Analysing droughts in the Magdalena-Cauca macro-basin in Colombia

Jesús María Zamora, *Río Magdalena* (1947)

Jip Grootveld

Universiteit Utrecht

Analysing droughts in the Magdalena-Cauca macro-basin in Colombia

MSc Thesis, GEO4-1520, 37.5 ECTS
May 12, 2019

Author: Jip Grootveld
Student number: 3921018
E-mail: j.grootveld@students.uu.nl

First supervisor: Geert Sterk
Second supervisor: Rens van Beek
External supervisor: Erasmo Alfredo Rodriguez Sandoval

MSc Programme Earth Surface and Water
Department of Physical Geography
Faculty of Geosciences
Utrecht University

Abstract

Droughts can have detrimental consequences for agriculture, hydropower generation and ecology. It is important to increase our understanding of this phenomenon to be able to assist in water management and to be able to take preventive measures. The Magdalena-Cauca macro-basin (MCMB) is located in the northwest of Colombia. The climate is mainly influenced by shifting of the Inter Tropical Convergence Zone, the north-easterly trade winds and the Chocó-jet. Droughts are known to occur in the MCMB during El Niño events.

Four large scale hydrological models (ORCHIDEE, HTESSEL, WaterGAP3 and PCR-GLOBWB) were evaluated for the MCMB, followed by a quantification of droughts using five different drought indices (SPI, SPEI, ETDI, SMDI and SDI) and the determination of the correlation between droughts and the El Niño Southern Oscillation (ENSO). The MSWEP precipitation data exhibited good performance. Therefore, this forcing dataset was used as forcing for the MCMB. Not all model products had a good performance, nevertheless, they were still valuable for the drought analysis because the absolute error is longer relevant for drought indices due to their computation procedure.

Droughts were found to occur most often in the southeast of the catchment, likely because this region is less affected by the westerly Chocó-jet, which supplies moist air. During El Niño events, the drought-affected area starts in the northeast of the MCMB and moves from there to the southwest. This finding agrees well with the fact that the Chocó-jet decreases in intensity during El Niño events, enabling the drier easterly trade winds to reach further to the west.

The correlation with ENSO is strongest in the west of the catchment, and is strongest for the 6-month model ensemble mean SPEI index. The correlation between drought indices and the Multivariate ENSO index (MEI) is weaker for a strongly positive MEI values.

This drought analysis of the MCMB can assist in future research using use drought indices, and their correlation to ENSO, to predict droughts in the MCMB and assist in water management.

Contents

1. Introduction

The Magdalena-Cauca macro-basin (MCMB) is the largest river catchment in Colombia. It is located in the west of country, where the climate is principally influenced by the shift of the Intertropical Convergence Zone (ITCZ). Variabilities in the climate are caused by the El Niño Southern Oscillation (ENSO) (Hoyos et al., 2013; Restrepo & Kjerfve, 2000). El Niño, the warm phase of ENSO, is known to cause droughts in the MCMB (Poveda et al., 2001).

Droughts can have many detrimental consequences for both ecosystems and societies. Crop yields can decrease and cause economic damage, water bodies with drinking water reserves shrink and ecosystems can be disturbed. Droughts are a complex hazard, and are relatively poorly understood. Analysing historical drought events is crucial to increase our understanding of the phenomenon and to assist in water management (Heim Jr., 2002; Keyantash & Dracup, 2002; van Loon, 2015; Mishra & Singh, 2011; Narasimhan & Srinivasan, 2005; Vicente-Serrano et al., 2010; Wilhite, 2000; Zargar et al., 2011).

Drought can be described as a temporary, below average availability of water resources due to variability of natural conditions (Heim Jr., 2002; van Loon, 2015; Werner & Gründermann, 2014). Generally, four types of drought can be distinguished: i) meteorological drought, ii) agricultural drought, iii) hydrological drought and iv) socio-economic drought (Figure 1) (Heim Jr., 2002; van Loon, 2015; Wilhite, 2000; Zargar et al., 2011). Meteorological drought is a lack of precipitation over an extensive area and period of time. Agricultural drought refers to a soil moisture deficit in the root zone, thus affecting the amount of moisture available for vegetation. Hydrological drought is used to describe water deficits in surface and subsurface water bodies. Finally, socio-economic drought refers to the impact caused by the previous three types of droughts, such as crop failure or drinking water shortages (van Loon, 2015).

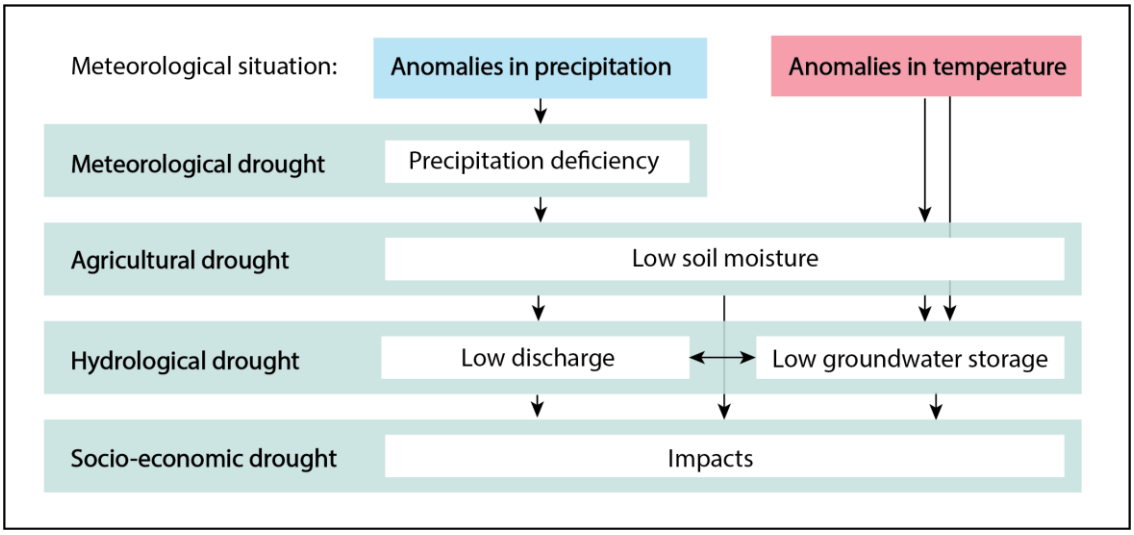

Figure 1. Schematic representation of the relations between different categories of drought (van Loon, 2015)

Through drought propagation, a drought starting as a meteorological drought can develop to become an agricultural drought, because the water deficit propagates through the hydrological system (Figure 2) (van Loon, 2015). This development has four distinguishable effects: i) pooling; ii) lag; iii) lengthening and iv) attenuation. First, the timescale on which the droughts act increases from meteorological drought to hydrological drought. This causes pooling: multiple, smaller meteorological droughts can result in one large hydrological drought. Second, the hydrological drought will occur later than the meteorological drought - an effect called lag - because it takes time for the water deficit to develop through the hydrological system. Third, the drought event will last longer; this is called lengthening. Finally, the intensity of the hydrological drought will be less compared to the meteorological drought, because the water deficit will be spread over a longer time. This effect is called attenuation.

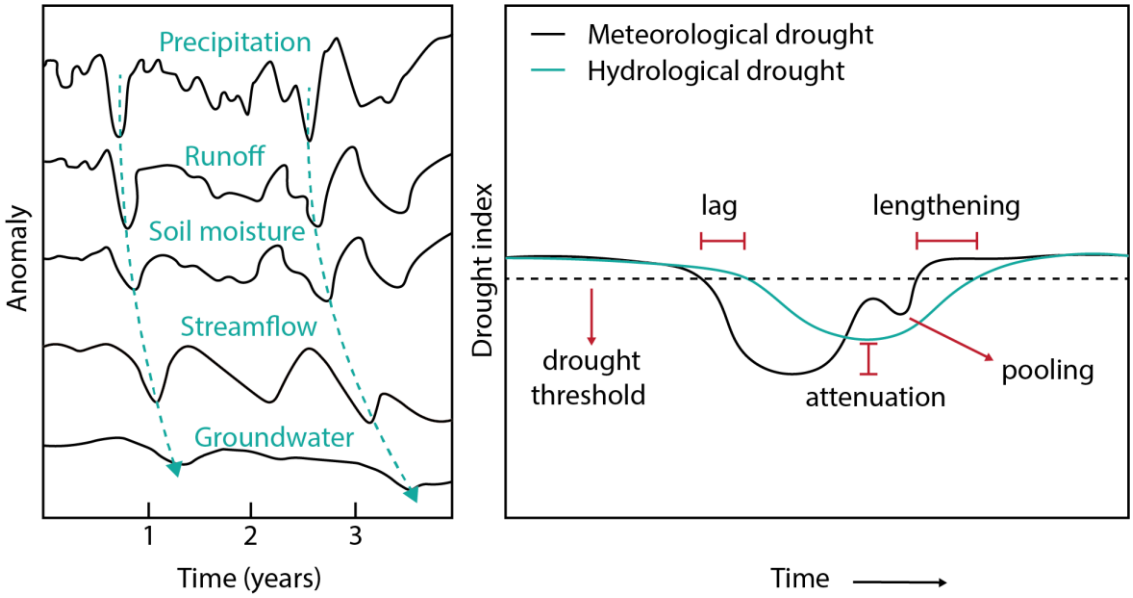

Figure 2. Schematic representation of drought development through the hydrological system (van Loon, 2015, van Loon & van Lanen, 2012)

Drought is often confused with aridity, water shortage and water scarcity. Aridity is, in contrast to drought, a more permanent condition and can be explained as the general characteristic of an arid climate (van Loon, 2015; Werner & Gründermann, 2014). A water shortage is a temporary situation where the human water demand cannot be met. When the water shortage is more permanent, one can speak of water scarcity (Werner & Gründermann, 2014).

Drought analysis is important for water management. In order to perform the analysis, drought indices have been developed. More than 100 different drought indices have been proposed, each taking into account different aspects of a drought and each with their own strengths and weaknesses. Drought indices can take many aspects of a drought into account, both quantitative and qualitative. For a quantitative analysis, the duration, intensity, timing and spatial extent can be incorporated. Qualitative aspects in drought indices are represented by using different hydro-climatic parameters, such as precipitation for meteorological drought, soil

moisture for agricultural drought and discharge for hydrological drought (Heim Jr., 2002; Keyantash & Dracup, 2002; Nalbantis & Tsakiris, 2009; Zargar et al., 2011). Drought indices are computed using hydro-meteorological data, for which in-situ observations can be used. When in-situ observations are lacking, Water Resources Reanalysis (WRR) data can be used. WRR data combines in-situ observations, earth observation data and global hydrological model output (Flörke et al., 2013; Van Loon, 2015; Werner & Gründermann, 2014). There are a lot of different hydrological models available to generate the hydrological model output for WRR datasets. For drought management it is important to use (a) suitable hydrological model(s) in order to generate a useful and meaningful WRR dataset.

Often, droughts can be correlated with sea surface temperature (SST) anomalies, such as the El Niño Southern Oscillation (ENSO). SST anomalies often have a periodic return, which makes that the next anomaly can be predicted with a certain confidence (Barnston et al., 1999; Chen et al., 2004; Zhao et al., 2010). When the correlation between SST anomalies and drought occurrence is strong, the predictive aspect of the SST anomaly can help predict the next drought event and thus assist in water management (Hoyos et al., 2017; Poveda et al., 2001, 2006)

The aim of this research is to perform a drought analysis for the Magdalena-Cauca macro-basin. This will be accomplished by fulfilling the following objectives:

I. Test the performance of four large scale hydrological models, for the precipitation forcing data, and the potential evapotranspiration and discharge model products, for the Magdalena-Cauca macro-basin;
II. Quantify droughts in the MCMB by computing five drought indices based on model products of the four models;
III. Determine the strength of the correlation between droughts and the El Niño Southern Oscillation (ENSO)

2. Study area

The Magdalena-Cauca macro-basin is located in the west of Colombia and comprises two main rivers: the Cauca river in the west and the Magdalena River in the east (Figure 3). The Andes mountain range reaches is most northern point in the MCMB, where it splits into three main branches: the Cordillera Occidental (forming the south-western border of the catchment) the Cordillera Central and the Cordillera Oriental (forming the eastern border of the catchment). Both the Cauca and Magdalena rivers flow from south to north. The Cauca River is, after the Magdalena River, the largest river in Colombia, and has its headwaters near the city of Popayán at 3125 m elevation. Its average annual discharge is 2347 m^3/s and the drainage basin area is 63.300 km^2 (Restrepo et al., 2006). The Cauca flows between the Cordillera Occidental (west) and the Cordillera Central (central) until it merges with the Magdalena river close to the city of Santa Cruz de Mompox. The Magdalena headwaters are a lake in the Huila department of Colombia at 3685 m elevation, and the river flows between the Cordillera Central (central) and the Cordillera Oriental (east). The length of the Magdalena from its headwaters to the sea is 1612 km. The combined Magdalena-Cauca macro-basin has an area of 257,438 km^2. The Magdalena eventually flows into the Caribbean Sea close to the city of Barranquilla (Restrepo et al., 2006; Restrepo & Kjerfve, 2000). The average annual discharge at Calamar, ± 100 km upstream of Barranquilla, is 7200 m^3/s, ± 5000 m^3/s.

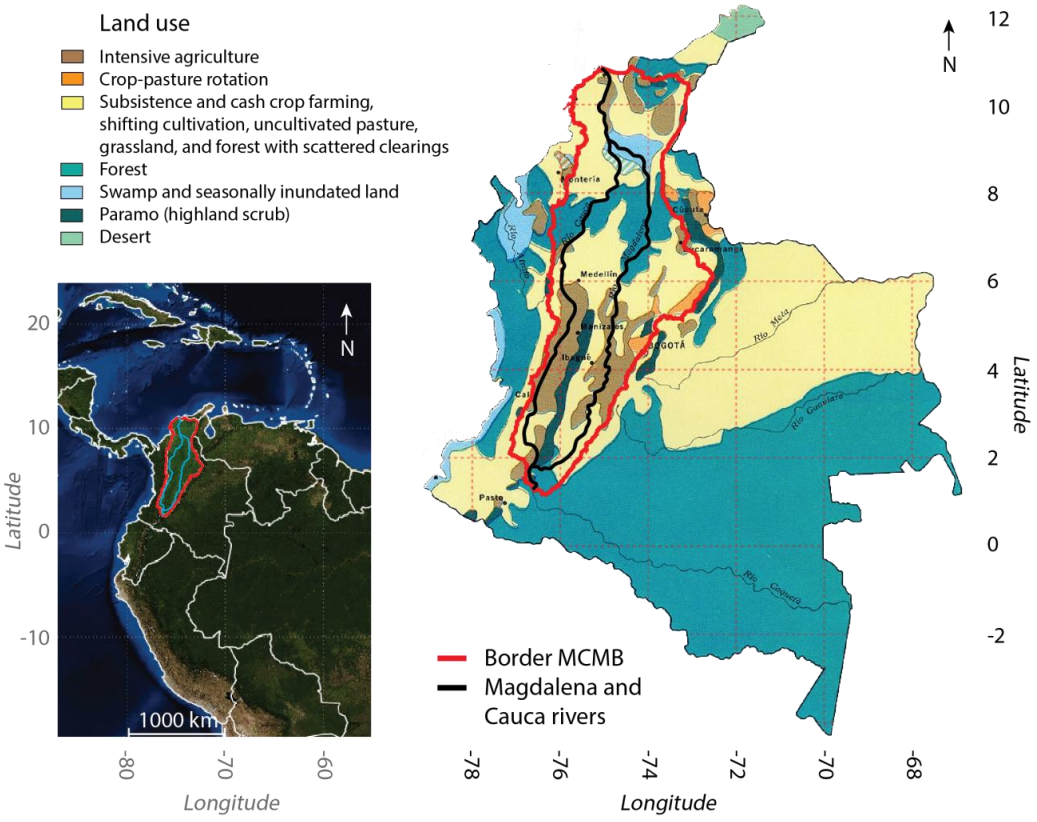

Figure 3. Location of the Magdalena-Cauca macro-basin in northern South-America (left), and the land use in Colombia (right). The catchment is located between -77.00° and -72.25° longitude and 1.75° and 11.25° latitude.

The Magdalena-Cauca river system is the largest river in the Andes (Restrepo et al., 2006) and it is the most important river basin in Colombia. It covers around 24% of the country's area and around 79% of the Colombian population lives in this area, which corresponds to 120 people/km2 (Restrepo et al., 2006). The river has the largest specific sediment yield in South America: it contributes 9% of the sediment discharged from eastern South America (Restrepo & Kjerfve, 2000). The specific sediment yield is defined as the sediment yield divided by the drainage area of a river. Sediment finds its origin in erosion and depositional processes, which are influenced by factors such as topography, climate, vegetation, soil and bedrock properties, and land use. In the past decades, the basin has seen large-scale land use changes (Restrepo & Syvitski, 2006). The forest cover of the Colombian Andes has decreased due to population increase and land use changes. Poor agricultural practices lead to increased erosion rates (Restrepo et al., 2006).

The area of the MCMB is tectonically active (Restrepo & Kjerfve, 2000). The Cordilleras in the MCMB can have steep hill slopes of over 45°, which, in combination with high precipitation rates, may consequently lead to landslides. The Cordilleras reach elevations over 4000 m (Figure 4) and therefore influence the meteorology in the MCMB significantly (Poveda et al., 2005).

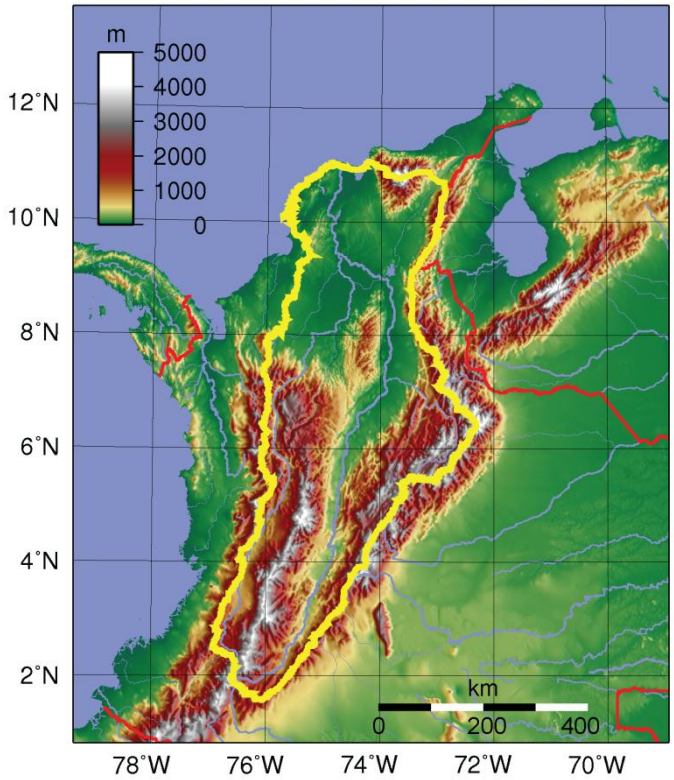

Figure 4. Topography of north-west Colombia, with the border of the MCMB indicated in yellow (Commons, 2019)

Because Colombia is located close to the equator, the annual hydrometeorological cycle in Colombia is affected by the shifting of the Intertropical Convergence Zone (ITCZ) (Poveda et al., 2007). Besides this, the climate of Colombia is also influenced by the north-easterly trade winds and the Chocó-jet. The north-easterly trade winds come from the Atlantic ocean. They get their southward direction from the Hadley circulation: warm air rises near the equator and moves

northwards at high altitudes. At ± 30° latitude, the air cools and sinks and travels southwards back to the equator while picking up moisture near the Earth's surface. The air gets a westwards direction due to the Coriolis effect. The Chocó-jet is a westerly jet at ± 5°N which brings moist air from the Pacific ocean into Colombia. The jet is driven by a sea-surface temperature gradient between the cold water in front of Ecuador and Peru and warm temperatures in Colombia and the Colombian Pacific (Poveda & Mesa, 2000). When the jet is blown into Colombia, it is blocked by the Andes mountain range and forced to rise due to orographic effects along the western flank of the Cordillera Occidental (Poveda, 1998). Here, the extremely high precipitation occurs with rates up to 12000 mm per year (Poveda & Mesa, 2000). On the top of the Cordillera Occidental, the Chocó-jet meets the north-easterly trade winds. This enhances deep convection in the Cauca and Magdalena valleys (Figure 5) (Poveda, 1998). The Chocó-jet is strongest from May to November, when the ITCZ is located on the northern hemisphere which enhances the eastward direction of the jet.

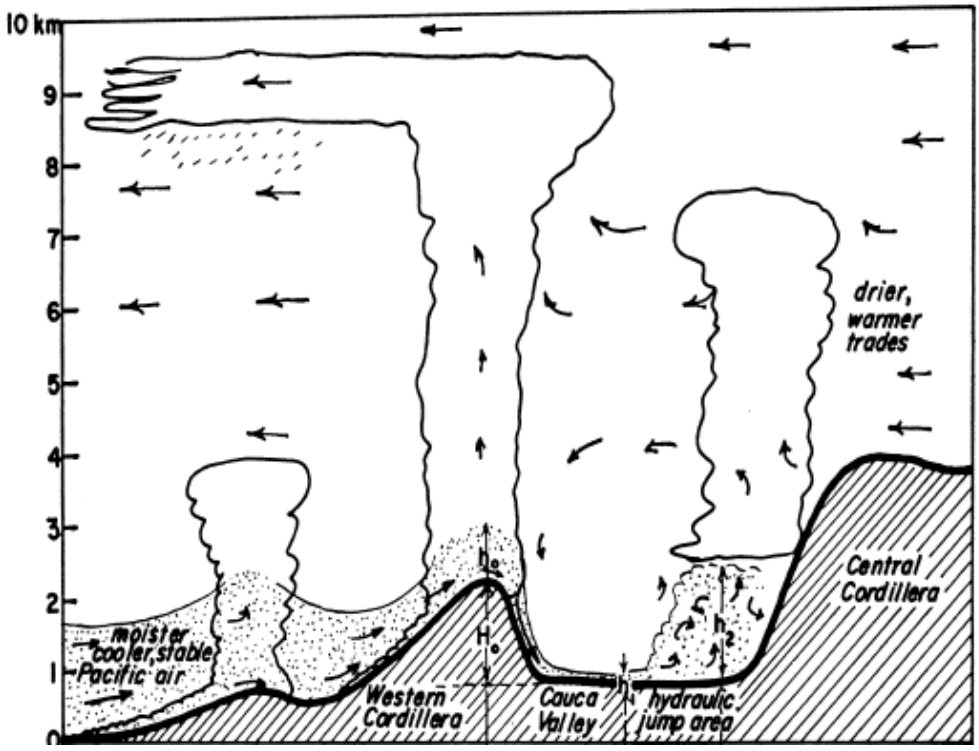

Figure 5. Schematic representation of the interactions between the westerly Chocó-jet and easterly trade winds in the Cauca valley, Colombia (López & Howell, 1967)

There are two wet seasons (March to May and October to November) and two dry seasons (December to February and June to September) in Colombia. Both wet seasons have roughly the same characteristics, however, in the upper part of the MCMB the first wet season of the calendar year is more intense (Restrepo et al., 2006). This intensification can be explained by a stronger Chocó-jet during this period: the Chocó-jet becomes dominant over the north easterly trade winds at lower elevations and penetrates into inland Colombia, bringing more moist air (Gallego et al.,

2018; Poveda et al., 2001, 2007). Rainfall amounts are generally around 2050 mm per year, but due to the mountain ranges variation is high.

The Colombian climate is influenced over longer timescales by the El Niño Southern Oscillation (ENSO) (Hoyos et al., 2013). During a normal situation, there is a warm pool in the west of the Pacific ocean, causing deep atmospheric convection and a high air pressure in the western Pacific and a low air pressure in the eastern Pacific. The low level trade winds blow from east to west. During an El Niño event (the warm phase of ENSO), the warm water from the western Pacific and the accompanying high air pressure system move to the east. The sea-surface temperature gradient decreases and this caused the trade winds to weaken. Because the Chocó-jet is driven by the sea-surface temperature gradient, the strength of the Chocó-jet also decreases during El Niño (Poveda et al., 2007). Poveda & Mesa (1997), found that in the north of South America, especially in Colombia, El Niño events are associated with negative rainfall anomalies whereas La Niña events are associated with positive rainfall anomalies. There are multiple consequences of droughts corresponding with El Niño events described in literature. The 1991-1992 El Niño event caused a reduction in precipitation, which consequently lead to reduced crop production and energy generation (Gutiérrez & Dracup, 2001). The 1997-1998 El Niño event also had severe economic consequences: coffee and other crop production failed due to the drought and large amounts of livestock were lost due to the heat waves caused by the El Niño event. The Colombia Reports, described that during the 2015 El Niño event, water levels in the Magdalena river reached a lowest level on record (Lander, 2015). According to (Poveda et al., 2004), total daily precipitation measured between 1960 and 1999 over 51 stations over the middle and south of the MCMB increases with an average of 20.8% during La Niña and decreases with an average of 22.2% during El Niño events. (Restrepo & Kjerfve, 2000b) found that 69% of the variability in stream flow in the Magdalena River can be explained by the Southern Oscillation Index (SOI), an index used to quantify the strength of ENSO. Their research showed higher discharges during La Niña events and lower discharges during El Ninõ events. The effects of ENSO are more prominent in rivers in the west and north than in the east and south of Colombia (Poveda et al., 2005; Restrepo & Kjerfve, 2000a). The phase for discharge anomalies were almost in phase with the SOI anomalies over three years, which implies that there is almost no lag between the SOI and river discharge (Restrepo & Kjerfve, 2000a).

3. Methods

This section loosely follows the order of the objectives described in the introduction. First, the data used for this study will be discussed, followed by a description of the four large scale hydrological models under evaluation and the methods of measuring their performance. This corresponds with objective 1 (Figure 6). Next, the different drought indices are described followed by the methods used for drought analysis, corresponding to objective 2. Finally, the methods for determining the strength of the correlation between droughts and ENSO are discussed, which corresponds to objective 3.

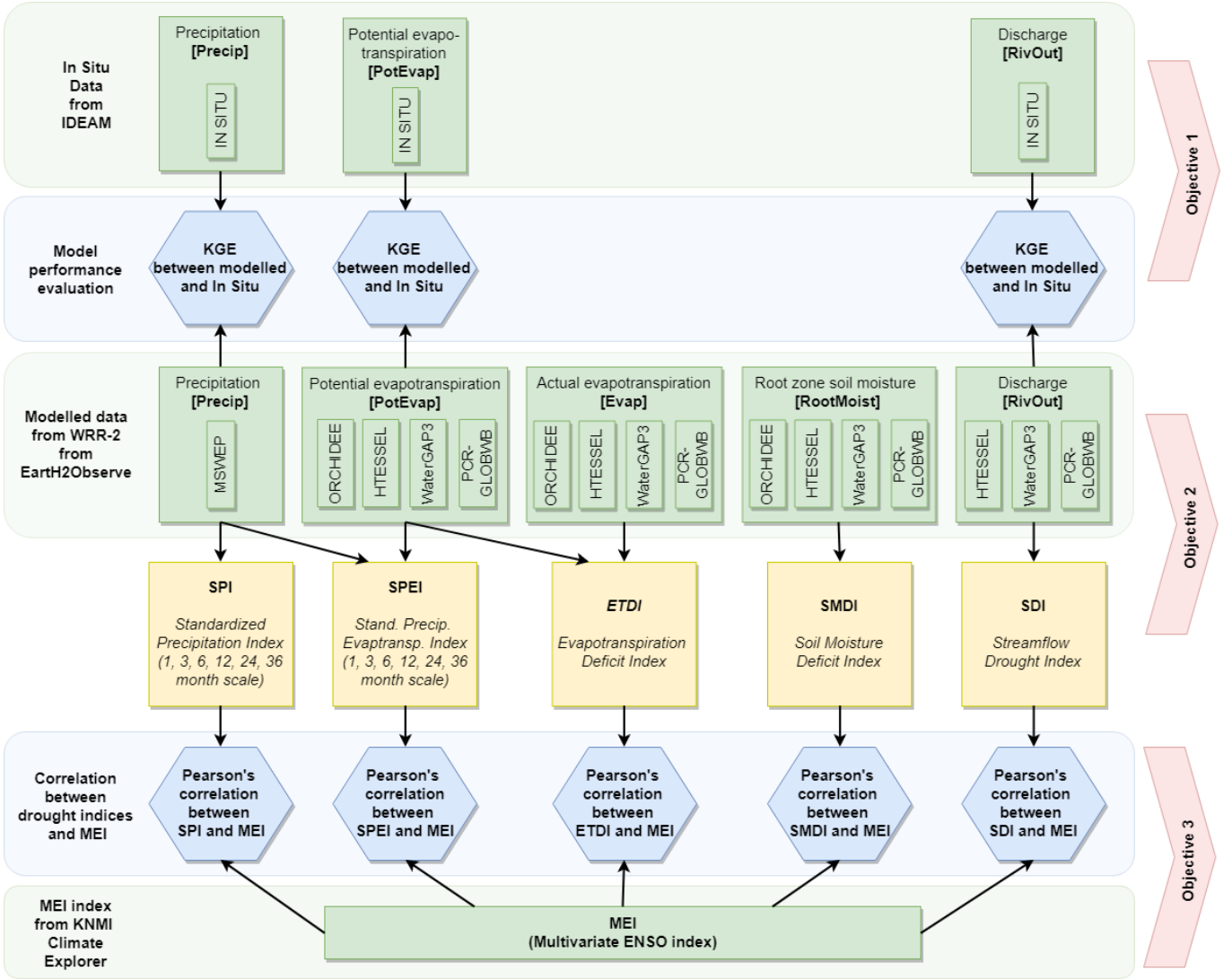

Figure 6. Workflow of this study. Input data are presented by green boxes, final products are presented by blue boxes and the objectives are presented by red boxes.

3.1 Data

The in-situ data for precipitation, temperature and discharge was made available by the Colombian Institute of Hydrology, Meteorology and Environmental Studies (IDEAM). The daily time series for precipitation were available for 2256 stations, and the daily time series for minimum and maximum temperature were available for 468 and 481 stations, respectively. For precipitation, stations with ≥ 85% complete observations were used for interpolation. For temperature, stations with ≥ 50% complete observations were used for interpolation, because temperature has a high correlation with elevation. Eventually, for precipitation 905 stations were used for interpolation and for minimum and maximum temperature 259 and 468 stations were used, respectively. Gaps in the data were not filled. Spatial interpolations in the MCMB were performed on a daily basis, using the stations with data available for that day. The Kriging with External Drift (KED) methodology (Wackernagel, 1998) was used to interpolate the precipitation data. For interpolation of the temperature data, the Co-kriging (CK) method was used (Matheron, 1971). The interpolations were validated through cross validation, using the data from the stations with less than 85% and 50% complete observations. For both datasets, the DEM was used as a covariance. The interpolation procedures resulted in daily maps with a spatial resolution of 0.1° and a temporal range from January 1980 up to and including December 2012. This data was converted to monthly data with 0.25° resolution for further processing.

The reference evapotranspiration (ET_0) was calculated from the temperature data using the Hargreaves equation (Hargreaves & Samani, 1985). To be able to compare the in-situ data with the modelled data, the in-situ precipitation data and the calculated reference evapotranspiration were re-scaled to a raster with a 0.25° resolution. Daily in-situ discharge data was available for 262 stations located along the Cauca and Magdalena rivers. The daily data was first converted to monthly data, where months with one or more missing values were given no data. Locations with more than 90% complete observations over time were converted to cells in a raster with a 0.25° resolution, in order to allow further analysis. Here, the cells were given the same value as the data point. These precipitation, reference evapotranspiration and discharge data sets were used as benchmark for the model performance.

The four evaluated models and their products are part of the eartH2Observe project . This project combined earth observations, global hydrological models and in-situ data to produce two water resources reanalysis (WRR) datasets: WRR-1 and WRR-2 (Table 1). WRR-2 used the Multi-Source Weighted-Ensemble Precipitation (MSWEP) dataset for model forcing (Beck et al., 2017). The MSWEP dataset has a spatial resolution of 0.25° and runs from 1979 up to and including 2014. Because the MSWEP precipitation dataset combines three different hydro-climatic data sources, has a higher spatial resolution, corrects for gauge under-catch and orographic effects (Beck et al., 2017) and because this dataset is known to give better results in the MCMB (Sterk & Rodriguez, 2017), WRR-2 was chosen over WRR-1 for this study. Since the MSWEP data is used as forcing for every model in WRR-2, the precipitation data is the same for every model.

The forcing data and all modelled data was obtained through the *thredds*-server of the E2O project in a NetCDF format for the period from January 1980 up to an including December 2014. Since the *thredds*-server provides the data for the entire world, the datasets were masked for the MCMB, using a shapefile of the MCMB from IDEAM. Next, the units were converted so the values would represent monthly data.

Table 1. Overview of main characteristics of the WFDEI and MSWEP datasets

Full name	Abbre-viation	Forcing for:	Data sources	Temporal coverage	Spatial resolution	Temporal resolution	Reference
WATCH Forcing Data applied to ERA-Interim data	WFDEI	WRR-1	• Gauge • Reanalysis	1979 – 2012	0.5°	3 hours	(Weedon et al., 2014)
Multi-Source Weighted-Ensemble Precipitation	MSWEP	WRR-2	• Gauge • Reanalysis • Satellite	1979 – 2015	0.25°	3 hours	(Beck et al., 2017)

3.2 Models

The eartH2Observe project (E2O) includes ten large scale hydrological models. For global water resources reanalysis, the use of multiple models can reduce the errors and uncertainties induced by individual models (Beven & Binely, 1992; Schellekens et al., 2017). These ten models produced, together with in-situ data and earth observation data, the WRR-2. From the ten models in E2O, six models are Global Hydrological Models (GHMs) and four models are Land Surface Models (LSMs) (Beck et al., 2017). GHMs simulate (sub-)surface water fluxes and storages, and solve only the water balance. LSMs simulate interactions between the soil, vegetation and atmosphere and are more physically based. They solve both the water and the energy balances (Bierkens, 2015). The six GHMs in eartH2Observe are: LISFLOOD, PCR-GLOBWB, SWBM, W3RA, WaterGAP3 and HBV-SIMREG. The four LSMs in eartH2Observe are: HTESSEL, JULES, ORCHIDEE and SURFEX. Four of the models in the eartH2Observe project have been calibrated in previous studies: HBV-SIMREG, SWBM, LISFLOOD and WaterGAP3 (Schellekens et al., 2017). Since none of the models is perfect, it is useful to define which models perform best in a certain-situation.

The four models were evaluated using the model products and five computed drought indices. Therefore, all parameters for calculating the drought indices needed to be included by the models used for the evaluation. For this study, the following models were used for evaluation because they include the necessary parameters: ORCHIDEE, HTESSEL, WaterGAP3 and PCR-GLOBWB (Table 2). Of these four models, ORCHIDEE and HTESSEL are LSMs and WaterGAP3 and PCR-GLOBWB are GHMs. Since the ORCHIDEE, HTESSEL and PCR-GLOBWB models were not

calibrated, this has to be accounted for when evaluating the model performances. Besides individual evaluation of the four models mentioned above, the ensemble means of these four models were also evaluated because the ensemble of multiple models can reduce the errors and uncertainties induced by individual models (Beven & Binely, 1992; Schellekens et al., 2017).

Table 2. Characteristics of the four models used for the evaluation (after: Schellekens et al., 2017)

Model	**ORCHIDEE**	**HTESSEL**	**WaterGAP3**	**PCR-GLOBWB**
Model type	LSM	LSM	GHM	GHM
Interception	Single reservoir Structural resistance to evaporation	Single reservoir, Potential evaporation	Single reservoir	Single layer, subject to open water evaporation
Evaporation	Bulk PET (Barella-Ortiz *et al.*, 2013)	Penman-Monteith	Priestley-Taylor	Penman-Monteith
Soil layers	11	9	1	2
Soil depth	2 m	3 m	1 m	1.5 m
Ground-water	Yes	No	Yes	Yes
Runoff	Green and Ampt infiltration	Saturation excess	Beta function	Saturation excess
Reservoirs/ lakes	No	No	Yes	Yes
Routing	Based on HydroSHED (Lehner and Grill, 2013)	CaMa-flood	Manning-Strickler	Travel time approach
Water use	Irrigation only	No	Yes	Yes
Time step	900 s energy balance, 3 hour routing	1 hour	1 day	1 day
Calibrated	No	No	Yes	No
Reference	(D'Orgeval et al., 2008)	(Balsamo et al., 2009)	(Döll et al., 2009, 2012; Flörke et al., 2013)	(van Beek et al., 2011; Wada et al., 2014)

3.2.1 ORCHIDEE (model provider: CNRS)

ORCHIDEE is a land surface model. As described by (D'Orgeval et al., 2008), it consists of three modules that operate at different scales. The first module is the energy balance module, and the second module is the hydrology module, which solves the hydrological balance with tile sizes that depend on the distribution of vegetation. This module splits surface infiltration and runoff using a time-splitting procedure; a method that allows for time steps smaller than 30 minutes. Vertical diffusion is solved using the Van Genuchten-Mualem parameters and the Fokker-Planck equation. 13 types of vegetation are included in the ORCHIDEE model, grouped into three main classes: bare soil, trees and grass and crop. For determining the root zone infiltration, the different root characteristics for different types of vegetation are accounted for. The total soil column has a constant depth of 2 meters. Potential evaporation is estimated using an unstressed surface energy balance (Barella-Ortiz et al., 2013). The third module is the river routing module. Rivers, surface and subsurface runoff are routed through three different reservoirs, each with their own residence time (Figure 7). Runoff is calculated by the Green and Ampt infiltration equation and is used as input for the fast reservoir (indicated by V2 in Figure 7). The slow reservoir (indicated by V3 in Figure 7) is fed by drainage and together with the fast reservoir it feeds the downstream stream reservoir. Irrigation is also included in this model (Schellekens et al., 2017).

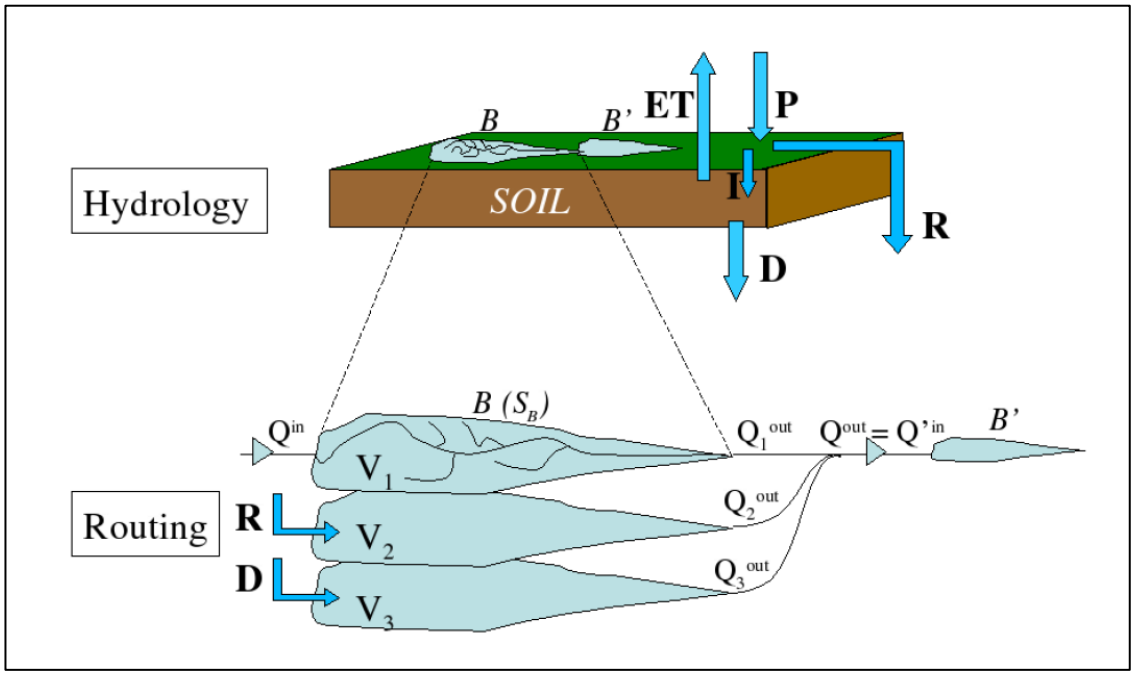

Figure 7. Schematic representation of the ORCHIDEE river routing scheme (D'Orgeval et al., 2008)

3.2.2 HTESSEL (model provider: ECMWF)

The Hydrology Tiled ECMWF Scheme for Surface Exchanges over Land (HTESSEL) is a land surface model with a single reservoir. The model is not calibrated and uses a priori parameter estimations. For determining bare ground evaporation it uses the Penman-Monteith equation (Schellekens et al., 2017). As the 'T' for 'Tiled' in the name 'HTESSEL' indicates, tiles are included in the model. These tiles consist of bare ground, low and high vegetation, intercepted water, shaded and exposed snow and open and frozen water (Figure 8). Each tile has its own water and energy balances. There are nine soil layers with a total depth of three meters, and additionally a layer of snow can be added. By using multiple soil layers, processes in the soil with different timescales between one day and one year can be represented. Darcy's law is used for sub-surface processes. For hydraulic conductivity and diffusivity, the Van Genuchten equation is used. Precipitation can be intercepted until saturation occurs, and once saturation occurs, water will fall as throughfall and is divided over surface runoff and infiltration. This is done according to a variable infiltration capacity based on soil type and local topography. Six soil texture classes based on the Food and Agriculture Organization from 2003 are used: coarse, medium, medium fine, fine, very fine and organic. Each soil class has its own volumetric moisture content for saturation, field capacity and wilting point (Balsamo et al., 2009). The routing is performed by the catchment-based, macro-scale floodplain model CaMa. The CaMa river routing model calculates the floodplain inundation dynamics by parameterizing the subgrid-scale topography of a floodplain (Yamazaki et al., 2011). The HTESSEL model is the only model of the four models used in this study that does not include groundwater (Schellekens et al., 2017).

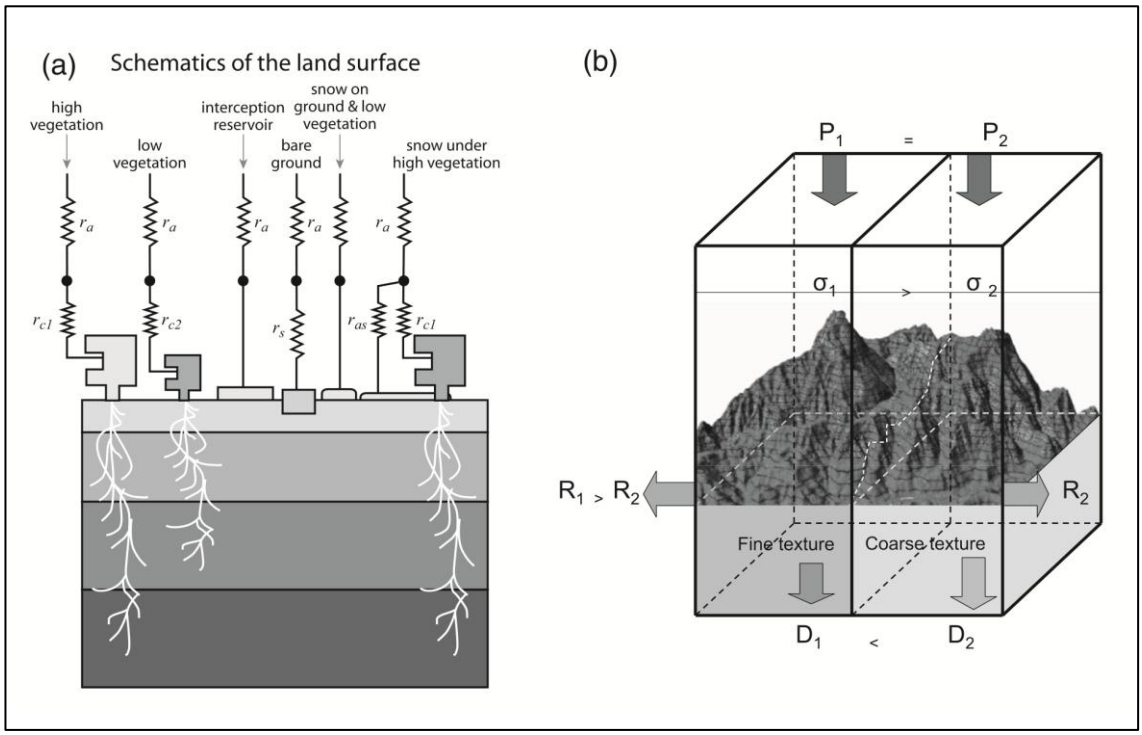

Figure 8. Schematic representation of the HTESSEL model (Balsamo et al., 2009)

3.2.3 WaterGAP3 (model provider: Kassel University)

Water Global Assessment and Prognosis 3 (WaterGAP3) is the only calibrated model in this study. It is a global hydrological model taking into account anthropogenic influences such as water abstractions and dams. It is grid based and the hydrological cycle is represented as a sequence of storage compartments, such as canopy, snowpack, soil, groundwater and surface water (Figure 9). The maximum canopy storage depends on the leaf area index. The Priestley-Taylor equation is used to determine evaporation. The snowpack is represented as a single layer and is modelled using a degree-day approach. Infiltration is modelled by representing the soil as a single layer of 1 m depth, where the infiltration depends on the effective precipitation, soil texture, soil saturation and land cover. Water that does not form runoff, will be infiltrated into the soil storage.

WaterGAP3 consists of both the WaterGAP Global Hydrology Model (WGHM) and five separate water abstraction models for irrigation, livestock, households, thermal power plants and manufacturing (Döll et al., 2009, 2012).

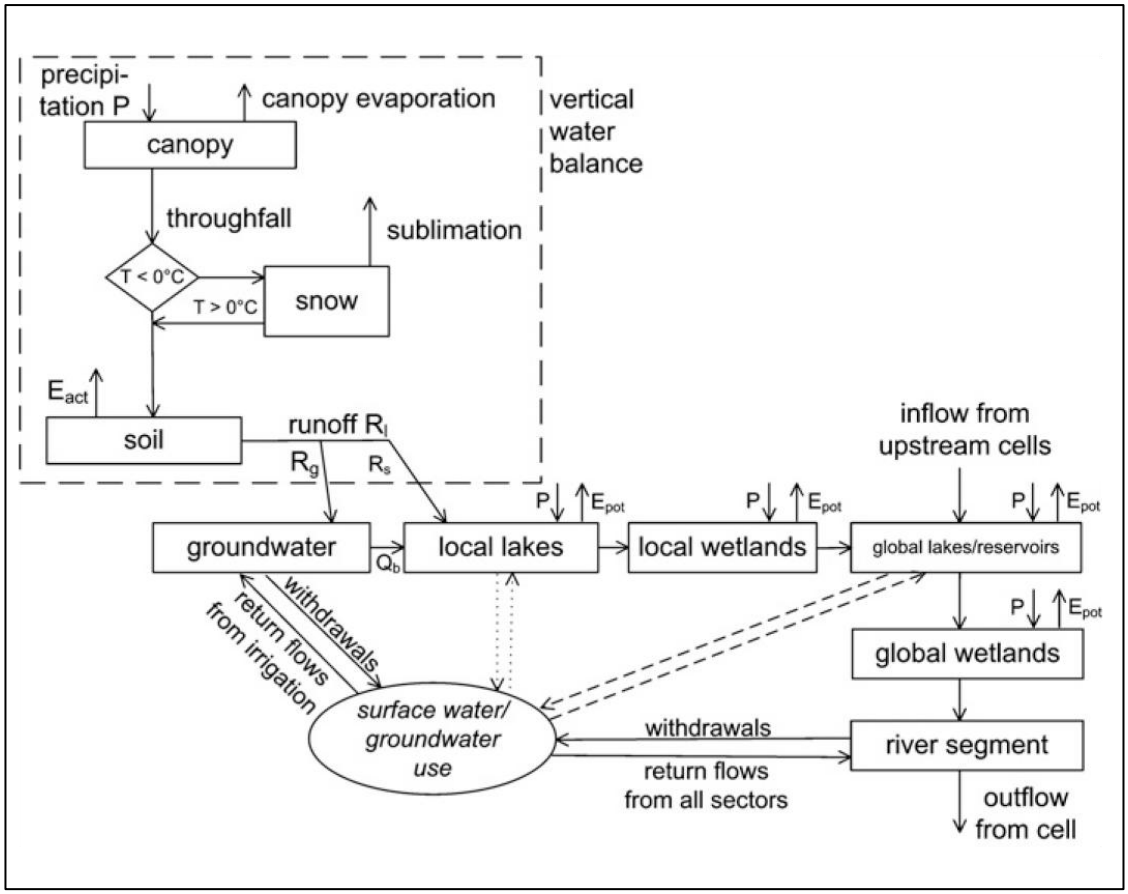

Figure 9. Schematic representation of the WaterGAP3 model (Döll et al., 2012)

3.2.4 PCR-GLOBWB (model provider: Utrecht University)

PCR-GLOBWB (PCRaster GLOBal Water Balance model) is a large scale, grid based, global hydrological model (van Beek & Bierkens, 2009). Each cell is composed of a soil layer, two underlying storage layers (with maximum depths of 0.3 and 1.2 respectively) and a groundwater layer (Figure 10). Small vegetation extracts water from only the upper soil layer, while large vegetation also extracts water from the bottom layer. PCR-GLOBWB is a leaky-bucket type model, meaning that the leaky-bucket principle is applied on each cell to determine the transport of moisture between the layers and between the surface and the atmosphere. Direct runoff, interflow and base flow are included even as canopy interception and snow storage. Precipitation falls as rain when temperatures are above freezing point, and as snow otherwise. Snow melt is modelled based on temperature, using the HBV model (Bergstrom, 1976). For surface flow, the kinematic wave approximation is used. Potential evapotranspiration can come from either canopies or from the bare soil. Actual evaporation can be assigned, or calculated by the model from the potential evaporation and the moisture content of the soil, using the Penman-Monteith equation. PCR-GLOBWB was not calibrated for eartH2Observe WRR-2.

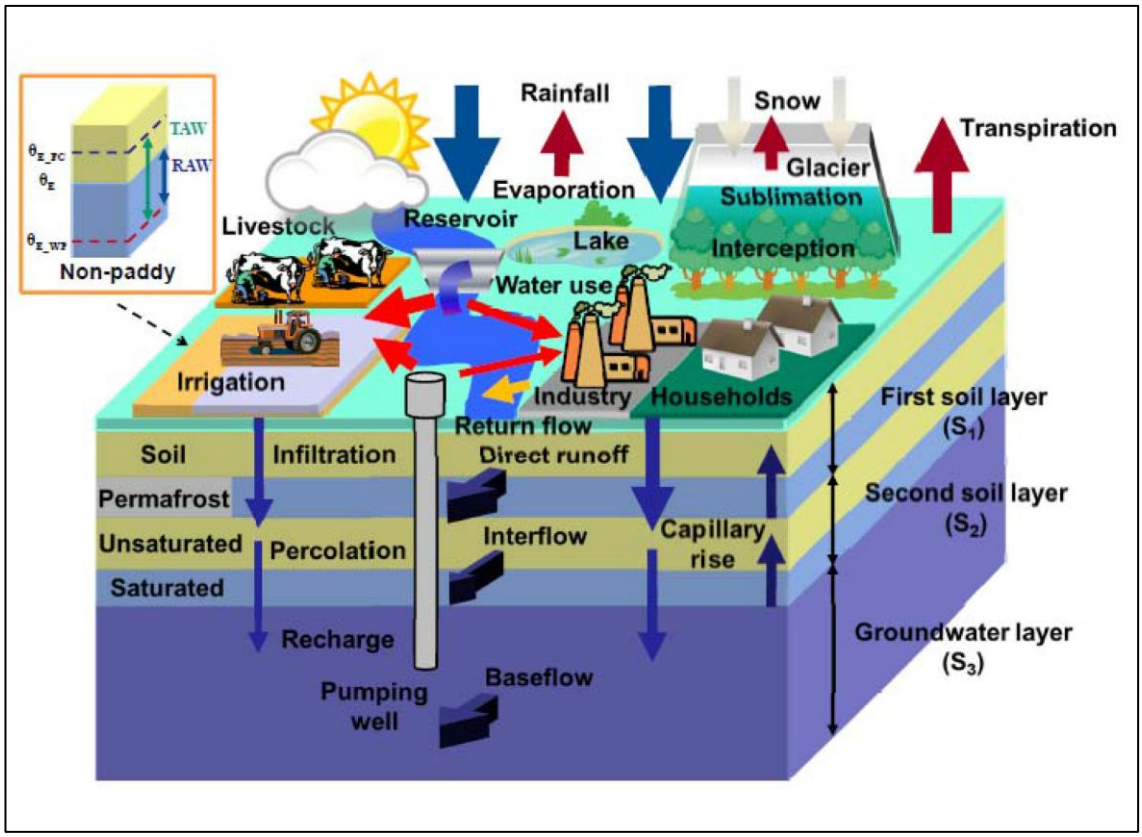

Figure 10. Schematic representation of the PCR-GLOBWB model (van Beek and Bierkens, 2009)

3.3 Model evaluation

There are several performance measures to quantify the performance of hydrological models, of which the mean squared error (MSE), the Nash-Sutcliffe efficiency (NSE) and the Kling-Gupta efficiency (KGE) are the most widely used. Although the MSE and NSE are used often, they have the disadvantage that the mean of the observations is used as a baseline. Using the mean can lead to an overestimation of the model skill when there is a strong seasonality in the dataset (Gupta et al., 2009), or an underestimation when the mean of the modelled data and observed data differ, while the standard deviations are very similar. Since the latter was likely to be the case in this research, the model performance was quantified using the Kling-Gupta efficiency (KGE). This performance criteria has the advantage that the Pearson correlation coefficient (r), the bias (β), and the variability ratio (α) are included. The KGE represents a Euclidian distance in a 3 dimensional space and is defined as follows (Gupta et al., 2009):

$$KGE = 1 - ED \quad \text{(eq. 1)}$$

$$ED = \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \quad \text{(eq. 2)}$$

$$r = \frac{Cov_{mo}}{\sigma_m \cdot \sigma_o} \quad \text{(eq. 3)}$$

$$\alpha = \frac{\sigma_m}{\sigma_o} \quad \text{(eq. 4)}$$

$$\beta = \frac{\mu_m}{\mu_o} \quad \text{(eq. 5)}$$

Cov_{mo} is the covariance between the modelled and the observed values; σ_m is the standard deviation of the modelled values and σ_o is the standard deviation of the observed values; and μ_m and μ_o are the mean of the modelled and observed values, respectively.

The KGE values range from $-\infty$ to 1, where 1 indicates a perfect model performance of the simulated dataset with regard to the observed dataset. A perfect Pearson correlation has a value of -1 for a perfect negative correlation and a value of 1 for a perfect positive correlation. For both the Beta and Alpha component, a value of 1 is ideal, but values can range between 0 and ∞.

For each model product of the four models, graphs for the time series were made. To illustrate the spatial differences, the time series of the model products are given for five cells in the MCMB (Figure 11). The particular five cells are chosen in order represent both upstream and downstream locations, and different proximities to major river branches and mountain ranges. For precipitation, potential evapotranspiration, actual evapotranspiration and root zone soil moisture, the following five locations were selected:

- Location A (cell 66) is located at (-74.875° W, 10.125° N), at a downstream part of the Magdalena-Cauca river.
- Location B (cell 295) is located in the central basin of the MCMB, at (-74.625° W, 7.125° N).

- Location C (cell 360) is located in the Cordilleras Oriental, the eastern mountain range forming the eastern border of the MCMB, at (-72.625° W, 6.375° N).
- Location D (cell 428) is located half way on the Magdalena river, at (-74.625° W, 5.375° N).
- Location E (cell 535) is located at (-76.375° W, 3.875° N) close to an upstream part of the Cauca river and close to the Cordilleras Central.

For the discharge time series, location B is replaced by location F and location E is replaced by location G, because location B and E are not located exactly on a major river branch.

- Location F (cell 164) is located at (-74.125° W, 8.875° N) at a downstream part of the Magdalena river, in a swampy area not far upstream from the meeting point of the Magdalena and Cauca rivers.
- Location G (cell 591) is located at (-76.625° W, 3.125° N), where conditions are similar to the conditions of location E.

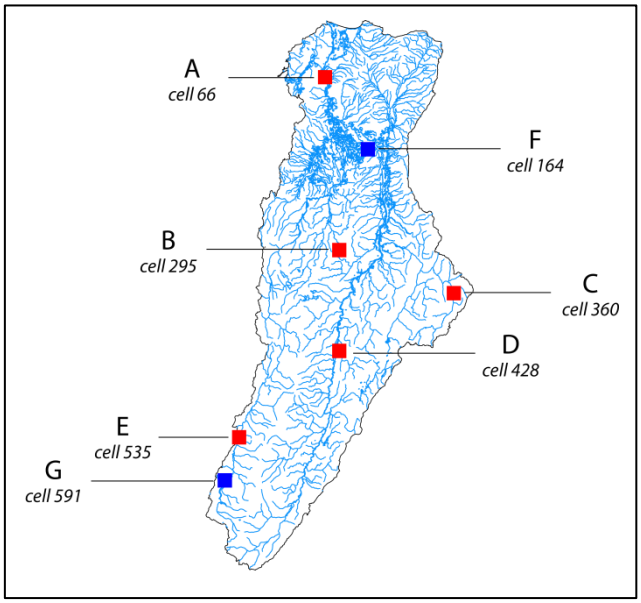

Figure 11. MCMB with the locations selected for analysis and the main rivers and wetlands (in light blue)

For precipitation, reference evapotranspiration and discharge, in-situ data was available. This allows for the quantification of the performance using the in-situ observations from IDEAM as benchmark. For precipitation, the KGE for the MSWEP forcing data was calculated using the in-situ precipitation. For potential evapotranspiration the KGE for the potential evapotranspiration data of the four models and the ensemble mean was calculated using the reference evapotranspiration. For discharge, the KGE for the discharge data of the three models and the ensemble mean was calculated using the in-situ discharge data. The calculations for precipitation and potential evapotranspiration were done for every 0.25° cell in the MCMB. The KGE of discharge was only calculated for cells in the catchment for which in-situ data was available. The spread of KGE values and of the Pearson's correlation, bias and variability ratio are illustrated in plots of the Cumulative Distribution Functions (CDFs). The CDFs give the frequency that the KGE (or one of its

components) will be equal or below a specific value. The spatial distribution of the KGE values is illustrated by raster maps of the MCMB.

For actual evapotranspiration and root zone soil moisture, in-situ data was not available, therefore, the KGE could not be computed. Box plots are given showing the distribution of all values in the population, over both space and time. The box plots give information on the median, indicated by the thick line in the middle of the box, standard deviation, indicated by the box, and the lower and upper 1.5 interquartile range, indicated by the whiskers. Outliers are values lower or higher than the lower and upper 1.5 interquartile range, respectively, and are indicated by small circles.

3.4 Drought indices

To give an overview of drought propagation through the hydrological system of the MCMB, different drought indices representing different types of droughts were used for the drought quantification in this study. For this study, five different drought indices were calculated: the Standardized Precipitation Index (SPI), the Standardized Precipitation Evapotranspiration Index (SPEI), the Evapotranspiration Deficit Index (ETDI), the Soil Moisture Deficit Index (SMDI) and the Streamflow Drought Index (SDI) (Table 3). The data needed for computation of these drought indices came from the eartH2Observe WRR-2. The indices were calculated for each of the four models and for the ensemble mean on a cell by cell basis, allowing to observe differences between models and to analyse spatial trends through time.

Table 3. The drought indices used in this study. Input data refers to Precipitation (P) from the MSWEP forcing data, Potential Evapotranspiration (PET), Actual Evapotranspiration (AET), Root Zone Soil Moisture (RM) and Discharge (Q).

Index	Type	Input data	Period	Abbr.	Reference
Standardized Precipitation Index	Meteorological	P	1, 3, 6, 12, 24, 36	SPI	(McKee et al., 1993)
Standardized Precipitation Evapotranspiration Index	Meteorological	P, PET	1, 3, 6, 12, 24, 36	SPEI	(Vicente-Serrano et al., 2010)
Evapotranspiration Deficit Index	Agricultural	PET, AET	1	ETDI	(Narasimhan & Srinivasan, 2005)
Soil Moisture Deficit Index	Agricultural	RM	1	SMDI	(Narasimhan & Srinivasan, 2005)
Streamflow Drought Index	Hydrological	Q	1	SDI	(Nalbantis & Tsakiris, 2009)

3.4.1 SPI

The SPI is based solely on precipitation and is a meteorological drought index (Zargar et al., 2011). The SPI was developed by McKee et al. (1993). This index compares observed rainfall with a rainfall probability distribution function. First, a monthly precipitation dataset of a period of preferably 30 years or longer is standardized for a certain period. This makes it possible to compare the index over different regions and points in time. This period can be 1, 3, 6, 12, 24 or 36 month(s), where each period represents different water sources affected by precipitation deficits. Each month, a new value is determined using the previous 1, 3, 6, 12, 24 or 36 month(s). The standardized precipitation is then fitted to a gamma function to determine the correlation between precipitation and probability. This probability is used to determine the precipitation deviation for a normally distributed probability density with a mean equal to zero and a standard deviation equal to unity. This is the SPI (McKee et al., 1993). Different intensities of drought can be defined based on the SPI, ranging from values just below zero (mild drought) to values below –2.00 (extreme drought) (Table 4). McKee et al. (1993) state an event with a negative SPI can be regarded as a drought event when the SPI falls below -1.0. The start of the event is the moment the SPI becomes negative and the event ends when the SPI becomes positive again.

Table 4. SPI values and corresponding drought categories (McKee et al., 1993)

SPI values	**Drought category**
0 to -0.99	Mild drought
-1.00 to -1.49	Moderate drought
-1.50 to -1.99	Severe drought
≤ -2.00	Extreme drought

An advantage is that the SPI can give information on drought on multiple time scales, depending on the time interval chosen for computation. When propagating from the 1-month to the 36-month SPI via the 3-, 6-, 12- and 24-month SPI, different aspects of the hydrological system are represented. According to McKee et al. (1993), these timescales “represent arbitrary but typical time scales for precipitation deficits to affect the five types of usable water sources”. The five usable water sources “include soil moisture, ground water, snowpack, streamflow and reservoir storage”. Zargar et al. (2011), state that the 1- and 3-month SPI reflect short-term conditions, such as seasonal estimation of precipitation and moisture conditions. A disadvantage of the SPI is that it only takes precipitation into account. Other factors that can contribute to a drought, e.g. temperature, evapotranspiration and soil moisture, are neglected (Vicente-Serrano et al., 2010).

The precipitation data is supplied by the MSWEP dataset, the forcing dataset for all models in the WRR-2 of the eartH2Observe project. Therefore, the SPI is the same for all models.

3.4.2 SPEI

The SPEI is also a meteorological drought index (Zargar et al., 2011). This drought index takes the role of precipitation and potential evapotranspiration into account. The SPEI is also sensitive to long term climate change because temperature is indirectly taken into account by the potential evapotranspiration, a factor which the SPI is not sensitive to (Vicente-Serrano et al., 2010). The method for computation is similar to the method for calculating the SPI, only instead of solely taking the precipitation, the difference between the precipitation and the potential evapotranspiration (PET) per month *i* is taken:

$$D_i = P_i - PET_i \quad \text{(eq. 6)}$$

D_i is a simple measure of the water surplus or deficit for a specific month. Determination of the potential evapotranspiration is a challenge, because additional data on wind speed, solar radiation, temperature and relative humidity is necessary. Empirical equations can be used to determine the potential evapotranspiration. According to (Mavromatis, 2007), the method chosen to determine the potential evapotranspiration does not make a significant difference in the drought index product.

3.4.3 ETDI

The evapotranspiration deficit index (ETDI) is an agricultural drought index developed by (Narasimhan & Srinivasan, 2005). For determining the ETDI, the water stress ratio is calculated first:

$$WS = \frac{PET - AET}{PET} \qquad \text{(eq. 7)}$$

WS is the monthly water stress ratio, PET the monthly potential evapotranspiration and AET the monthly actual evapotranspiration. The long term water stress ratio for each month in a year is determined by taking the median of the water stress ratio for that month during a long term period of multiple decades (e.g. 30 years). The maximum and minimum water stress ratio are also obtained for each month of the year from the long term period dataset. The monthly water stress anomaly (WSA) is then calculated by subtracting the water stress ratio from the median long term water stress, and dividing this by the median minus the minimum in case the water stress ratio is equal to the median water stress ratio, or by dividing by the maximum minus the median in case the soil water is larger than the median water stress ratio. The result is multiplied by 100 to give the WSA in percentages. By this method, the seasonality is removed from the water stress anomaly. Eventually, the ETDI is obtained by equation (8):

$$ETDI_t = 0.5 * ETDI_{t-1} + \frac{WSA_t}{50} \qquad \text{(eq. 8)}$$

3.4.4 SMDI

The SMDI was developed by Narasimhan & Srinivasan (2005) as well. It is an agricultural drought index that makes use of the amount of water in the root zone. The amount of moisture in the soil is, among other things, influenced by temperature, precipitation and evapotranspiration. The type of land use and land cover is of importance, since some types of vegetation demand more water and have a higher transpiration rate in contrast to other crops, which will influence the amount of water in the soil. The computation procedure is very similar to the computation procedure of the ETDI, only instead of the WS, the monthly available soil water in the root zone is used. The long term soil moisture for each month is determined by taking the median of available soil water for that month during a long term period of multiple decades. The maximum and minimum soil water are also obtained for each week from the long term period dataset. The soil moisture deficit (SMD) is then calculated by subtracting the median soil water from the monthly soil water, and dividing this by the median minus the minimum in case the soil water is equal to the mean soil water, or by dividing by the maximum minus the median in case the soil water is larger than the median soil water. The result is multiplied by 100. The SMDI is obtained by equation (9):

$$SMDI_t = 0.5 * SMDI_{t-1} + \frac{SMD_t}{50} \qquad \text{(eq. 9)}$$

3.4.5 SDI

The Streamflow Drought Index (SDI) is a hydrological drought index developed by Nalbantis & Tsakiris (2009). Hydrological drought is concerned with the amount of water in water bodies, such as rivers and lakes. Hydrological drought indices therefore make use of discharge data. To calculate the SDI, monthly streamflow volumes are needed. From the monthly streamflow volumes, the cumulative streamflow is computed for a reference period. The SDI is computed by dividing the cumulative streamflow minus the mean cumulative streamflow by the standard deviation of the cumulative streamflow. The SDI is then equal to the standardized streamflow volume. Next, the SDI can be transformed into a normal distribution, by taking the natural logarithms of cumulative streamflow.

3.5 Drought analysis

Spatial differences between drought index time series are relatively small due to the computation methods of the drought indices, which often involve a form of standardization or normalization. Therefore, for each drought index, one time series per model is given, showing per month the average drought index values of all cells in the catchment.

To illustrate the spatial distribution of the frequency of the occurrences of dry months in the MCMB, the ratio of dry months to all months was computed: the number of months with a drought index value below -1 were divided by the total number of months for which a drought

index value is available. Note that for drought indices with a larger aggregation period, the total number of months with available data is smaller.

3.6 Correlation between droughts and ENSO

The Multivariate ENSO Index (MEI) was used in this study to quantify ENSO. The data for the MEI were obtained from the Climate Explorer of the Royal Dutch Meteorological Institute (KNMI). Because the MEI includes six atmospheric and oceanographic parameters, it is regarded as a very comprehensive index (Wolter & Timlin, 2011). Positive MEI values coincide with El Niño events and negative MEI values coincide with La Niña events. Values above an index of 2 or below an index of -2 can be regarded as a strong El Niño or La Niña event, respectively. The strongest El Niño events in the period between 1980 and 2015 occurred in the winters of 1982-1983 and 1997-1998 and two smaller El Niño events occurred in the winters of 1986-1987 and 1991-1992 (IDEAM, 2014), as can be observed in Figure 12. El Niño events can often be correlated with droughts in Colombia (Hoyos et al., 2013, 2017) (Poveda et al., 2007). A strong negative correlation between a drought index and the MEI indicates a strong positive correlation between droughts and El Niño.

The correlation between droughts and ENSO is determined using the Pearson's correlation between the drought indices and the MEI (equation 3). The Pearson correlation coefficient was computed for each cell in the MCMB and for every drought index. For each drought index, a CDF is given showing the distribution of the Pearson's correlation values. Also, a scatter plot is given showing the monthly spatial means of the drought indices plotted against the monthly MEI values. The spatial distribution of the Pearson's correlation is illustrated by raster maps of the MCMB.

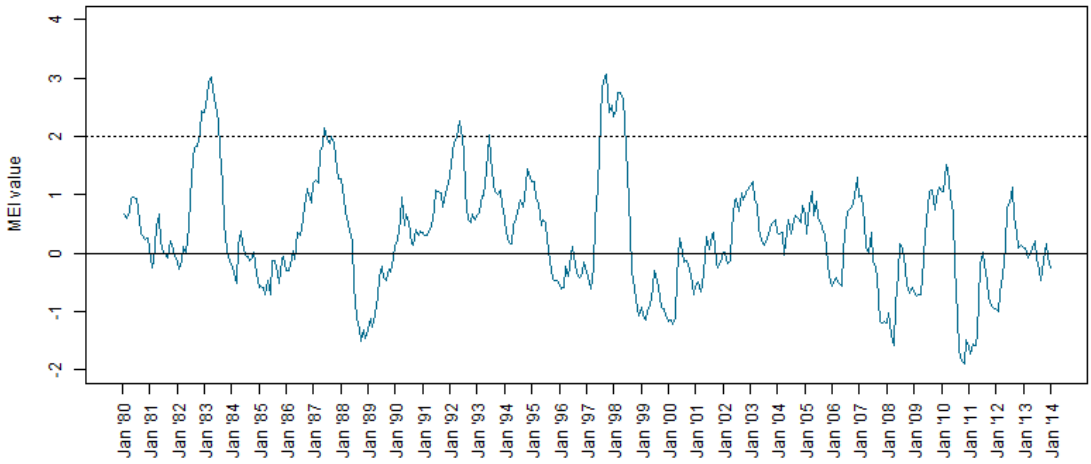

Figure 12. Multivariate ENSO index (KNMI, 2019).

4. Results

4.1 Performance of the selected models for the MCMB

4.1.1 Precipitation

The precipitation data is equal to the MSWEP forcing data and is therefore the same for all four models. The MSWEP dataset combines gauge, reanalysis and satellite data and corrects for gauge under-catch and orographic effects. From the time series it is evident that the MSWEP forcing data performs well (Figure 13). An annual cycle can be observed with two peaks of high precipitation per year: one around May and one around October. This corresponds with the timing of wet seasons due to the shifting of the ITCZ, as described by Restrepo et al. (2006). Precipitation amounts vary per location. Locations B and D show the highest amount of precipitation: maxima are around 400 mm per month. These two cells are located at lower elevations in the basin of the catchment. Locations A, C and E have almost halve of this precipitation amount (around 200 mm per month), and are located on the mountain ranges in the west and east (location C and E, respectively) and more north, near the coast (location A).

The high KGE values for most cells in the catchment indicate a good performance (Figure 14). Locations A, B, C, D and E have KGE values of 0.86, 0.78, 0.90, 0.94 and 0.93, respectively. 89% of the cells in the catchment have a KGE of 0.50 or higher, and 62% have a KGE of 0.75 or higher. Some cells in the north and along the border of the MCMB show lower KGE values (Figure 15). The KGE comprises three components: the Pearson's correlation (eq. 3), bias (eq. 4) and the variability ratio (eq. 5), which each have the same weight in the KGE computation. The Pearson's correlation is strong for all cells in the catchment (99% of the cells have a Pearson's correlation above 0.75), indicating good timing of the minima and maxima of the MWSEP dataset. Therefore, the Pearson's correlation is not responsible for the lower KGE values. The bias and variability ratio are good for most cells in the catchment, except for cells in the north and along the border. This indicates that the mean and the standard deviation of the MSWEP data and in-situ data do not agree well for these cells in the catchment, which results in lower KGE values in these areas.

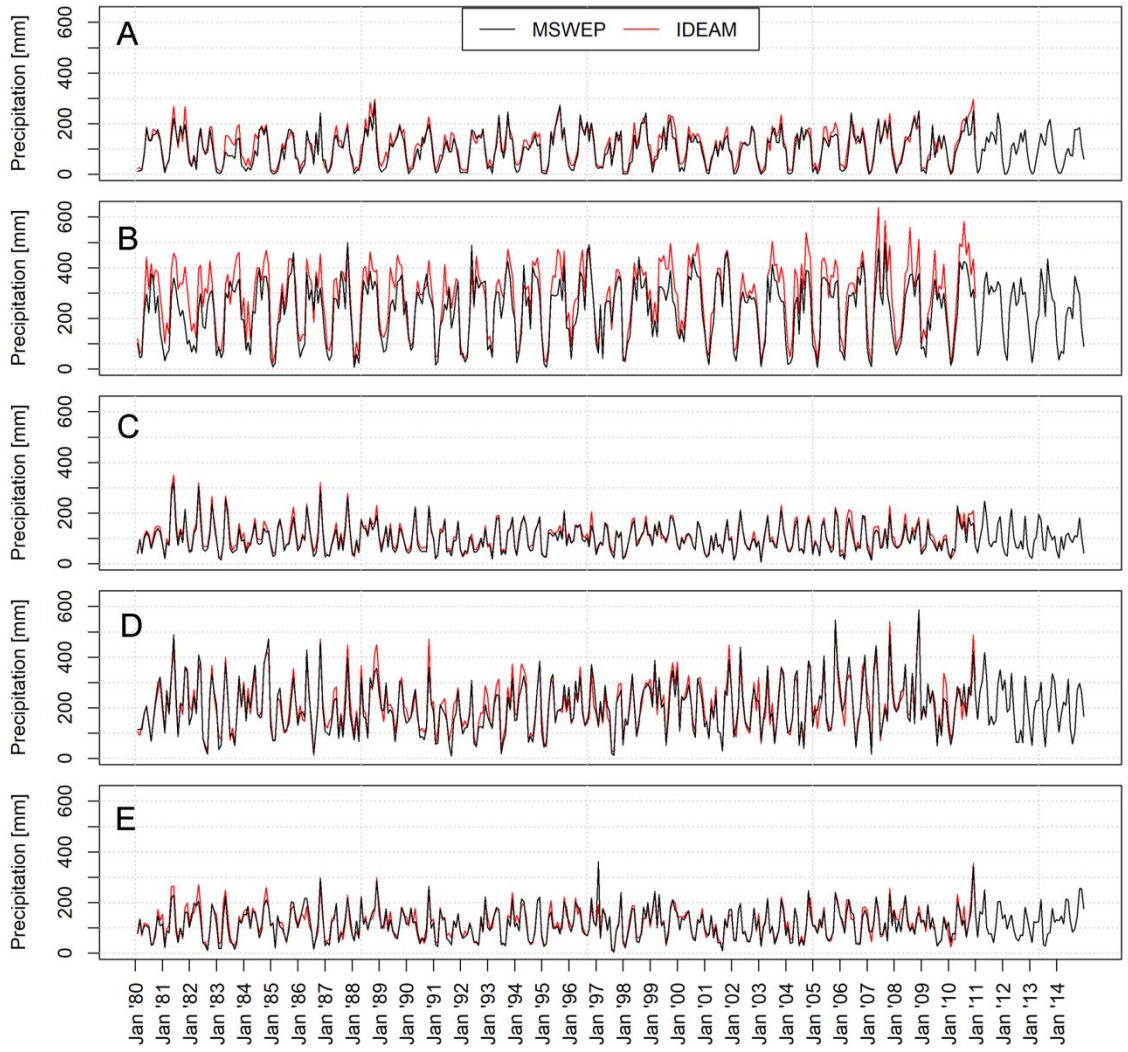

Figure 13. Precipitation of MSWEP forcing data and IDEAM for cell 66 (A), 295 (B), 360 (C), 428 (D) and 535 (E) in the MCMB.

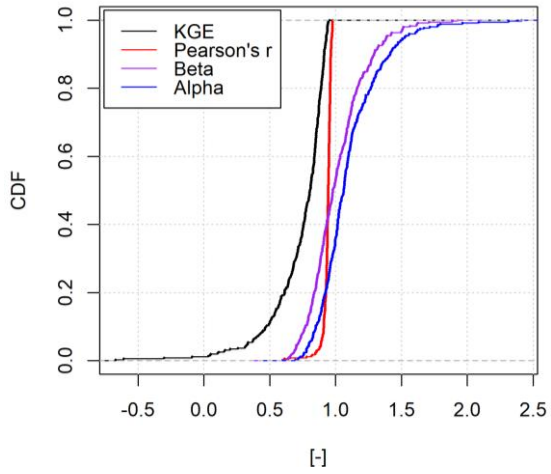

Figure 14. CDF plot of KGE, Pearson's correlation, bias and variability ratio for the MSWEP dataset

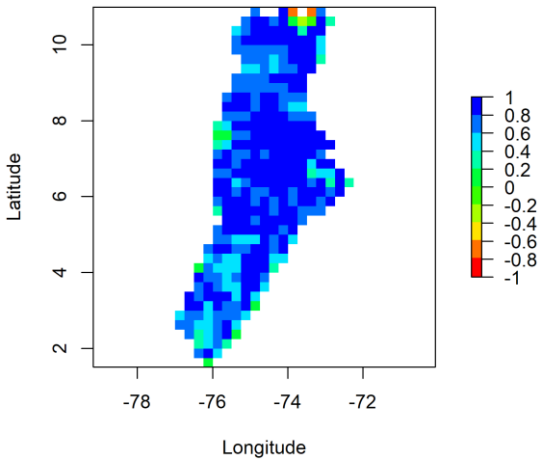

Figure 15. Spatial distribution of the KGE for the MSWEP dataset over the MCMB

4.1.2 Potential evapotranspiration

The potential evapotranspiration is negative, because the fluxes are directed downward in eartH2Observe. The KGE was calculated for the potential evapotranspiration of the four models and the ensemble mean using the reference evapotranspiration (ET_0) based on in-situ temperature data. The potential evapotranspiration is the maximum achievable evapotranspiration for a specific crop type. The reference evapotranspiration is calculated for a reference grass crop (Thornthwaite, 1948). If a model uses a crop type that is different from the reference grass crop, this will result in a different mean potential evapotranspiration. Besides this, the models use different methods to calculate the potential evapotranspiration. HTESSEL and PCR-GLOBWB use the Penman-Monteith equation (van Beek & Bierkens, 2009; Schellekens et al., 2017), whereas ORCHIDEE uses the Bulk method (Barella-Ortiz et al., 2013) and WaterGAP3 uses the Priestley-Taylor equation (Schellekens et al., 2017).

For the four most northern locations, the models – especially the HTESSEL model – show larger potential evapotranspiration amounts and a larger standard deviation than the reference potential evapotranspiration (Figure 16). The models disagree strongly on the maximum potential evapotranspiration, but agree relatively well on the minimal evapotranspiration. At location E, the potential evapotranspiration amounts are more in line with the reference evapotranspiration in terms of magnitude, and the different models agree well. Again, the annual cycle with two minima and two maxima per year can be observed.

Table 5. KGE values for ORCHIDEE, HTESSEL, WaterGAP3 and PCR-GLOBWB and the ensemble mean of these four models, for cells 66 (A), 295 (B), 360 (C), 428 (D) and 535 (E) in the MCMB

	A	**B**	**C**	**D**	**E**
ORCHIDEE	-1.99	-0.76	-1.33	-0.78	-2.00
HTESSEL	-4.84	-4.81	-8.36	-7.18	-1.52
WaterGAP3	0.49	-0.49	0.15	-0.46	-0.93
PCR-GLOBWB	-0.04	0.21	-0.16	-0.06	-0.56
Ensemble mean	-1.51	-1.30	-2.21	-2.01	-1.22

The KGE values of the time series shown in Figure 16 is given in Table 5. For the potential evapotranspiration, the KGE is generally lower than for the precipitation: no model has cells with a KGE over 0.75 (Figure 17). This low KGE value is mainly influenced by a large variability ratio. Especially the HTESSEL model has much larger deviations from the mean and is more negative compared to the other models and the in-situ data. The GHMs (PCR-GLOBWB and WaterGAP3) perform the best: 12% and 1% of the cells in the catchment have a KGE of 0.50 or higher, respectively. The Pearson's correlations are moderately strong, especially for the WaterGAP3 and PCR-GLOBWB models. For the ORCHIDEE, HTESSEL, WaterGAP3 and PCR-GLOBWB models and the ensemble mean, 35%, 24%, 73%, 69%, and 38% of the cells have Pearson's correlations of 0.50 or higher, respectively.

The ORCHIDEE, WaterGAP3 and PCR-GLOBWB models show a spatial pattern of KGE values where the KGE is best just above the middle of the catchment, and worst in the upper rim and in the south and east of the catchment (Figure 18). The HTESSEL model shows a different pattern, where the KGE is best at the south and north eastern border. The HTESSEL model shows a highly irregular pattern with very poor KGE values spread throughout the catchment.

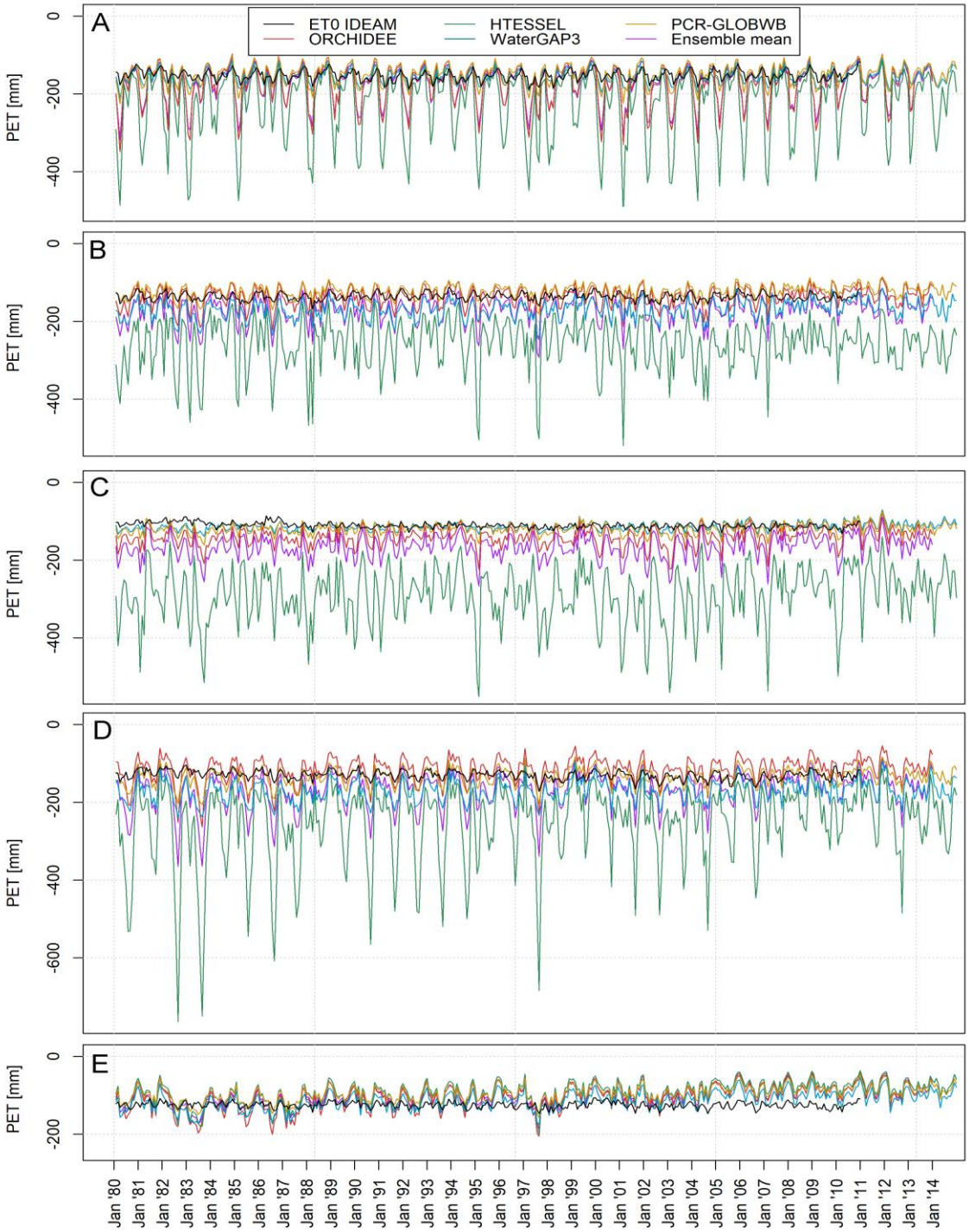

Figure 16. Reference evapotranspiration (ET0) of IDEAM and potential evapotranspiration for ORCHIDEE, HTESSEL, WaterGAP3 and PCR-GLOBWB models and the ensemble mean of these four models, for cell 66 (A), 295 (B), 360 (C), 428 (D) and 535 (E) in the MCMB

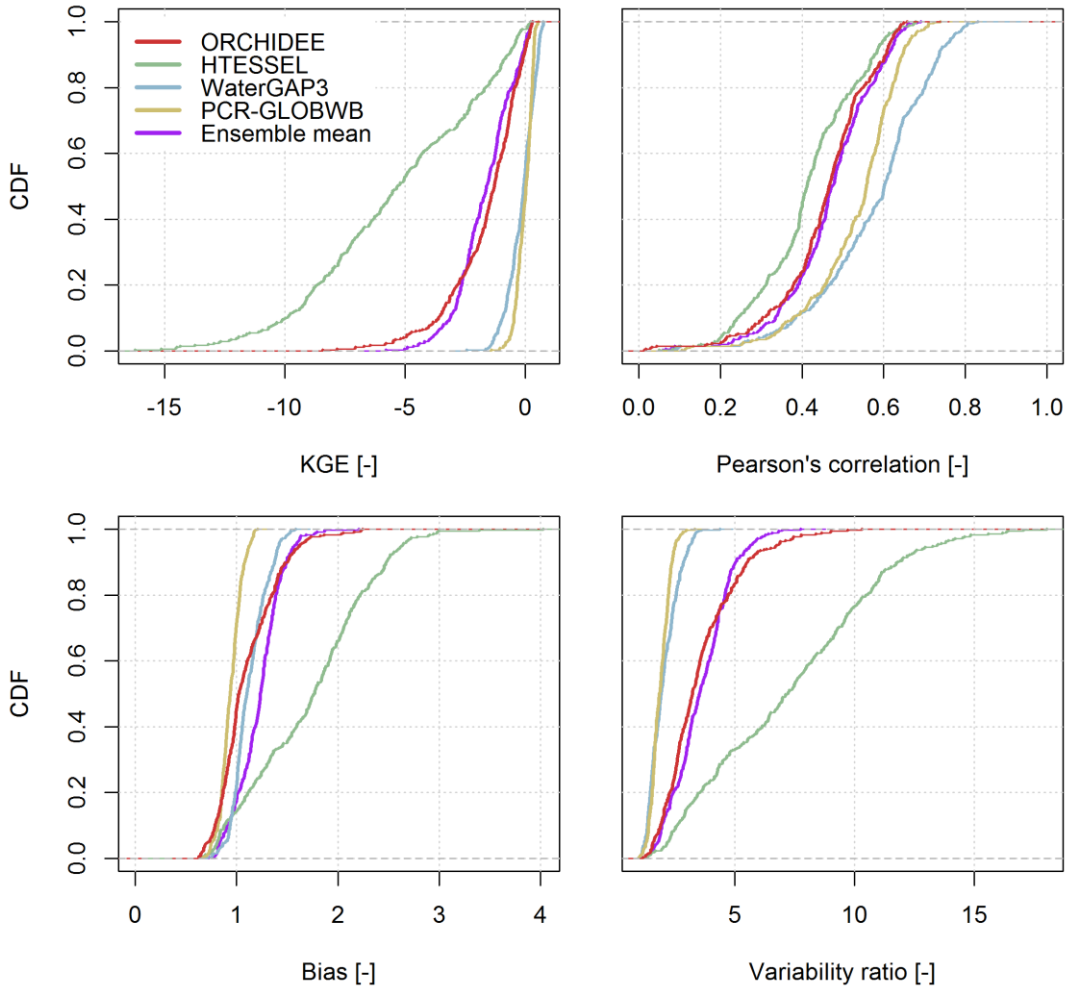

Figure 17. KGE (A), Pearson's correlation (B), bias (C) and variability ratio (D) of the potential evapotranspiration for the ORCHIDEE, HTESSEL, WaterGAP3 and PCR-GLOBWB models and the ensemble mean of these four models

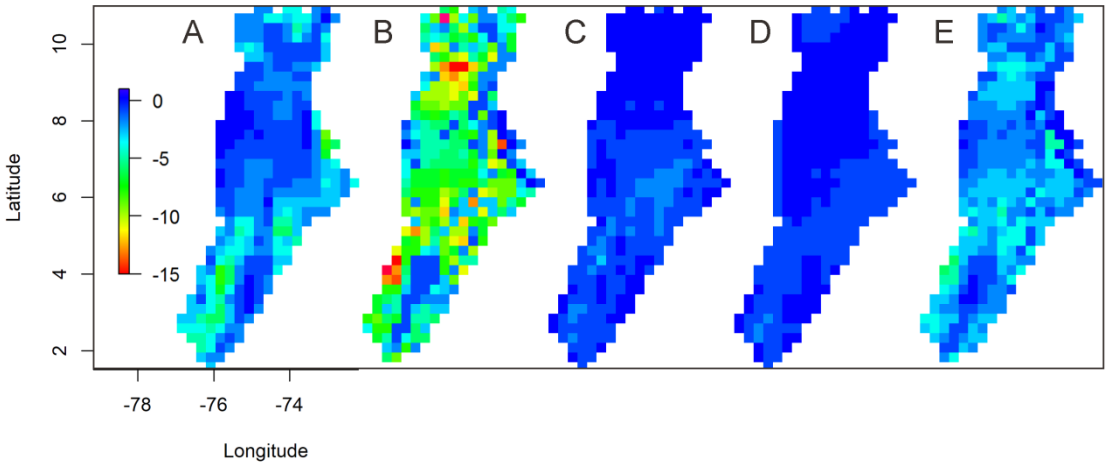

Figure 18. Spatial distribution of the KGEs of the potential evapotranspiration for the ORCHIDEE (A), HTESSEL (B), WaterGAP3 (C) and PCR-GLOBWB (D) models and the ensemble mean of these four models (E)

4.1.3 Actual evapotranspiration

For actual evapotranspiration, the values are also negative because of the downward directed fluxes in eartH2Observe. Because there was no in-situ data available for the actual evapotranspiration, a performance test was not possible. To illustrate the differences between the models, model statistics will be discussed briefly.

The standard deviations of the actual evapotranspiration of the WaterGAP3 and PCR-GLOBWB models show a larger standard deviation compared to the ORCHIDEE and HTESSEL models (Figure 19). The timing of maxima and minima is sometimes off for the WaterGAP3 model compared to the other models, which correspond relatively well with each other (Figure 20).

Between May and October, a maximum of evapotranspiration occurs, often with a small minimum in the middle. This period corresponds with the main growing season of rice in Colombia (FAO, 2002). Around February, a minimum in total evaporation can be observed, which corresponds with the end and harvesting of the second growing season of rice.

All models show a similar spatial pattern throughout the catchment. More evapotranspiration occurs in the centre of the catchment and along the Magdalena river towards the southeast. In the southeast of the basin, rice paddies and other forms of agriculture are located, and there is dense vegetation due to a humid climate. Less actual evapotranspiration occurs along the borders of the catchment, where the mountain ranges are located.

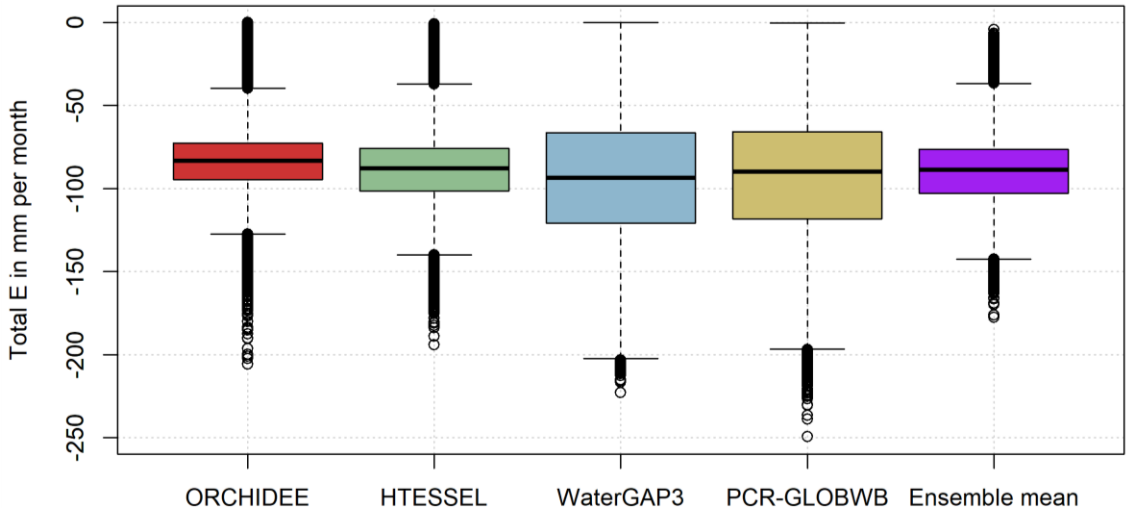

Figure 19. Box plots of actual evapotranspiration values for all cells in the MCMB and for all months from January 1980 up to and including December 2014, for the ORCHIDEE, HTESSEL, WaterGAP3 and PCR-GLOBWB models and the ensemble mean of these four models

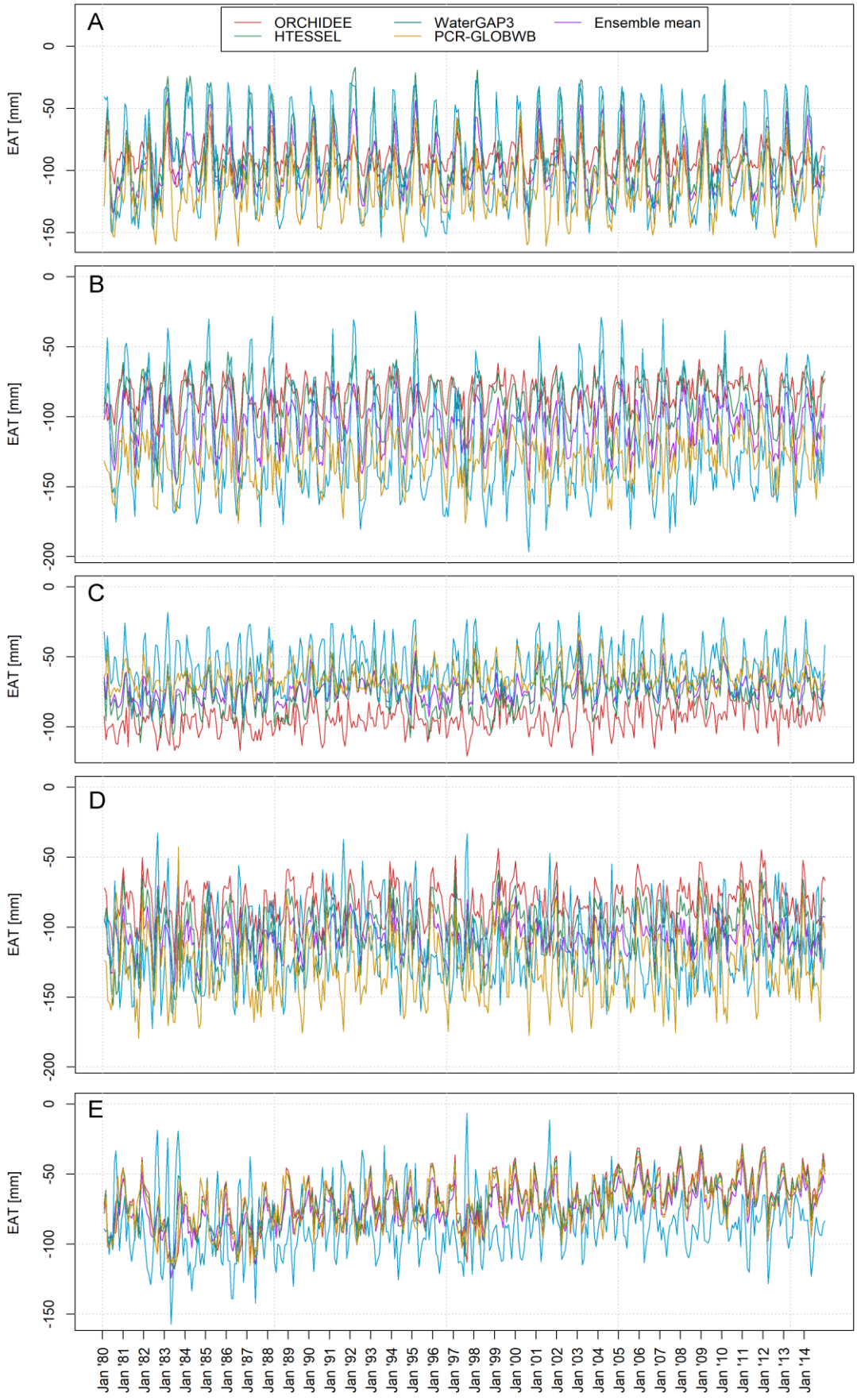

Figure 20. Actual evapotranspiration for ORCHIDEE, HTESSEL, WaterGAP3 and PCR-GLOBWB and the ensemble mean of these four models, for cells 66 (A), 295 (B), 360 (C), 428 (D) and 535 (E) in the MCMB

4.1.4 Root zone soil moisture

For the root zone soil moisture, no in-situ was data available and only modelled values are compared. The root zone is not defined in the same way for every model. According to the eartH2Observe literature, the root zone soil moisture includes the total soil moisture available for evapotranspiration according to the specific model. When this is not defined, the soil moisture up to 1 m depth is used (Beck et al., 2017). The number of soil layers differs as well. This influences the absolute root zone soil moisture amount, as can be observed in the time series. The models also use different (amounts of) soil types with different parameter values, and different types of vegetation that influence the root zone soil moisture content. This leads to differences in the standard deviation of root zone soil moisture.

From the box plot, it is directly evident that the ORCHIDEE model produces relatively much root zone soil moisture and the WaterGAP3 model produces relatively little root zone soil moisture (Figure 21). PCR-GLOBWB, HTESSEL and the ensemble mean fall in between.

All models show a yearly cycle with a maximum around January and a mininum around April and June (Figure 22). For locations C and E, located at the Cordilleras Central and Oriental, respectively, root zone soil water amounts are smaller. At these locations, the minima and maxima per year often show a dent in the middle of the peak..

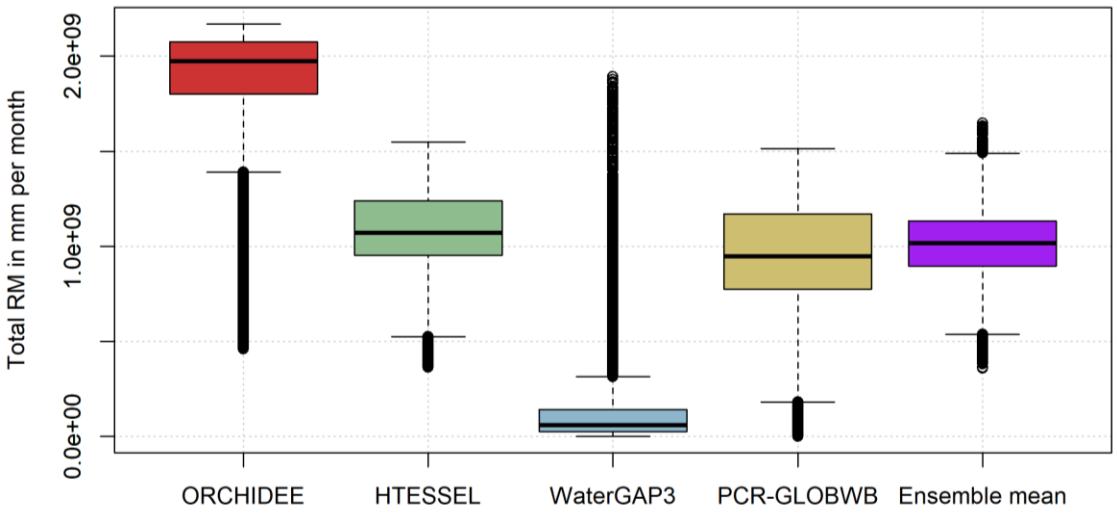

Figure 21. Box plots of root zone soil moisture values for all cells in the MCMB and for all months from January 1980 up to and including December 2014, for ORCHIDEE, HTESSEL, WaterGAP3 and PCR-GLOBWB and the ensemble mean of these four models

When the spatially mean time series are normalized, all models show similar temporal patterns (Figure 23). Through the process of normalization, absolute differences in the soil moisture amounts are no longer relevant because the monthly spatial mean root zone soil moisture values have been re-scaled between the minimum and maximum of the time series.

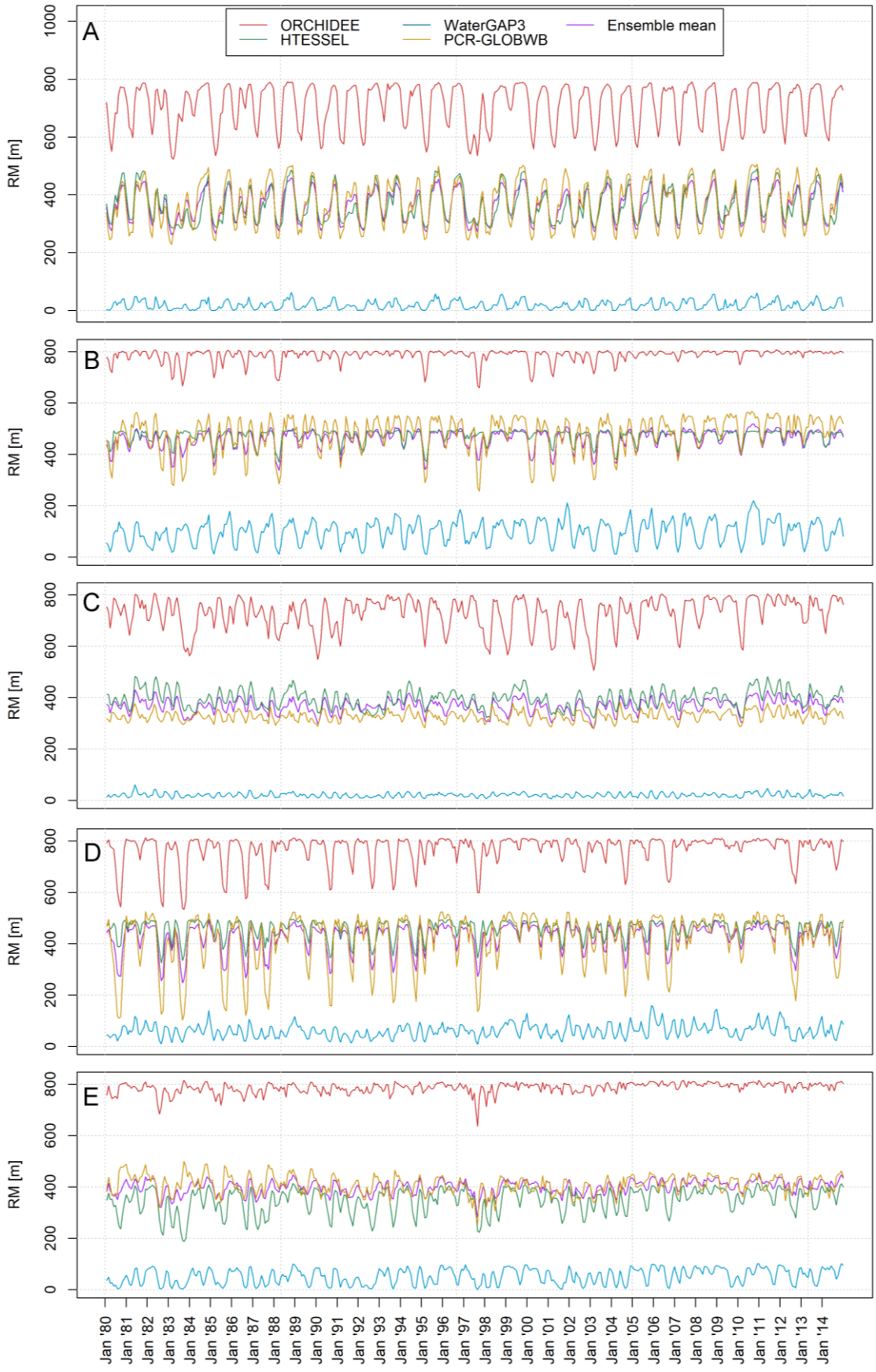

Figure 22. Root zone soil moisture for ORCHIDEE, HTESSEL, WaterGAP3 and PCR-GLOBWB and the ensemble mean of these four models, for cells 66 (A), 295 (B), 360 (C), 428 (D) and 535 (E) in the MCMB

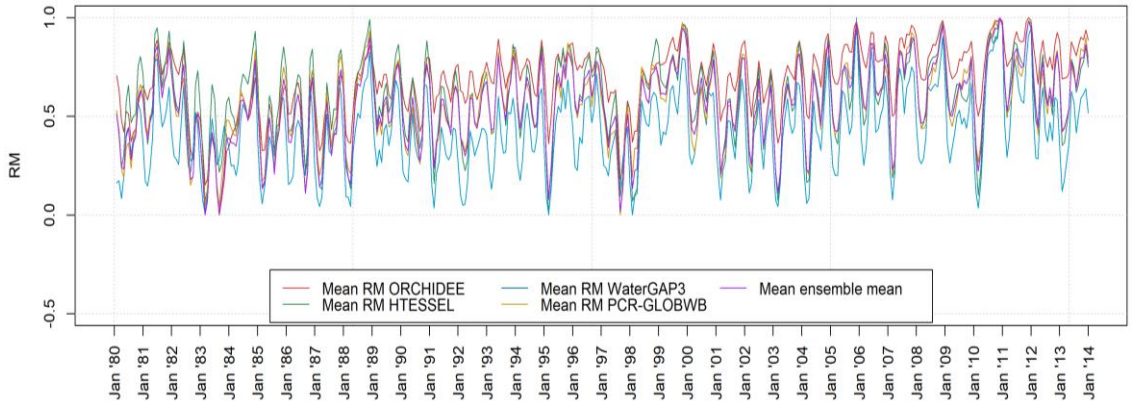

Figure 23. Spatial mean of the normalized root zone soil moisture of ORCHIDEE, HTESSEL, WaterGAP3 and PCR-GLOBWB and the ensemble mean of these four models

4.1.5 Discharge

Within the eartH2Observe project, there was no modelled discharge data available for ORCHIDEE model. Therefore, this model is not included in the discharge model performance. The model evaluation for discharge is based on point locations of discharge measurements along the Magdalena and Cauca river and their tributaries. These points have been converted to a raster with a resolution of 0.25° x 0.25°, to be able to compare the in-situ with the modelled data. The value of the raster cell is equal to the corresponding data point.

The difference between the modelled discharge time series and the in-situ discharge time series is relatively large for some locations in the MCMB (Figure 25). HTESSEL and WaterGAP3 tend to significantly overestimate the discharge, whereas PCR-GLOBWB slightly underestimates the discharge. For location A, located downstream of the Magdalena-Cauca river, all models agree well with the IDEAM discharge data. For location C, discharge amounts for the HTESSEL model are almost zero for the entire time series, while the other models and the IDEAM data do show a significant amount of discharge. The HTESSEL model also shows very little discharge for location E, but this is in line with the IDEAM data. Here, the other models overestimate discharge. Per year, two maxima can be observed at April and October, corresponding with the timing of wet seasons due to the shifting of the ITCZ (Restrepo et al., 2006).

The performance for discharge is very poor (Table 6, Figure 24). Most cells do have a positive KGE value, but some cells have extremely negative KGE values (Figure 26). The ensemble mean has the best KGE score – 4% of the cells have a KGE over 0.75 and 20% of the cells have a KGE over 0.50 – since in the ensemble mean the over- and underestimations of the different models are averaged out. All models have a relatively strong Pearson's correlation with the in-situ data. For the HTESSEL, WaterGAP3, PCR-GLOBWB models and the ensemble mean, 82%, 77%, 24% and 80% of the cells have Pearson's correlations of 0.5 or higher.

The low KGE values are therfore largely influenced by strong biases and variability ratios: the mean standard devation of the model products for discharge deviate significantly form the observed discharge data.

Table 6. KGE values for the HTESSEL, WaterGAP3 and PCR-GLOBWB models and the ensemble mean of these three models, for cells 66 (A), 164 (F), 360 (C), 428 (D) and 591 (G) in the MCMB

	A	F	C	D	G
HTESSEL	0.40	-0.07	-0.47	-23.59	0.16
WaterGAP3	0.55	0.15	0.09	-26.20	-14.95
PCR-GLOBWB	0.55	0.38	0.22	-7.71	-4.13
Ensemble mean	0.65	0.52	0.41	-18.74	-5.36

When the KGE is plotted against the upstream area of the corresponding discharge stations, it becomes apparent that for the cells for which the upstream area is available, the performance is generally better for stations that have a larger upstream area (Figure 27).

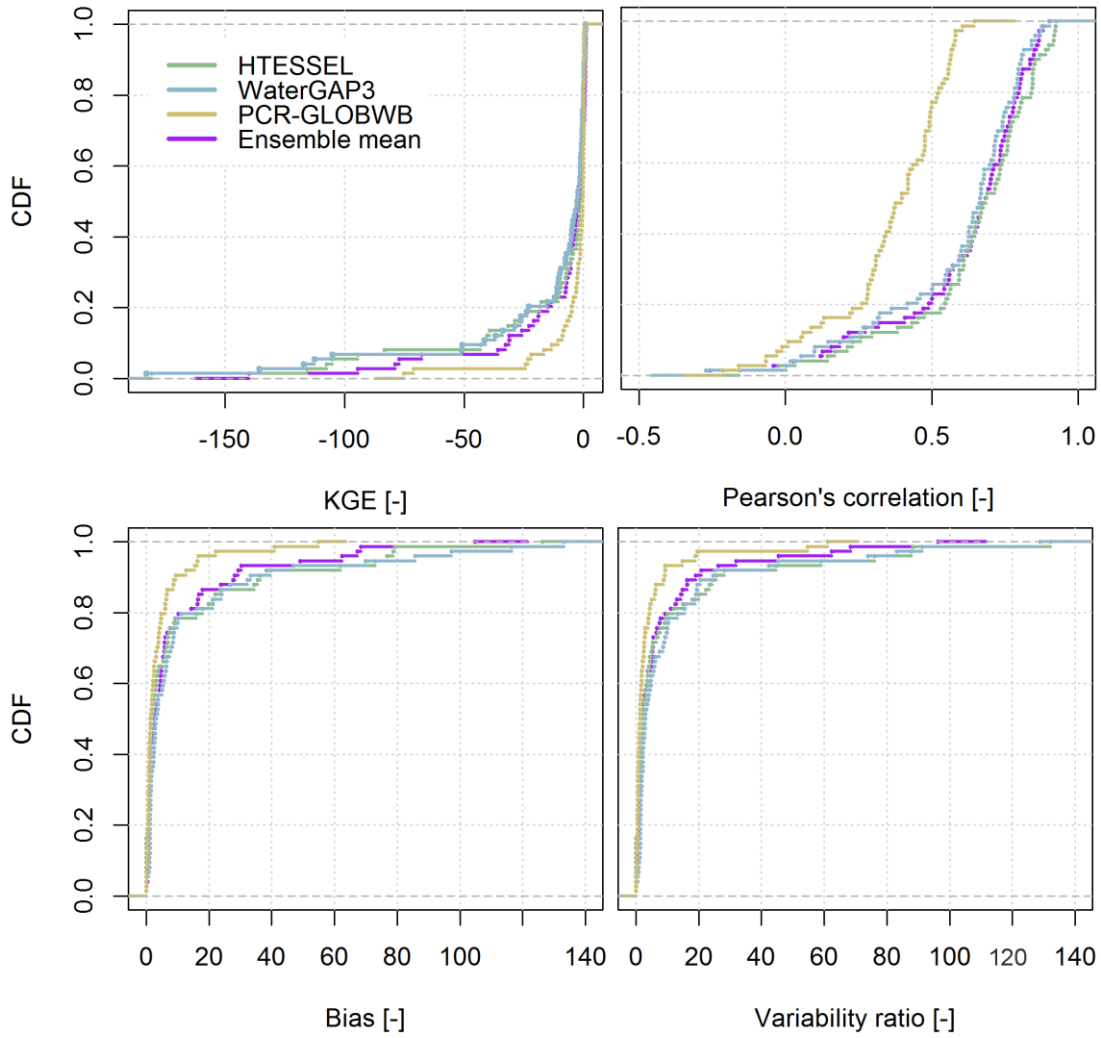

Figure 24. KGE, Pearson's correlation, bias and variability ratio of the discharge for the HTESSEL, WaterGAP3 and PCR-GLOBWB models and the ensemble mean of these three models

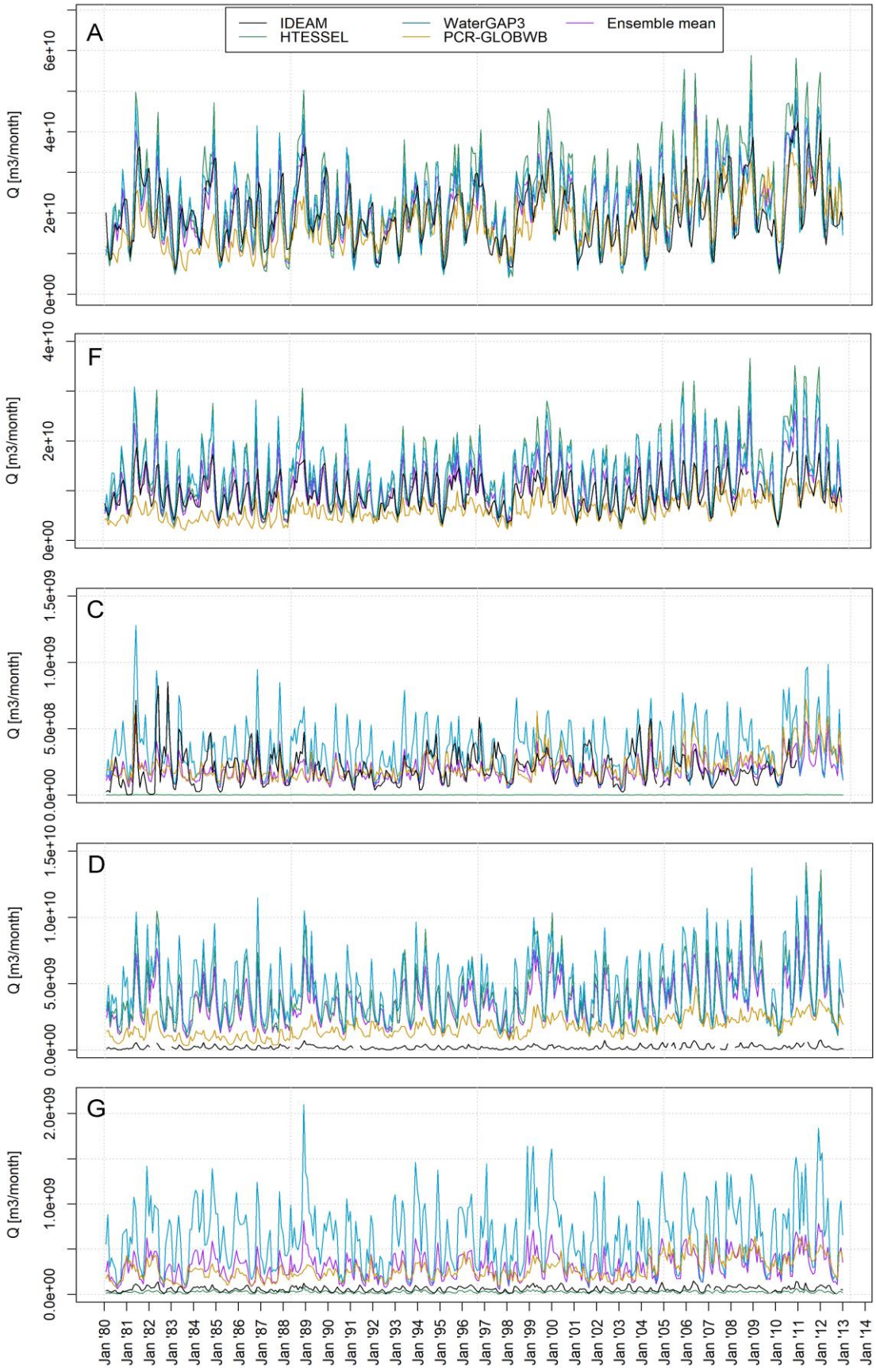

Figure 25. Discharge time series of IDEAM and the HTESSEL, WaterGAP3 and PCR-GLOBWB models and the ensemble mean of these three models, for cells 66 (A), 164 (F), 360 (C), 428 (D) and 591 (G) in the MCMB

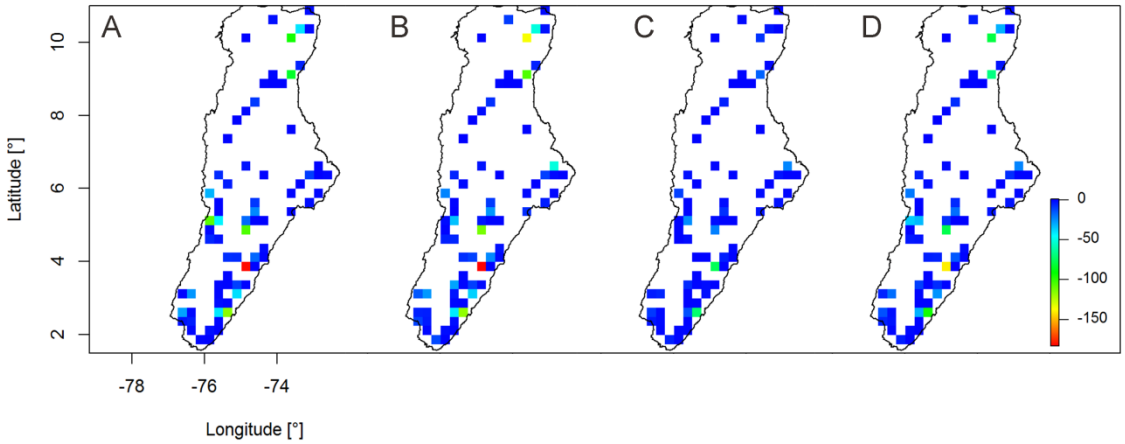

Figure 26. Spatial spread of the KGE values of the discharge for the HTESSEL (A), WaterGAP3 (B) and PCR-GLOBWB (C) models and the ensemble mean of these three models (D)

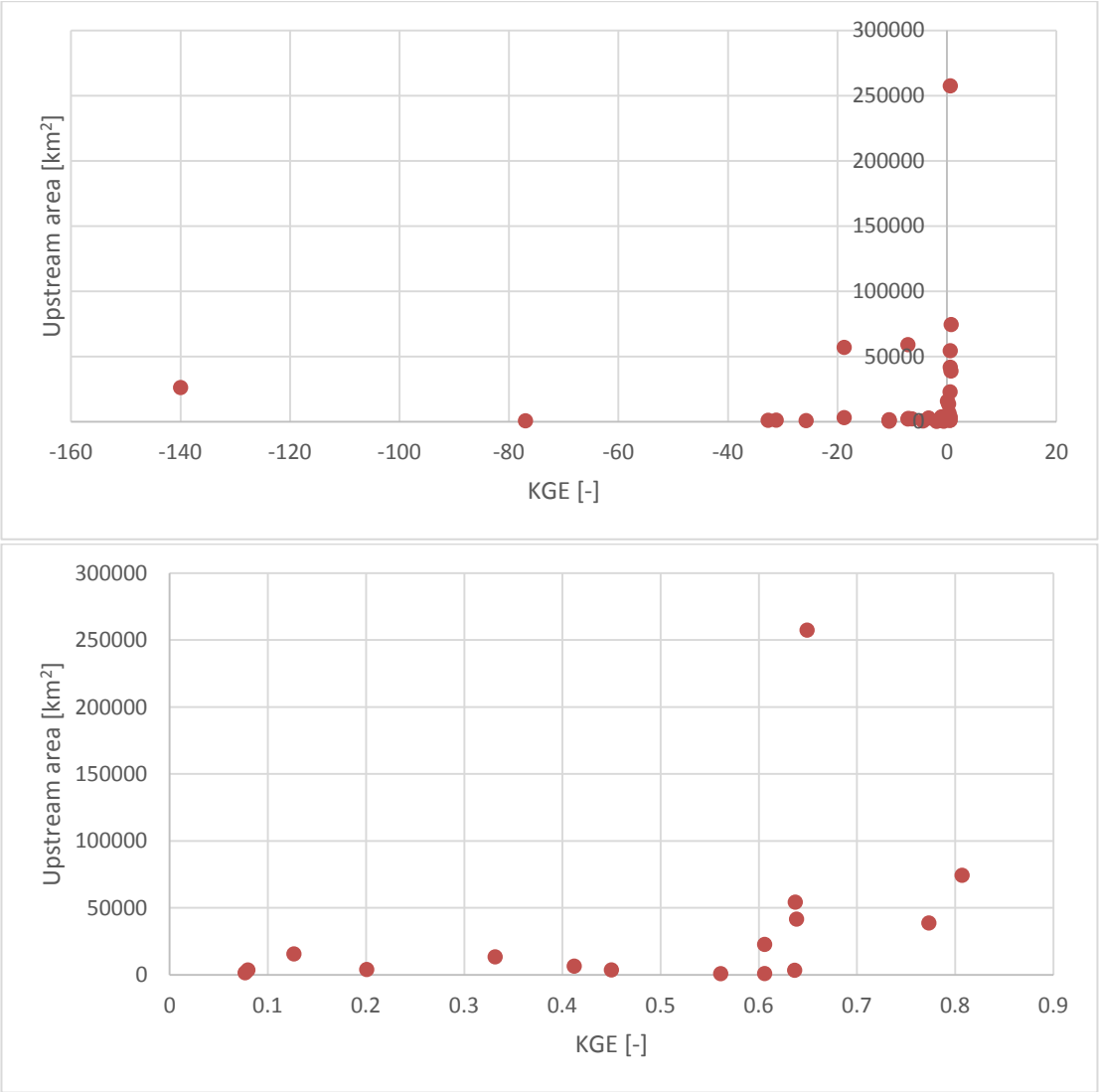

Figure 27. Upstream area plotted against the KGEs of discharge of the ensemble mean of the HTESSEL, WaterGAP3 and PCR-GLOBWB models, for all KGE values (above), and only for KGE values above 0 (below).

4.2 Drought analysis

4.2.1 SPI

Since the SPI is solely based on precipitation data, which is equal to the MSWEP forcing data, this drought index is the same for all four models. The SPI is calculated for temporal aggregations of 1, 3, 6, 12, 24, and 36 months (n), meaning that the SPI of month x takes the previous n months into account.

Pooling, lengthening, attenuation and lag can be observed in Figure 28. When increasing the time scale of the SPI from 1-month to 36-month, the number of drought events (where SPI < -1) decreases from a spatial average of 34 events to only 1 event (pooling). The length of each individual drought event increases from a spatial average between 1 and 8 months to 53 months (lengthening). The intensity of the events decreases when the time scale is increased (attenuation). Lag can be observed by the later timed onsets of each drought even when increasing the aggregation period from 1-month to 36-months.

A SPI of -1 is reached first around ’82-’83. This drought event is relatively short and is followed by multiple other small drought events until the end of ’83. The next drought event occurs in ’85. This event is more pronounced in the SPIs with shorter aggregation periods. This indicates that over a longer period, no abnormal low amount of precipitation has fallen. In ’92-’93, one of the two most severe drought events occurs. This event lasts at least until ’84, which is relatively long. The next drought event is the other most severe event, but this event lasts rather short - from ’97 to ’98 – but is relatively intense, especially for shorter aggregation periods, with SPI values dropping below -2. The final significant drought event occurs around January ’10, which lasts relatively short again. This drought event does not occur in SPIs with longer aggregation periods.

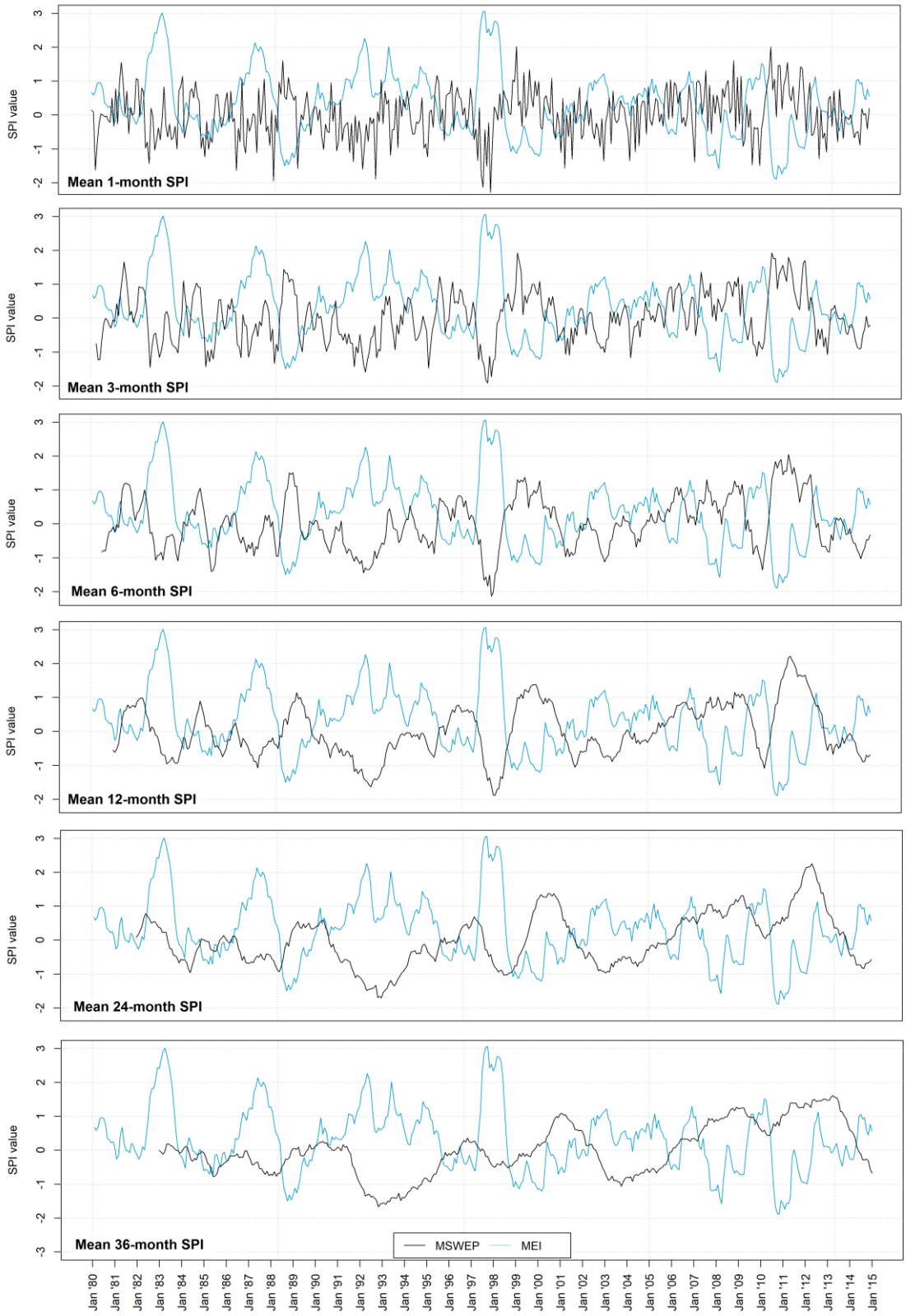

Figure 28. Time series of the spatial mean per month of the 1, 3, 6, 12, 24 and 36-month SPI

Droughts occur everywhere in the catchment (Figure 29). It differs per month, year and temporal aggregation period where the most negative SPI values can be found. However, large differences within the catchment in a single month are unlikely to occur. Generally, there are slightly more dry

months in the southeast of the catchment and at -74° longitude, between 6° to 8° latitude . There are less dry months in the central basin of the catchment, around location B at (-74.625°, 7.125°). The ratio of dry months to all months becomes bigger when the aggregation period is increased.

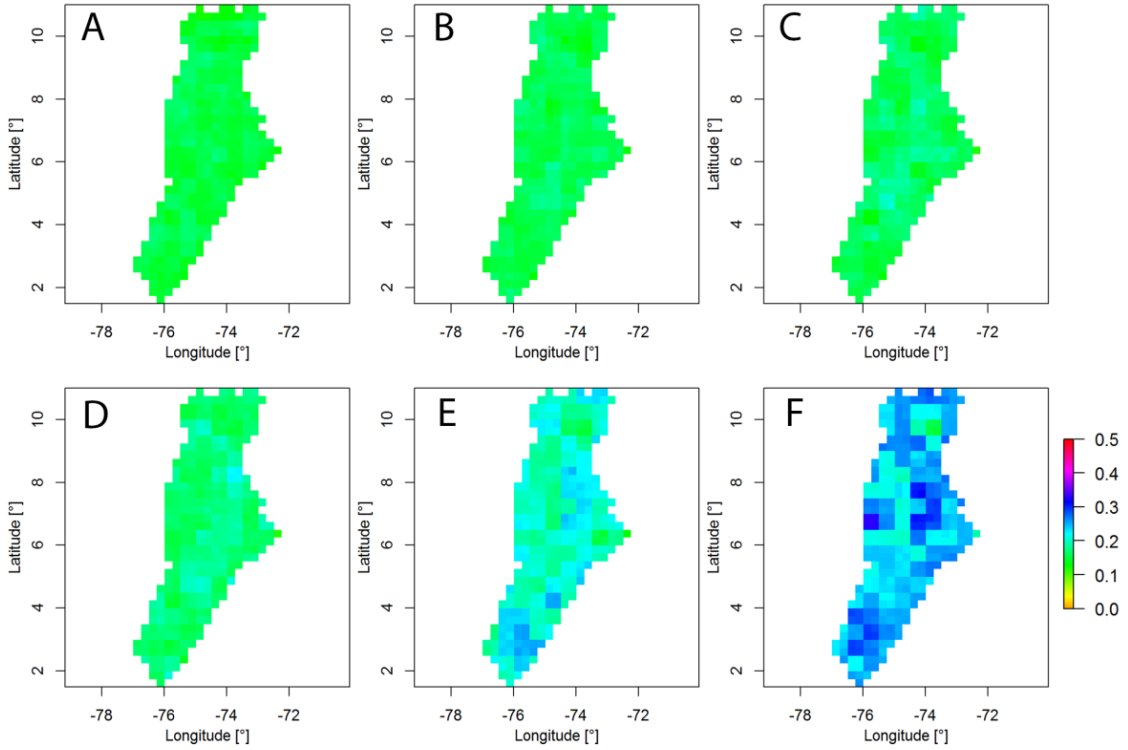

Figure 29. Ratio of months with a SPI value below -1 to all months, for every cell in the MCMB for the 1-month (A), 3-months (B), 6-months (C), 12-months (D), 24-months (E) and 36-months (F) SPEI

For the 1-month SPI, a single month can show very negative SPI values while the preceding and succeeding months show positive SPI values and a different spatial pattern. For SPIs with an aggregation period of six months or more, a clearer pattern can be observed: generally, at the start of a drought, more negative SPEI values are located in the northeast, whereas towards the end of the drought more negative values are located in the southwest. This pattern is present for the droughts between October 1982 and January 1984, between September 1986 and July 1987, between September 1991 and December 1992 and between April 1997 and June 1998 (Figure 30). The drought between December 2009 and April 2010 is mainly located in the south of the MCMB.

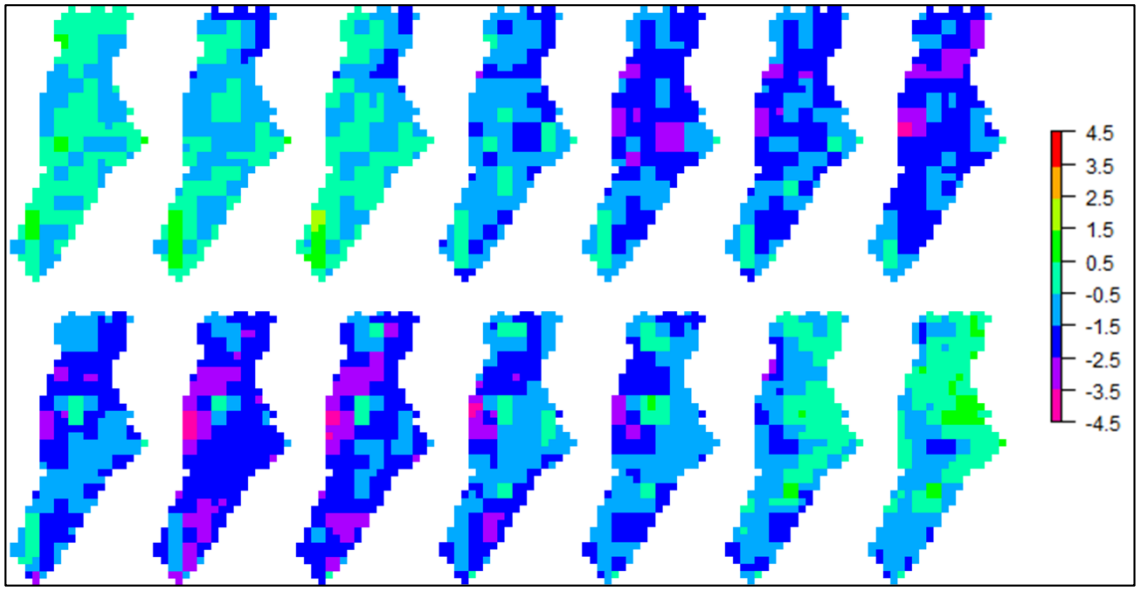

Figure 30. Spatial distribution of the 6-month SPI values for(from left to right and from top to bottom) every month from April 1997 to June 1998

4.2.2 SPEI

The Standardized Precipitation Evapotranspiration Index is based on both the precipitation data and the potential evapotranspiration data. Again, the precipitation data is the same for all four models, but the potential evapotranspiration differs per model, which results in different SPEI indices for each model. The SPEI shows behaviour that is very similar to the SPI, both in space and time, for all aggregation periods. Therefore only the 6-month SPEIs of all models are given (Figure 31).

The timing of increase and decrease of the SPEI is very similar per model, but the maxima and minima show differences for each model. The differences between the models increase with increasing temporal aggregation. It is remarkable that it is not the same model that gives the lowest SPEI values all the time. This might indicate that it is not the same model parameter that causes a minimum or maximum in the SPEI, but rather that different parameters take the overhand per minimum or maximum. During the drought events following the 1982-1983, 1986-1987 and 1997-1998 El Niño events, the HTESSEL model shows the lowest SPEI values and both the WaterGAP3 and the PCR-GLOBWB models show the highest SPEI values. In contrast, after the 1991-1992 El Niño event the WaterGAP3 and PCR-GLOBWB models show the lowest SPEI values and the HTESSEL model shows the highest SPEI values. The SPEI values of the ORCHIDEE model and the ensemble mean seem to fall between the values of these three afore mentioned models.

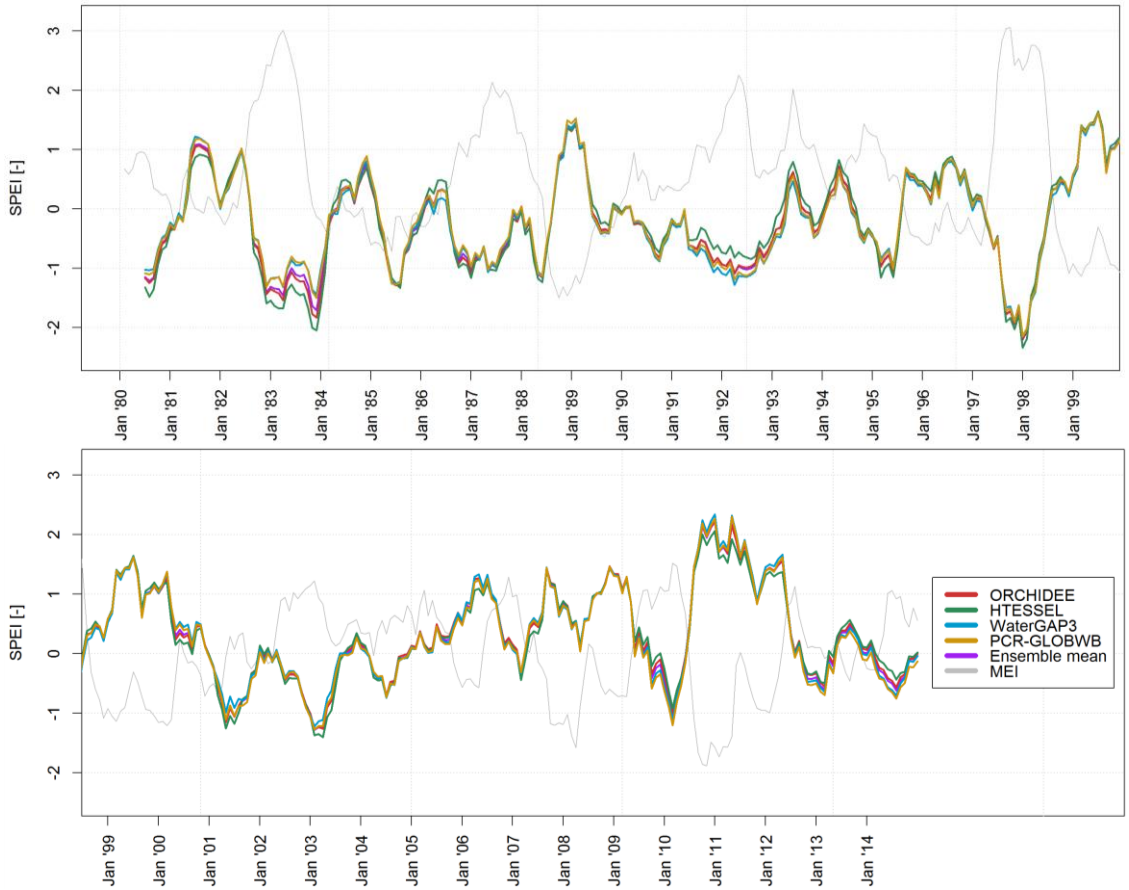

Figure 31. Time series of the spatial mean of the 6-month SPEI for ORCHIDEE, HTESSEL, WaterGAP3, PCR-GLOBWB and the ensemble mean of these four models

For the SPEI, droughts can be found everywhere in the catchment, and distributional patters are similar to the SPI (Figure 32, Figure 33). The pattern with the droughts shifting from the north of the catchment to the south of the catchment can also be observed for the same months as for the SPI.

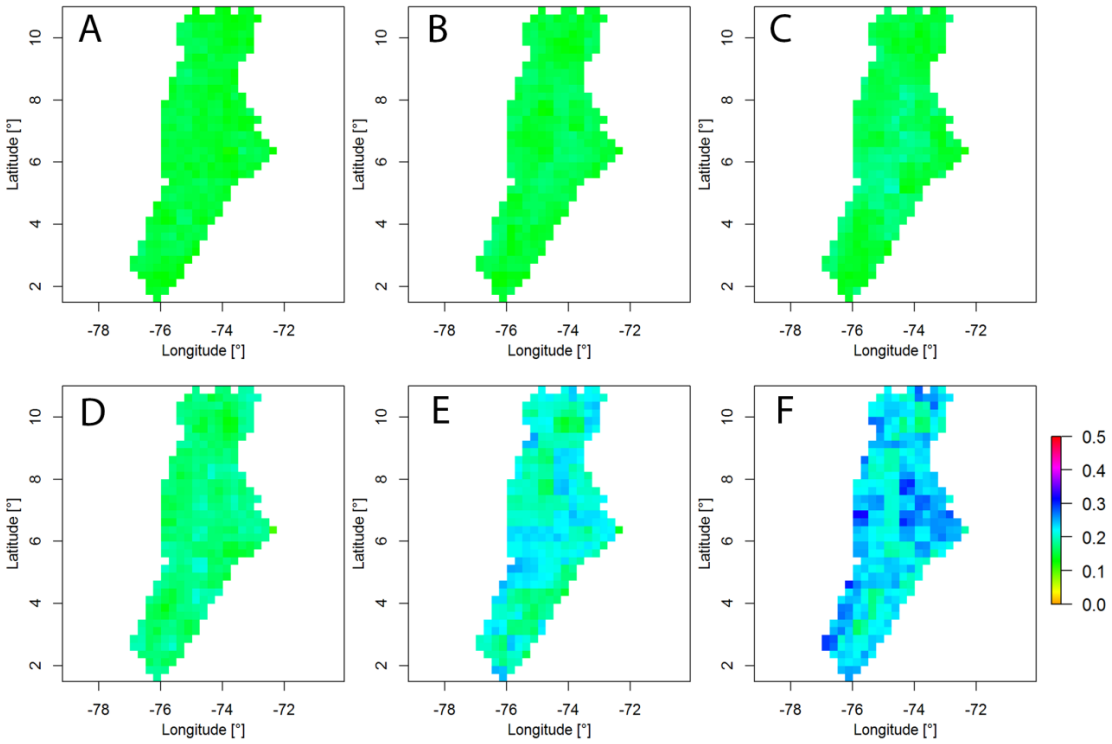

Figure 32. Ratio of months with a SPEI value below -1 to all months, for every cell in the MCMB for the 1-month (A), 3-months (B), 6-months (C), 12-months (D), 24-months (E) and 36-months ensemble mean SPEI (F)

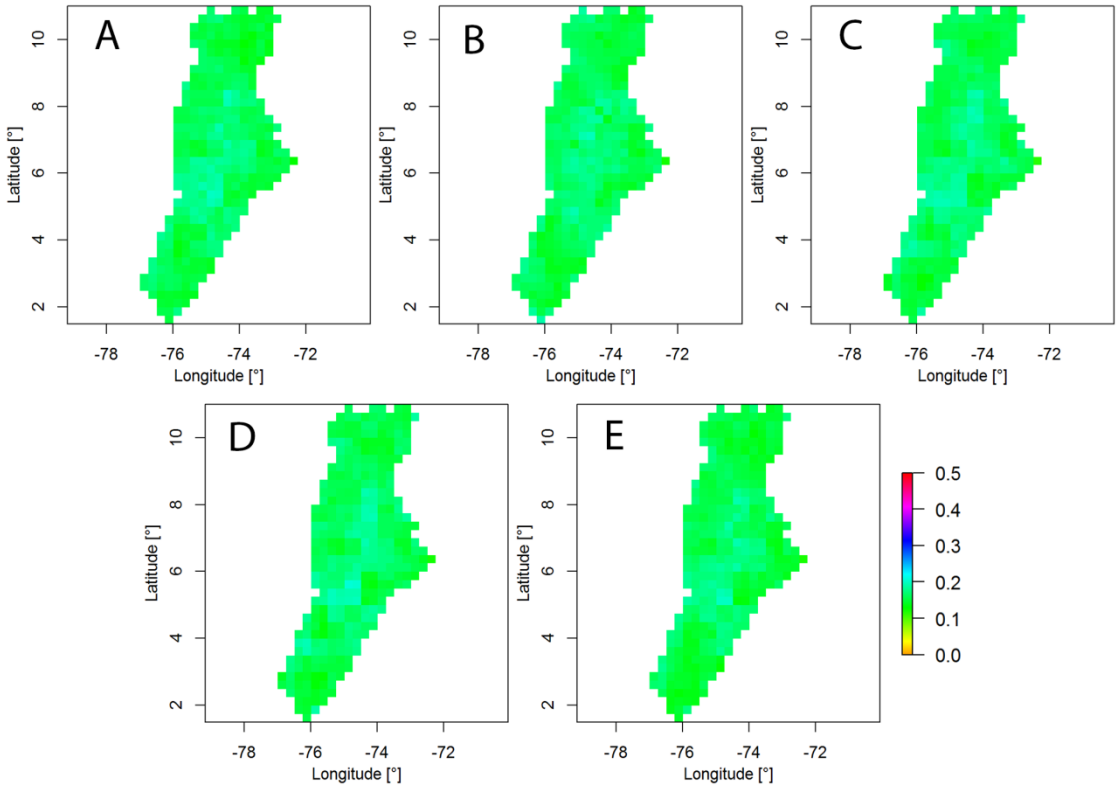

Figure 33. Ratio of months with a SPEI value below -1 to all months, for every cell in the MCMB for the 6-month SPEI of ORCHIDEE (A), HTESSEL (B), WaterGAP3 (C), PCR-GLOBWB(D) models and the ensemble mean of these four models (E)

4.2.3 ETDI

The evapotranspiration deficit index is based on the potential and actual evapotranspiration. What becomes directly apparent when looking at the mean ETDI's time series for all models is that all models give different mean ETDI time series (Figure 34). Especially the WaterGAP3 model has very different ETDI values compared to the other three models and the ensemble mean. WaterGAP3 often showed significantly larger actual evapotranspiration amounts, but the potential evapotranspiration minus the actual evapotranspiration did not get below zero.

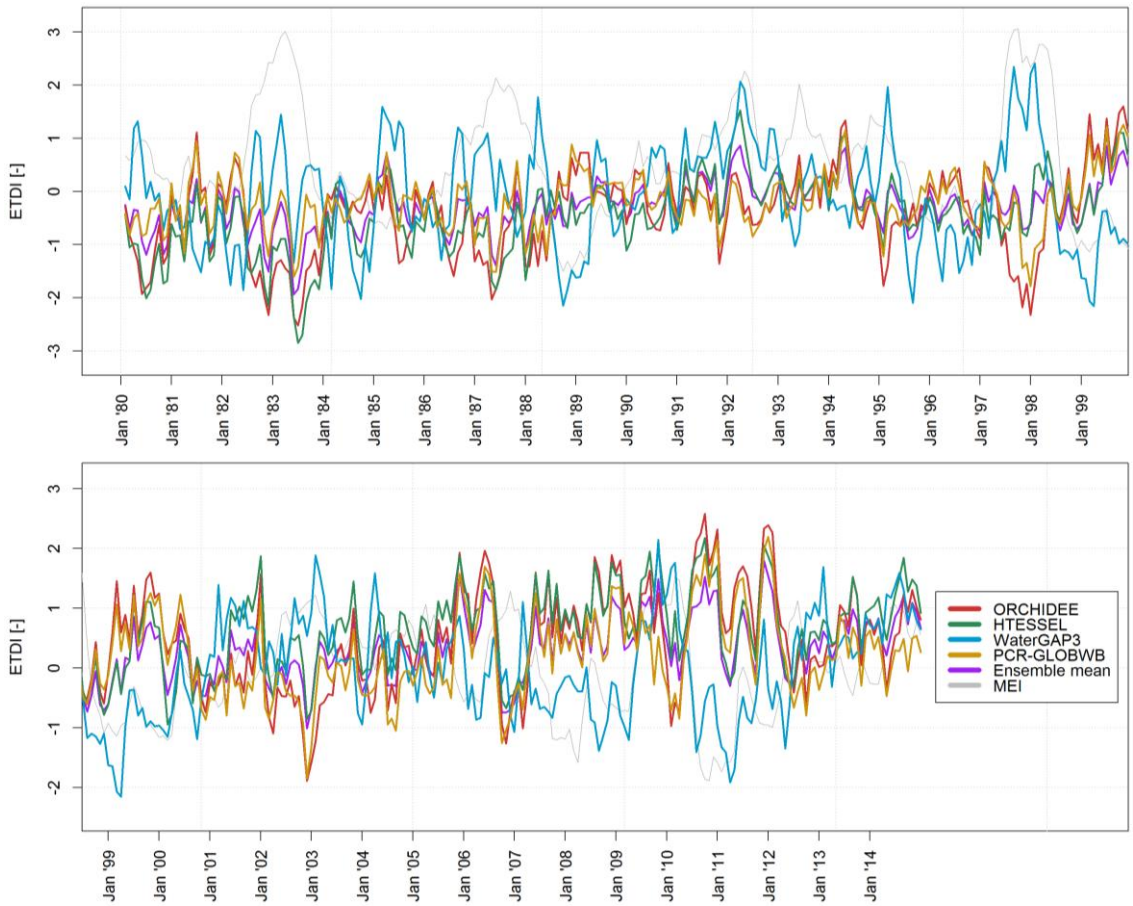

Figure 34. Time series of the spatial mean of the ETDI for ORCHIDEE, HTESSEL, WaterGAP3, PCR-GLOBWB and the ensemble mean of these four models

In contrast to the SPI and SPEI, very negative and very positive ETDI values can occur in the MCMB in the same month. ETDI values are especially low between April 1997 to June 1998, mainly in the southwest of the catchment. Generally, more negative ETDI values occur in the south and the north (Figure 35). The same months show negative values as for the SPI and SPEI. These trends are less pronounced or absent for the WaterGAP3 ETDI values. For this model, little islands with positive or negative ETDI values are spread over the catchment and the timing of strongly negative values is different. The shift of droughts from north to south is not observed for the ETDI.

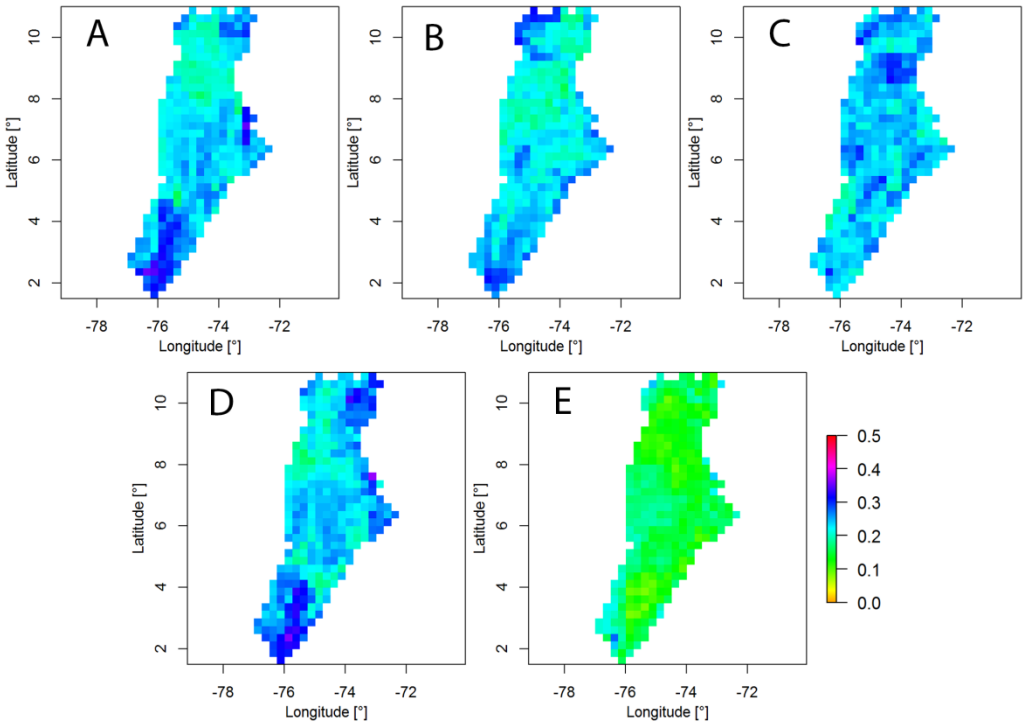

Figure 35. Ratio of months with a ETDI value below -1 to all months, for every cell in the MCMB for the ORCHIDEE (A), HTESSEL (B), WaterGAP3 (C), PCR-GLOBWB(D) models and the ensemble mean of these four models (E)

4.2.4 SMDI

The differences between the four models are pronounced in the Soil Moisture Deficit Index, but the models do correspond in timing of the minima and maxima (Figure 36). The large differences between models that could be observed for the root zone soil moisture data are not present in the SMDI, due to the computation methods for this index. The values range from -3 to 3 and vary a lot over time. Again, it differs per drought event which model gives the most negative SMDI values. For the '83-'84 drought event, both the ORCHIDEE and PCR-GLOBWB models give the most negative SMDI values. For the '91-'92, '97-'98 and 2013 and 2014 drought events, the WaterGAP3 model produces the most negative SMDI values. This is also evident in Figure 37, where the WaterGAP3 model shows the highest ratio of months with a sub -1 SMDI value to all months. The smallest ratio of dry months to all months can generally be found in the centre of the catchment above 6°N.

The same months show droughts for the SMDI as for the 3-month and 6-month SPI and SPEI, and ETDI, but the SMDI does have a higher standard deviation than the 3-month and 6-month SPI and SPEI. The same shifts of droughts from north to south with time as is the case for the SPI and SPEI can also be observed for the SMDI. Very negative SMDIs tend to occur more often in the north, east and south.

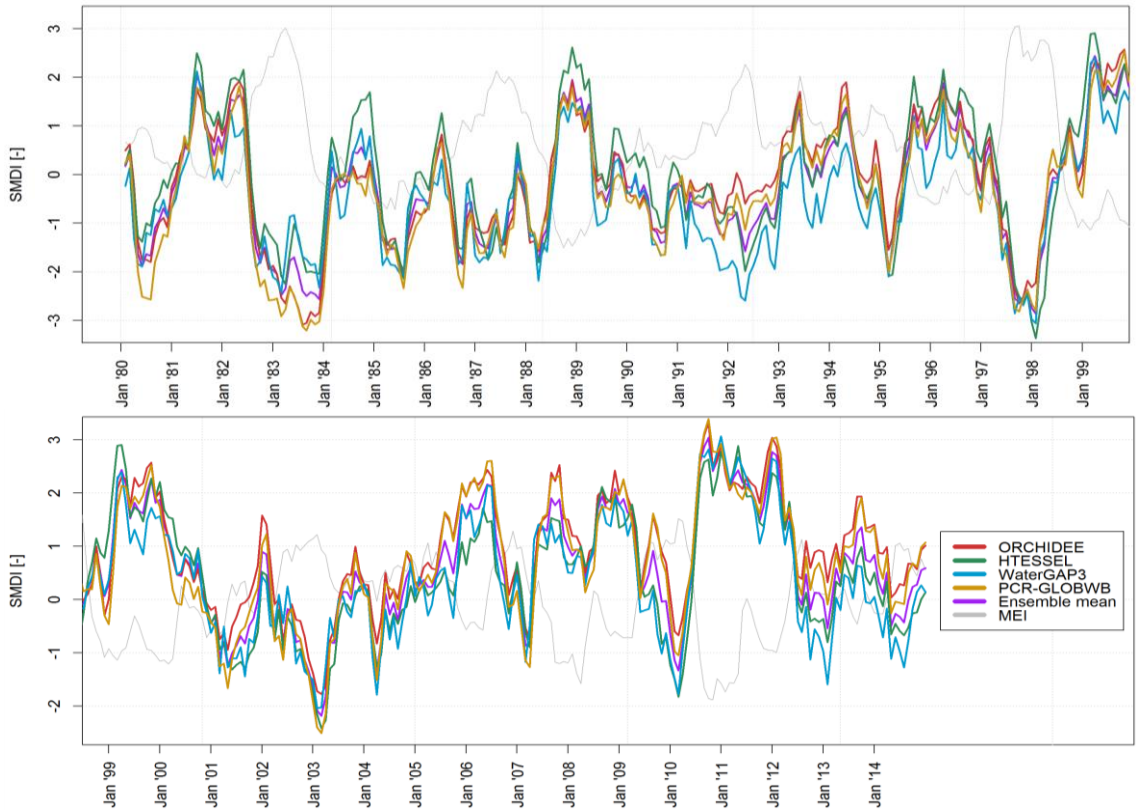

Figure 36. Time series of the spatial mean of the SMDI for the ORCHIDEE, HTESSEL, WaterGAP3, PCR-GLOBWB and the ensemble mean of these four models

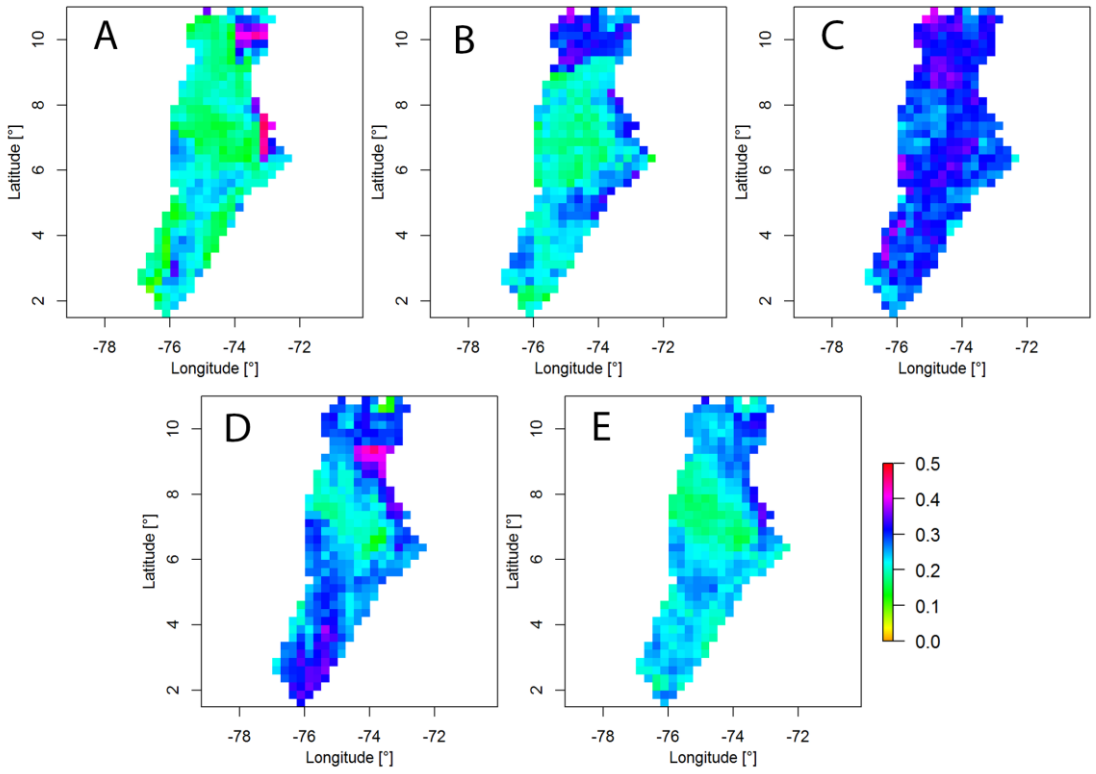

Figure 37. Ratio of months with a SMDI value below -1 to all months, for every cell in the MCMB for the ORCHIDEE (A), HTESSEL (B), WaterGAP3 (C), PCR-GLOBWB(D) models and the ensemble mean of these four models (E)

4.2.5 SDI

The Streamflow Deficit Index is based on discharge data of the models. Because there is no discharge data available for the ORCHIDEE model in the WRR-2, this model is not included in the drought analysis.

All models and all cells show similar temporal patterns, even though the discharge data showed big differences between models for some locations (Figure 38). This can be explained by the standardization and transformation to a normal distribution performed in the computation of the SDI. The pattern of the SDI is similar to the 3-month and 6-month SPI and SPEI. The models agree relatively well, however, the PCR-GLOBWB model seems to give lower SDI values at the beginning of the time series and higher values towards the end compared to the other models. The same pattern could be observed for the discharge values.

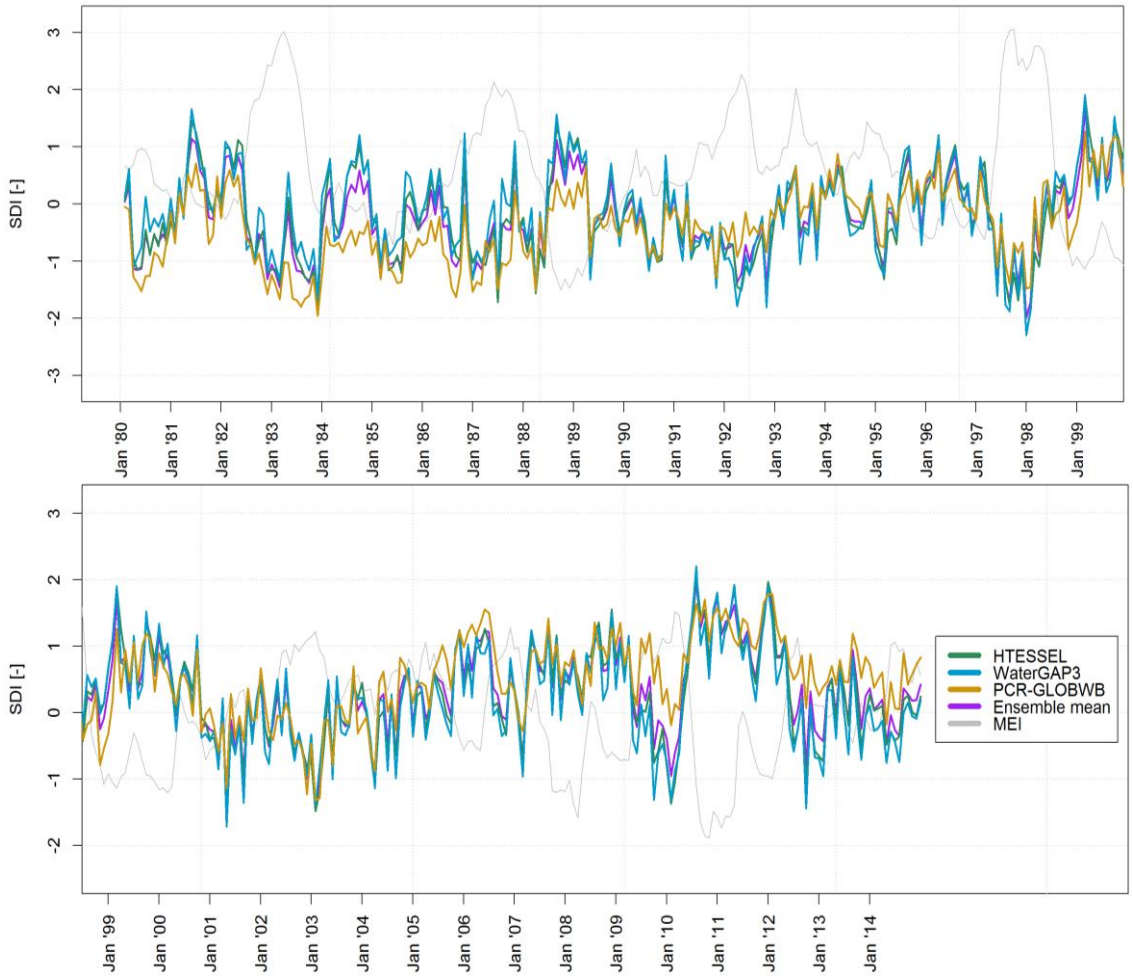

Figure 38. Time series of the spatial mean of the SDI for HTESSEL, WaterGAP3, PCR-GLOBWB and the ensemble mean of these three models

In the spatial distribution of ratio of months with a SDI below -1 to all months, the river networks of the Magdalena and Cauca rivers can be distinguished (Figure 39). This is especially clear for the PCR-GLOBWB model. Because the cells containing a major river branch will experience stronger fluctuations in discharge, it is logical that these cells will experience more dry months.

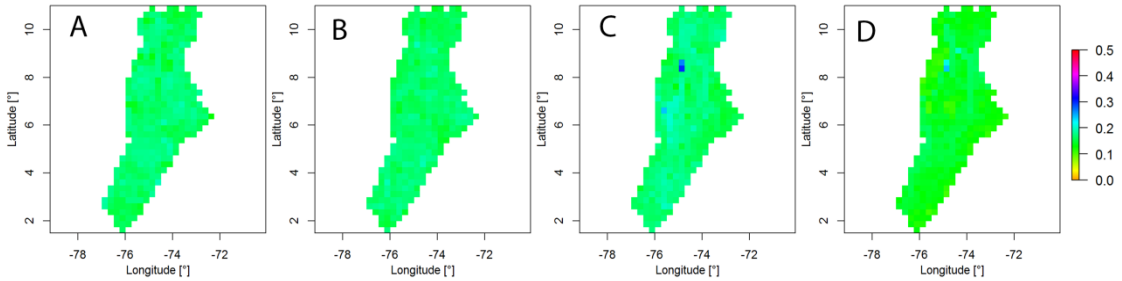

Figure 39. Ratio of months with a SDI value below -1to all months, for every cell in the MCMB for HTESSEL (A), WaterGAP3 (B), PCR-GLOBWB(C) and the ensemble mean of these three models (D)

4.3 Correlation between drought and ENSO

4.3.1 SPI

The strength of the negative Pearson's correlation between the SPI and the MEI strongly depends on the chosen temporal aggregation (Figure 40). The correlation is strongest for the 6-month SPI, followed by the 12-month SPI. It can also be observed that the points in the SPI vs. MEI plot tend to be located around the 1:1 line, except for the area with a negative SPI and positive MEI (Figure 41). Here, the points are located above the 1:1 line, which means that for a strongly positive MEI, the drought index it not necessarily strongly negative. A strong El Niño event is thus not per definition a severe drought. This could also be observed in Figure 28: for the '82-'83 and the '86-'87 El Niño events, the SPI does not drop far below -1.

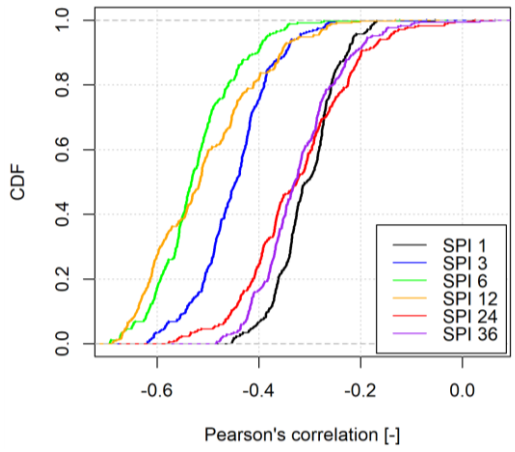

Figure 40. CDF plot of the Pearson's correlation between the 1, 3, 6, 12, 24, and 36-month SPI and the MEI

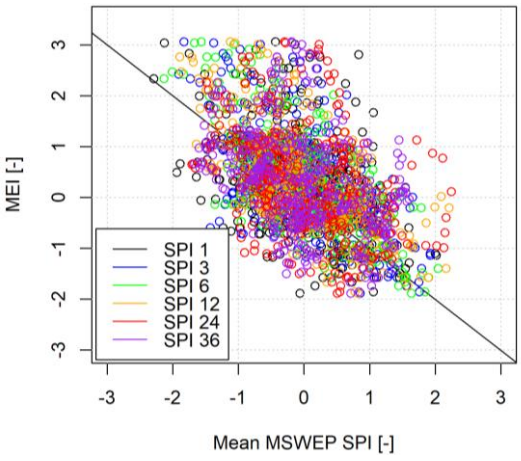

Figure 41. Scatter plot of the spatial mean of the 1, 3, 6, 12, 24 an 36-month SPI values versus the MEI values

The spatial distribution of the strength of the correlation is roughly the same for all SPI time scales, with stronger correlations north of the middle of the catchment and the weakest correlations along the southeast border of the catchment, roughly between coordinates (-76,2) and (-72,6), where the mountains are located (Figure 42). The correlation is also weaker in the absolute north of the catchment, close to the Atlantic Ocean.

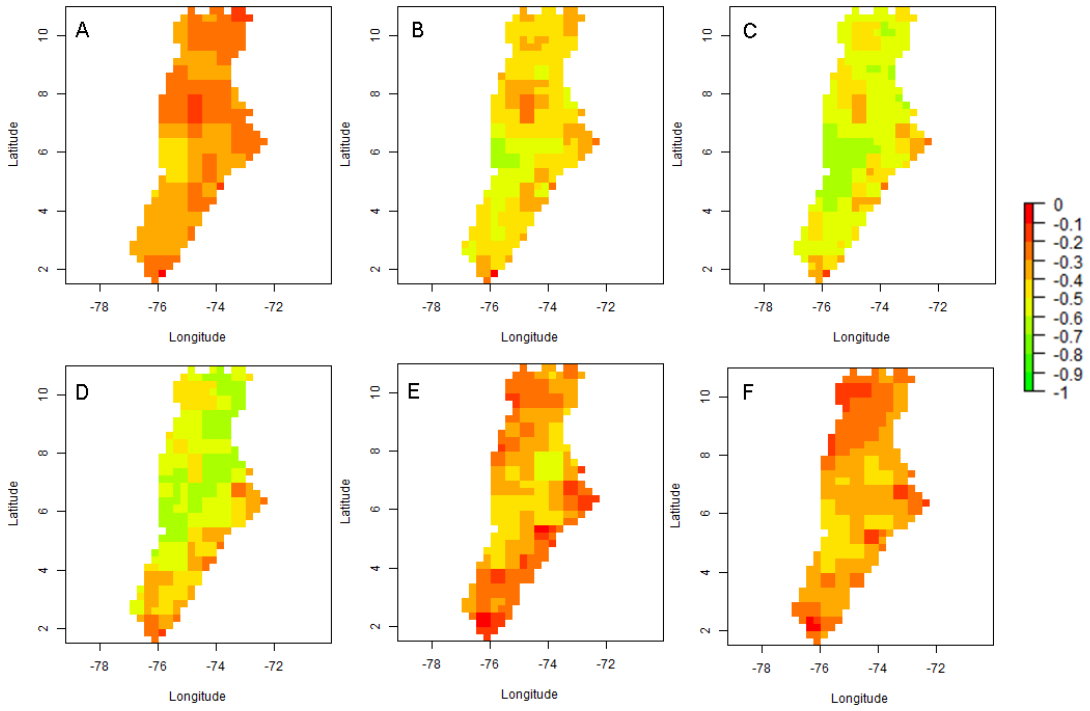

Figure 42. Spatial distribution of the Pearson's correlation between the 1, 3, 6, 12, 24 and 36-month SPI and the MEI

4.3.2 SPEI

The 6-month SPEI shows the strongest correlation with the MEI (Figure 43). Again, the correlation is weakest for the 1-month and 36-month SPEI. Therefore, only the figures for the 6-month SPEI are included. The correlation between the 6-month SPEI and the MEI is a stronger negative correlation than for the 6-month SPI, with at least 81% of the cells for all models and the ensemble mean having a Pearson's correlation of -0.50 of lower, with respect to 68% of the cells for the 6-month SPI.

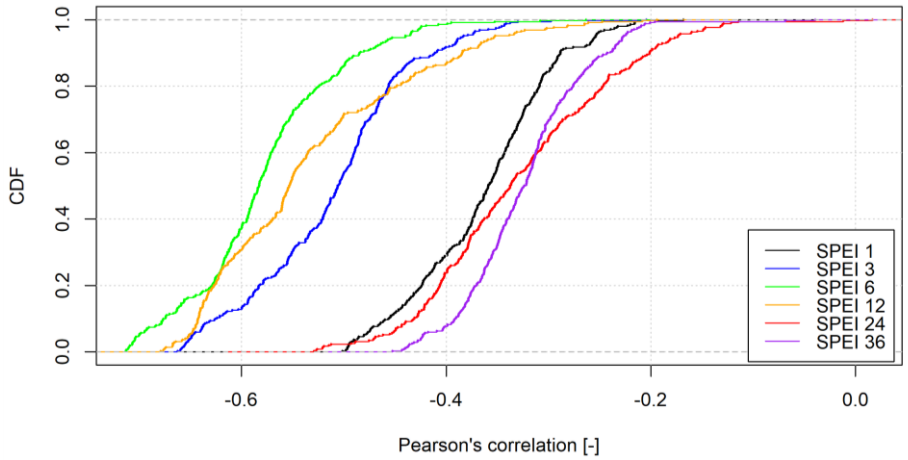

Figure 43. CDF plot of the Pearson's correlation between the ensemble mean of the 1, 3, 6, 12, 24, and 36-month SPEI and the MEI

The four different models agreed well on the SPEI values, which is reflected in similar Pearson's correlation values (Figure 44). For the 6-month SPEI, the ensemble mean shows the strongest correlation with the MEI: 86% of the cells show a Pearson's correlation of -0.50 or lower. Similar to the correlation between the SPI and MEI, the correlation is weaker for positive MEI and negative SPEI values (Figure 45).

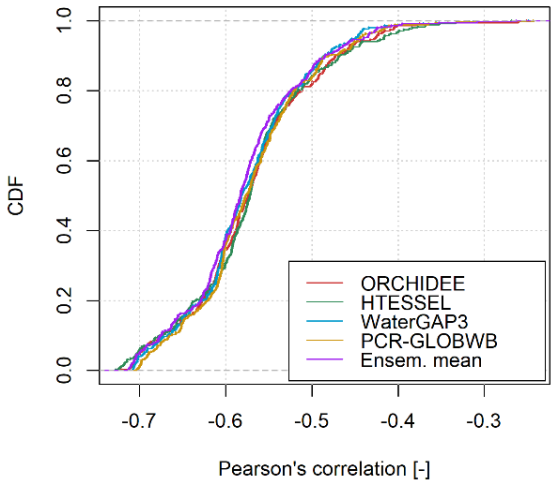

Figure 44. CDF plot of the Pearson's correlation between the 6-month SPEI and the MEI for ORCHIDEE, HTESSEL, WaterGAP3, PCR-GLOBWB and the ensemble mean

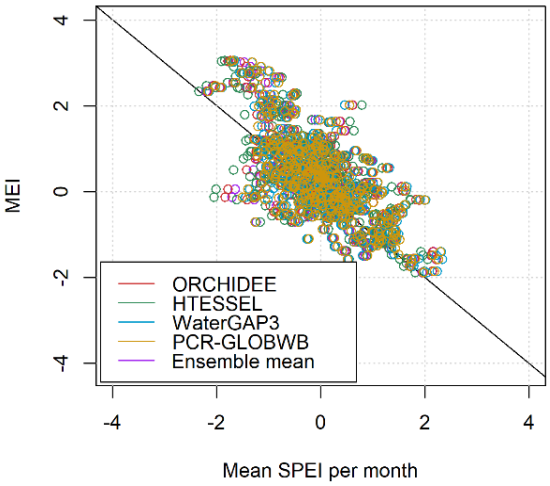

Figure 45. Scatter plot of the spatial mean 6-month SPEI values versus the MEI values for ORCHIDEE, HTESSEL, WaterGAP3, PCR-GLOBWB and the ensemble mean

The spatial patterns are also rather similar to the spatial pattern of Pearson's correlations for the SPI. The weakest correlations can again be found along the south-eastern border of the catchment along the line between coordinates (-76°, 2°) and (-72°, 6°) (Figure 46). The strongest correlations can be found along above this diagonal line, ranging from the southwest to the north of the catchment. A difference between the spatial pattern for the SPEI and the SPI is that the SPEI shows also strong correlations in the north, whereas this is not the case for the SPI.

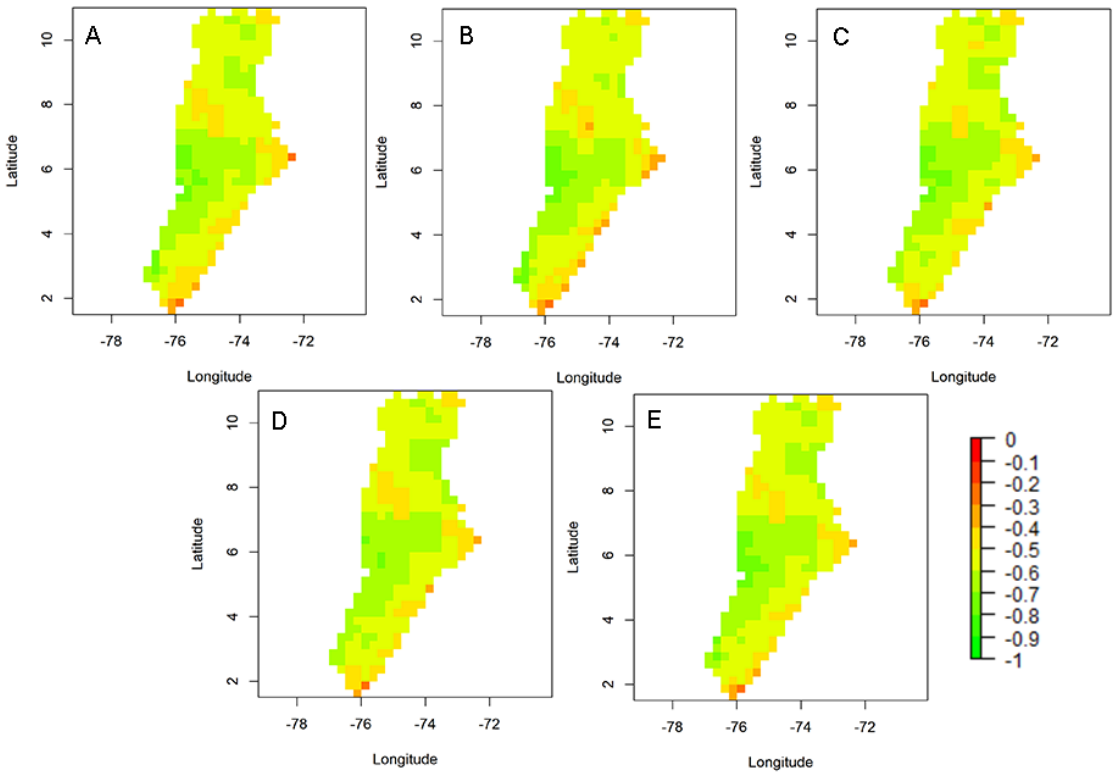

Figure 46. Spatial distribution of the Pearson's correlation between the 6-month SPEI and the MEI for the ORCHIDEE (A), HTESSEL (B), WaterGAP3 (C) and PCR-GLOBWB (D) models and the ensemble mean of these four models (E)

4.3.4 ETDI

All models show relatively large ranges in the CDF plot between the Pearson's correlation's minimum and maximum (Figure 47, Figure 48). WaterGAP3 has a positive relationship with the MEI. The correlation between the ETDI and the MEI is strongest for the ORCHIDEE model, where 32% of the cells has a Pearson's r below -0.50. The PCR-GLOBWB model also has a stronger correlation: 29% of the cells have a Pearson's correlation below -0.50. The WaterGAP3 models shows the weakest negative correlation with the MEI: there are no cells with a value below -0.50, and only 0.28% of the cells has a value below -0.25. The WaterGAP3 models does show the strongest positive correlation, with 80% of the cells having an r-vale > 0.25 and 6.8% of the cells having an r-value >0.5.

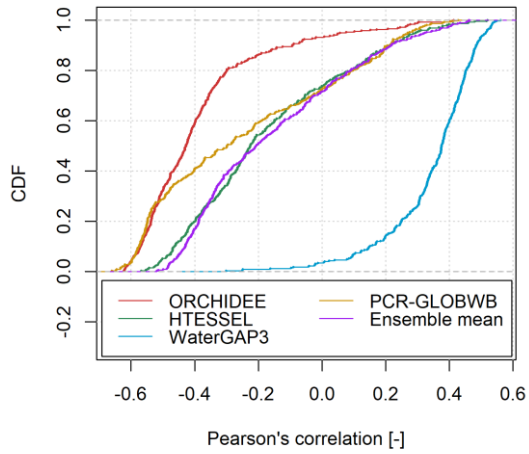

Figure 47. CDF plot of the Pearson's correlation between the ETDI and the MEI for ORCHIDEE, HTESSEL, WaterGAP3, PCR-GLOBWB and the ensemble mean

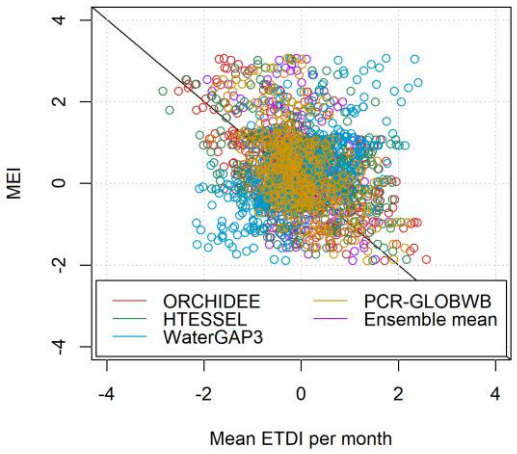

Figure 48. Scatter plot of the spatial mean ETDI values versus the MEI values for ORCHIDEE, HTESSEL, WaterGAP3, PCR-GLOBWB and the ensemble mean

The spatial distribution of the Pearson's correlation is roughly similar for the ORCHIDEE, HTESSEL and PCR-GLOBWB models and the ensemble mean, with the strongest negative correlations located along the western border of the catchment and in the middle of the catchment around 6.5 ° latitude (Figure 49). The Pearson's r correlation is generally low or sometimes even positive in the east and north of the catchment. For the WaterGAP3 model, the correlation is strongest and most positive in the north and east, and weakest in the west, which is in contrast with the other models.

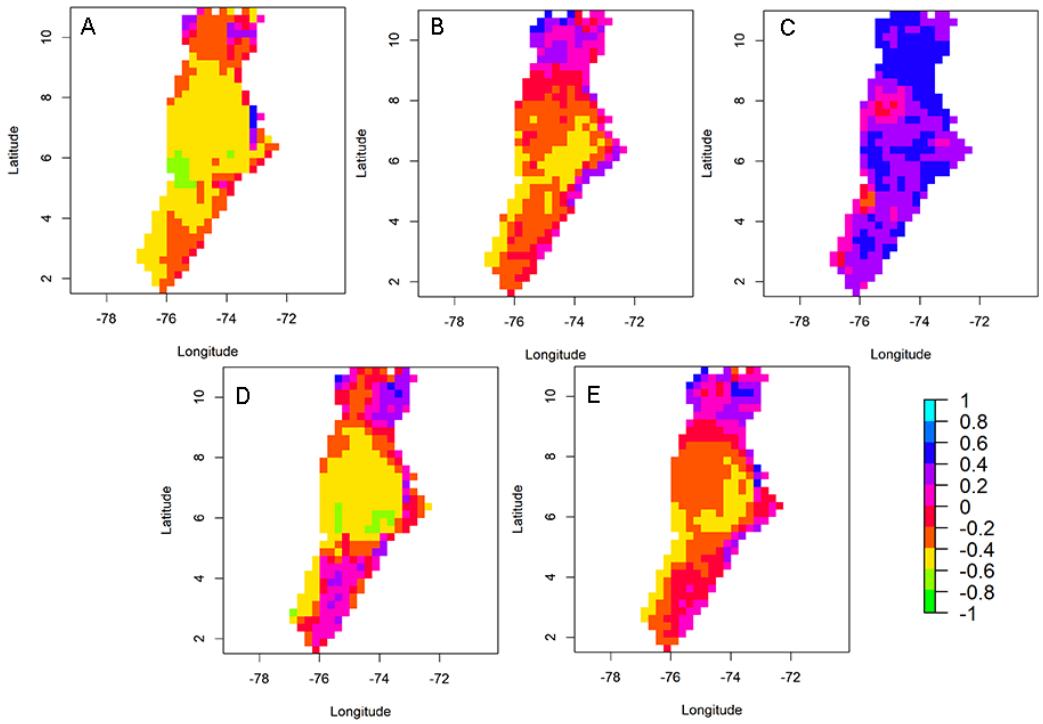

Figure 49. Spatial distribution of the Pearson's correlation between the ETDI and the MEI for the ORCHIDEE (A), HTESSEL (B), WaterGAP3 (C) and PCR-GLOBWB (D) models and the ensemble mean of these four models (E)

4.3.3 SMDI

The correlation with the MEI is negative, and strongest for the ensemble mean, where 78% of the cells have a Pearson's correlation value of -0.50 or higher (Figure 50). The model with the strongest correlation with the MEI is the HTESSEL model: 66.86% of the cells have an r value of -0.50 of higher.

When the MEI is plotted against the SMDI, it becomes apparent that there is stronger Pearson correlation for values around a SMDI of 0 and a MEI of 0, but when values become more positive or negative, the correlation becomes weaker (Figure 51). The fact that the correlation is weaker for minimum and maximum SMDI values, indicated that the MEI has less influence on the SMDI.

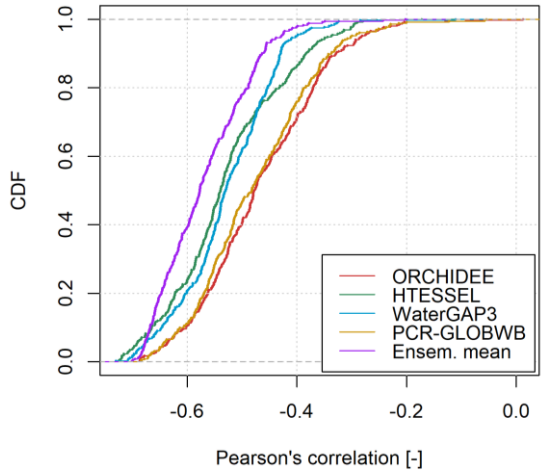

Figure 50. CDF plot of the Pearson's correlation between the SMDI and the MEI for ORCHIDEE, HTESSEL, WaterGAP3, PCR-GLOBWB and the ensemble mean

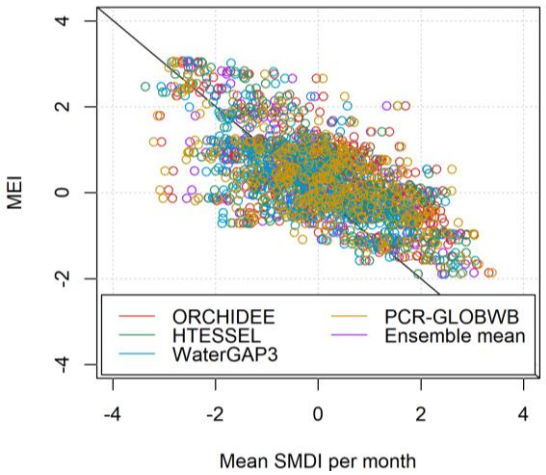

Figure 51. Scatter plot of the spatial mean SMDI values versus the MEI values for ORCHIDEE, HTESSEL, WaterGAP3, PCR-GLOBWB and the ensemble mean

The spatial distribution of the Pearson's correlations is roughly similar per model (Figure 52). The correlation for all models is strongest in the west of the catchment between 2.5° and 6.5° latitude. Also in the north of the catchment between 9° and 10° latitude the correlation is strong. At 6° latitude, the area of stronger correlations spreads almost to the east of the catchment. All models show a generally weak correlation in the far eastern corner of the catchment at 6.5° latitude and in the west of the catchment around 8.25° latitude. Besides these locations, the correlation is also weak at the southeast border of the catchment, which is also the case for the other drought indices. Although the spatial distribution of correlations between the SMDI and the MEI per model are similar, the strength of the correlations per location do vary per model.

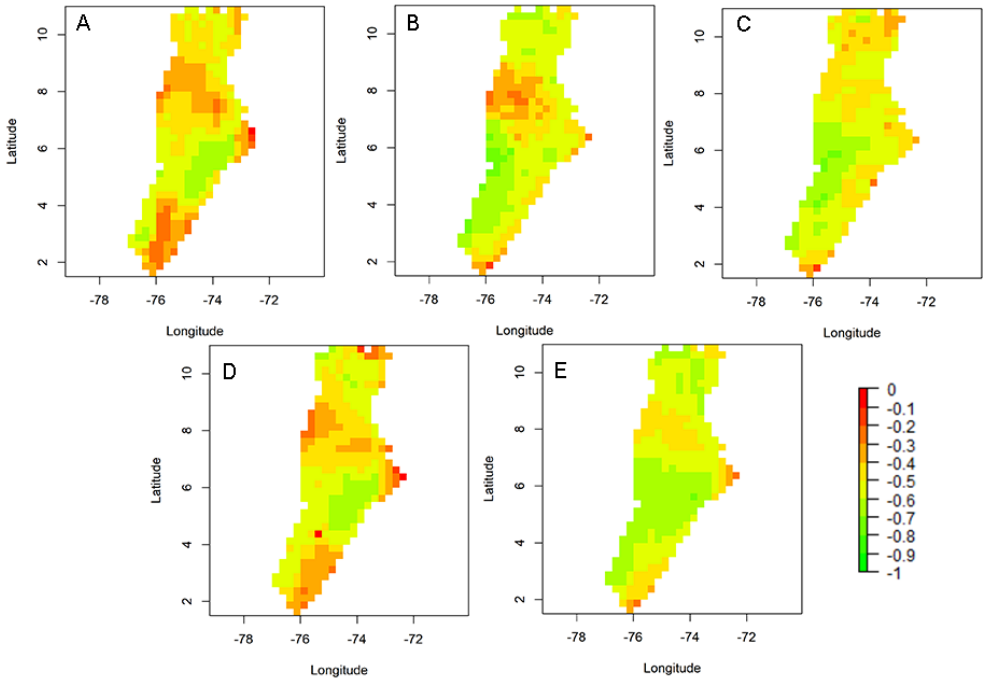

Figure 52. Spatial distribution of the Pearson's correlation between the SMDI and the MEI for the ORCHIDEE (A), HTESSEL (B), WaterGAP3 (C) and PCR-GLOBWB (d) models and the ensemble mean of these four models (E)

4.3.5 SDI

The ensemble mean shows the strongest correlation with the MEI: 49% of the cells have a correlation below -0.50 (Figure 53). The HTESSEL model shows a strong correlation as well, with 48% of the cells having a correlation below -0.50. In the SDI versus MEI plot, all data points are located in a denser cloud compared to the other drought indices, because the variability of the SDI is smaller compared to the other drought indices (Figure 54). This is caused by spatially averaging, because most of the area of the MCMB is not river, and therefore the variations in runoff are lower.

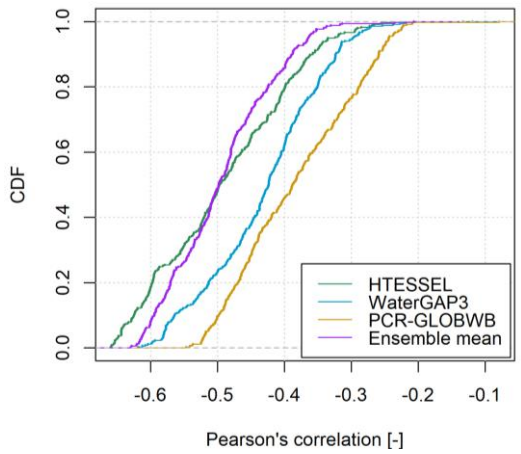

Figure 53. CDF plot of the Pearson's correlation between the SDI and the MEI for HTESSEL, WaterGAP3, PCR-GLOBWB and the ensemble mean

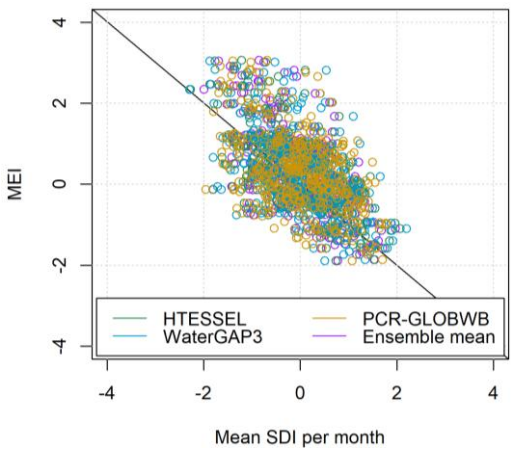

Figure 54. Scatter plot of the spatial mean SDI values versus the MEI values for HTESSEL, WaterGAP3, PCR-GLOBWB and the ensemble mean

The spatial pattern of the Pearson's correlation values is more or less similar to the spatial pattern of the other drought indices (Figure 55). Along the south eastern border of the catchment, the correlation is generally weaker. This is also the case along the borders in the north of the catchment; especially for the north eastern corner and the at the west border at 8 ° latitude, where no major river branches are located. Besides these areas, there is also a small area where the correlation is weaker located in the middle of the catchment around -74.5° longitude and 7.5° latitude. This area corresponds with the most northern part of the Cordillera Central.

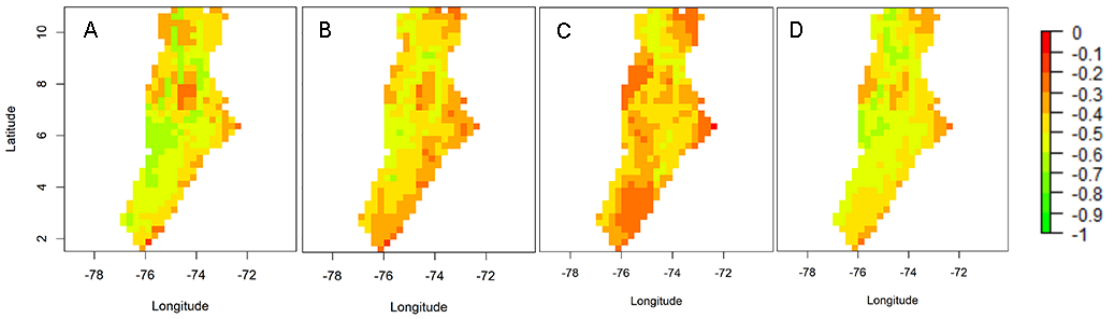

Figure 55. Spatial distribution of the Pearson's correlation between the SDI and the MEI for the HTESSEL (A), WaterGAP3 (B) and PCR-GLOBWB (C) models and the ensemble mean of these three models (D)

5. Discussion

In order to perform a drought analysis four models within the eartH2Observe project WRR-2 were analysed for the Magdalena-Cauca macro-basin in Colombia: ORCHIDEE, HTESSEL, WaterGAP3 and PCR-GLOBWB. This was followed by a drought analysis using drought indices and exploring their relation to ENSO. This study was performed for the period from January 1980 up to and including December 2014 on a monthly basis, with a 0.25° resolution using in-situ data from IDEAM.

5.1 Model performance

The MSWEP forcing data performs well in the MCMB (62% of the cells have a KGE of 0.75 or higher) and therefore this dataset is suitable as forcing for the four large scale hydrological models in this study. Because the forcing dataset has a good performance, errors in, and differences between the model products are unlikely to be caused by the forcing data. At the edges of the catchment, KGE values are generally lower. There are several factors possibly contributing to this reduces performance. The reduced performance in the southeast could be caused by the presence of the Cordilleras Oriental mountain ranges. Although the MSWEP dataset is corrected for gauge under-catch and orographic effects, gauge observations are less dense in mountainous areas (Beck et al., 2017; Viviroli et al., 2007). This can result in a lesser performance of the MSWEP data over mountainous areas. The reduced performance in the north of the catchment could be explained by the proximity to the Atlantic ocean. Oceanic areas are not included in the MSWEP forcing data. Therefore, terrestrial hydrology for coastal areas is not influenced by the adjacent ocean. This can result in a lesser performance in coastal areas (Beck et al., 2018). The interpolation of the precipitation data by IDEAM (Kriging with External Drift (KED), 3.1 Data) can also be a cause for a lower performance.

The performance for the potential evapotranspiration is relatively poor, due to the comparison made with the reference evapotranspiration and due to the different methods used by the models to compute the potential evapotranspiration. Therefore, not too much emphasis should be placed on the bias and variability ratio components of the KGE: the potential evapotranspiration amounts are larger than the reference evapotranspiration amounts because water limitations are not taken into account. Also, the use of a different crop type instead of the reference grass crop or a different computation method will result in different mean potential evapotranspiration values and different standard deviations. Both the GHMs perform better than the LSMs: they are the only two models which produce cells that have a KGE over 0.50, and over half of the cells in the catchment have a Pearson's correlation above 0.50. A potential explanation could be that LSMs compute the potential evapotranspiration through the surface temperature whereas GHMs do not (Barella-Ortiz, 2015). The very poor model performance of the HTESSEL model cannot directly be explained by the use of a different equation to calculate the PET, since both PCR-GLOBWB and HTESSEL use the same equation, and PCR-GLOBWB does not perform as poorly. PCR-GLOBWB

does use a crop coefficient that takes crop-specific transpiration and bare soil evaporation into account (Wada et al., 2014). The HTESSEL model mainly focuses on bare ground evaporation (Trambauer et al., 2013), but improvements on evaporation for temperate climates have been made. Balsamo et al. (2011) state that a point for improvement is the evaporation over free-water surfaces.

Both the performance and the Pearson's correlation of the models for potential evapotranspiration are generally better in the Magdalena valley and the north of the catchment, compared to the rest of the catchment, which can be explained by the land cover and topography. In the Magdalena valley and the north of the catchment, the land cover mainly consists of grassland and crop farming (Figure 3). The topography is also relatively flat and elevation is low; below 500 m (Figure 4). Since the vegetation types are more similar to the crop type used for calculation of the reference evapotranspiration and because the topography is more homogeneous, this explains the good performance of the models in this area.

The model performances for discharge are relatively poor, most likely due to the limited spatial resolution. The models tend to over or underestimate the discharge. For some of the locations more upstream, some models give discharge values close to zero, whereas IDEAM gives significantly larger discharge values, and vice versa. An explanation for this phenomenon might be that due to the cell size, cells contain a mix of rivers and other types of land covers. Some river branches can therefore be assigned to one cell in one model and to another cell in another model, depending on the set threshold. Differences could also be explained by the fact that the models compute discharge over the entire cell, and this value is compared to the discharge at a point location from IDEAM to compute the KGE. If there are multiple river branches in one cell, this will lead to an overestimation of the models. Locations in the MCMB with a larger upstream area often have a higher KGE, most likely because a larger area limits the effect of a local error.

The performance of the MSWEP precipitation dataset is very good for the MCMB. Although for the potential evapotranspiration and for the discharge the KGEs indicate a lesser performance, these model products are still useful for drought analysis because the Pearson's correlations between the IDEAM and modelled data are strong (above 0.6 for a majority of the cells for the best performing models). Because the data is normalized or standardized in the computation procedures of the drought indices, the bias and variability ratio are no longer relevant in a drought analysis using drought indices.

5.2 Drought indices

For all drought indices, the ratio of dry months to all months becomes bigger when the aggregation period is increased. This can either be explained by the lengthening of droughts with increasing aggregation period, and/or by the lack of severe droughts at the start of the time series and the shortening of the observation period for SPIs with a longer aggregation period (because the first months of the time series cannot return a SPI value yet).

For the ETDI, the WaterGAP3 model shows a very different pattern from the other models and the ensemble mean. It could be caused by the relatively large total evaporation values of this model. Due to the calculation method of the ETDI, this results in a relatively small water scarcity ratio (3.4.3 ETDI). This indicates that a lot of evaporation takes place, and has as a consequence that the ETDI of the current month is similar to the ETDI of the previous month, except for when the water stress ratio is large. For the HTESSEL model, the ETDI pattern seems similar to the ORCHIDEE and PCR-GLOBWB models (even though the magnitude of the potential evapotranspiration amounts deviated significantly from the reference evapotranspiration and the other models, whereas this was not the case for the actual evapotranspiration) because the water stress ratio does not become very small.

The difference in the SMDI between models can be explained by the different soil moisture parameters and different soil moisture depths (Table 2). When the soil depth in the model is doubled, the water holding capacity is also doubled. This can have consequences during dry seasons, because deeper soils have a larger buffer of root zone soil moisture (D'Orgeval et al., 2008). During periods with little precipitation, this can have as a consequence that the onset of a drought according to the SMDI is timed later for a model with a deeper soil compared to a model with a shallower soil. It might also have as a consequence that the drought is less intense, meaning that the minimum SMDI is less negative.

The SMDI shows a temporal pattern that is similar to the 3-month and 6-month SPI and SPEI. This finding is supported by the research of Vicente-Serrano et al. (2010) and Zargar et al. (2011), who state that the medium or 3-month SPI, respectively, relates to soil water content. The pattern of the SDI time series is similar to the 3-month and 6-month SPI and SPEI. This confirms that longer timescales SPIs and SPEIs reflect deficits in streamflow, as stated by McKee et al. (1993).

Even though not all models have a good skill, they can be of value when analyzing droughts, because the different models gave different discharge values, the SDIs are relatively similar for all models.

5.3 Drought analysis

For all drought indices except the ETDI, the area affected by drought moves from the northeast to the southwest of the catchment over a period of multiple months. This pattern coincides with El Niño events and can be explained by the effects of El Niño on the Chocó-jet. The reduced sea-surface temperature gradient over the Pacific Ocean occurring during El Niño causes the Chocó-jet to become weaker and reach less far inland every month, while the north-easterly trade winds, which supply drier air, reach further to the west (Poveda et al., 2005). Therefore, the northeast of the catchment is the first to experience a reduced moisture supply. The consecutive year, the Chocó-jet is no longer under the influence of El Niño and it can regain its full strength again from

May to November. Since most precipitation falls north of 6° N, the north of the catchment recovers first, and the south of the catchment recovers later.

The ratio of dry months to all months is higher in the southeast of the catchment. This area receives the least amount of moist air from the Chocó-jet (Poveda et al., 2014), due to the Cordilleras Occidental and Central forming a barrier in the east. The area is also on the leeward side of the north-easterly trade winds, and therefore also receives little moisture from these winds.

More potential and actual evapotranspiration is modelled for the central north of the catchment (where the climate is generally warmer and wetter) and along the main rivers, and less actual evapotranspiration is modelled for the mountains. The spatial pattern of the ETDI also clearly follows the topography. Topography is also recognized as an important influences of (potential) evapotranspiration, due to lower temperature in the mountains, and the aspects of slopes determining the amount of solar radiation received (Kafle & Yamaguchi, 2009).

For the SPI and the SPEI, it stands out that the basin of the catchment does not experience a severe drought following the 1982-1983 El Nino event, even though this region experiences reduced precipitation. This can be explained by the higher precipitation rates in the months preceding the El Niño event, which are also experienced exclusively at this location, forming a buffer to the reduced precipitation amounts after the El Nino event. The increased precipitation rates in this area preceding the El Niño event can be explained by the surrounding topography. The Cordillera Occidental reaches its most northern point just south of this area. Therefore, there is no barrier present blocking the moist air coming from the west transported by the Chocó-jet.

5.4 ENSO

The MEI is calculated based on six components: sea-level pressure, zonal and meridional components of the surface wind, sea surface temperature, surface air temperature, and total cloudiness fraction of the sky. When looking at the time series of all six components, it can be observed that most components have a period of either 6 or 12 months (ICOADS, 2016). This can explain why the drought indices with a temporal aggregation period of 6 months show the strongest correlation with the MEI.

A strong El Niño event is not per definition a severe drought, partly due to the ’82-’83 and the ’86-’87 strong El Niño events, where the SPI does not drop far below -1. This makes it difficult to predict the intensity of future droughts using ENSO.

The correlation between the drought indices and the MEI is the weakest along the south-eastern border of the catchment and strongest at the western border for all drought indices. The most likely explanation is that the Andes acts as a barrier for the effects of ENSO to enter inland Colombia. This observation is also described in literature: Poveda (1998) states that the effect of ENSO decreases when moving in eastern direction as a result of the interactions with the branches of the Colombian Andes. According to Poveda & Mesa (1997), the correlation between the SOI and precipitation are strongest near the Pacific ocean, and correlations decrease towards the east.

For all drought indices except the ETDI (where the strength of the negative correlation for the ensemble mean is weakened by the positive correlation of WaterGAP3), the ensemble mean has the strongest correlation with the MEI. The drought index with the strongest correlation with the MEI is the 6-month ensemble mean SPEI. The ensemble mean corrects for extremes of the individual models, which explains this high correlation. Several other studies also conclude that the ensemble mean outperformed the individual models (Alfieri et al., 2012; van Loon et al., 2012; Velázquez et al., 2013).

6. Conclusion

This study performed a drought analysis for the Magdalena-Cauca macro-basin by I) evaluating four different global hydrological and land surface models; II) quantifying droughts using five different drought indices based on the model products and III) determining the correlation with ENSO.

Overall, the performance is best for the MSWEP precipitation forcing data and lower for the potential evapotranspiration and discharge model products. This is reasonable, because the MSWEP precipitation dataset used multiple comprehensive data sources. Due to the good performance of the forcing data, the MSWEP forcing data is suitable as forcing in the MCMB. For potential evapotranspiration the WaterGAP3 model performs best, however, the comparison between the modelled potential evapotranspiration and the reference evapotranspiration is not ideal. For discharge, the ensemble mean has the most cells with a good performance. The performance is generally better if the upstream area is bigger. If the spatial resolution could be improved, this might increase the performance because it would be better possible to distinguish rivers in the models. Even though the performance of some model products was not adequate, the model products still had value for drought analysis, due to the removal of absolute differences in the drought indices computations.

Droughts occur everywhere in the catchment, but occur most often in the southeast of the catchment. This area is often not reached by the westerly Chocó-jet supplying moist air and is on the leeward side of the Cordilleras Occidental. Therefore, moisture supply is limited. During El Niño events, the area affected by droughts moves from the northeast to the southwest of the catchment, due to weakening of the Chocó-jet and dominance of the easterly trade winds at lower elevations.

The correlation between drought indices and the MEI is weaker for a strongly positive MEI values. The correlation is strongest for the 6-month ensemble mean SPEI. The correlation is strongest in the west of the catchment.

In all, this study paves the way for future research which may use drought indices, and their correlation to ENSO, to predict droughts in the MCMB and assist in water management.

References

Alfieri, L., Thielen, J., & Pappenberger, F. (2012). Ensemble hydro-meteorological simulation for flash flood early detection in southern Switzerland. *Journal of Hydrology*, *424–425*, 143–153. https://doi.org/10.1016/j.jhydrol.2011.12.038

Balsamo, G., Viterbo, P., Beljaars, A., van den Hurk, B., Hirschi, M., & Betts, A. K. (2009). A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System. *Journal of Hydrometeorology*, *10*(3), 623–643. https://doi.org/10.1175/2008JHM1068.1

Balsamo, G., Dutra, E., Beljaars, A., & Viterbo, P. (2011). Evolution of land-surface processes in the IFS. *ECMWF Newsletter*, *127*(127), 6. https://doi.org/10.21957/x1j3i7bz

Barella-Ortiz, A. (2015). *Analysis and modelling of soil moisture and evaporation processes , implications for climate change*. Université Pierre et Marie Curie – Paris VI.

Barella-Ortiz, A., Polcher, J., Tuzet, A., & Laval, K. (2013). Potential evaporation estimation through an unstressed surface-energy balance and its sensitivity to climate change. *Hydrology and Earth System Sciences*, *17*(11), 4625–4639. https://doi.org/10.5194/hess-17-4625-2013

Barnston, A. G., Glantz, M. H., & He, Y. (1999). Predictive skill of statistical and dynamical climate models in SST forecasts during the 1997-98 El Niño episode and the 1998 La Niña onset. *Bulletin of the American Meteorological Society*, *80*(2), 217–243. https://doi.org/10.1175/1520-0477

Beck, H.E., van Dijk, A. I. J. M., De Roo, A., Dutra, E., Fink, G., Orth, R., & Schellekens, J. (2017). Global evaluation of runoff from 10 state-of-the-art hydrological models. *Hydrology and Earth System Sciences*, *21*, 2881–2903. https://doi.org/10.5194/hess-2016-124

Beck, H.E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., & De Roo, A. (2017). MSWEP: 3-hourly 0.25° global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences*, *21*(1), 589–615. https://doi.org/10.5194/hess-21-589-2017

Beck, Hylke E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., van Dijk, A. I. J. M., et al. (2018). Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. *Hydrology and Earth System Sciences Discussions*, (September), 1–23. https://doi.org/10.1016/j.cub.2014.12.055

van Beek, L. P. H., & Bierkens, M. F. P. (2009). *The global hydrological model PCR-GLOBWB: conceptualization, parameterization and verification*. *Report Department of Physical Geography*. Utrecht, The Netherlands. Retrieved from http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf

van Beek, L. P. H., Wada, Y., & Bierkens, M. F. P. (2011). Global monthly water stress: 1. Water balance and water availability. *Water Resources Research*, *47*(7). https://doi.org/10.1029/2010WR009791

Bergstrom, S. (1976). *Development and application of a conceptual runoff model for Scandinavian catchments*. Norrkoping.

Beven, K., & Binely, A. (1992). The future of distributed models: Calibration and uncertainty prediction. *Hydrological Processes*, *6*, 279–298.

Bierkens, M. F. P. (2015). Global hydrology 2015: State, trends, and directions. *Water Resources Research*, *51*(February), 4923–4947. https://doi.org/10.1002/2015WR017173.Received

Chen, D., Cane, M. A., Kaplan, A., Zebiak, S. E., & Huang, D. (2004). Predictability of El Niño over the past 148 rears. *Nature*, *428*(15 April 2004), 4.

Commons, W. (2019). Colombia Topography. Retrieved April 29, 2019, from https://upload.wikimedia.org/wikipedia/commons/3/3b/Colombia_Topography.png

D'Orgeval, T., Polcher, J., & De Rosnay, P. (2008). Sensitivity of the West African hydrological cycle in ORCHIDEE to infiltration processes. *Hydrology and Earth System Sciences*, *12*(6), 1387–

1401. https://doi.org/10.5194/hess-12-1387-2008

Döll, P., Fiedler, K., & Zhang, J. (2009). Global-scale analysis of river flow alterations due to water withdrawals and reservoirs. *Hydrology and Earth System Sciences*, *13*(12), 2413–2432. https://doi.org/10.5194/hess-13-2413-2009

Döll, P., Hoffmann-Dobrev, H., Portmann, F. T., Siebert, S., Eicker, A., Rodell, M., et al. (2012). Impact of water withdrawals from groundwater and surface water on continental water storage variations. *Journal of Geodynamics*, *59–60*, 143–156. https://doi.org/10.1016/j.jog.2011.05.001

FAO. (2002). *FAO Rice Information*. Rome. Retrieved from http://www.fao.org/3/Y4347E/y4347e0h.htm

Flörke, M., Kynast, E., Bärlund, I., Eisner, S., Wimmer, F., & Alcamo, J. (2013). Domestic and industrial water uses of the past 60 years as a mirror of socio-economic development: A global simulation study. *Global Environmental Change*, *23*(1), 144–156. https://doi.org/10.1016/j.gloenvcha.2012.10.018

Gallego, D., García-Herrera, R., Gómez-Delgado, F. de P., Ordoñez-Perez, P., & Ribera, P. (2018). Tracking the Choco jet since the 19th Century by using historical wind direction measurements. *Earth System Dynamics Discussions*, (July), 1–23. https://doi.org/10.5194/esd-2018-54

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Gutiérrez, F., & Dracup, J. A. (2001). An analysis of the feasibility of long-range streamflow forecasting for Colombia using El Niño–Southern Oscillation indicators. *Journal of Hydrology*, *246*, 181–196.

Hargreaves, G. H., & Samani, Z. A. (1985). Reference crop evapotranspiration from temperature. *Applied Engineering in Agriculture*, *1*(2), 96–99.

Heim Jr., R. R. (2002). A Review of Twentieth-Century Drought Indices Used in the United States. *Bulletin of the American Meteorological Society*, *22*(August), 1149–1165.

Hoyos, N., Escobar, J., Restrepo, J. C., Arango, A. M., & Ortiz, J. C. (2013). Impact of the 2010-2011 La Niña phenomenon in Colombia, South America: The human toll of an extreme weather event. *Applied Geography*, *39*, 16–25. https://doi.org/10.1016/j.apgeog.2012.11.018

Hoyos, N., Correa-Metrio, A., Sisa, A., Ramos-Fabiel, M. A., Espinosa, J. M., Restrepo, J. C., & Escobar, J. (2017). The environmental envelope of fires in the Colombian Caribbean. *Applied Geography*, *84*, 42–54. https://doi.org/10.1016/j.apgeog.2017.05.001

ICOADS. (2016). Release 2.5 Data characteristics. Retrieved March 3, 2019, from https://icoads.noaa.gov/r2.5.html

Kafle, H. K., & Yamaguchi, Y. (2009). Effects of topography on the spatial distribution of evapotranspiration over a complex terrain using two-source energy balance model with ASTER data. *Hydrological Processes*, *23*(16), 2295–2306. https://doi.org/10.1002/hyp.7336

Keyantash, J., & Dracup, J. A. (2002). The quantification of drought: An evaluation of drought indices. *American Meteorological Society*, (August), 1167–1180.

KNMI. (2019). Time series monthly MEI. Retrieved January 28, 2019, from https://climexp.knmi.nl/getindices.cgi?WMO=NOAAData/mei&STATION=MEI&TYPE=i&id=someone@somewhere

Lander, R. (2015, September 22). Colombia suffering worst drought in recorded history. *Colombia Reports*. Retrieved from https://colombiareports.com/colombia-suffering-worst-drought-in-recorded-history/

van Loon, A. F. (2015). Hydrological drought explained. *WIREs Water*, *2*, 359–392. https://doi.org/10.1002/wat2.1085

van Loon, A. F., van Huijgevoort, M. H. J., & van Lanen, H. A. J. (2012). Evaluation of drought

propagation in an ensemble mean of large-scale hydrological models. *Hydrology and Earth System Sciences*. https://doi.org/10.5194/hess-16-4057-2012

López, M. E., & Howell, W. E. (1967). Katabatic winds in the equatorial Andes. *Journal of the Atmospheric Sciences*, *24*, 29–35.

Matheron, G. (1971). *The theory of regionalized variables and their applications*. Fontainebleau, Paris: Centre de Geostatistique.

Mavromatis, T. (2007). Drought index evaluation for assessing future wheat production in Greece. *International Journal of Climatology*, *4*(27), 911–924. https://doi.org/10.1002/joc.1444

McKee, T. B., Nolan, J., & Kleist, J. (1993). The relationship of drought frequency and duration to time scales. In *Eighth Conference on Applied Climatology, American Meteorology Society*.

Mishra, A. K., & Singh, V. P. (2011). Drought modeling - A review. *Journal of Hydrology*, *403*(1–2), 157–175. https://doi.org/10.1016/j.jhydrol.2011.03.049

Nalbantis, I., & Tsakiris, G. (2009). Assessment of hydrological drought revisited. *Water Resources Management*, *23*(5), 881–897. https://doi.org/10.1007/s11269-008-9305-1

Narasimhan, B., & Srinivasan, R. (2005). Development and evaluation of Soil Moisture Deficit Index (SMDI) and Evapotranspiration Deficit Index (ETDI) for agricultural drought monitoring. *Agricultural and Forest Meteorology*, *133*(1–4), 69–88. https://doi.org/10.1016/j.agrformet.2005.07.012

Poveda, G. (1998). Retroalimentación dinámica entre El Niño Oscilación del Sur y la hidrología de Colombia. Disertación Ph.D., Universidad Nacional de Colombia. Posgrado en Aprovechamiento de Recursos Hidráulicos, Medellín.

Poveda, G., & Mesa, O. (2000). On the existence of Lloró (the rainiest locality on Earth): enhanced ocean-land-atmosphere interaction by a low-level jet. *Geophysical Research Letters*, *27*(11), 1675–1678.

Poveda, G., & Mesa, O. J. (1997). Feedbacks between hydrological processes in tropical South America and large-scale ocean-atmospheric phenomena. *Journal of Climate*, *10*(10), 2690–2702. https://doi.org/10.1175/1520-0442(1997)010<2690:FBHPIT>2.0.CO;2

Poveda, G., Jaramillo, A., Gil, M. M., Quiceno, N., & Mantilla, R. I. (2001). Seasonality in ENSO-related precipitation, river discharges, soil moisture, and vegetation index in Colombia. *Water Resources Research*, *37*(8), 2169–2178. https://doi.org/10.1029/2000wr900395

Poveda, G., Mesa, O. J., Salazar, L. F., Arias, P. A., Moreno, H. A., Vieira, S. C., et al. (2005). *The Diurnal Cycle of Precipitation in the Tropical Andes of Colombia*.

Poveda, G., Waylen, P. R., & Pulwarty, R. S. (2006). Annual and inter-annual variability of the present climate in northern South America and southern Mesoamerica. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *234*(1), 3–27. https://doi.org/10.1016/j.palaeo.2005.10.031

Poveda, G., Vélez, J. I., Mesa, O. J., Cuartas, A., Barco, J., Mantilla, R. I., et al. (2007). Linking long-term water balances and statistical scaling to estimate river flows along the drainage network of Colombia. *Journal of Hydrologic Engineering*, *12*(1), 4–13. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:1(4)

Poveda, G., Jaramillo, L., & Vallejo, L. F. (2014). Seasonal precipitation patterns along pathways of South American low-level jets and aerial rivers. *Water Resources Research*, *50*(1), 98–118. https://doi.org/10.1002/2013WR014087

Restrepo, J. D., & Kjerfve, B. (2000a). Magdalena river: Interannual variability (1975-1995) and revised water discharge and sediment load estimates. *Journal of Hydrology*, *235*(1–2), 137–149. https://doi.org/10.1016/S0022-1694(00)00269-9

Restrepo, J. D., & Kjerfve, B. (2000b). Water discharge and sediment load from the western slopes of the colombian andes with cocus on Rio San Juan. *The Journal of Geology*, *108*(1), 17–33. https://doi.org/10.1086/314390

Restrepo, J. D., & Syvitski, J. P. M. (2006). Assessing the effect of natural controls and land use

change on sediment yield in a major Andean River: the Magdalena drainage basin, Colombia. *Ambio*, *35*(2), 65–74. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16722251

Restrepo, J. D., Kjerfve, B., Hermelin, M., & Restrepo, J. C. (2006). Factors controlling sediment yield in a major South American drainage basin: The Magdalena River, Colombia. *Journal of Hydrology*, *316*(1–4), 213–232. https://doi.org/10.1016/j.jhydrol.2005.05.002

Schellekens, J., Dutra, E., Martínez-De La Torre, A., Balsamo, G., Van Dijk, A. I. J. M., Sperna Weiland, F., et al. (2017). A global water resources ensemble of hydrological models: The eartH2Observe Tier-1 dataset. *Earth System Science Data*, *9*(2), 389–413. https://doi.org/10.5194/essd-9-389-2017

Sterk, G., & Rodriguez, E. (2017). Colombia. In *Global Earth Observation for integrated water resource assessment: Applicability of Earth Observation and Global Hydrological Data for Local Applications* (pp. 59–78).

Thornthwaite, C. W. (1948). An approach toward a rational classification of climate. *Geographical Review*, *38*(1), 55–94. https://doi.org/10.2307/210739

Trambauer, P., Maskey, S., Winsemius, H., Werner, M., & Uhlenbrook, S. (2013). A review of continental scale hydrological models and their suitability for drought forecasting in (sub-Saharan) Africa. *Physics and Chemistry of the Earth*. https://doi.org/10.1016/j.pce.2013.07.003

Velázquez, J. A., Schmid, J., Ricard, S., Muerth, M. J., Gauvin St-Denis, B., Minville, M., et al. (2013). An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources. *Hydrology and Earth System Sciences*, *17*(2), 565–578. https://doi.org/10.5194/hess-17-565-2013

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *Journal of Climate*, *23*(7), 1696–1718. https://doi.org/10.1175/2009JCLI2909.1

Viviroli, D., Dürr, H. H., Messerli, B., Meybeck, M., & Weingartner, R. (2007). Mountains of the world, water towers for humanity: Typology, mapping, and global significance. *Water Resources Research*. https://doi.org/10.1029/2006WR005653

Wackernagel, H. (1998). *Multivariate geostatistics: an introduction with applications*. Springer-Verlag.

Wada, Y., Wisser, D., & Bierkens, M. F. P. (2014). Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources. *Earth System Dynamics*, *5*(1), 15–40. https://doi.org/10.5194/esd-5-15-2014

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., & Viterbo, P. (2014). The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resources Research*, *50*, 7505–7514. https://doi.org/10.1002/2014WR015638.Received

Werner, M., & Gründermann, G. (2014). Lesson 1: Using the EartH2Observe data portal to analyse drought indicators. *Using the EartH2Observe Data Portal to Analyse Drought Indicators*, 1–30.

Wilhite, D. A. (2000). Drought as a natural hazard: Concepts and definitions. In *Drought: A Global Assessment* (pp. 3–18). New York: Routledge. https://doi.org/10.1177/0956247807076912

Wolter, K., & Timlin, M. S. (2011). El Niño Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI. ext). *International Journal of Climatology*, *31*(7), 1074–1087.

Yamazaki, D., Kanae, S., Kim, H., & Oki, T. (2011). A physically based description of floodplain inundation dynamics in a global river routing model. *Water Resources Research*, *47*(4), 1–21. https://doi.org/10.1029/2010WR009726

Zargar, A., Sadiq, R., Naser, B., & Khan, F. I. (2011). A review of drought indices. *Environmental Reviews*, *19*, 333–349. https://doi.org/10.1139/a11-013

Zhao, M., Held, I. M., & Vecchi, G. A. (2010). Retrospective forecasts of the hurricane season using a global atmospheric model assuming persistence of SST anomalies. *Monthly Weather Review*,

138(10), 3858–3868. https://doi.org/10.1175/2010MWR3366.1