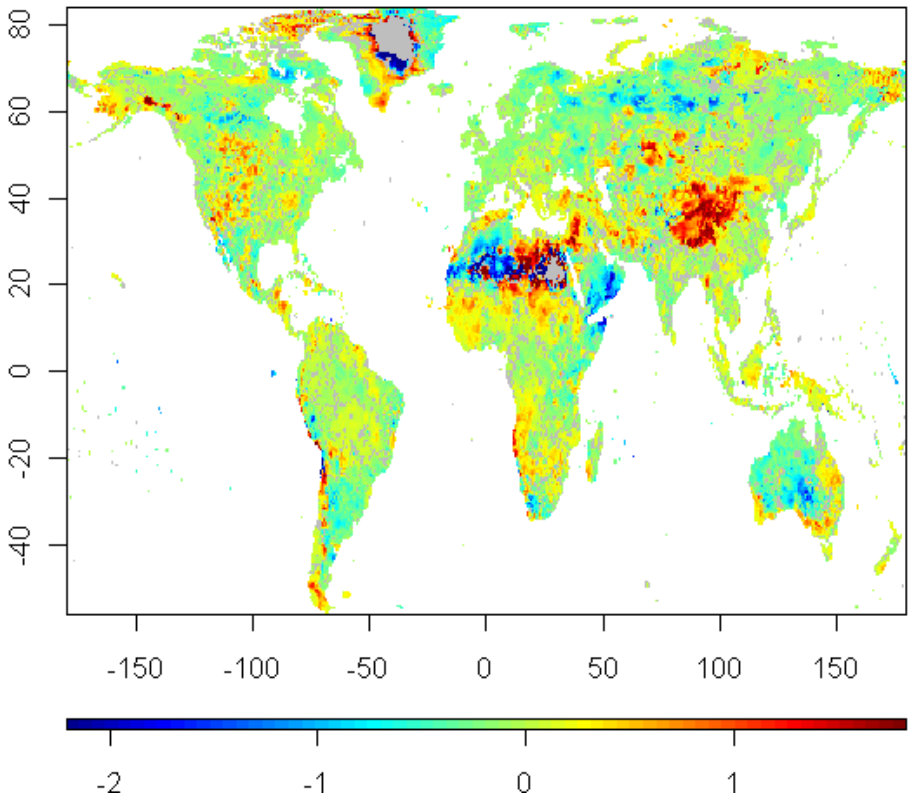


Figure 3-2. Scaled magnitude of shift of daily runoff scaled by the mean of daily runoff of the whole simulation period.

The results obtained from the t-test for daily runoff show that about 85% of land surface cells are found to have statistically significant trend shift between the first and the second half of the 20th century. It is expected that this is due to large fluctuations in daily runoff values between the different inter- and intra-yearly periods. Hence, the mean of such values is expected to vary substantially. A Statistical summary of the scaled magnitude of shift in mean daily runoff is given in **Error! Reference source not found.**. A scaled magnitude of 1 implies a shift in mean equivalent to the mean of the total runoff of a given cell for the whole 20th century. Positive scaled in magnitude means the mean in the first half is higher than that of the second half (drying condition) and negative scaled magnitude indicates the mean of the first half is lower than the second half (wetting condition). From **Error! Reference source not found.** it can be seen that majority (70%) of the land surface grids have a scaled mean of in the range of -0.5 to +0.5.

Table 3-1. Statistical summary of the ranges of scaled magnitude of shift in mean daily runoff between the first and the second d half of the twenty century

Scaled magnitude of shift in mean daily runoff	Percentage of Land surface grids (%)
<-1	2.2
-1 - -0.5	5.1
-0.5 - -0.25	10.9
-0.25 - -0.1	18.3
-0.1 – 0.1	20.2
0.1 - 0.25	11.7
0.25 - 0.5	*8.6*
0.5 - 1	5.4
>1	2.7

3.1.2 T-test using monthly mean runoff data

In order to smooth the high fluctuations of daily runoff values and improve computational efficiency, monthly mean values of total runoff are computed for each land point and the two tailed t-test is performed again for each cell of the grid. When monthly mean values are used, the number of cells showing statistically significant shift decreased drastically from 85% to 48%. The spatial location of cells that show significance differences in mean monthly runoff between the two periods of the century is shown in **Error! Reference source not found.**(left). The absolute magnitude of shift for each cell that has significant differences in mean is also shown in **Error! Reference source not found.** (right).

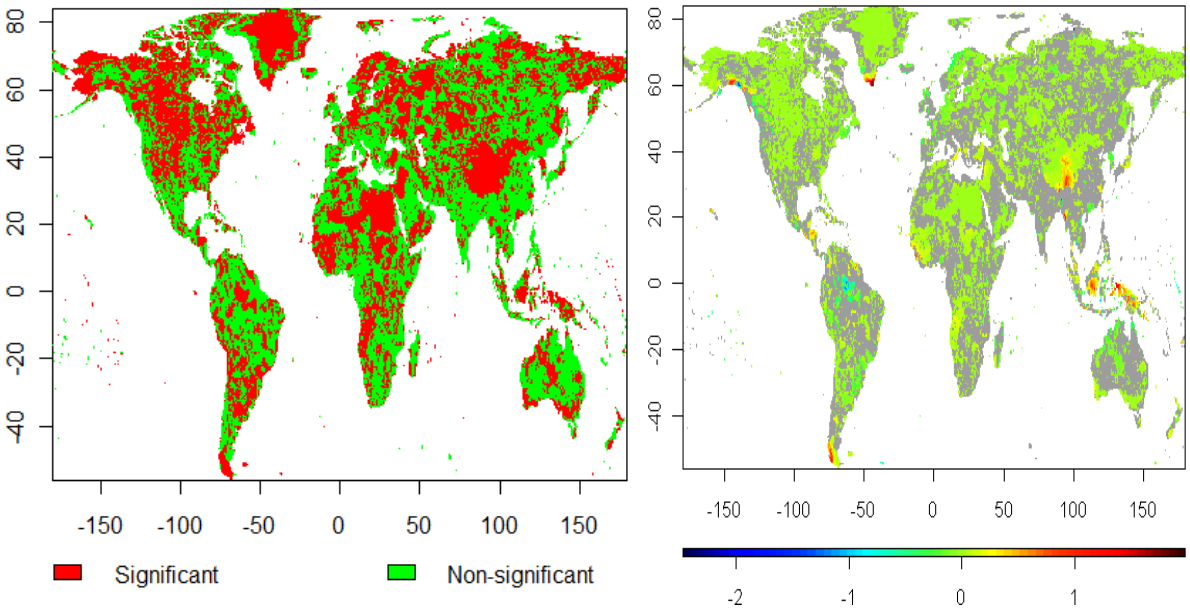


Figure 3-3. Left: Spatial location of cells that show statistically significant shift (red) and non-significant shift (blue) in mean monthly runoff. Right: The absolute magnitude of shift in mean monthly runoff

Since it is difficult to clearly see the difference in magnitude of shift due to the spatial climatic variability, the same approach (Equation 2.2) used for the daily runoff t-test is used

again to scale the magnitude of shift and the result is shown in Figure 3-4. A statistical summary of the scaled magnitude of shift for monthly mean runoff is given in Table 3-2.

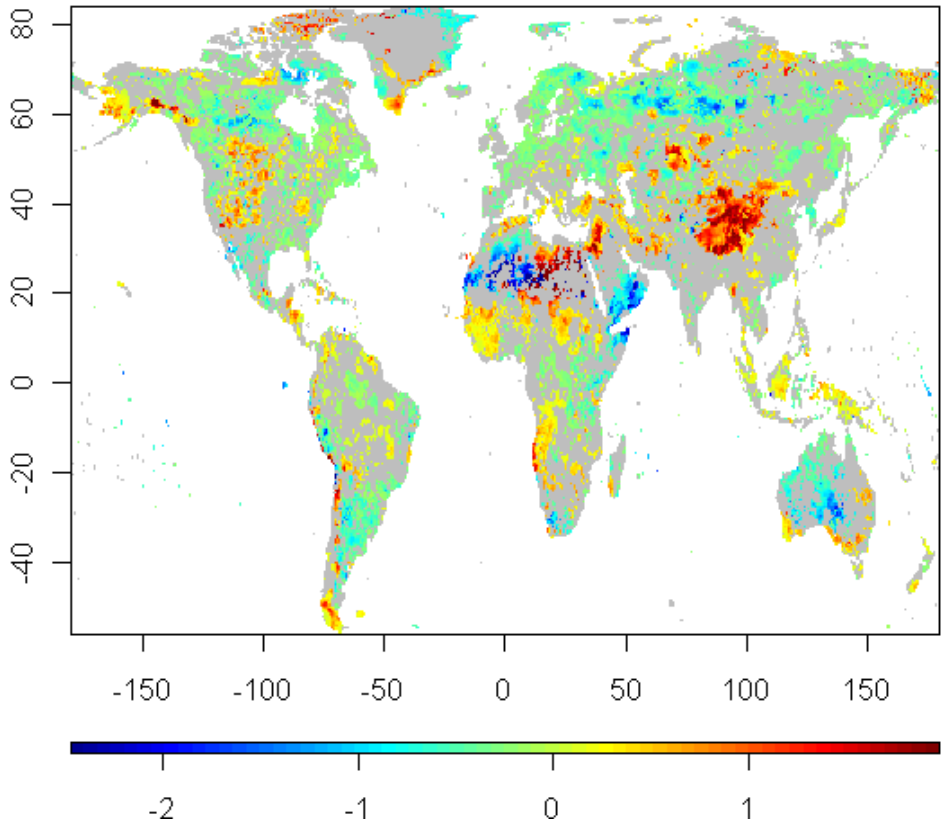


Figure 3-4. Scaled magnitude of shift of mean monthly runoff between the first and second half of the 20th century for WaterGAP model

Apart from the t-test for the monthly mean runoff values, the t-test is also performed for three lower percentiles and median of the monthly values, namely 5th, 10th, 20thand 50th percentile, as the purpose of this study is analysis of drought which is associated with low flows (where nth percentile corresponds to a value below which n% of observations fall). The magnitude of shift of each percentile is then scaled using the mean runoff value of the corresponding percentile for the whole 20th century. The spatial variability of the scaled magnitude of shift in mean between the first and the second half of the 20th century for the four monthly percentile values mentioned above is shown in Figure 3-5. From the figures it is observed that generally there is a similar trend in spatial distribution as well as the magnitude of shift for the four percentile values. The detail of the statistical summary of the scaled magnitude of shift for each percentile is given in Table 3-2. As already observed from the t-test using daily data (Table 3-1), majority of cells that have significant shift have scaled magnitude in the range of -0.5 to 0.5 and this pattern persists for the mean as well as for the four percentiles.

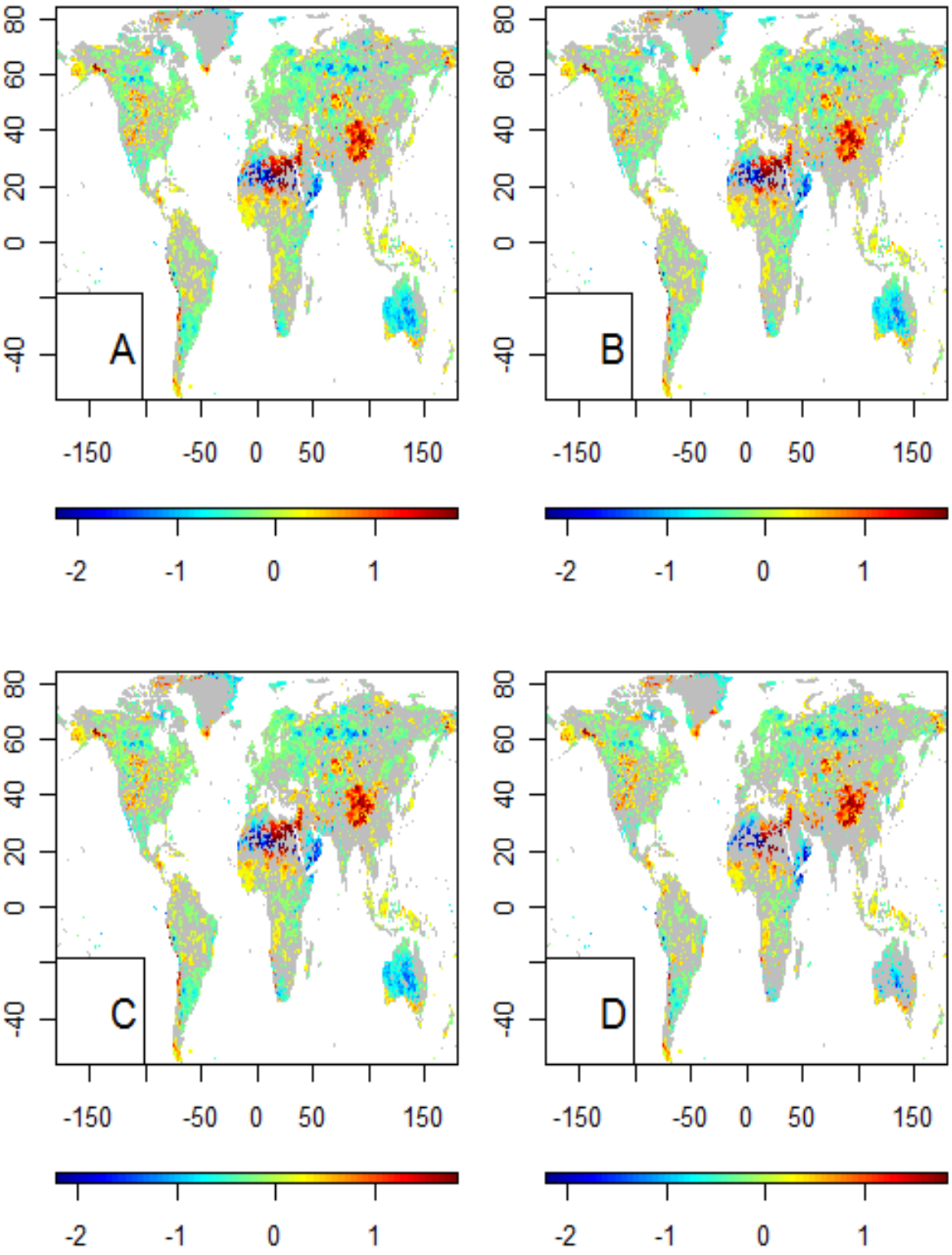


Figure 3-5. Scaled magnitude of shift for: A) 5th monthly percentile, B) 10th monthly percentiles, C) 20th monthly percentile, and D) the 50th monthly percentile

Table 3-2. Statistical summary of the relative magnitude shift for the four monthly percentile values and the mean monthly runoff

Scaled magnitude of shift	Percentage(%) of land point cells				
	Mean monthly runoff	5th monthly percentiles	10th monthly percentiles	20th monthly percentiles	50th monthly percentiles
<1	1.6	1.8	1.9	2.1	1.5
-1 - -0.5	4.4	6.0	5.9	5.7	4.4
-0.5 - -0.25	8.4	10.0	10.0	9.8	7.9
-0.25 - -0.1	9.6	12.1	11.8	11.2	9.1
-0.1 – 0.1	1.4	1.6	1.5	1.4	1.3
0.1 - 0.25	4.1	4.8	4.6	4.4	3.6
0.25 - 0.5	6.0	5.3	5.3	5.3	5.1
0.5 - 1	4.5	3.6	3.7	3.7	4.1
>1	2.2	2.4	2.4	2.4	2.1
Total	42.2	47.6	47.1	46.2	39.1

Instead of considering the complete time series in order to identify a shift in the runoff between the two different periods of the 20th century, another possibility is to perform a similar analysis for a shorter period. The benefit of performing such an analysis is that results of less influence by large scale climate shift influence both periods. For this purpose, four time windows are considered. These different time windows include five, ten, fifteen and twenty years mean monthly runoff of both periods centred around 1957/58. For instance, a time window of ten year means t t-test is performed to assess the trend shift in monthly runoff ten year prior 1957/58 and ten year after. The results of the spatial patterns of scaled magnitude of shift in mean between the two periods of the century each time window is shown in Figure 3-6. To analyse the differences among the four time windows further, statistics of ranges of the scaled magnitude of shift for each time window is given in Table 3-3.

Table 3-3. Statistical summary of the scaled magnitude shift by considering five, ten, fifteen and twenty years mean monthly runoff centred around 1957/58.

Scaled magnitude of shift	Percentage(%) of land point cells			
	5 years	10 years	15 years	20 years
<1	3.0	2.4	1.6	1.4
-1 - -0.5	5.8	5.4	4.4	3.9
-0.5 - 0.5	7.5	13.1	16.6	19.2
0.5 - 1	4.8	4.8	5.7	5.3
>1	3.6	2.7	1.9	1.8
Total	24.7	28.4	30.0	31.84

Generally it is observed that when more years of data are used in the analysis, the percentage of land surface cells that have statistically significant differences in mean increases this because dry and wet periods are clustered and hence adding or subtracting a few years of data in the time series may lead to a substantial shift in the trend (Hisdal *et al.*, 2001). Moreover, the scaled magnitude of shift in mean is higher when less years of data are considered.

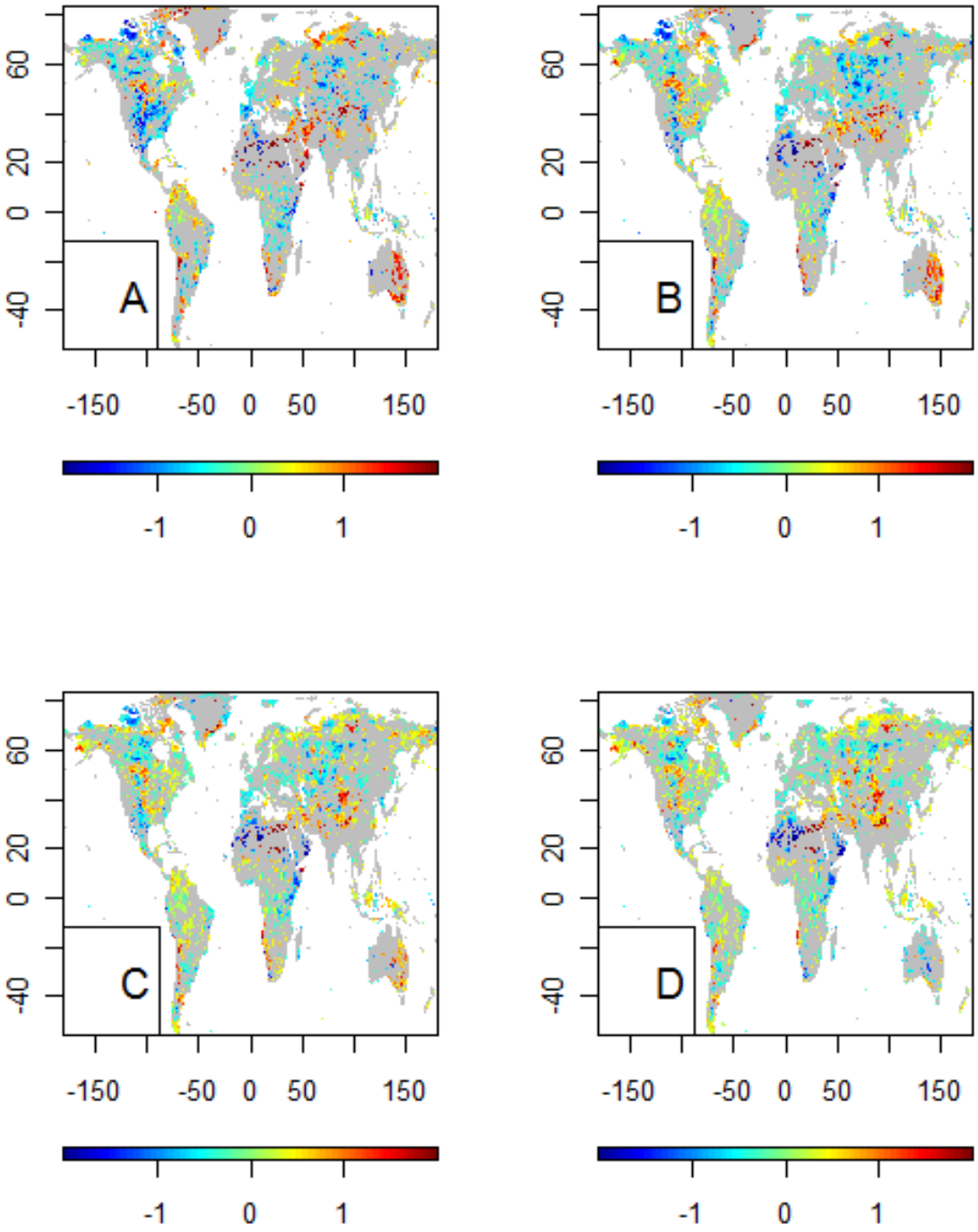


Figure 3-6. Scaled magnitude of shift in mean between the first and the second half of the 20th century considering: A) 5 years , B)10 years , C) 15 years D) 20 years of mean monthly runoff that are centred around (1957/58)

In order to show the differences in time series runoff values for significant and non-significant cells, the time series plots of randomly selected significant and non-significant cells is made for Africa (Figure 3-7), North America (Annex I), and South America (Annex II**Error! Reference source not found.**). It is observed that for those cells that have scaled magnitude of shift of greater than 1 and less than 1, the shift in mean is quite remarkable in mean monthly runoff plots.

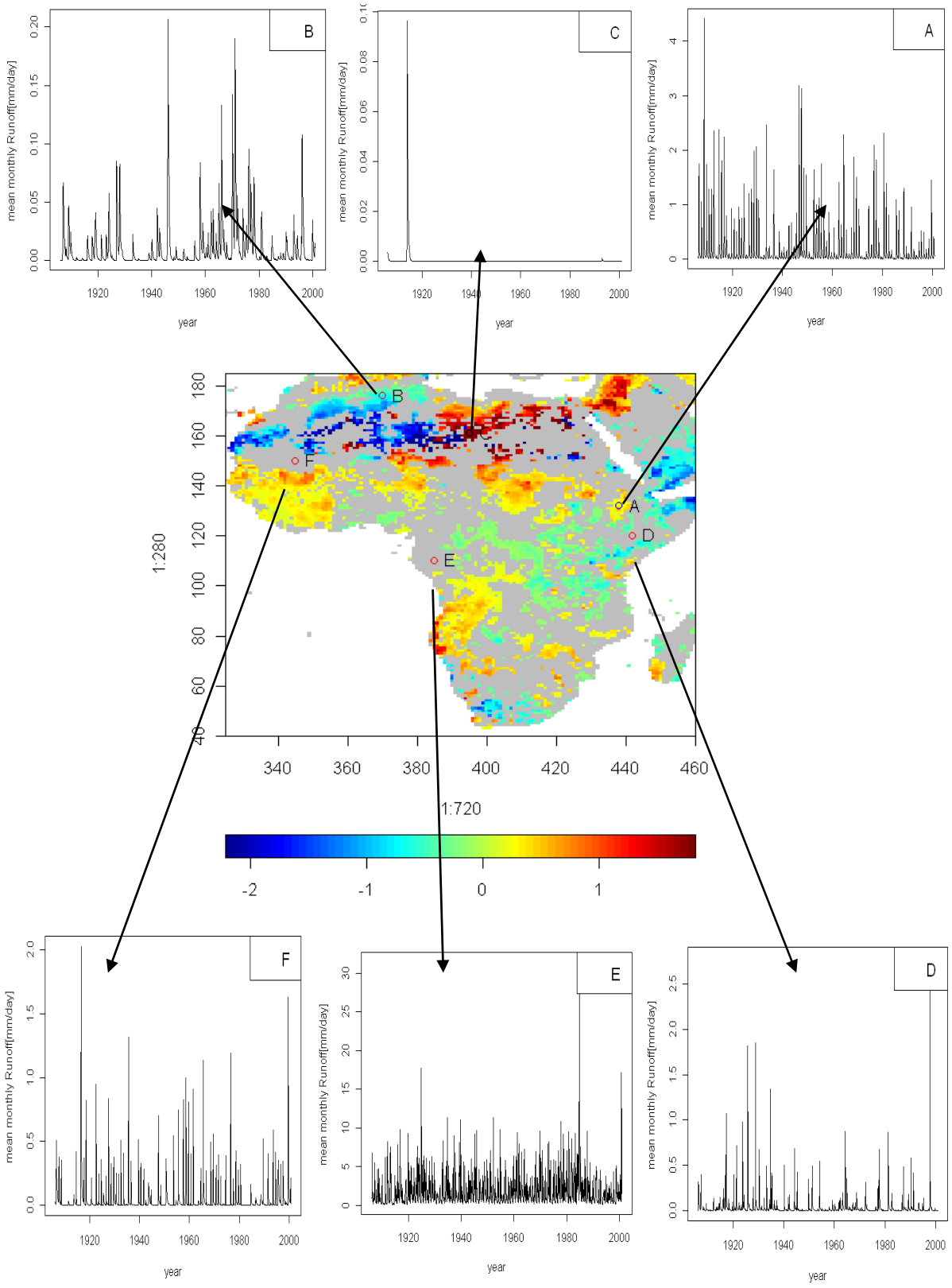


Figure 3-7. Time series plots of three cells that have statstically significant shift in mean between the two period of the century (A,B,C) and three cells which have non-significant mean shift (D,E, F) from Africa

3.2 Common procedures for other models

Based on the comprehensive analysis of the trend shifts that are carried out using WaterGAP model, it is observed that monthly t-test values give an overview of trends by smoothing the fluctuations of daily runoff and improving the computational efficiency. From the analysis of the t-test for the 5th, 10th, 20th and 50th percentiles, it is observed that for all the percentiles similar spatial patterns and scaled magnitudes of shift persist as that of the t-test for the mean monthly runoff. The t-test for using different time windows demonstrate that the percentage of land surfaces that show statistically significant trend shifts increase as more number of years of data is used with the maximum being when the whole period is considered. Therefore, the mean monthly t-test using the whole period of both periods (i.e. 1906-1957 and 1958-2000) is chosen to further investigate trend shifts for the other models

3.3 Other models

Apart from WaterGAP, two models (GWAVA and HTESSEL) are tested for trend shifts using the monthly mean t-test. The results indicate that nearly half of the land points of the globe show statistically significant trend shifts; though there are differences in locations of cells (see Figure 3-8).

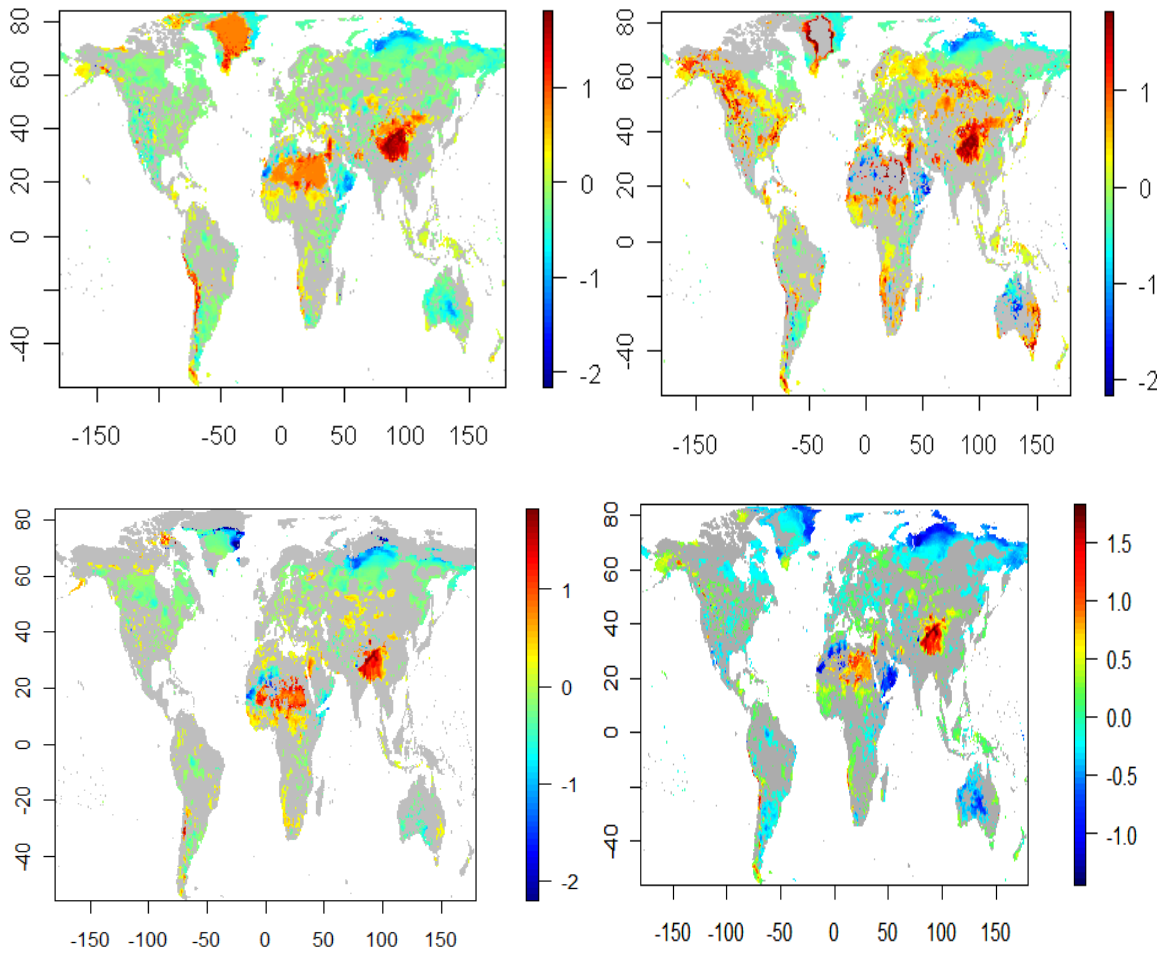


Figure 3-8. Scaled magnitude of shift of mean monthly runoff between the first and second part of the 20th century: Top left: GWAVA, Top right: HTESSEL, Bottom left: Orchidee, and Bottom right: Multi-model ensemble mean

The higher magnitude of shifts (positive or negative)) are observed in the Sahara and Green land for most of the models (Figure 3-8). For the Sahara, this higher magnitude of shift is associated with one or two rainfall events that occur once in a while (e.g. Figure 3-7(c)). In addition due to their extreme hydrological characteristic, the models might not well capture the hydrology of these regions. Therefore, it is decided to exclude these two regions from our drought analysis.

4 Drought Characterization: 1st and 2nd part of the 20th century

In this chapter the influence of the different time windows on determination of thresholds for drought occurrence is presented using WaterGAP for some contrasting sample cells from different continents. Moreover, the overall influence of these thresholds on the occurrence of drought at global scale is assessed. T he global as well as continental occurrence of drought is analyzed using the different large-scale models and the consistence and differences among them are described.

4.1 Influence of time windows on threshold determination

As discussed in Chapter 3 of this report, nearly half of the land point cells show statistically significant trend shifts in simulated runoff between the two periods of the century for the four models as well as for the multi-model ensemble mean. Therefore, three different time windows have been tested to define the threshold for drought occurrence using WaterGAP. The first part of this section illustrates the difference between the three thresholds derived from three different time windows and the second part presents the influence of these three thresholds on characterization of global drought occurrence for WaterGAP model as an example.

4.1.1 Sample cells for illustrating the three different thresholds

Three different time windows are used to determine the threshold for drought: 1) monthly variable thresholds are determined based on the time series of runoff for the whole 20th century (i.e. 1906-2000), 2) monthly variable thresholds are determined separately for the two periods of the century (i.e. 1906-1957 and 1958-2000), and 3) seven monthly variable thresholds are determined using seven 40 years time windows sliding every ten year (i.e. 1906-1940, 1911-1950,..., 1961-2000). A combination of variable monthly threshold of 20 percentile and a consecutive dry period fixed threshold of 20 percentile discussed in Section 2.4 is used for drought characterization. To illustrate the difference in thresholds for different time windows, three sample cells that behave differently are selected. One cell that has a negative trend shift (Figure 4-1) between the two periods (wetting condition), another cell that has a positive trend shift (drying condition) (Figure 4-2), and a third cell with no statistically significant trend shift (Figure 4-3) are illustrated in the following section.

Figure 4-1(top) shows the time series values of simulated runoff for a pixel located in North America (location of the pixel in the continent is shown in Annex-I , point C). Figure 4-1 (middle) shows difference in monthly thresholds when derived from the whole period and separately for the two periods. It can be seen that the thresholds derived from the first part are lower than those derived from the whole period whereas that of the second part is higher. This is consistent with the trend shifts identified by the statistical tests. Figure 4-1(bottom) shows differences in threshold when derived from the whole period and those derived using sliding time windows. The magnitude of the difference among the thresholds varies throughout the year with the maximum difference being in higher runoff period of the year (July – November). Using the threshold from sliding time windows smooth the drastic shift in threshold due to the trend shift between the two periods and it captures the effect of inter-decadal climatic variability on the threshold.

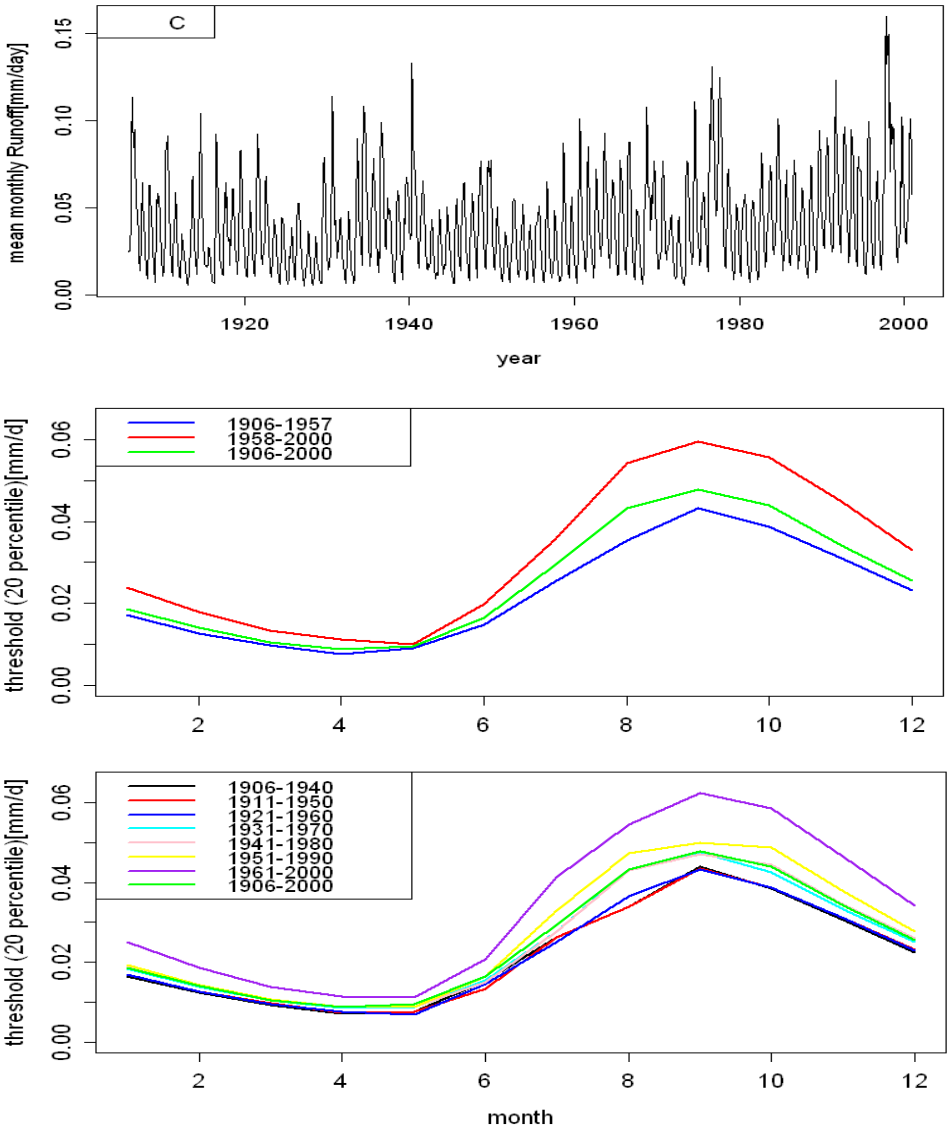


Figure 4-1. Monthly thresholds for different time windows for a cell showing a negative trend shift (wetting condition towards the second part of the 20th century. Top: time series monthly mean simulated runoff (with WaterGAP model) of a cell from North America. Middle: monthly threshold derived from whole period and the two periods. Bottom: monthly threshold derived from whole period and seven sliding time windows

Figure 4-2 presents the differences in thresholds for a cell from South America (location of the pixel in the continent is shown in Annex-II , point A) that shows a positive trend shift in simulated runoff. For such cells, using the threshold derived from whole period as well as that of the first period of the century will lead to more droughts since the trend shift is quite large. On the other hand, using thresholds derived from the second half will lead to fewer droughts in the second half. Thresholds derived from sliding windows will smooth these differences in thresholds and will give more realistic results.

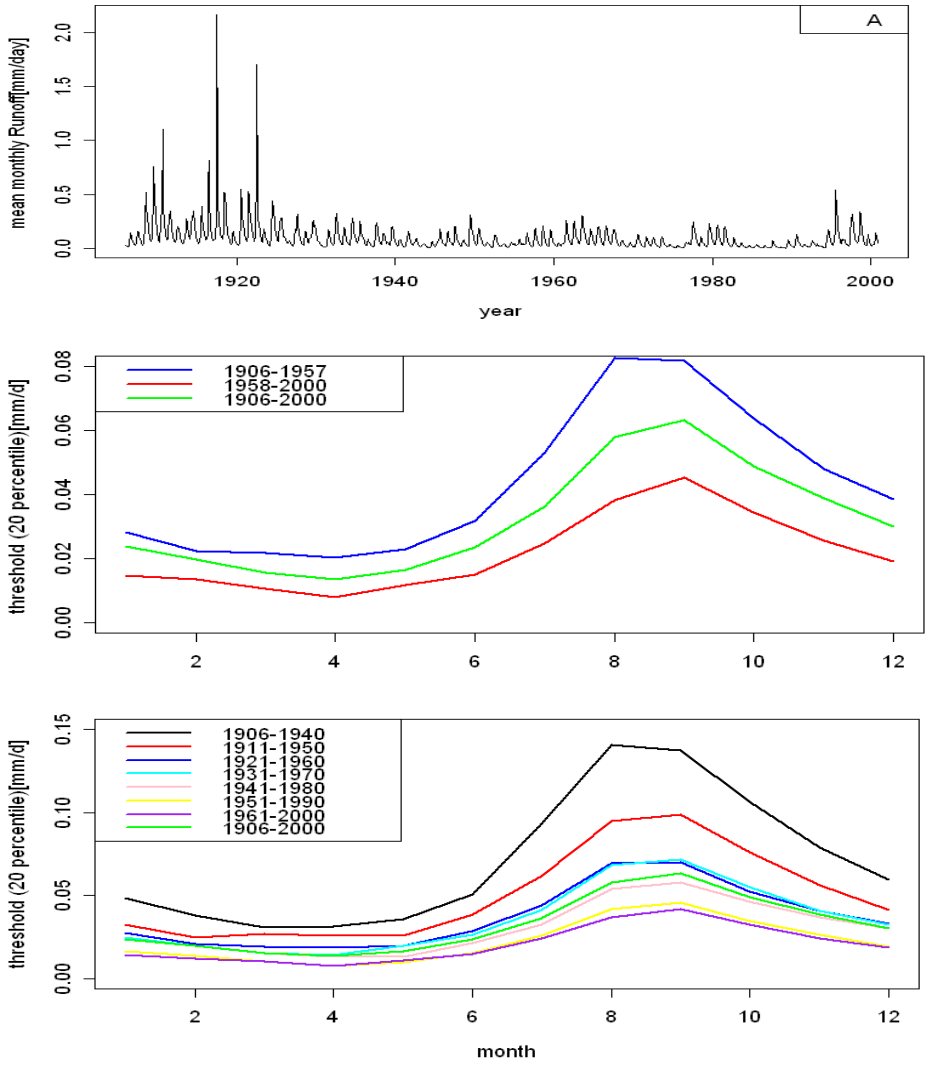


Figure 4-2. Monthly thresholds for different time windows for a cell showing a positive trend shift (drying condition towards the second part of the 20th century. Top: time series monthly mean simulated runoff (with WaterGAP model) of a cell from South America. Middle: monthly threshold derived from whole period and the two periods. Bottom: monthly threshold derived from whole period and seven sliding time windows.

Figure 4-3 shows the differences in thresholds for a cell from Africa (location of the pixel in the continent is shown in Figure 3-7, point D) that does not show statistically significant trend shift among the two periods of the century. It can be seen from the figure that for such cells the difference in thresholds from the three time windows is relatively small compared to those cells that show statistically significant trend shift. However, differences in thresholds for the sliding time windows are larger for high flow periods of the year which could be attributed to inter-decadal climatic variability.

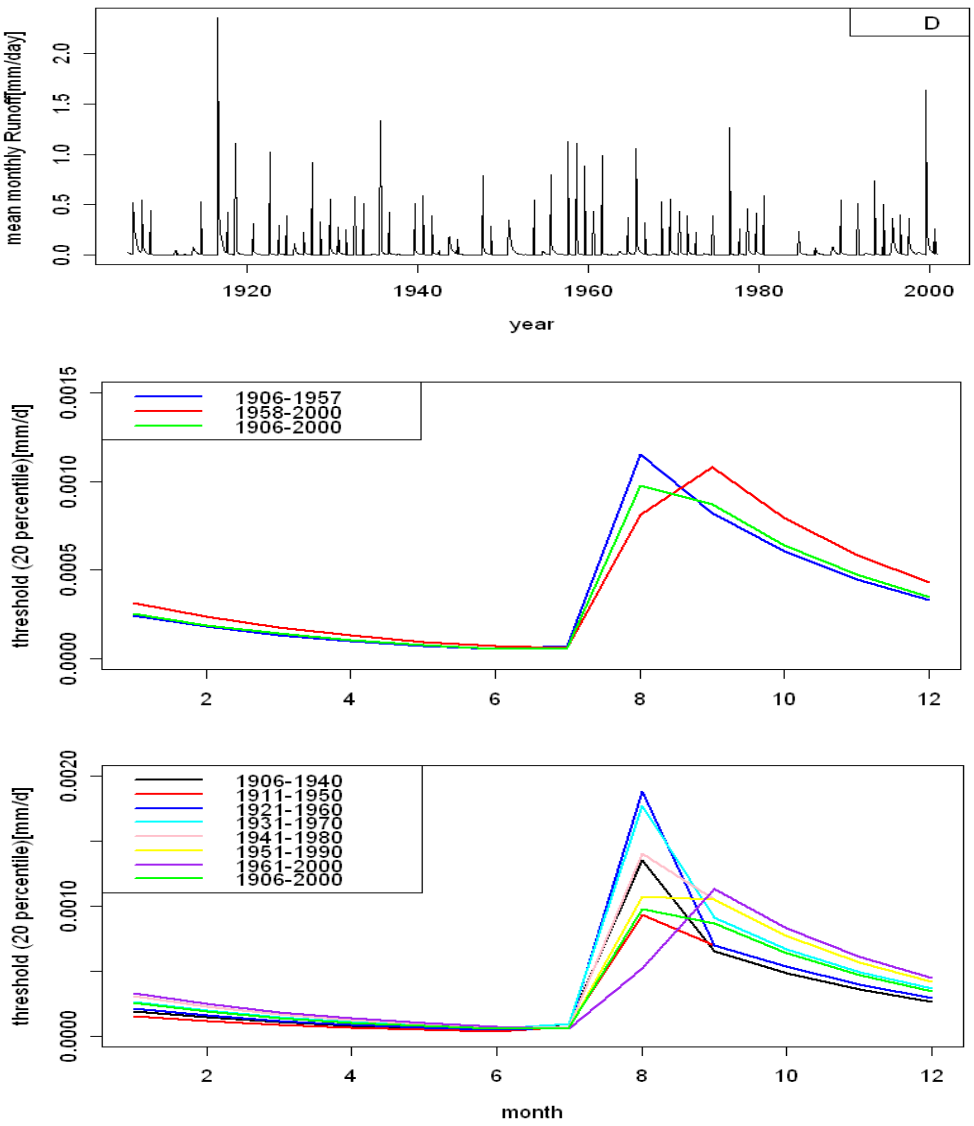


Figure 4-3. *Monthly thresholds for different time windows for a cell showing no statistically significant trend shift between the two periods of the 20th century. Top: time series monthly mean simulated runoff (with WaterGAP model) of a cell from Africa. Middle: monthly threshold derived from whole period and the two periods. Bottom: monthly threshold derived from whole period and seven sliding time windows*

4.1.2 Influence of the three thresholds on global drought characterization

The effect of using the three different thresholds defined based on three time windows on is drought characterization is shown in Figure 4-4, where the temporal variation of percentage of globe in hydrological drought is given for the WaterGAP model. From Figure 4-4, it can be observed that the three thresholds determined using three different time windows have resulted in different percentages of area in drought for the same threshold value of 20 percentile.

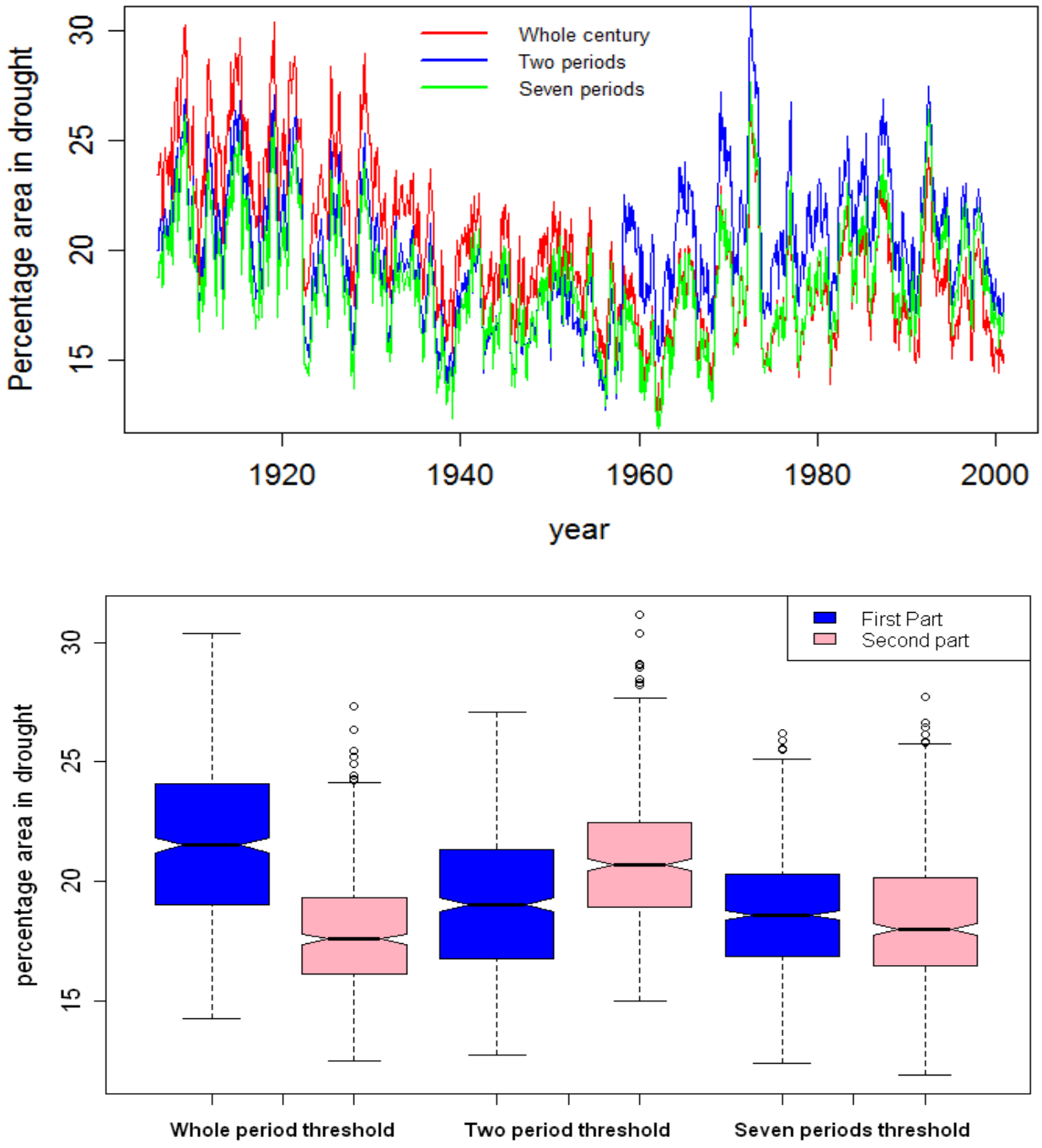


Figure 4-4. Percentage area of the globe in drought (based on the WaterGAP model). Top: monthly time series of percent area of the globe in hydrological drought using three different thresholds. Bottom: Box plots of the time series distribution of percentage of area in drought between the two periods using the three threshold determination approaches (whole period, two periods and seven sliding periods.

The analysis of spatial extent of drought determined using thresholds based on the whole simulation period resulted in the highest percentage of globe in drought for the first part of the century and the lowest percentage of globe in drought for the second part of the century (see Figure 4-4, bottom). This could be attributed to the long term climatic variability between the

two periods of the century. In addition, the differences in the approaches followed to derive the WFD for the two periods also plays a role since the availability of observation for bias correction is limited for the first period of the century.

To investigate the effect of the trend shift between the two periods of the century, monthly thresholds are determined separately for the two periods of the century and percentage of the globe in drought is computed accordingly as shown in Figure 4-4. This approach has resulted in a relatively lower percentage of area in drought for the first period of the century, where as more on the second part of the century as compared to using the threshold from the whole century (compare the first four boxes in the bottom of Figure 4-4). To capture this contradicting effects of these thresholds, another thresholds that are defined from seven sliding time windows of each 40 years overlapping every ten year. This method resulted in an intermediate result since it smooth the effect of trend shifts and inter-decadal climatic variability. Therefore, the seven sliding time windows of 40 years each sliding every ten year are used for determining thresholds for multi model analysis of drought in the following Sections. .

4.2 Global drought analysis

In this section, the spatio-temporal distribution of global and continental drought occurrence will be discussed in detail. The results focus on WaterGAP but the outcomes of the other models including the multi-model ensemble mean are also provided. The discrepancy and consistence of drought occurrence among the different models are discussed. The second part of this section describes the possible trends in drought for the two periods of the 20th century at global as well as continental scale.

4.2.1 Multi-model Analysis

Based on the variable threshold determined for each cell in the grid, the absolute values of runoff of all cells are replaced with the corresponding percentiles. Using monthly variable thresholds (20 percentile), those cells that fall below the threshold are identified as cells in drought for each month of the simulation period. To assess the overall spatial occurrence of drought, percentage of the globe in drought for each month are explored using WaterGAP model as shown in the colour coded map in Figure 4-5. Similar analysis is carried out for all models including the multi-model ensemble mean (see Figure 4-7).

Figure 4-5 shows the yearly and inter-monthly variation of the percentage of the globe in drought for the whole simulation period. The maximum percentage of area in drought (using WaterGAP) is found to be 26 % and the minimum 12%. Corzo Perez et al. (2011) performed the same analysis with the WaterGAP model for the period 1963-2000 using daily subsurface runoff, and obtained a maximum of 30% and a minimum of 12% of the globe to be in drought. The difference in the maximum area percentage can be attributed to the smoothing of the daily runoff values when averaged to monthly time scale, the difference in the period for analysis, and the variable used to define drought. For our analysis, the maximum percentages of globe in drought appear in the year 1972/73 and 1992 whereas for analysis made by Corzo

Perez et al. (2011), the higher percentages appear in the year 1982/83 and 1992/93. The years mentioned above were identified to be strong El Nino years (Corzo Perez et al., 2011,

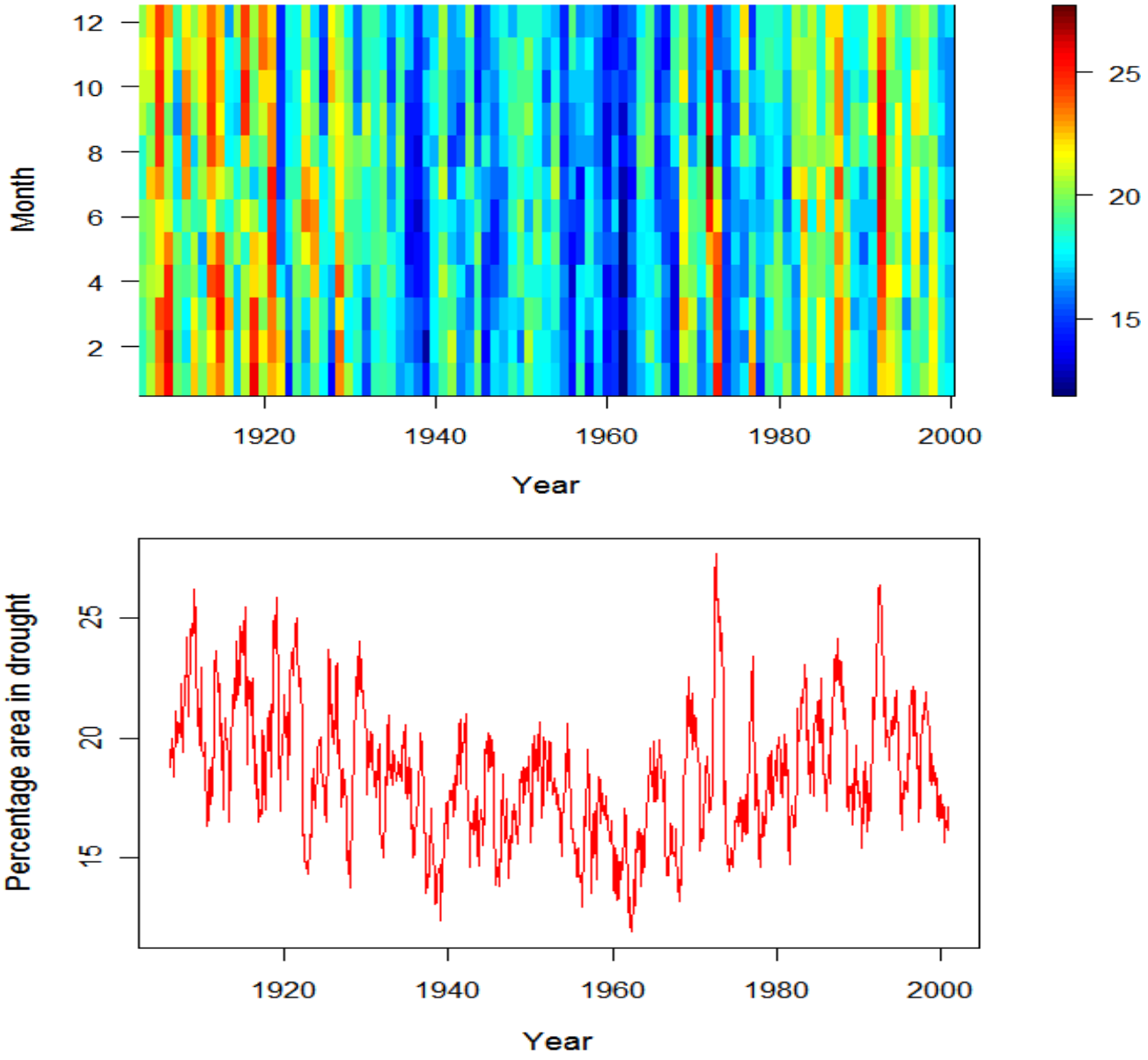


Figure 4-5. Percentage area of the globe in drought. Top: Percentage of globe in hydrological drought derived from total runoff simulated with WaterGAP global hydrological model for the period 1906-2000. Bottom: Time series plots of percentage of globe in drought

To give an overview of the spatial distribution of global drought occurrence and its variation among models, the spatial locations of cells in drought are provided in Figure 4-6 for the month of June, 1992. The 1992 high percentage has also been observed in the other models though monthly variations are observed among the models (see Figure 4-7).

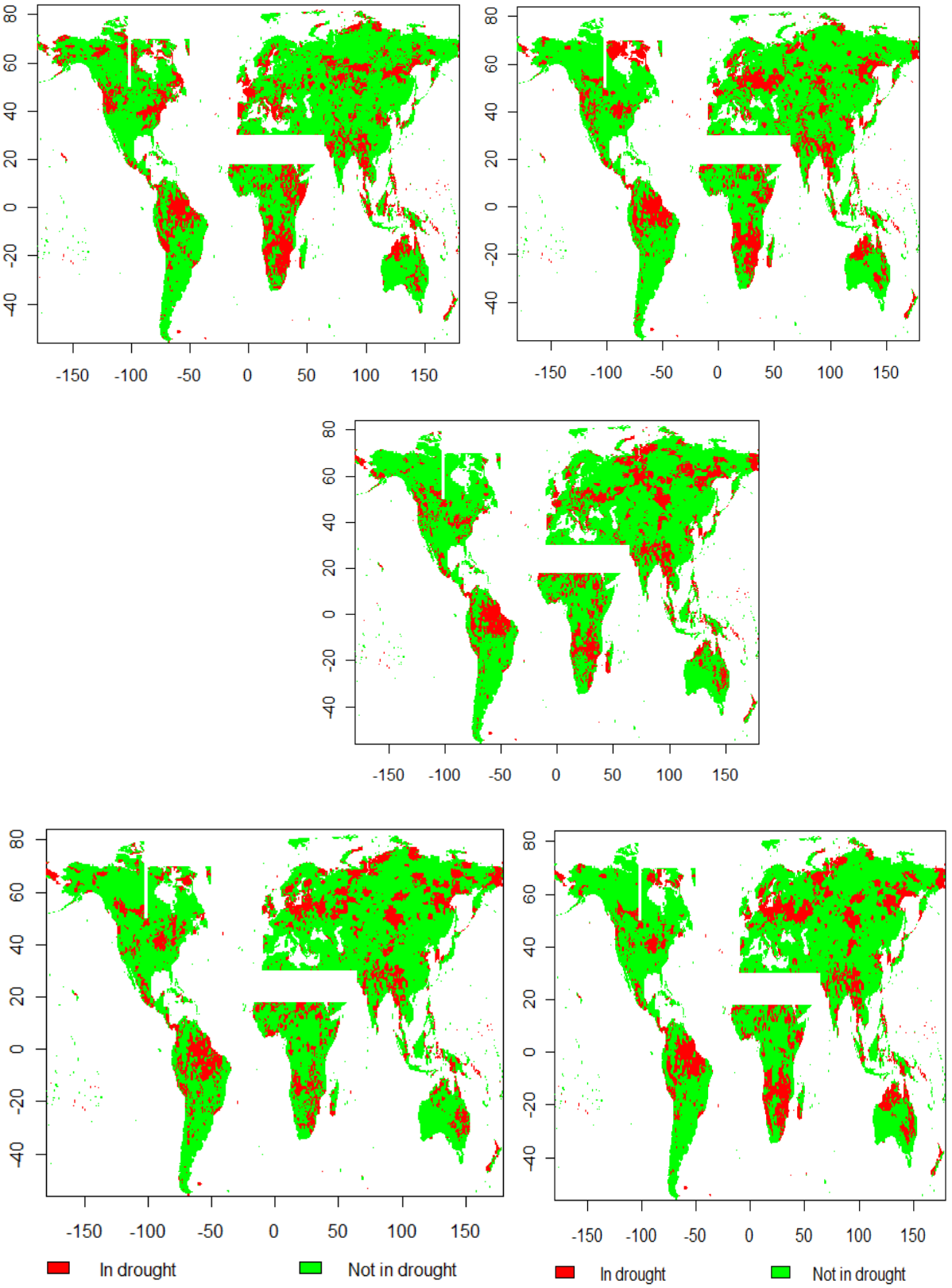


Figure 4-6. Example of spatial distribution of drought for an ElNino year (1992, June): WaterGAP (top-left), GWAVA (top-right), HTESSEL (centre), Orchidee (Bottom-left), and Multi-model ensemble mean (bottom-right).

For a particular drought event in June 1992, it can be observed from Figure 4-6 that there exist spatial patterns among all the models for South America, south and eastern Africa, North and eastern Australia and also to some extent in Asia. Although the 1972/73 drought is also weakly observed in the other models, it is over estimated by WaterGAP as compared with the other models (Figure 4-7). Highest percentages in drought are usually preceded by dry years for all models except WaterGAP which indicate the persistency among years. (Corzo Perez et al., 2011) have identified that historical high percentage persist in the months between April to May, however, in our analysis with WaterGAP, no clear monthly trend is observed in high as well as low percentages . Among the other models used in this study, monthly consistency of higher percentages of area in drought has been clearly observed for the HTESSEL model, and the months identified coincide with mentioned by (Corzo Perez et al., 2011)(Figure 4-7). However, the results obtained from HTESSEL are rather uncertain since the high percentages are persistently observed throughout the whole simulation period. When the same analysis is performed at continent scale with HTESSEL, these monthly persistent high percentages are observed in North America and Asia as well (see Annex III and IV). This persistent high percentage for April is mainly because of the winter period in the northern hemisphere (Asia and North America). HTESSEL unlike the other model assign runoff of zero during the winter period hence lead to persistent high percentage of area in drought for this particular month for the whole simulation period.

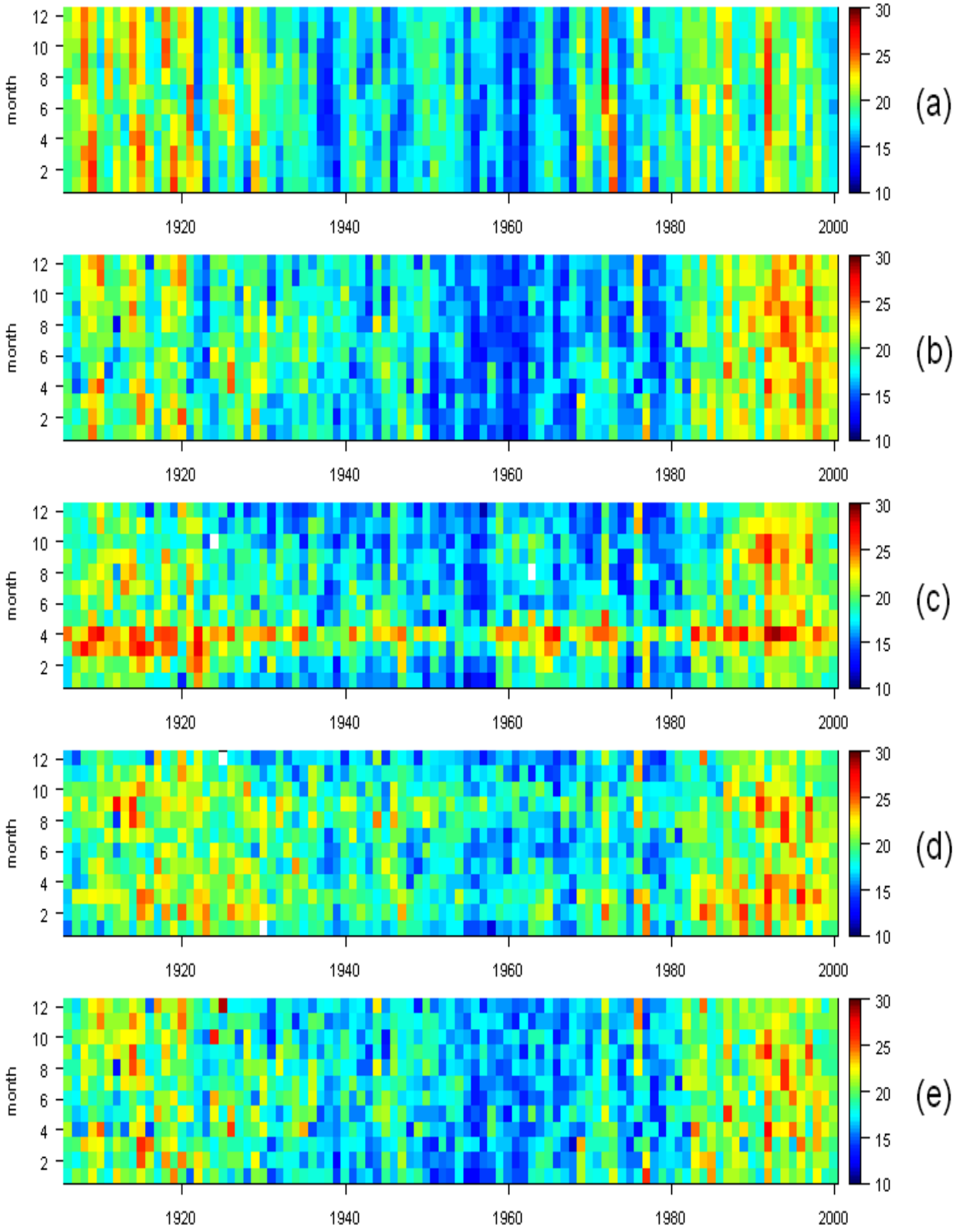


Figure 4-7. Percentage of area of the globe in drought based on the analysis of: (A) WaterGAP, (B) GWAVA, (C) HTESSEL, (D) Orchidee, and (E) model ensemble mean.

Percentage of the globe in drought provides the overall spatial extent of drought and helps to identify critical dry periods of the century; however it does not show spatial distribution of drought events since it is aggregated globally. To get better understanding of the spatial occurrence of drought, the temporal evolution of drought occurrence is also assessed at continent scale. The result for Africa is given in Figure 4-8, whereas for the other continents they are provided in AnnexesIII-VIII.

Generally it is observed from **Error! Reference source not found.** that higher percentages are persistently observed at the beginning of the simulation period (1906-1920) for whole globe as well as for the continents. The reason for this can be attributed to: 1) the threshold method used for the analysis, and 2) limited observations for bias correction of the meteorological forcing data (Weedon *et al*., 2011). The thresholds from sliding window are usually an approximation of two or more thresholds which tend to smooth the fluctuations of thresholds due to trend shift as well as climatic variability hence provide realistic results (Section 4.1). However, for the period 1906-1920, one threshold is used because of the inherent nature of the sliding window with no or less overlapping periods at the beginning of the period hence influence the characterization of drought. In addition, limited availability of observations during this period leads to uncertainty in the meteorological forcing data which consequently affects the reliability of the simulations and subsequent drought characterizations.

One approach to validate the performance of the models at continental scale is looking at historic drought events (eg. Sheffield and Wood, 2008; Corzo Perez *et al*., 2011). For example for North America the 1930s and 1950s drought were identified as the most severe drought (Andreadis *et al*., 2005; Dai, 2011). According to our analysis, in the beginning of the 1930s higher percentages that are preceded by dry years are observed in all the models though there exist differences in magnitude among the models (see Annex IV). The 1950s droughts in North America are also observed in our analysis though the percentages are not as high as that of the beginning of the 1930s and persist for longer duration. For Africa, persistent high percentages are observed from 1982 to 2000 among all models, though the percentages get relatively lower through time for WaterGAP (Figure 4-8). Sheffield and Wood (2008) and Corzo Perez *et al*. (2011) also mention an increase in spatial extent of drought in Africa during the last quarter of the century. The 1983/84 historical drought in eastern Africa is clearly observed with the similar pattern in WaterGAP, GWAVA and the model ensemble mean whereas for HTESSEL and Orchidee substantial inter-monthly variation is observed (Figure 4-8). In addition the 1972/73 drought, which is caused by the ENSO phenomena, is clearly observed for Africa in all the models though the magnitude is rather weak compared to that of 1982 onwards. The 1976 European drought (Tallaksen and Van Lanen, 2004; Corzo Perez *et al*., 2011) is clearly revealed in GWAVA and HTESSEL and also to some extent by the other models (see Annex V).

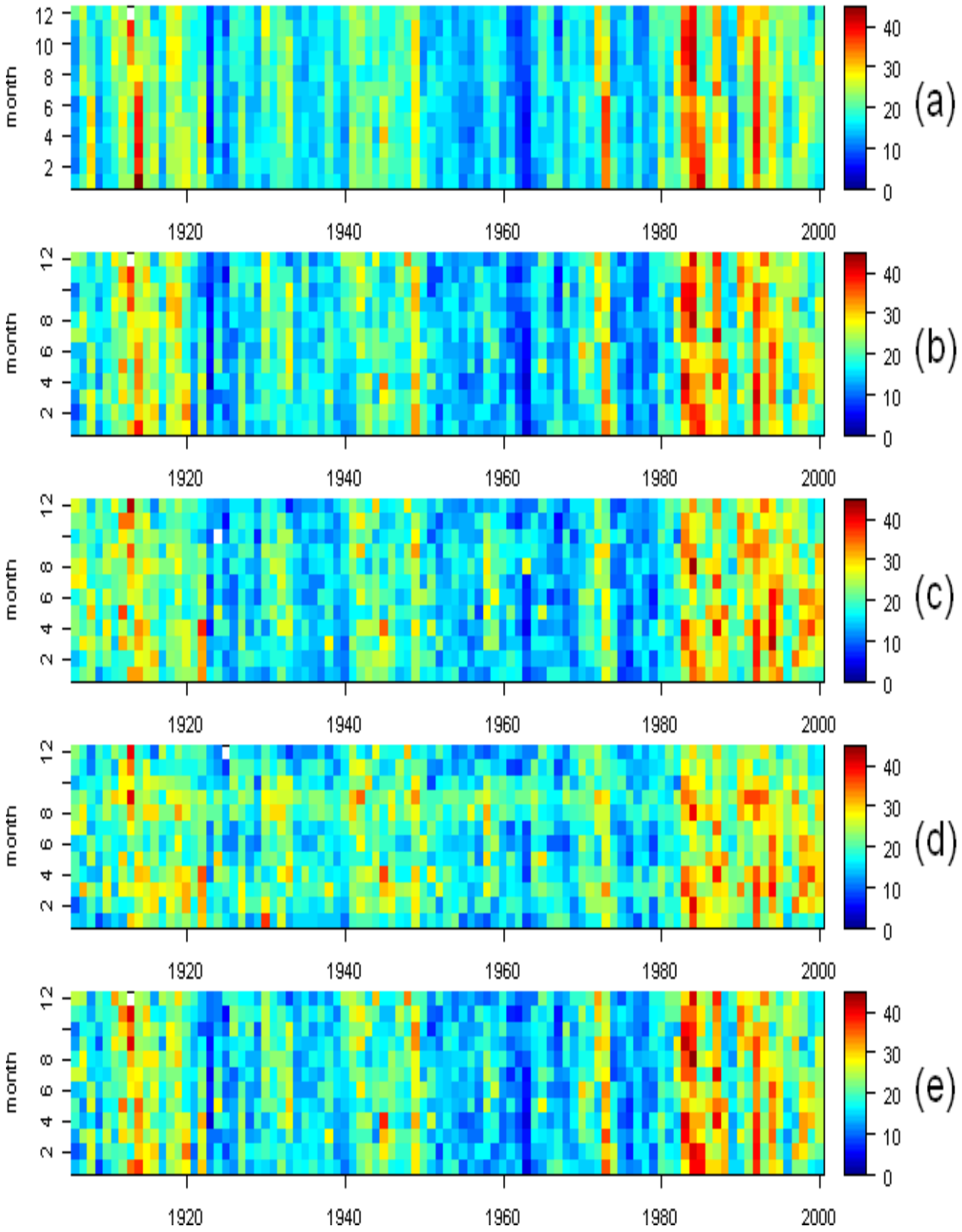


Figure 4-8. Percentage area of Africa in drought based on the analysis of: (A) WaterGAP, (B) GWAVA, (C) HTESSEL, (D) Orchidee and (E) model ensemble mean.

4.2.2. Trends in spatial extent of Ggobal and regional drought

4.2.1.1 Trends in Global Drought

One of the objectives of this study is to assess differences in the occurrence of hydrological drought between the first (1906-1957) and second part (1958-2000) of the century because of the different approaches used to drive the WFD for the two periods. For this purpose the Mann-Kendall's non- parametric trend test (Mann, 1945; Kendall, 1970) is performed for the whole century and separately for the two periods of the century assuming individual drought events to be independent (Sheffield and Wood, 2008). Table 4-1 provides the trends in spatial extent of global drought expressed as percentage area of the globe in drought: 1) for the first part, 2) for the second part, and 3) for the whole period of the century for the different models and the multi-model ensemble mean. The analysis based on WaterGAP model indicated that there is an overall decreasing trend in spatial drought extent of -0.0886 %/yr for the whole century which is statistically significant at 0.05 level. On the other hand, the analysis based on Orchidee indicates a statistically significant an overall increasing trend of 0.0313 %/yr for the whole century. The other models and the multi-model ensemble mean generally also show an increasing trend for the analysis for the whole century, although the trends are not statistically significant at 0.05 level. Comparing the two periods of the century, statistically significant consistent decreasing trends are observed for the first part of the century, whereas increasing trends are observed for the second part of the century among all models. For WaterGAP and the ensemble mean relatively higher decreasing trends are obtained for the first half of the century, whereas for GWAVA and the multi-model ensemble mean higher increasing trends are found for the second part of the century.

Table 4-1. Trends in spatail extent of global drought for the four different models and the multi-model ensemble mean using Mann-Kendall test. Trend values in bold are those that are statstically significant at 0.05 level.

Models	Trends in drought spatial extent (% / year)		
	1906-1957	1958-2000	1906-2000
WaterGAP	**-0.373**	**0.308**	**-0.0886**
GWAVA	**-0.346**	**0.577**	0.0172
HTESSEL	**-0.365**	**0.267**	-0.0006
Orchidee	**-0.240**	**0.461**	**0.0313**
Ensemble model mean	**-0.36**	**0.461**	-0.0106

The shifts in the spatial extent of drought between the two periods of the century for the different models and the multi-model ensemble mean have been further investigated using parallel box plots of the percentage of the globe in drought obtained from each model for the two periods of the century (Figure 4-9). It can be observed from Figure 4-9 that all the models except WaterGAP show a shift in median, quartiles and lower and upper bounds. Inter-comparison of the spatial extent of drought using the four different models and the multi-model ensemble mean shows that GWAVA estimate higher percentage of drought for the second part the 20th century, whereas WaterGAP estimate the lowest.

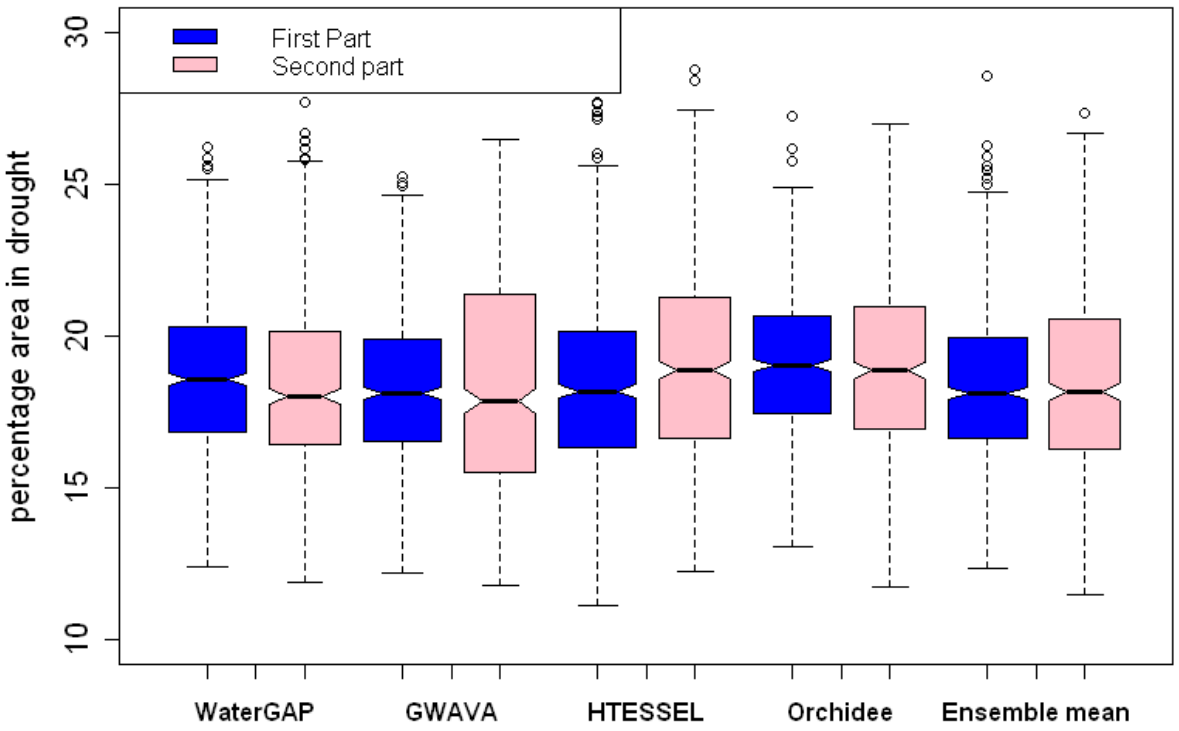


Figure 4-9. Distributions of the spatial extent of global drought for the first and the second part of the century as obtained from the different models and the multi-model ensemble mean.

Comparing the distributions of spatial extent of drought for the two periods of the century (Figure 4-9), no substantial difference is observed for WaterGAP (the inter-quartile range of the two periods is almost the same). However, for the other models the inter-quartile range shows a shift between the two periods is remarkable, with the highest shift being observed for GWAVA which also shows the highest increasing trend of the percentage of the globe in drought for the second part of the century(according to Mann-Kendall test, Table 4-1).

4.2.1.2 Trends in continental drought

The analysis adopted at the global scale is also performed at the continent scale. Table 4.2 gives the summary of the trends in spatial extent of drought for the two periods of the century for the seven continents based on the analysis of the different models. Asia consistently shows statistically significant decreasing trends in spatial extent of drought for the first part of the century for all models with trends ranging between -0.247 and -0.335 %/yr. For the

second part of the century, statistically significant increasing trends ranging from 0.285 to 0.637 %/yr are observed. North America shows a statistically significant decreasing trend of spatial extent of drought in the first half of the century for all the models. For the second part of the century, GWAVA, Orchidee and ensemble mean show statistically significant increasing trend in spatial extent of drought, whereas HTESSEL shows a decreasing trend and WaterGAP shows no statistically significant trend in spatial extent of drought. For Europe no clear statistically significant trend is observed consistently among the models for the first part the century, whereas an increasing trend is observed for the second part of the century. Africa shows an increasing trend in the spatial extent of drought for the second half of the century with all models except WaterGAP where no statistically significant trend is observed. For Australia, all models except WaterGAP show a statistically significant decreasing trend in spatial extent of drought for the second part of the century.

Table 4-2. Trends in spatail extent of drought in the seven continents for the four different models and the multi-model ensemble mean using Mann-Kendall test. Trend values in bold are those that are statstically significant at 0.05 level.

Continent	WaterGAP		GWAVA		HTESSEL		Orchidee		Ensemble mean	
	1906-1957	1958-2000	1906-1957	1958-2000	1906-1957	1958-2000	1906-1957	1958-2000	1906-1957	1958-2000
Asia	**-0.335**	**0.285**	**-0.323**	**0.637**	**-0.316**	**0.348**	**-0.247**	**0.426**	**-0.326**	**0.443**
North America	**-0.155**	-0.069	**-0.081**	**0.137**	**-0.184**	**-0.077**	**-0.168**	**0.155**	**-0.105**	**0.086**
Europe	**-0.182**	**0.34**	-0.007	**0.206**	**-0.054**	**0.161**	0.015	**0.162**	-0.031	**0.15**
Africa	-0.042	0.049	**-0.159**	**0.326**	**-0.239**	**0.39**	**-0.096**	**0.337**	**-0.155**	**0.335**
South America	-0.162	**-0.008**	-0.17	**0.025**	-0.156	-0.136	-0.074	**-0.02**	-0.113	**-0.015**
Oceania	**-0.044**	0.119	**-0.005**	**0.044**	-0.056	**0.0225**	**0.029**	**0.044**	**0.007**	0.059
Australia	**0.111**	**0.007**	**0.024**	-0.126	0.054	-0.128	0.073	-0.145	0.03	-0.107

Similar box plots as for the globe (Figure 4-9) are also plotted to compare the distribution of the spatial extent of drought between the two periods of the century for the seven continents separately (Annex IX). WaterGAP is found to be the most out liar model in this aspect. For example for Africa increase in droughts for the second part of the century are clearly observed among all models. However, WaterGAP shows a decrease in the inter-quartile range for Africa (Figure 4-10 (top)). In addition, most of the models do not show remarkable difference in spatial extent of drought between the two periods of the century for Europe, whereas WaterGAP shows a clear increase in the inter quartile range of percentage of area in drought for Europe (Figure 4-10 (bottom))..

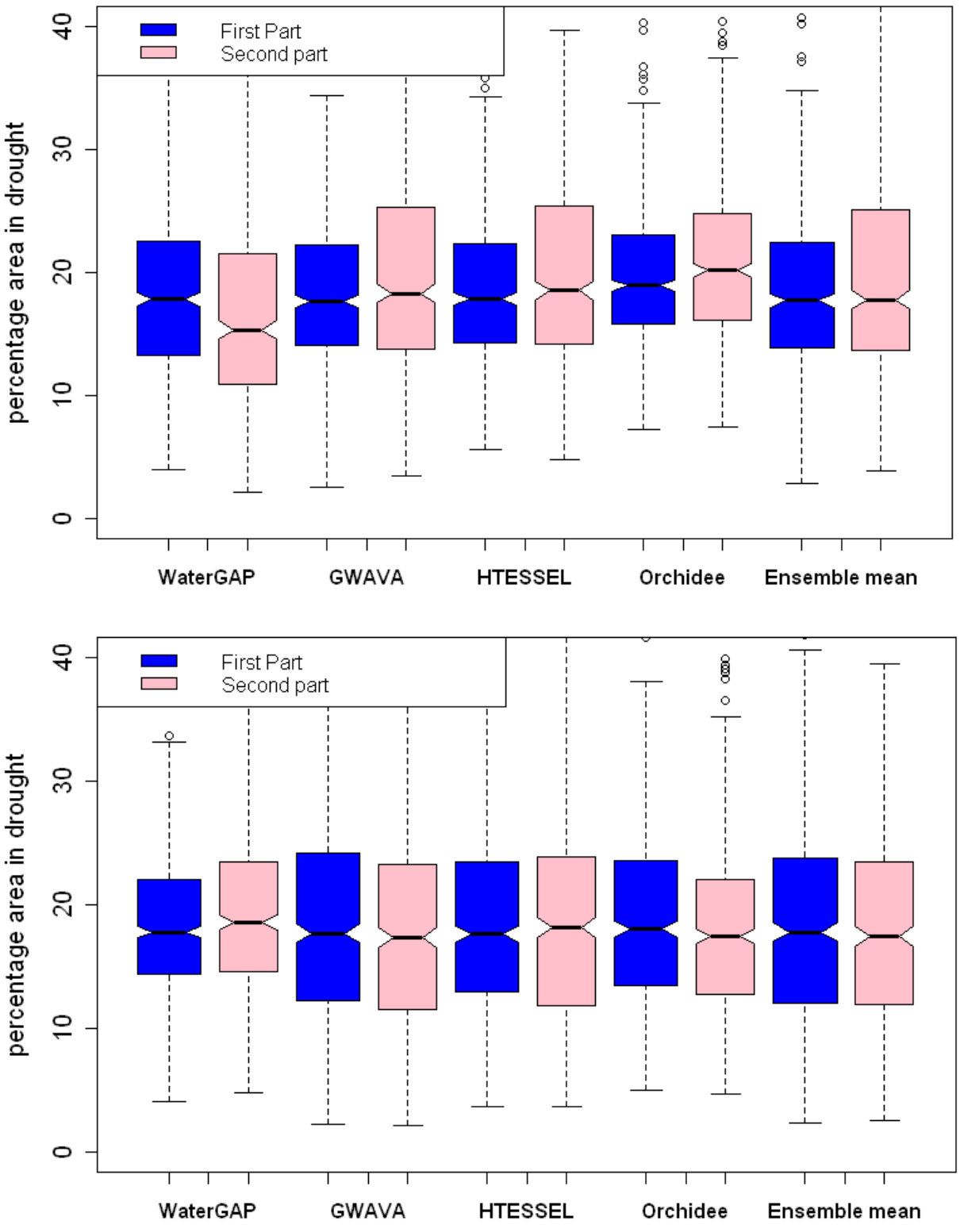


Figure 4-10. *Distributions of the spatial extent of drought in Africa (top) and Europe (bottom) for the first and the second part of the century as obtained from the four different models and the multi-model ensemble mean*

5 Conclusions and recommendations

5.1 Conclusions

- Statistical tests reveal that nearly half of the land points of the globe show statistically significant shift in the simulated runoff between the first(1906-1957) and the second part (1958-2000) of the century for the four models as well as the multi-model ensemble mean
- The choice of the time windows for threshold greatly affects the value of the threshold which consequently influences the occurrence of drought. Using sliding windows provides more realistic results since it helps to capture the influence of trend shift between the two periods of the century as well as inter-decadal climatic variability.
- The range of the percentage of the globe in drought for the simulation period is found to be in consistent among all models (10-30 %)
- Most of the world wide droughts, that are often associated with the ENSO phenomena, are identified by the models though there exist differences in percentage of area in drought among models. Though WaterGAP tends to overestimate the spatial extent of such events compared to the other models. However, HTESSEL reveals higher percentage of area in drought for April and May that is persistent throughout the simulation period which makes rather unrealistic. This patterns are also clearly observed in the continental analysis for North America and Asia .
- Most of the models capture the documented major historical droughts at continent scale, though there exist differences in percentage of area in drought
- Inter-Comparing the two periods of the century, consistent statistically significant decreasing trends are observed in area of the globe in drought for the first part of the century, whereas increasing trends are observed for the second part of the century among all models.
- The increasing trend in spatial extent of global drought for the second part of the century are found to be higher for GWAVA and Ensemble mean whereas lower for HTESSEL
- With respect to trends at Continent scale, WaterGAP seems to behave against the other models. For example, decreasing trends in the second half of the century for Africa and increasing trend in the second half of the century for Europe

5.2 Recommendations

- The global as well as regional analysis in this study focuses on the spatial extent of drought in terms of percentage of area in drought. However, the duration, intensity and severity of drought need to be assessed in order to get a better overview f the performance of the model for extreme hydrological events.
- The models are very sensitive to the forcing datasets. Therefore investigating the variation of the meteorological forcing data (precipitation and temperature) spatially as well temporally and assessing their correlation with drought occurrence will give more insight about the performance of each model for predicting extreme events.

References

Alcamo, J., Doll, P., Henrichs, T., Kaspar, F., Lehner, B., Rosch, T., Siebert, S., 2003. Development and testing of the WaterGAP 2 global model of water use and availability. Hydrological Sc. J. 48, 317-337.

Andreadis, K.M., Clark, E.A., Wood, A.W., Hamlet, A.F., Lettenmaier, D.P., 2005. Twentieth-Century Drought in the Conterminous United States. J. Hydrometeor 6, 985-1001.

Corzo Perez, G.A., van Huijgevoort, M.H.J., Voss, F., van Lanen, H.A.J., 2011. On the spatio-temporal analysis of hydrological droughts from global hydrological models. Hydrology and Earth System Sciences Discussions 8, 619–652.

Dai, A., 2011. Drought under global warming: a review. WIREs Clim Change 2, 45-65.

Deni, S. M., & Jemain, A. A., 2009. Mixed log series geometric distribution for sequences of dry days. Atmospheric Research, 92(2), 236-243.

ERA-40 Atlas [WWW Document], 2011. . URL http://www.ecmwf.int/research/era/ERA-40_Atlas/docs/index.html

Fleig, A.K., Tallaksen, L.M., Hisdal, H., Demuth, S., 2006. A global evaluation of streamflow drought characteristics. Hydrol. Earth Syst. Sci. 10, 535-552.

Groisman, P. Ya., & Knight, R. W. 2008. Prolonged dry episodes over the conterminous United States: New tendencies emerging during the last 40 years. Journal of Climate, 21(9), 1850-1862.

Haddeland, I., Clark, D.B., Franssen, W., Ludwig, F., Voß, F., Arnell, N.W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S.N., Hagemann, S., Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koirala, S., Oki, T., Polcher, J., Stacke, T., Viterbo, P., Weedon, G.P., Yeh., P., 2011. Multi-Model Estimate of the Global Terrestrial Water Balance: Setup and First Results. J. Hydrometeor 110531121709055.

Hisdal, H., Stahl, K., Tallaksen, L.M., Demuth, S., 2001. Have streamflow droughts in Europe become more severe or frequent? Int. J. Climatol. 21, 317-333.

Hisdal, H., Tallaksen, L. M., Clausen, B., Peters, E., & Gustard, A. 2004. Hydrological Drought Characteristics. Pages pp. 139-198 of: Tallaksen, L. M., & van Lanen, H.A.J. (eds), Hydrological Drought Processes and Estimation Methods for Stream

Kendall, M.G., 1970. Rank Correlation Methods. Fourth Edition. Hafner Press, New York.

Mann, H.B., 1945. Nonparametric Tests against Trend. Econometrica 13, 245-259.

Rodionov, S.N., 2005. A brief overview of the regime shift detection methods, in: Large-scale Disturbances (regime Shifts) and Recovery in Aquatic Ecosystems: Challenges for Management Toward Sustainability (eds Velikova V., Chipev N.). Varna, Bulgaria: UNESCO-ROSTE/BAS Workshop on Regime Shifts. Http://biocore. Ecolab. Bas. Bg/events/past/unesco-ws.

Sheffield, J., Goteti, G., Wood, E.F., 2006. Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling. J. Climate 19, 3088-3111.

Sheffield, J., Wood, E.F., 2007. Characteristics of global and regional drought, 1950–2000: Analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle. J. Geophys. Res. 112.

Sheffield, J., Wood, E.F., 2008. Global Trends and Variability in Soil Moisture and Drought Characteristics, 1950–2000, from Observation-Driven Simulations of the Terrestrial Hydrologic Cycle. J. Climate 21, 432-458.

Sheffield, J., Andreadis, K. M., Wood, E. F., & Lettenmaier, D. P. 2009. Global and Continental Drought in the Second Half of the Twentieth Century: Severity-Area-Duration Analysis and Temporal Variability of Large-Scale Events. Journal of Climate, 22(8), 1962-1981.

Tallaksen, L.M., Lanen, H.A.J. van, 2004. Hydrological drought: processes and estimation methods for streamflow and groundwater. Elsevier.

Tallaksen, L.M., Lanen, H.A.J. van, 2005. Hydrological Drought, Volume 48: Processes and Estimation Methods for Streamflow and Groundwater. Elsevier Science.

Uppala, S.M., KAllberg, P.W., Simmons, A.J., Andrae, U., Bechtold, V.D.C., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R.P., Andersson, E., Arpe, K., Balmaseda, M.A., Beljaars, A.C.M., Berg, L.V.D., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B.J., Isaksen, L., Janssen, P.A.E.M., Jenne, R., Mcnally, A.P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N.A., Saunders, R.W., Simon, P., Sterl, A., Trenberth, K.E., Untch, A., Vasiljevic, D., Viterbo, P., Woollen, J., 2005. The ERA-40 re-analysis. Q.J.R. Meteorol. Soc. 131, 2961-3012.

Vincent, LA, & Mekis, E. 2006. Changes in daily and extreme temperature and precipitation indices for Canada over the twentieth century. Atmosphere-Ocean, 44(2), 177-193

Weedon, G.P., Gomes, S., Viterbo, P., Österle, H., Adam, J.C., Bellouin, N., Boucher, O., and Best, M., 2010. The WATCH forcing data 1958-2001: a meteorological forcing dataset for land surface-and hydrological-models., Tech. rep., WATCH Technical Report no. 22, available at: http://www.eu-watch.org/nl/25222760-Technical Reports.html.

Weedon, G.P., Gomes, S., Viterbo, P., Shuttleworth, W.J., Blyth, E., Österle, H., Adam, J.C., Bellouin, N., Boucher, O., Best, M., 2011. Creation of the WATCH Forcing Data and its use to assess global and regional reference crop evaporation over land during the twentieth century. Journal of Hydrometreology.

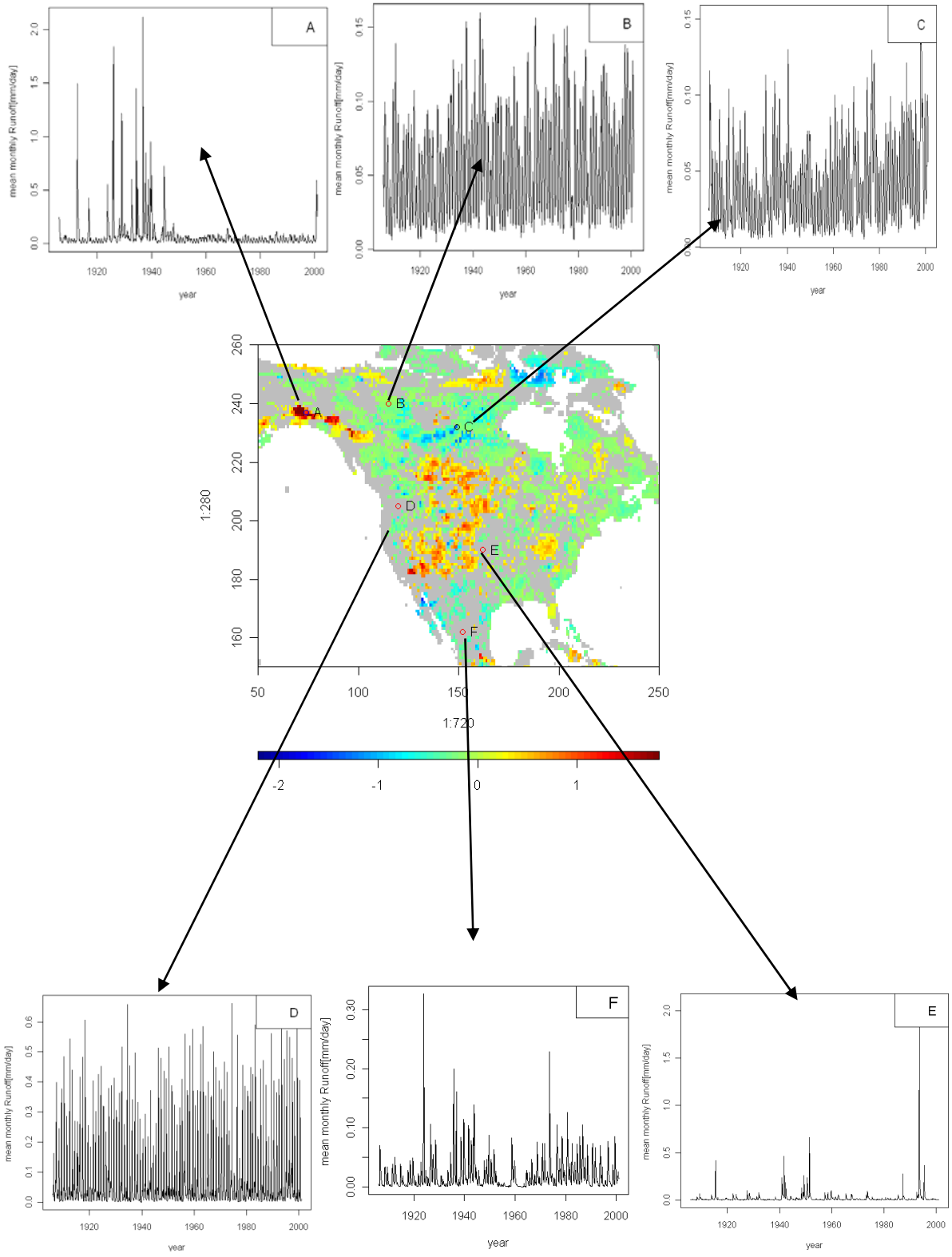
Weiland, F.C.S., van Beek, L.P.H., Kwadijk, J.C.J., Bierkens, M.F.P., 2010. The ability of a GCM-forced hydrological model to reproduce global discharge variability. Hydrology and Earth System Sciences 14, 1595–1621.

Xiong, L., Guo, S., 2004. Trend test and change-point detection for the annual discharge series of the Yangtze River at the Yichang hydrological station. Hydrological Sc. J. 49, 1-112.

Yevjevich, V. 1967. An objective approach to definition and investigations of continental hydrologic droughts. Colorado State University.

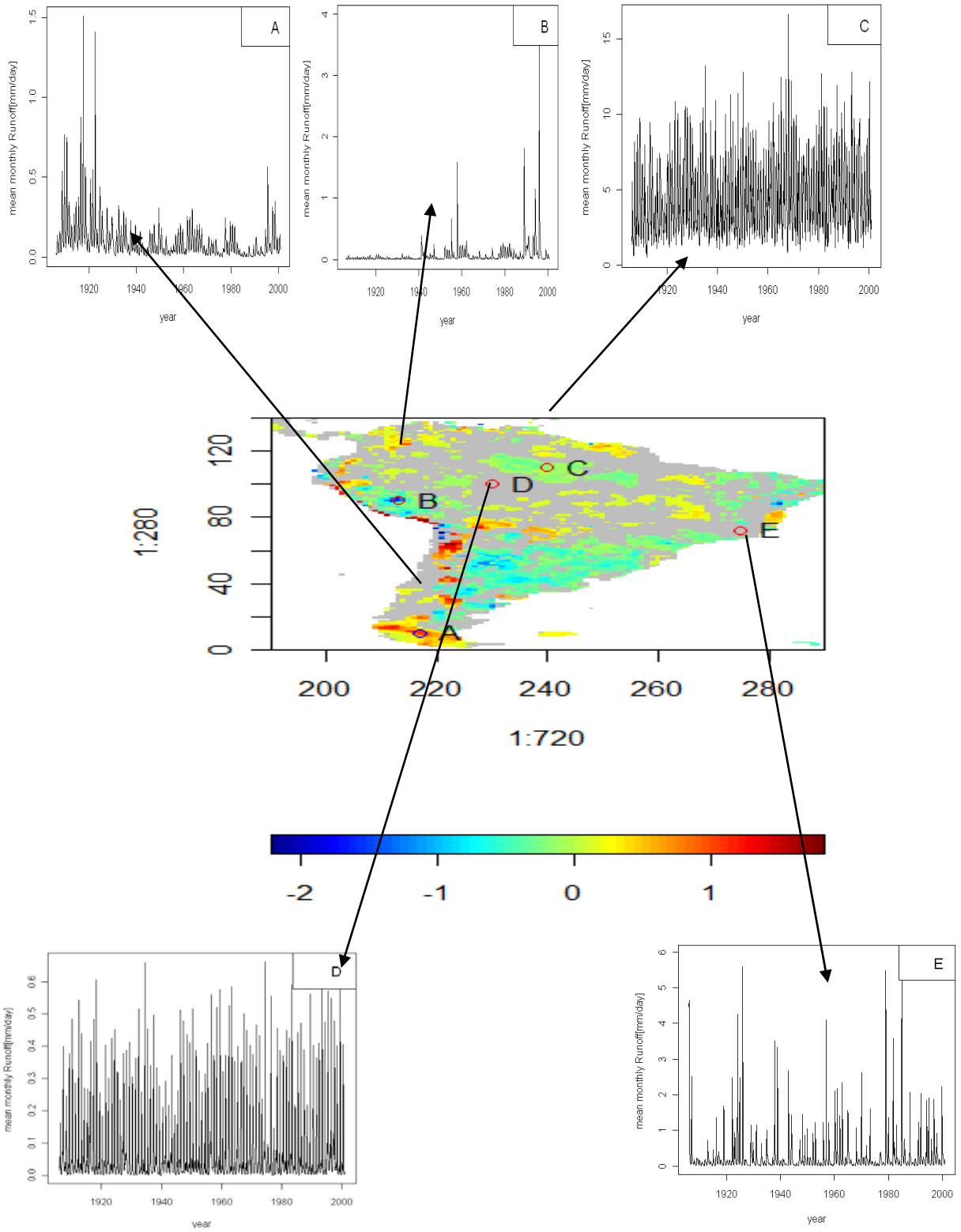
Annex I - Time series runoff plots of cells from North America

Time series plots of three cells that have statistically significant shift in mean between the two period of the century (A,B, C) and three cells which have non-significant shifts (D,E, F) from North America



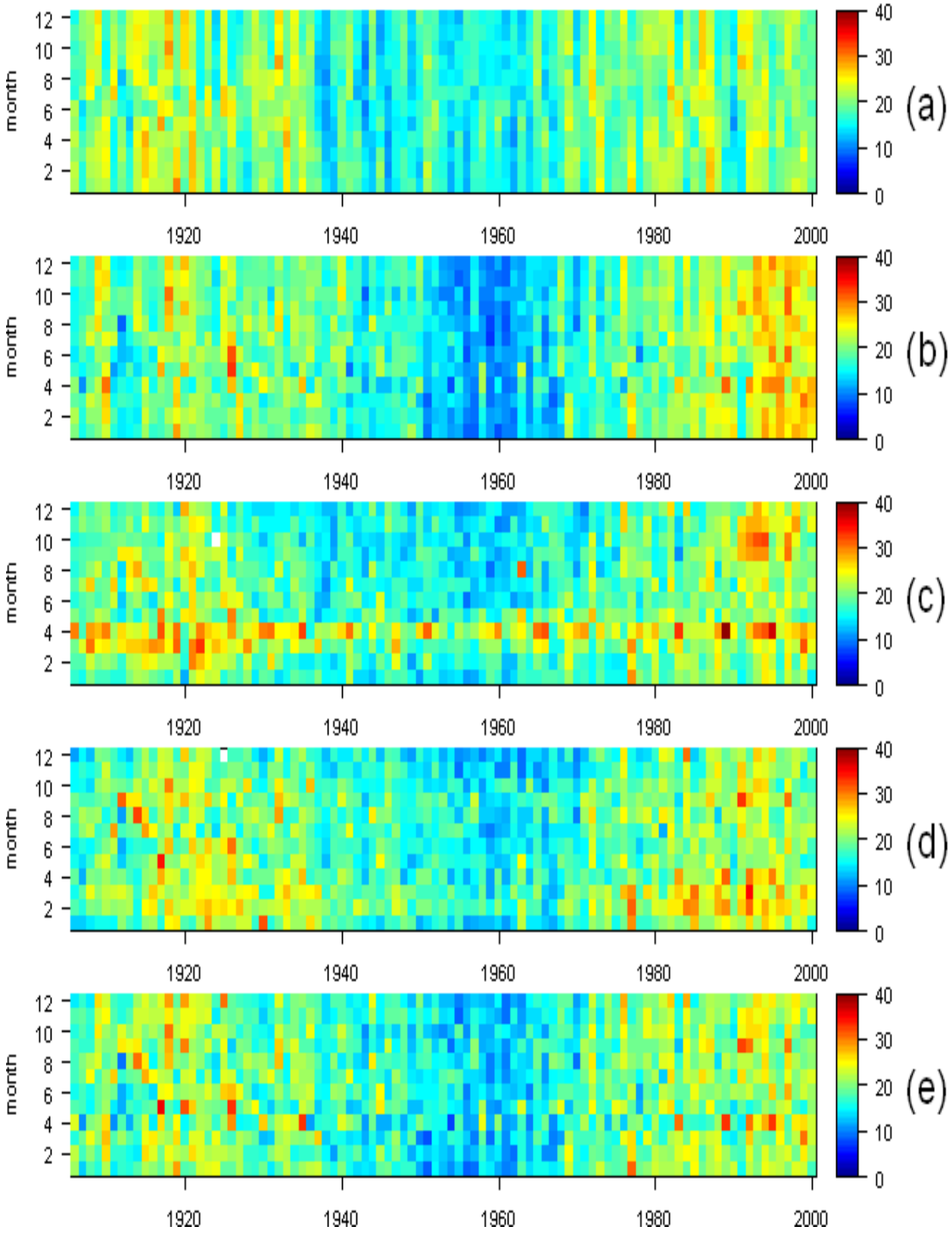
Annex II -Time series runoff plots of cells from South America

Time series plots of three cells that have statistically significant shift in mean between the two period of the century (A,B, C) and two cells which have non-significant shift (D,E) from South America



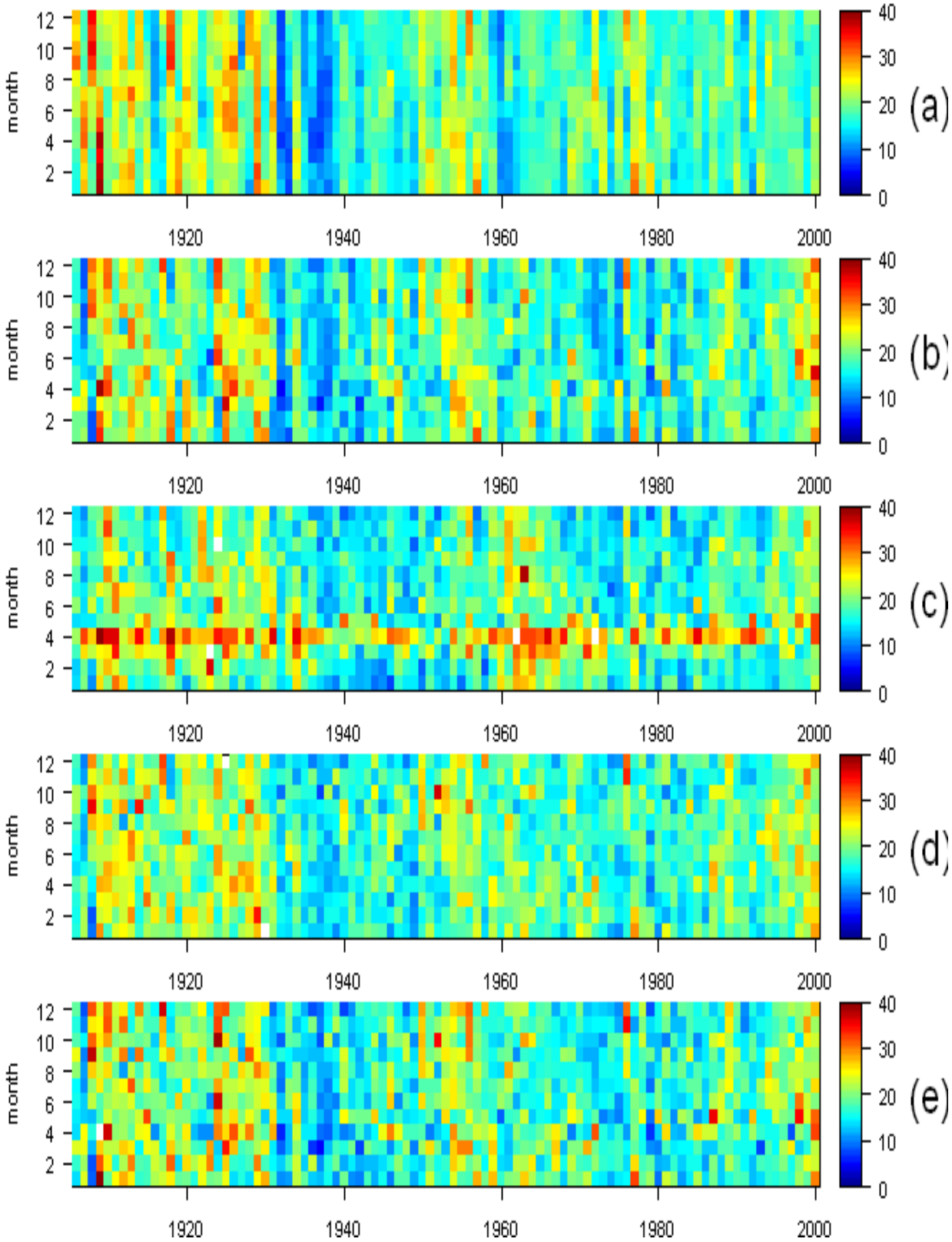
Annex III - Percentage area of Asia in drought

a) WaterGAP, b) GWAVA, c), HTESSEL, d) Orchidee, and (e) model ensemble mean



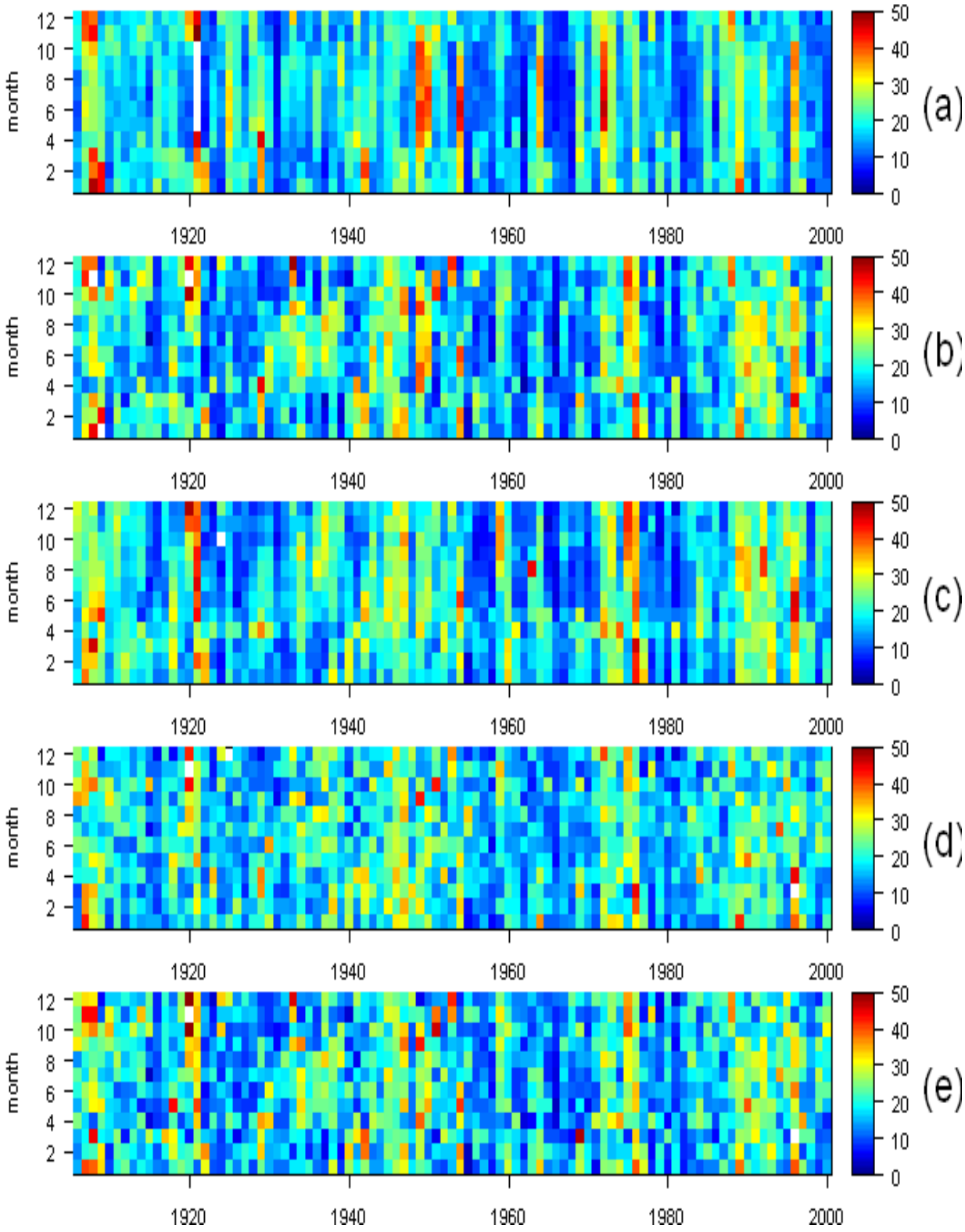
Annex IV - Percentage area of North America in drought

a) WaterGAP, b) GWAVA, c), HTESSEL, d) Orchidee and e) model ensemble mean



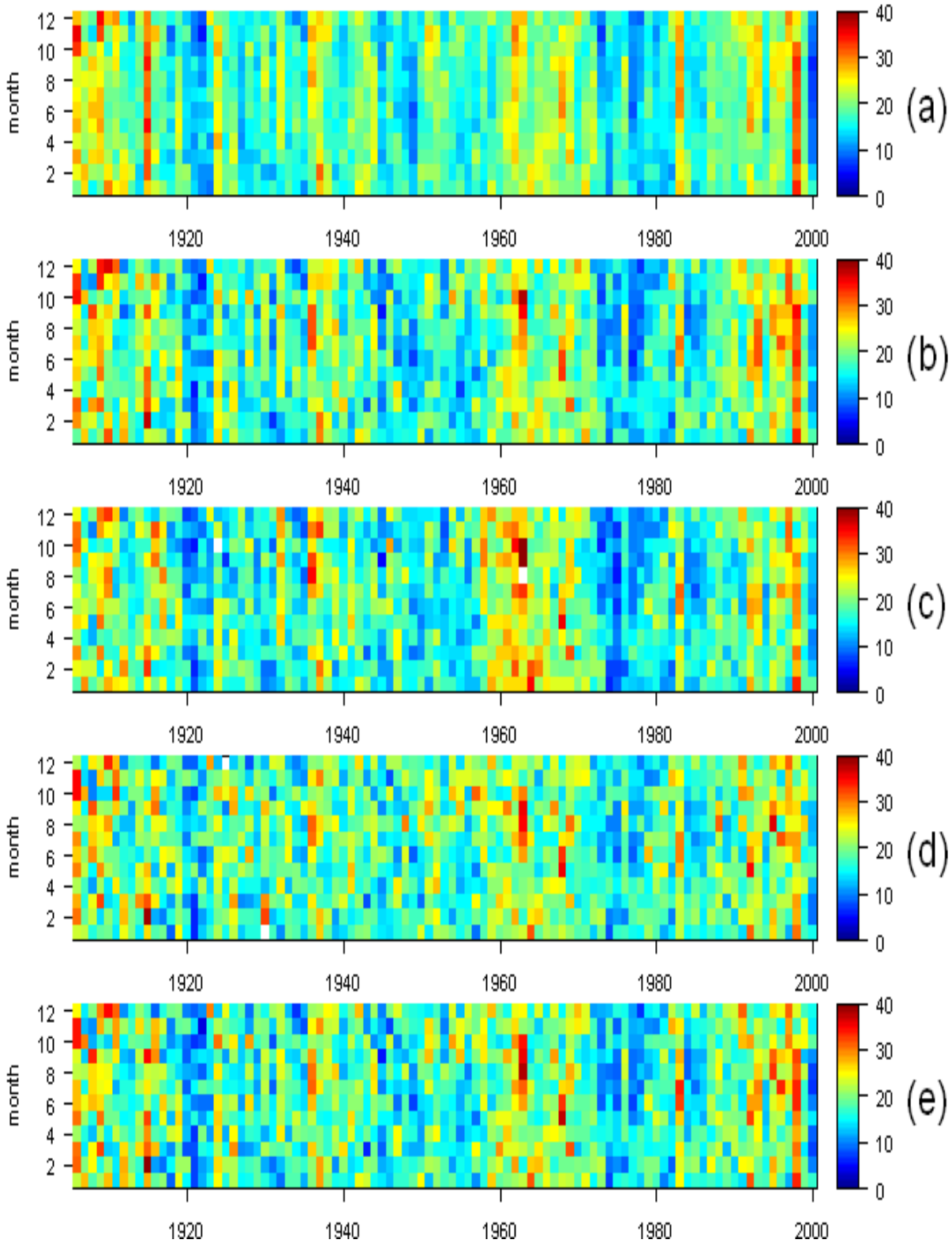
Annex V - Percentage area of Europe in drought

(a) WaterGAP, (b)GWAVA, (c), HTESSEL, (d) Orchidee, and (e) model ensemble mean



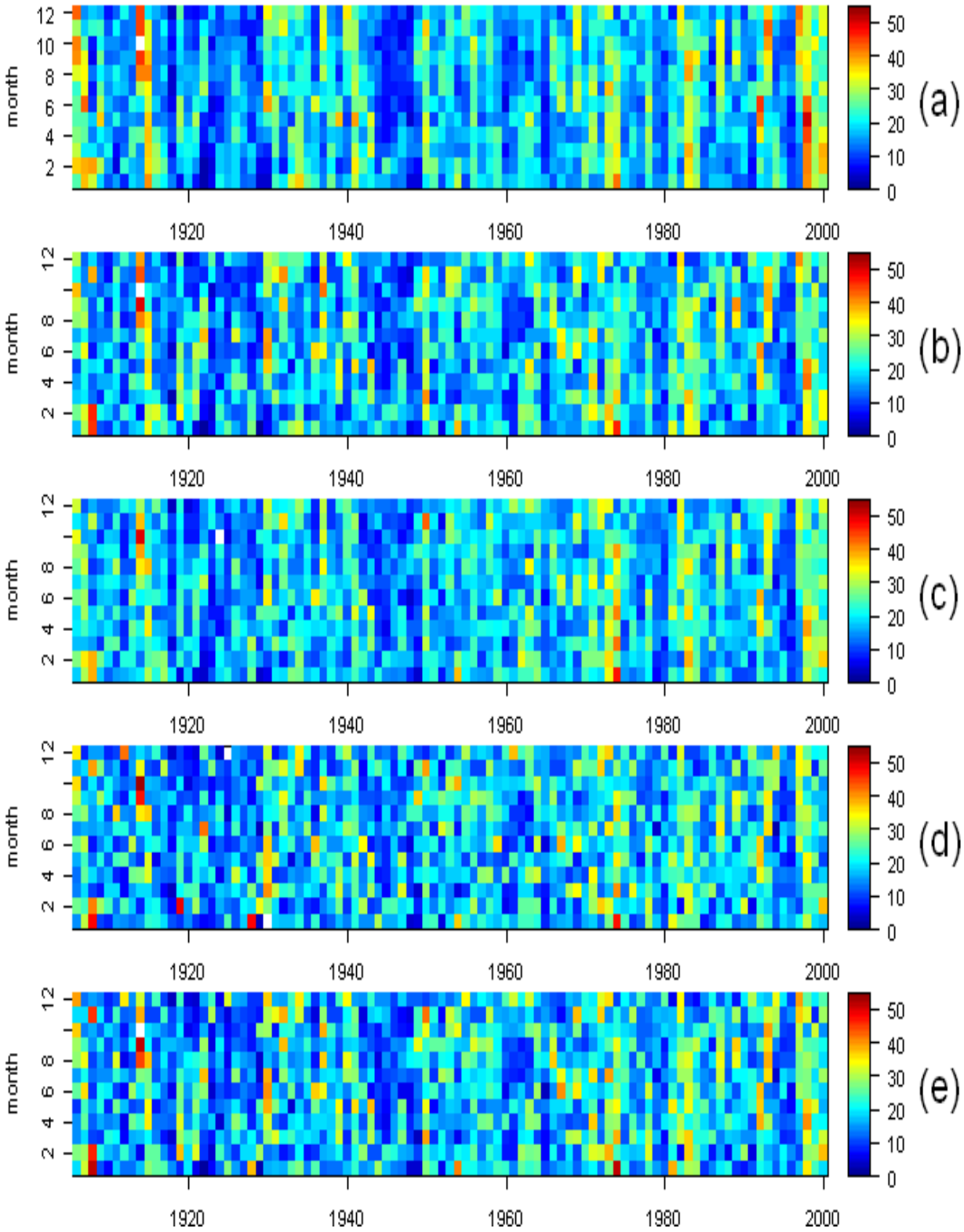
Annex VI - Percentage area of South America in drought

(a) WaterGAP, (b)GWAVA, (c), HTESSEL, (d) Orchidee, and (e) model ensemble mean



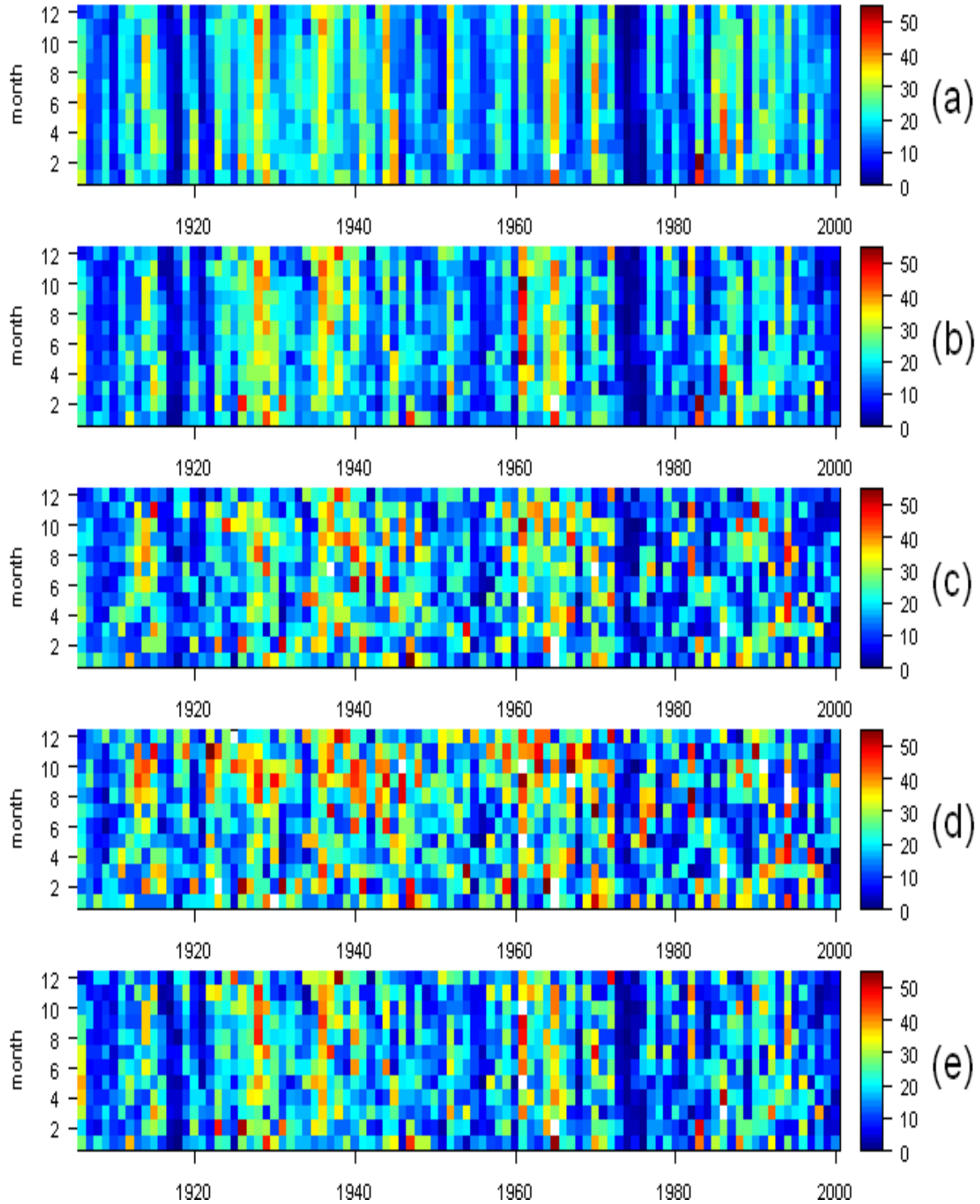
Annex VII - Percentage area of Oceania in drought

(a) WaterGAP, (b)GWAVA, (c)HTESSEL, (d) Orchidee, and (e) model ensemble mean



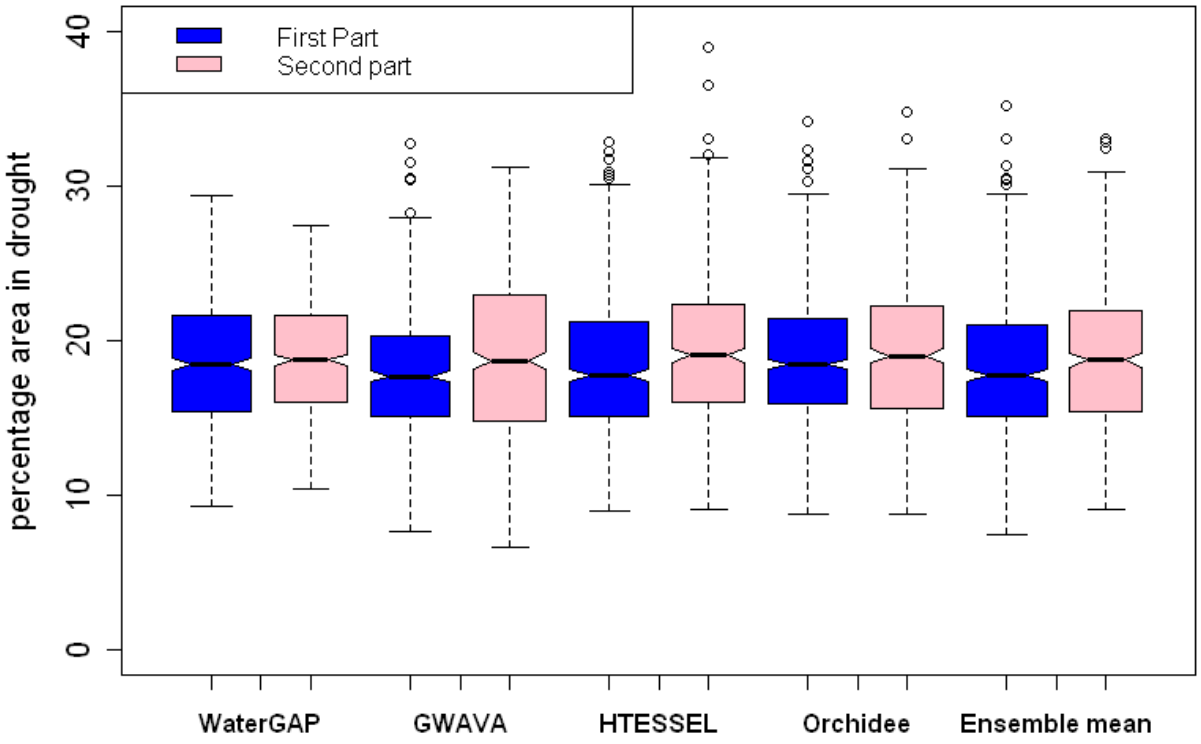
Annex VIII - Percentage area of Australia in drought

(a) WaterGAP, (b)GWAVA, (c), HTESSEL, (d) Orchidee, and (e) model ensemble mean

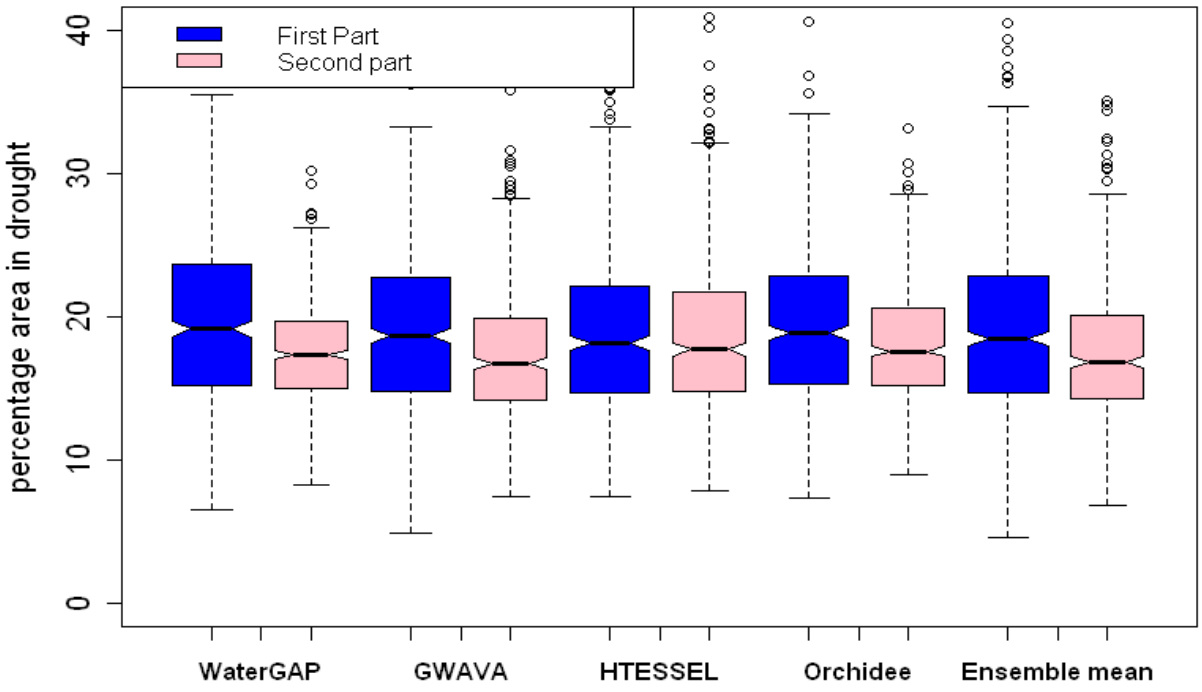


Annex IX- Comparison of trends between the two periods for each continent

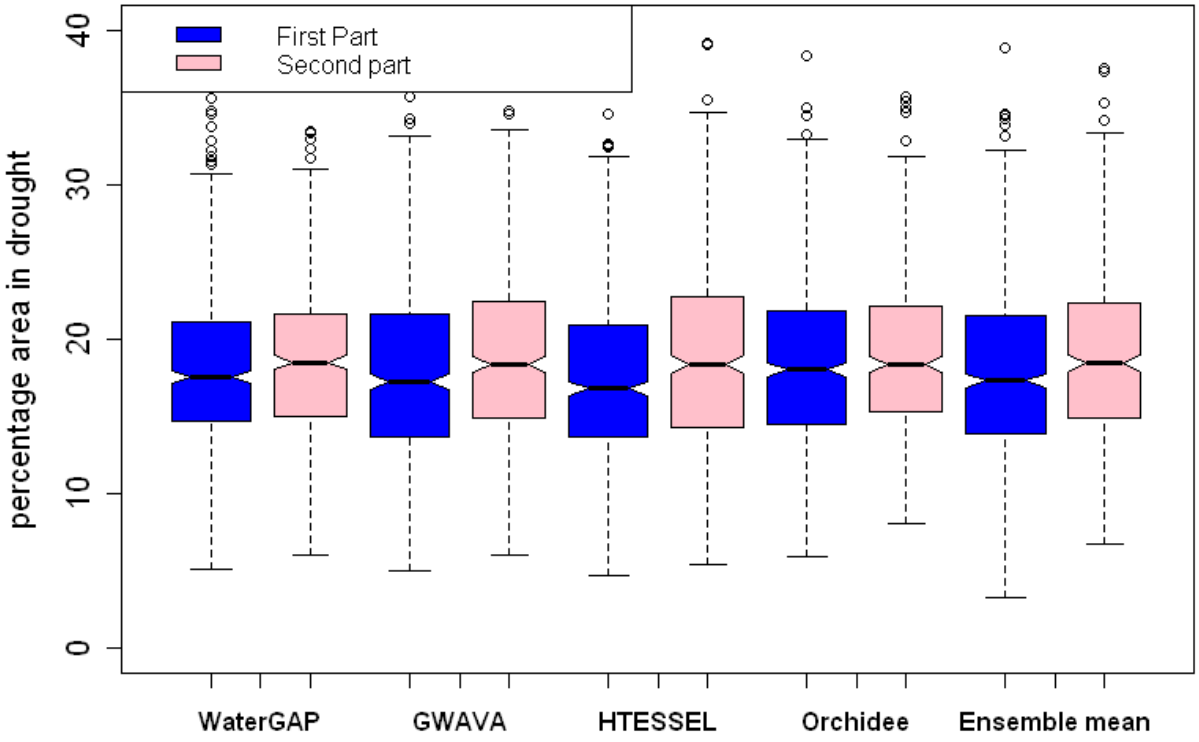
Asia



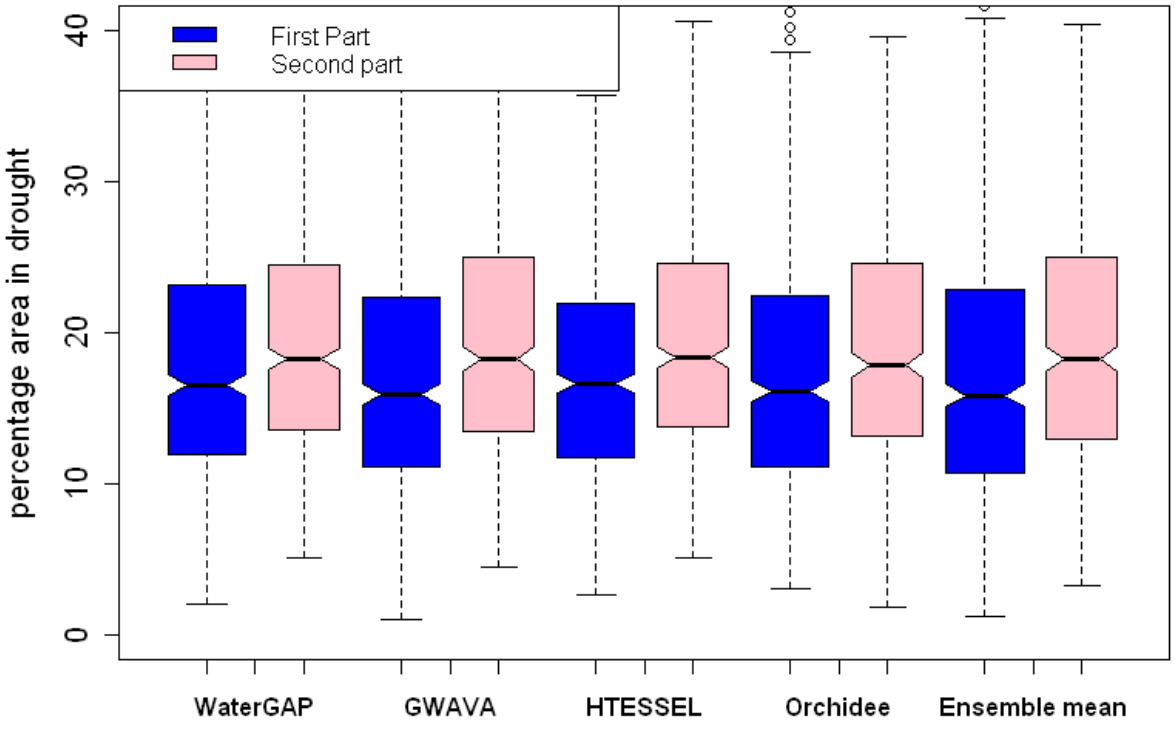
North America



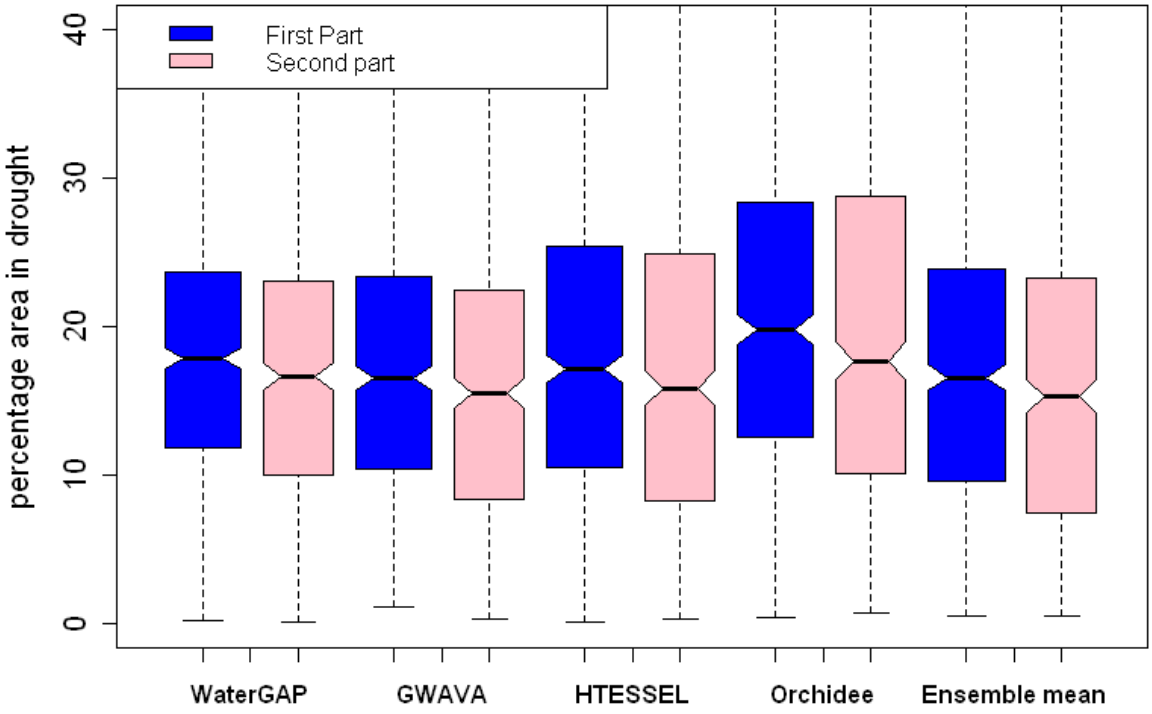
South America



Oceania



Australia


First Part
Second part
percentage area in drought
0
10
20
30
40
WaterGAP
GWAVA
HTESSEL
Orchidee
Ensemble mean