EUPORIAS

THEME ENV.2012.6.1-1

EUPORIAS

(Grant Agreement 308291)

EUPORIAS

European Provision Of Regional Impact Assessment on a

Seasonal-to-decadal timescale

Deliverable D21.2

Report on the added value of dynamical and statistically downscaled data

EUPORIAS

Deliverable Title	*Report on the added value of dynamical and statistically downscaled data*
Brief Description	*Report on the added value of dynamical and statistically downscaled data: Report on the added value of dynamical and statistically downscaled data based on anomaly-initialized GCM hindcasts compared to basic seasonal GCM hindcasts*
WP number	21
Lead Beneficiary	*Grigory Nikulin (Work Package Leader, **SMHI**)*
Contributors	***SMHI***: Ulf Hansson, Michael Kolax and Shiyu Wang; ***UC***: Jesus Fernández, Rodrigo Manzanas, María Eugenia Magariño, María Dolores Frías; ***DWD***: Shakeel Asharaf, Barbara Früh and Kristina Fröhlich; ***UL-IDL***: Rita Cardoso, Ricardo Tome, Pedro Soares; ***ENEA***: Sandro Calmanti; ***Met Office***: Simon Tucker and Carlo Buontempo.
Creation Date Version Number Version Date	2015.05.29 1 2015.05.29
Deliverable Due Date Actual Delivery Date	2015.04.30 2015.05.31
Nature of the Deliverable	x	*R - Report*
		P - Prototype
		D - Demonstrator
		O - Other
Dissemination Level/ Audience	x	*PU - Public*
		PP - Restricted to other programme participants, including the Commission services
		RE - Restricted to a group specified by the consortium, including the Commission services
		CO - Confidential, only for members of the consortium, including the Commission services

Version	Date	Modified by	Comments
1	2015.05.29	Grigory Nikulin	Finalised based on review comments

EUPORIAS

Table of Contents

List of Tables

List of Figures

EUPORIAS

1 Executive Summary

This report details an ensemble of seasonal hindcast generated by a global model EC-EARTH and then downscaled by five regional models and by two statistical methods over eastern Africa with focus on Ethiopia. The hindcast includes 15 members, initialised on May 1st and covers the period 1991-2012. There are two sub-regions where the global hindcast has some skill in predicting June-September precipitation (northern Ethiopia - North-East Sudan and southern Sudan - northern Uganda), while there is no skill elsewhere in the region. The RCMs are able to capture and reproduce the signal evident in the driving EC-EARTH seasonal hindcast over Ethiopia in June-September showing about the same performance as their driving GCM. Statistical downscaling, in general, loses a part of the EC-EARTH skill and shows a weaker performance in terms of predictability. At the same time there are no clear evidences that the RCMs provide the added value compared to the driving EC-EARTH if we define the added value as a higher forecast skill in the RCM hindcast. This conclusion is only for Ethiopia in the June-September season and cannot be generalised for other regions and seasons.

2 Project Objectives

With this deliverable, the project has contributed to the achievement of the following objectives (DOW, Section B1.1):

No.	Objective	Yes	No
1	Develop and deliver reliable and trusted impact prediction systems for a number of carefully selected case studies. These will provide working examples of end to end climate-to-impacts-decision making services operation on S2D timescales.		x
2	Assess and document key knowledge gaps and vulnerabilities of important sectors (e.g., water, energy, health, transport, agriculture, tourism), along with the needs of specific users within these sectors, through close collaboration with project stakeholders.		x
3	Develop a set of standard tools tailored to the needs of stakeholders for calibrating, downscaling, and modelling sector-specific impacts on S2D timescales.	x	
4	Develop techniques to map the meteorological variables from the prediction systems provided by the WMO GPCs (two of which (Met Office and MeteoFrance) are partners in the project) into variables which are directly relevant to the needs of specific stakeholders.	x	

5	Develop a knowledge-sharing protocol necessary to promote the use of these technologies. This will include making uncertain information fit into the decision support systems used by stakeholders to take decisions on the S2D horizon. This objective will place Europe at the forefront of the implementation of the GFCS, through the GFCS's ambitions to develop climate services research, a climate services information system and a user interface platform.		x
6	Assess and document the current marketability of climate services in Europe and demonstrate how climate services on S2D time horizons can be made useful to end users.		x

3 Detailed Report

3.1   Introduction

In the last decades a significant progress has been achieved in the prediction of seasonal mean states of weather and, therefore, seasonal forecasting has become an operational activity in a number of national weathers services worldwide (Graham et al. 2011). Global seasonal prediction systems are being used increasingly, operating at a 50-200km range of resolution, while many users require seasonal forecast at regional to local scales. A common approach for providing high resolution climate information in the climate projection framework is to supplement global models by empirical-statistical or dynamical downscaling techniques (ESD or DD respectively). To derive regional climate information ESD applies a statistical relationship between information from global models and local-scale processes while DD uses regional climate models (RCMs) driven by global models. ESD is more widely used in seasonal predictions as a computationally efficient approach when large amount of hindcasts and forecasts are downscaled but depends on the availability of observations and is limited to a few variables. Dynamical downscaling using RCMs is computationally expensive and requires much more recourses than ESD, saving 6-hr boundary conditions from GCMs for example. However, in contrast to the ESD approach RCMs can provide a larger number of variables in a physically consistent way, including regional and local feedbacks which can be important in seasonal forecasting.

Due to its simplicity, ESD is applied in seasonal forecasting more often than RCMs, although there are only a few institutions where ESD is used operationally. There are a number of studies applying RCMs for downscaling seasonal forecast but such studies are more experimental (not operational activities) and usually include a single RCM (Díez et al. 2011; Castro et al. 2012; Diro et al. 2012; Cheneka et al. 2016;). The added value of RCM in seasonal predictions is still not clearly evident, whether RCMs can provide more skilful forecast compared to their driving GCM or not, especially considering their higher cost. While some studies show benefits of downscaled seasonal forecasts other studies indicate no benefits.

One of EUPORIAS activities is the provision of downscaled and bias-corrected seasonal forecasts for use in EUPORIAS impact and climate service applications. The first focus area

in the downscaling EUPORIAS activities is Europe where high-quality observations exist and ESD methods are applied (see Deliverable D21.1). A second focus in the downscaling activities being eastern Africa, where temperature and precipitation exhibit better predictability at seasonal timescale than in the extra-tropics, potentially allowing to apply forecast data in sectors such as: food security, drought early-warning systems and human health. In addition, East Africa has a complex orography and thus could provide a perfect test bench for downscaling methodologies. The EUPORIAS eastern Africa activities assess the utility of downscaling techniques to provide seasonal forecast data for impact models over the region. In particular, the World Food Programme (WFP) was planning to assess the utility of downscaled data in their Livelihoods, Early Assessment and Protection (LEAP) system in Ethiopia. In EUPORIAS we therefore investigate the ability of RCMs and ESD methods to downscale global seasonal hindcasts over East Africa, assessing the level of added-value compared to the driving GCM hindcasts.

3.2   Data and Methods

3.2.1   Global hindcasts

At the beginning of the EUPORIAS project in 2011, after consultations with the WFP, it was decided to focus on the Kiremt rainy season (June-September) in Ethiopia using seasonal hindcast initialised in May, which can be used as input to the LEAP system. This was a trade-off between user needs, more keen on rainy season forecasts when impacts of water deficits on agriculture are larger, and forecast skill, which peaked in November-January, associated to ENSO variability. We finally opted for addressing the end-user needs, focusing on JJAS.

The first step was to provide boundary conditions from a global seasonal forecast to regional climate modelling groups for subsequent downscaling. A straightforward approach was to downscale the ECMWF System 4 (S4) seasonal hindcast (Molteni et al. 2011) which became available in 2011. However, the S4 model levels necessary for downscaling were not archived for all members of the seasonal hindcast ensemble and only every second model level is saved.

To provide consistent boundary conditions for downscaling SMHI reran the first 15 members of the S4 hindcast, initialised on May 1st, by a coupled global climate model, EC-EARTH (v. 3.1) in the atmospheric-only mode. EC-EARTH is a consortium model (http://www.ec-earth.org/) contributed to the CMIP5 activities and based on the ECMWF Integrated Forecast System (Hazeleger et al. 2010). One additional advantage is that EC-EARTH can be configured to use exactly the same resolution (T255) and the same 91 vertical levels as in the System 4. Atmospheric initial conditions (ICs) on May 1st were generated at the ECMWF using a methodology similar to the one applied in S4. The same atmospheric ICs are used for all 15 members since influence of the atmospheric initial conditions is small beyond two weeks and does not impact the skill of seasonal forecasts. Temperature and soil moisture ICs are taken directly from S4 and they both are the same for all 15 members as in S4. Finally, since the drift in seasonal forecasts in the tropics is mainly related to sea surface temperature (SST) a bias correction replacing the S4 monthly mean SST climatology with the ERA-Interim reanalysis SST climatology (but preserving anomalies) has been applied. Thus, the members of the EC-EARTH hindcast are different only in SST. Following this approach an ensemble of 5-month (May-September) global seasonal hindcast has been generated at SMHI by EC-EARTH taking the above initial conditions and the bias-corrected

SST from the S4 hindcast. The ensemble includes 15 members, initialised on May 1st and covers the period 1991-2012. Note that this global forecast procedure depends on the availability of S4 SST and cannot be extended operationally. An operational forecast system potentially including dynamical downscaling is developed at Met Office and described in section 3.2.4.

3.2.2   Regional Climate Models

Five groups contribute by downscaling the EC-EARTH hindcast using four different RCMs listed in Table 1. Two groups use the Weather Research and Forecasting Model (WRF) but applying different versions and configurations originated in the Euro-CORDEX activities. The WRF Euro-CORDEX community shows that an ensemble of WRF simulations with different combinations of parameterizations generates a spread across the ensemble member similar to a multi RCM ensemble (Katragkou et al. 2015, García-Díez et al. 2015). A number of domain configurations for eastern Africa, different in size and resolution, were tested. Considering the computational costs, a common domain at 0.22° resolution has been finally set up (Figure 1) and used by all groups. Two streams of downscaling, depending on resources available, have been defined. Namely:

i)   Full hindcast – all 15 members and all years;
ii)   A subset of the full hindcast – 15 members for four preselected years (two wet years - 2006/2007 and two dry years - 2002/2009 in Ethiopia) and the first three members for all years in order to establish the downscaled hindcast climatology.

The full hindcast was downscaled by DWD and SMHI and the subset by ENEA, UCAN and UL-IDL. No nudging toward EC-EARTH was applied in any of the RCMs within the model domain. For a downscaling seasonal forecast system developed at Met Office see section 3.2.4.

Table 1. List of RCMs and their details.

Institution	RCM (short name)	Stream	Full domain	Reference
DeutscherWetterdienst (DWD)	CCLM4-8-21 (CCLM4)	full	24°E - 65°E 9°S - 27°N	Rockel et al. 2008
Swedish Meteorological and Hydrological Institute (SMHI)	RCA4 (RCA4)	full	-24°E - 65°E -14°S - 38°N	Strandberg et al. 2015
Italian National Agency for New Technologies, Energy (ENEA)	RegCM-4-3 (RegCM4)	subset	10°S -23°N 19°E-68°E	Giorgi et al. 2012
Universidad de Cantabria (UCAN)	WRF341I (WRF341)	subset	24°E - 64°E 9°S - 27°N	Skamarock et al. 2008; Katragkou et al. 2015

Universidade de Lisboa (UL-IDL)	WRF360D (WRF360)	subset	24°E- 64°E 9°S - 27°N	Skamarock et al. 2008; Katragkou et al. 2015

3.2.3   Empirical Statistical downscaling

In addition to the dynamical downscaling, two ESD methods were applied by UCAN to both S4 and EC-EARTH hindcasts (all members and 1991-2010). Both methods use a Generalized Linear Model (GLM), which is based on large-scale information at the four nearest grid points for the first method (4NN) and on the first 15 principal components for the second method (15PC). The latter considers a predictor spatial domain similar to the computational domain of the RCMs (27°E-57°E, 2.5°S-20°N). The two methods were calibrated and cross-validated using daily precipitation from WFDEI (in particular its GPCC-calibrated version) as predictand dataset and the ERA-Interim reanalysis as predictor dataset. Based on the results from a screening process to select informative predictors, a combination of zonal wind at 850 and 250 hPa, specific humidity at 500 hPa, and temperature at 500 hPa over the East Africa EUPORIAS domain was considered. Given that these fields are available also for S4, ESD was also applied to this global forecast in addition to the EC-EARTH one.

3.2.4   GloSea5 Downscaling Experimental Setup for EUPORIAS

In parallel with the above downscaling activities based on the EC-EARTH hindcast the Met Office uses the Global Seasonal Forecast System, version 5 (GloSea5) to generate boundary conditions for high resolution regional climate modelling over East Africa. GloSea5 uses the Met Office Hadley Centre Global Environment Model, version 3 (HadGEM3), implementing The Met Office Global Coupled model 2.0 (GC2) configuration (Williams et al. 2015), which consists of the following components and their configurations:

- **Atmosphere**: MetUM (Brown et al., 2012), Global Atmosphere 6.0 (Walters et al., 2016)
- **Land surface**: Joint UK Land Environment Simulator (JULES; Best et al., 2011), Global Land 6.0 (Walters et al., 2016)
- **Ocean**: NEMO (Madec, 2008), Global Ocean 5.0 (Megann et al., 2014)
- **Sea-ice**: The Los Alamos Sea Ice Model (CICE; Hunke and Lipscomb, 2010), Global Sea-Ice 6.0 (Rae et al. 2015)

The MetUM consists of the ENDGAME dynamical core, which solves the equations of motion of a fluid on a semi-Lagrangian, three-dimensional grid. The above components are coupled, i.e. they are not treated separately, but instead information is exchanged between them. Additionally, the ocean and atmosphere are discretized using a tripolar grid, which does not contain singularities at the North and South poles. The grid resolutions are:

- **Atmosphere**:
  - Horizontally: N216. This corresponds to a horizontal resolution of 0.83 degrees longitude by 0.56 degrees latitude,
  - Vertically: resolved into 85 levels,
- **Ocean**:
  - Horizontally: ORCA 0.25, which corresponds to one quarter of a degree,

o Vertically: resolved into 75 levels.

For additional technical details, refer to MacLachlan et al. 2014 and Arribas et al. 2011.

Here GloSea5 is used first to generate the regional climate model (RCM) climatology for the JJAS season of interest with a three member stochastic physics ensemble for the years 1991-2012. Secondly a 15 member ensemble for four years of interest: two wet years (2006 and 2007) and two dry years (2002 and 2009) is generated. These hindcasts are dynamically downscaled by passing boundary conditions to a 12 km RCM, HadGEM3-RA, which uses the same GA6 atmosphere configuration as GloSea5.

At the Met Office, the GloSea5 system is run operationally by initialising two 120 day and two 60 day simulations every day. Should downscaling be shown to be of value then it would be theoretically possible to include it in the operational system.

3.2.5 Observations

The main focus of the study is on precipitation and there are a large number of gridded precipitation observational datasets covering Africa at various temporal (from hourly to monthly) and spatial resolutions (from 0.0375° to 2°). However, even if the observational datasets agree quite well with respect to large-scale precipitation pattern, significant deviations across them can occur locally (Nikulin et al. 2012). To estimate observational uncertainties, we include in our analysis a number of gridded precipitation products (Table 2). Three of them (GPCC, CRU and UDEL) are gauge-based only datasets while the rest (TARCAT, ARC, FEWS and CHIRPS) are satellite-gauge combinations. The WFDEI dataset is a quasi-observational product and presents the bias-corrected 3-hourly ERA-Interim reanalysis where the CRU or GPCC observations are used as reference for adjustment. Two global reanalyses, ERA-Interim and MERRA, are also included, mainly for analysis of large-scale circulation and the HadISST dataset provides sea surface temperature.

Table 2. List of observational datasets and their details.

Dataset	Dataset acronym	Version	Resolution	Source	Reference
Climate Research Unit Time-Series	CRU	3.23	0.5°	stations	Harris et al. 2013
Global Precipitation Climatology Centre	GPCC	7	0.5°	stations	Schneider et al. 2015
University of Delaware	UDEL	4.01	0.5°	stations	Legates and Willmott
Tropical Applications of Meteorology using SATellite data and ground-based observations	TAMSAT or TARCAT	2.0	0.0375°	satellite and stations	Maidmentet al. 2014
African Rainfall Climatology	ARC	2.0	0.1°	satellite and stations	Nicholas et al. 2012

African Rainfall Estimation Algorithm (from the Famine Early Warning System)	FEWS	2.0	0.1°	satellite and stations	http://www.cpc.noaa.gov/products/fews/RFE2.0_tech.pdf
Climate Hazards Group InfraRed Precipitation with Stations	CHIRPS	2.0	0.05°	satellite and stations	Funk et al. 2015
WATCH-Forcing-Data-ERA-Interim	WFDEI	N/A	0.5°	bias corrected ERA-Interim reanalysis (CRU and GPCC as reference)	Weedon et al. 2014
ECMWF Interim Reanalysis	ERAINT	N/A	0.75°	reanalysis	Dee et al. 2011
NASA Modern Era Reanalysis for Research and Applications	MERRA	2	0.625°×0.5°	reanalysis	Rienecker et al. 2011
Met Office Hadley Centre's sea ice and sea surface temperature	HadISST	1.1	1°	ocean measurements	Rayner et. al. 2003

3.2.6   Sub regions

Four observational datasets (ARC, TAMSAT, WFDEI and REF), used as input to LEAP, are taken for definition of geographical sub-regions in Ethiopia. Grid points are clustered according to the Euclidean distance between the associated average seasonal cycles of the monthly cumulated rainfall. The Ward hierarchical clustering with a cut-off at 4 branches of the dendogram was applied (Figure 2.). Depending on the underlying dataset, clusters have different patterns, although large-scale features are about the same. Taking into account that the WFDEI is scaled by the GPCC and GPCC is used as one of the main reference observational datasets it was decided to use the four clusters based on the WFDEI rainfall (Figure 2. d). Seasonal cycle of rainfall is pretty close for the clusters 2, 3 and 4 (see Figure 3) and an additional cluster (5) is defined as a combination of clusters 2, 3 and 4.

3.2.7   Verification metrics

A number of deterministic and probabilistic verifications metrics are applied and a short summary of the used metrics is as follows:

a. *Anomaly correlation coefficient (ACC):* It is a traditional deterministic approach for forecast verifications. As the name suggest, *ACC* is simply correlation between forecast and observed anomalies. The *ACC* measures how well a forecast captures the magnitudes of anomalies from reference time series. It varies from –1 to +1. If the

forecast is perfect, the score of *ACC* is equal to +1. For a worst forecast the score is -1, whereas if the score is zero then there is no skill in the forecast. Here we used the Pearson correlation method for the *ACC* metric. Since we focus on seasonal mean anomalies the ACC metric represents simply *interannual correlation* and calculated for the ensemble mean.

b. *Brier SkillScore (BSS)*: It is based on the Brier Score (*BS*). The *BS* measures the mean squared error of probability forecasts for a binary event. Based on the *BS* and considering the climatology as the bench forecast, *BSS* is a relative measure of probabilistic skill (Wilks, 2011) and is defined as the following,

$$BSS = 1 - \frac{BS}{BS_{c\,lim.}}$$

where *BS_{clim.}* is the Brier score of a climatological forecast. The BSS value ranges from –inf to +1. For a perfect (worst) forecast BSS is equal to +1 (negative), while a zero BSS indicates no skill and equivalent to the climatology forecast. It should be noted that the classic *BSS* is sensitive to ensemble size and negatively oriented for a small ensemble size (Müller et al 2005; Weigel et al. 2007). In order to overcome this discrepancy, we used a strictly proper fair BSS (Ferro 2014). The Fair BSS is already available in the R package SpecsVerification[1]. We used the inbuilt function 'FairBSs' for the *BSS* calculation.

c. *Attributes or reliability diagram*: It is a graphical summary of important statistical-attributes such as reliability, resolution, uncertainty. It provides useful information about the performance of a prediction system. Reliability or so-called the attributes diagram measures how well the forecast probabilities of an event are in line with the equivalent observed frequencies. For instance, a forecast probability of 0.8 is called perfect reliable if and only if the event is true for all 80% of the actual cases. In the reliability diagram, the diagonal line generally reflects a perfect reliability line. Reliability of a forecast is high if it closely follows the diagonal line, inferring a good correspondence between the forecast probabilities and the observed frequencies for a binary event.

d. *ROC Skill Score (ROCSS):* This skill score is based on the area beneath the Relative Operating Characteristic (ROC) curve (Joliffe and Stephenson 2003). This curve measures forecast discrimination for a binary event. The skill score compares this area with that of a climatological forecast. Numerically, ROCSS ranges from 1 (perfect forecast system) to -1 (perfectly bad forecast system). Zero indicates no skill compared to a forecast based on the climatological frequency of the event.

In the present analysis, two binary rainfall events (above-normal and below-normal) were considered based on upper and lower terciles, respectively. Given the small sample size (20 years), all terciles and anomalies were calculated in a 'one-year out cross-validation' mode, i.e. leaving the target year out from the computation in order to avoid the overfitting (Wilks 2011; Weigel et al. 2008). In order to combine all RCMs for the multi-model ensemble (MME), we used a simple 'pool' approach. MME ensembles were produced by simply

---

[1] https://cran.r-project.org/web/packages/SpecsVerification/index.html

pooling all RCM ensembles together, considering each model has equal weight (Weigel et al 2008).

For the reliability diagram, we used ten equidistant bins to discretize the forecast probabilities. As suggested by Doblas-Reyes et al. (2008), the maximum number of bins can go up to the number of ensemble members plus one. In present model experiments, we have minimum of 3-members for the WRF341, WRF360, and RegCM4 models and thus we did not take into account these three individual models for the reliability analysis. Instead we added these models in MME. As a consequence we have a total of 39 members for MME and 15-members for the ESDs and the dynamical models (including GCMs and the two individual RCMs: RCA4 and CCLM4).

In a reliability diagram, the data points (i.e. forecast probabilities vs. observed frequencies) usually do not stay along a line and therefore following Weisheimer and Palmer (2014) we performed a weighted linear regression as best guess estimate on all data bins. Here the forecast size in each bin was considered as weights of the regression model. Since the probability classification was based on terciles, the weighted linear regression line always passes through the climatological intersection point, i.e. one-third, as can be illustrated from Figures 28 and 29. The reliability slope, hereafter, is assessed as a comprehensive measure of the reliability of the prediction system.

3.3   Observational uncertainties

We first evaluate the spread among the precipitation datasets over eastern Africa to get an estimate of consistency and uncertainty across the observations. Figure 4 shows GPCC precipitation for the June - September (JJAS) season and deviations from GPCC in the other observational datasets and ERAINT. Figure 4 indicates that both TARCAT and ARC have a large dry bias over the Ethiopian Highlands compared to GPCC, with relative differences reaching of as much as 50% locally (not shown). The other precipitation datasets (CRU, UDEL, WFDEI and CHIRPS) and GPCP show much better consistency, although some significant differences can still be seen on smaller spatial scales. The WFDEI precipitation, showing the smallest difference, is scaled by an older version of GPCC (v. 5) and is expected to be consistent with GPCC. The ERAINT precipitation has mostly a wet bias if the GPCC is used as reference.

Consistency among different observations in reproducing interannual variability is even more critical issue for verification of seasonal hindcasts than reproducing seasonal mean climatology. If observational datasets do not agree on wet/dry years over a region for example, the verification result can be very uncertain and depends on observational datasets chosen as reference. Figure 5 displays anomaly correlation coefficient (interannual correlation) for detrended time series (a linear trend is removed) between GPCC, taken as reference, and other precipitation products. One can see that there are large discrepancies in interannual precipitation variability across the observational datasets. A common feature is almost zero correlation in dry regions (eastern Ethiopia and Somalia) in all datasets. Even WFDEI, scaled to GPCC, do not correlate with GPCC in large part of Somalia. Small amount of precipitation in JJAS (only rare sporadic rains) and low station density (or simply no stations) lead to large observational uncertainties of interannual variability in precipitation over this region. Another distinct feature is absence of correlation over western Ethiopia and southern Sudan (climatologically wet region in JJAS) for TARCAT, ARC and in less degree in CRU, UDEL and CHIRPS. This region has enough precipitation in JJAS, especially in the

western Ethiopian Highlands and such difference may be explained by low station density. On average, the best agreement over large part of Ethiopia is found between GPCC and CHIRPS. The large uncertainties in interannual variability of precipitation among the observational datasets in eastern Africa can put serious limitations on our ability to verify seasonal forecasts at grid box scale over this region.

3.4 Global hindcast climatology and interannual variability

Here, we present a short overview about ability of the S4 and EC-EARTH models to simulate climatology and interannual variability globally. ERAINT has a good large-scale agreement with CRU for temperature, although local differences can reach a few degrees (Figure 6). Both S4 and EC-EARTH exhibit a similar bias pattern underestimating temperature in northern Africa and overestimation in northern America but biases are amplified in EC-EARTH. All three datasets - ERAINT, S4 and EC-EARTH represent too wet climatology over the ITCZ position in JJAS compared to GPCC with almost the same spatial pattern of biases (Figure 7). Specific humidity in the lower troposphere at 850 hPa is, on average, is underestimated in S4 over the tropics (Figure 8). EC-EARTH has some similarities (Africa and South America) but the negative bias is reduced or becomes slightly positive in many regions (the tropical Pacific, northern Africa and South Asia). A distinct feature of the EC-EARTH hindcast is strong overestimation of specific humidity in southern part of the Arab Peninsula and the surrounding seas. Both S4 and EC-EARTH show too strong zonal wind at 850 hPa west of the African continent - the Somali Jet forming in summer with a smaller bias in S4 (Figure 9). EC-EARTH strongly overestimates the Southern Hemisphere Jet while S4 simulates the jet very accurately if we take ERAINT as reference dataset. The MERRA reanalysis for example shows a positive difference of the same magnitude as the EC-EARTH hindcast.

Interannual variability of temperature in the extratropics is simulated pretty well by both S4 and EC-EARTH ensembles taking CRU as reference (Figure 10). Spatial pattern of ACC is similar between ERAINT and the global hindcasts, but the later show lower values of ACC. The lowest ACC, even close to zero, is found in the tropics (South and South-East Asia, eastern and central Africa, Central and South America) contrasting higher correlation in the extratropics. Compared to temperature ACC strongly drops for precipitation (Figure 11) and interannual variability of precipitation is partly reproduced in S4 and EC-EARTH only over some regions (e.g. the Sahel, South Asia, and South America). At the same time the spatial pattern of ACC is almost the same in both S4 and EC-EARTH.

3.5 Climatology and interannual variability over eastern Africa

As shown in Data and Methods section, two streams (full hindcast and its subset) were downscaled. One of the assumptions was that the first 3 members are enough to establish the hindcast climatology in order to estimate anomalies in the four years chosen. Figure 12 illustrates the precipitation climatology for all 15-members. S4 has mixed biases of both signs over the Ethiopian Highlands that somehow might be expected as observational uncertainties at grid box scale in regions with complex topography can be large. EC-EARTH in turn generates too wet climate over the Highlands and the wet bias stretching to the Arab Peninsula. Such wet bias may be related to too far north propagation of the ITCZ rain belt and to too strong Somali Jet in the EC-EARTH hindcast. The RCA4 and CCLM4 downscaling reduce the EC-EARTH wet biases, for example in southern Sudan but still precipitate too much over the Highlands, although showing mixed positive/negative biases.

All four hindcasts (S4, EC-EARTH, RCA4 and CCLM4) underestimate rainfall intensity in south-west of the domain (Uganda and part of DRC). Both ESD methods (15PC and 4NN), as expected, show the smallest biases, especially 4NN. Similar analysis but for the first 3 members and three more RCMs (WRF341, WRF360 and RegCM4) is shown in Figure 13. One can see that the bias pattern and magnitude is almost the same in the 1-3 and 1-15 member ensembles for the datasets with all 15 members. Therefore, in the present experiment, precipitation climatology based on the first three members is a representative estimate of climatology based on all 15 members. Among the RCMs downscaling only the three first members for the entire period, both WRF versions strongly overestimate rainfall over large part of the domain while RegCM4 shows biases similar to RCA4 and CCLM4.

In the S4 15 member hindcast the ACC pattern has two distinct regions with higher correlation: northern Ethiopia and North-East Sudan and southern Sudan - northern Uganda (Figure 14). These two spots are pretty well reproduced by EC-EARTH and by RCA4 and CCLM. Statistical downscaling loses a part of this signal (weaker ACC) but the spot of high positive ACC in southern Sudan - northern Uganda is a robust feature in all ESD hindcasts. As expected, taking only the first three members of the hindcasts makes the ACC patterns noisier due to smaller ensemble size (Figure 15). Two regions with higher ACC become weaker and are not evident in all hindcasts. The WRF360 hindcast shows negative or zero ACC over the entire Ethiopia while the RegCM4 hindcast has some signs of an improvement of its driving EC-EARTH. In general even if the ACC patterns look noisy there are two sub-regions where higher positive ACC is evident across the global and part of regionally downscaled hindcasts.

3.6   Relative operating characteristic Skill Score (ROCSS)

First, focus is on the probabilistic verification of the forecasts, computing the ROCSS of the models that downscaled the full 15 member stream: System4, EC-EARTH, RCA4, CCLM4 and two ESD methods. The period of analysis is the longest common available period for all models, 1991 to 2010 (limited by ESD). Figure 16 provides a quantitative measure of the forecast skill by representing ROCSS values for the upper and lower terciles. The highest ROCSS values are found over North-West Ethiopia.

To visualize the forecast skill for the two regions considered (the highlands and the southern part of the country), we have used a tercile-validation-plot (Figure 17). In it, the probabilistic forecast for each year, computed as the number of members falling within each category for that particular year, is represented. White circles correspond to the observed precipitation tercile. Moreover, the ROCSS of the spatial mean is also shown for each tercile and region.

However, regardless the observational dataset considered (see Figure 18), all models have generally a low skill. Compared to EC-EARTH, System4 seems to better capture the signal. Moreover, the overall skill, as measured by ROCSS, is better preserved by the dynamical downscaling than by the statistical one. The latter shows low sharpness in the forecasts, i.e. intermediate probabilities are usually assigned to all terciles. The most outstanding feature of the tercile plots is the 1997 El Niño year. All models correctly predicted it over the Ethiopian Highlands. If we focus on the four years, 2002, 2006, 2007 and 2009, chosen as reference of dry or wet years in this season, some findings are worth mentioning. EC-EARTH and the downscaling models nested into it seem to correctly predict the dry season of 2002. 2009 presents a clearer signal, especially for the RCMs. In 2007 June to September season was

wetter than normal in both regions, and all the models forecasted it properly. But for 2006, all the different systems seem to have missed the event.

Figures 19-22 show the forecast of the most likely tercile. It is possible to compare the skill in predicting the tercile for the different probabilistic forecasts, according to the GPCC observational dataset. In each map, the transparency of the circle is associated to the ROCSS (past skill forecasting this event), the size represents, for a given year, forecast probability for the most likely tercile, which is coded with colors (blue refers to above normal tercile, grey is associated to the normal tercile and red corresponds to below normal tercile). Black circles indicate that it was not possible to compute the ROCSS on that position due to a lack of observations or due to missing values in the ESD models. White circles represent constant values of the observation and dashes indicate negative values of ROCSS (forecast system not useful). System4, EC-EARTH and the RCMs present some spots of predictability in the north west of Ethiopia, especially RCA4. However, for the ESD models, the signal, still present, is lower. For the 2007 season (Figure 20), the signal is stronger over Kenya and South Sudan, correctly forecasting a wetter than normal season. In 2009, EC-EARTH and CCLM4 are the models that better captured the signal over the northern Ethiopia and Sudan.

For the last part of the analysis we have additionally considered 3 RCMs, WRF341, WRF360 and RegCM4 which have generated the subset of the full hindcast (see Table 1). In this stream, three members are available for all models to compute a background climatology. The terciles of this background climatology are depicted as boxes in Figure 22 and record maximum/minimum values in the climatology are depicted as horizontal lines. Therefore, the boxes represent the normal conditions. Seasonal anomalies from their respective climatologies are calculated for all the models and observations. Each circle in Figure 22 represents the seasonal anomaly of each member of the ensemble, scaled by their standard deviation (derived from the climatology). Left column corresponds to the seasonal anomalies forecasts of the Highlands whereas right column values belong to the southern part of Ethiopia. This figure disaggregates, for a given year, the members that entered the tercile plot in Figure 17. Obviously, the results are consistent the previous findings.

Additionally, the figure shows the observational uncertainty, which is quite apparent. Only 2007 shows a consistent wet anomaly across all observational datasets and regions. For other selected years, some of the observational databases show them as normal years or even opposite anomalous years (see e.g. the ARC2 above-normal anomaly in 2009 for the Highlands, or in 2001 over southern Ethiopia).

Considering the 2002 event, System4 and EC-EARTH correctly predict the large rainfall deficit over the south of Ethiopia, although this is not the case for the north of the country. In the 2006 event, the signal is not well captured by the models and the ESD methods and some RCMs worsened it, especially RCA4 in the north and WRF360 in the south. However, for the 2007 and 2009 events, almost all models seem to capture the signal correctly over the two regions.

3.7   Time series and anomaly correlation

Figure 24 shows rainfall anomaly time-series for the downscaled and the GCM models for the five clusters. One can see a great interannual variability in all models that falls mostly in the observations range. However, the observational spread is found to be substantially large in all clusters and generally comparable to the models uncertainty range. Extreme years with an exception of the year 1997 and the wet 2006 year are generally well captured by all

models. The multi-model ensemble (MME) shows an improved performance over the driving EC-EARTH model. Figure 25 shows averaged anomaly correlation coefficient with the WFDEI observation for the ensemble mean GCMs, RCMs, MME, and ESDs for the five clusters. In C1, all models show relatively low correlation. The dynamical models generally outperform the statistically models.

3.8 Brier skill score (BSS) and Reliability diagram

Figures 26 and 27 show fair Brier skill score for the two rainfall terciles over Ethiopia. The skill score indicates mixed performances in Ethiopia for both GCMs and the downscaled hindcasts. A relatively higher skill is found over the north-west part of Ethiopia, while all models show poor skill in south-east part of Ethiopia which is generally characterized by very little rainfall during JJAS.

From Figure 28 one can see that all dynamical models generally outperform the statistical models. Cluster C1 has negative score and thus indicates a poor forecast skill over the region. Overall the maximum skill is found in C2 that contains mostly the highland regions. MME has generally higher positive score than the driving EC-EARTH model. ESDs show a relatively smaller positive skill than the dynamical models in most of the cases.

Figures 29 and 30 present the attribute or so-called reliability diagrams for the upper and lower terciles for the five cluster regions. The results that fall in the BSS area are quite consistent with the fair BSS results of Figure 28. In the lower and upper tercile cases, all models show overconfidence, i.e., they are below the diagonal line, though one can observe that the dynamical models are closer to the diagonal line than the statistical models with some exceptions in C1 and C3. Interestingly, MME outplays the driving GCM model and the individual RCM.

3.9 Summary and Conclusions

A 5-month seasonal hindcast generated by the ECMWF System4 was rerun by a coupled global climate model, EC-EARTH in the atmospheric-only mode in order to provide boundary conditions for dynamical downscaling. The hindcast ensemble includes 15 members, initialised on May 1st and covers the period 1991-2012. Atmospheric initial conditions (ICs) on May 1st were generated at the ECMWF using a methodology similar to the one applied in S4 while temperature and soil moisture ICs are taken directly from S4. See surface temperature (SST) from S4 was bias corrected by replacing the S4 monthly mean SST climatology with the ERA-Interim reanalysis SST climatology but preserving anomalies. The EC-EARTH seasonal hindcast has been downscaled over eastern Africa by 5 RCMs from about 80km resolution to 25km. Two streams of downscaling, depending on resources available, have been defined i) full hindcast and ii) a subset consisting of the first three members for all years and of all 15 members for four preselected years (wet years - 2006/2007 and dry years - 2002/2009 in Ethiopia). The full hindcast was downscaled by CCLM4 and RCA4 and the subset by RegCM4, WRF341 and WRF360. In addition to the dynamical downscaling, two ESD methods were applied to both S4 and EC-EARTH hindcasts (all members and 1991-2010). The EUPORIAS downscaled seasonal hindcast ensemble is the largest ensemble ever downscaled over eastern Africa in a consistent and coordinated way.

EUPORIAS

Both global and downscaled seasonal hindcast ensembles were analysed and verified using a number of different observational datasets and a number of deterministic and probabilistic metrics. Here we summarise our main funding:

i) **Observational uncertainties**. There are large discrepancies in interannual precipitation variability across different observational datasets at regional scale (e.g. Somalia and eastern Ethiopia). In Somalia, small amount of precipitation in JJAS and low station density leads to high signal to noise ratio resulting in large observational uncertainties. In western Ethiopia, a wet region in JJAS, different number of gauge stations used for creation or calibration of gridded precipitation datasets can explain almost zero interannual correlation between some observations. The large uncertainties in interannual variability of precipitation among the observational datasets in eastern Africa can be a serious limitation for verifying seasonal forecasts at grid box scale.

ii) **Global systems**. Both S4 and EC-EARTH shows almost the same interannual correlation pattern in East Africa when deterministic verification metric as anomaly correlation is used. There are two distinct regions with higher correlation: northern Ethiopia - North-East Sudan and southern Sudan - northern Uganda while there is no skill elsewhere. Probabilistic metrics (ROCSS and BSS) applied only over Ethiopia also show that the global systems have some predictability skill in northern Ethiopia. At the same the signal in northern Ethiopia is sensitive to observational datasets chosen and most pronounced if GPCC or WFDEI-GPCC is used as reference. Compared to EC-EARTH, System4 seems to better capture the signal.

iii) **Dynamical downscaling.** When the full hindcast (15 members) is downscaled, RCMs (CCLM4 and RCA4) are able to capture the EC-EARTH signal pretty well. The anomaly correlation pattern shows the same two regions with high correlation as in the driving EC-EARTH hindcast. Probabilistic metrics reveals similar behaviour of the global hindcasts and RCMs in Ethiopia. All models can correctly predict dry (1997, 2002 and 2009) and wet (2007) years over the Ethiopian Highlands. However, all the different systems seem to have missed the wet 2006 summer. Including the RCMs downscaling the subset of the full hindcast, in general, supports the above findings, although some results can be noisier due to the only 3 member ensemble, anomaly correlation patterns for example.

iv) **Statistical Downscaling.** Statistical downscaling cannot capture the anomaly correlation pattern evident in the global hindcast and RCMs showing much weaker or no correlation in northern Ethiopia - North-East Sudan. At the same time the spot of high positive ACC in southern Sudan - northern Uganda is a robust feature in all ESD hindcasts. In Ethiopia, the overall skill, as measured by ROCSS and BSS, is better preserved by the dynamical downscaling than by the statistical one. The latter shows low sharpness in the forecasts, i.e. intermediate probabilities usually are assigned to all terciles.

We can conclude that the RCMs are able to capture and reproduce the signal evident in the driving EC-EARTH seasonal hindcast over northern Ethiopia in June-September showing about the same performance as their driving GCM. However, on average, the RCM hindcast show no added value compared to the driving GCM if we define the added value as a higher skill in the RCM hindcast. Statistical downscaling, in general, loses a part of the EC-EARTH

skill and shows a weaker performance in terms of predictability. We also should note that these conclusions are only for Ethiopia in the June-September season and cannot be generalised for other regions and seasons. Additionally, large observational uncertainties can potentially prevent us from accurate verification of the high-resolution RCM hindcast.

3.10   Reference

Arribas, A., Glover, M., Maidens, A., Peterson, K., Gordon, M., MacLachlan C., Graham R., Fereday D., Camp J., Scaife A. A., Xavier P., McLean P., Colman A. and Cusack S. 2011. The GloSea4 ensemble prediction system for seasonal forecasting. Mon. Weather Rev. 139: 1891–1910, doi: 10.1175/2010MWR3615.1.

Best M. J., Pryor M, Clark D. B., Rooney G. G., Essery R. L. H., M´nard C. B., Edwards J. M., E Hendry MA, Porson A, Gedney N, Mercado LM, Sitch S, Blyth E, Boucher O, Cox PM, Grimmond CSB, Harding RJ. 2011. The Joint UK Land Environment Simulator (JULES), model description–Part 1: Energy and water fluxes. Geosci. Model Dev. 4: 677–699, doi: 10.5194/gmd-4-677-2011.

Brown A, Milton S, Cullen M, Golding B, Mitchell J, Shelly A. 2012. Unified modeling and prediction of weather and climate: A 25-year journey. Bull. Am. Meteorol. Soc. 93: 1865–1877, doi: 10.1175/BAMS-D-12-00018.1.

Castro, C. L., H. Chang, F. Dominguez, C.Carrillo, J. K.Schemm, and H. M. H.Juang, 2012: Can a regional climate model improve the ability to forecast the North American monsoon? J. Climate, 25, 8212–8237, doi:10.1175/JCLI-D-11-00441.1

Cheneka,B. R.,Brienen,S.,Fröhlich,K., Asharaf,S. and B.Früh, 2016: Searching for an Added Value of Precipitation in Downscaled Seasonal Hindcasts over East Africa: COSMO-CLM Forced by MPI-ESM, Advances in Meteorology, v. 2016, Article ID 4348285, doi:10.1155/2016/4348285

Díez, E., B. Orfila, M. D. Frías, J. Fernández, A. S. Cofiño, and J. M. Gutiérrez. "Downscaling ECMWF Seasonal Precipitation Forecasts in Europe Using the RCA Model." Tellus A 63, no. 4 (2011): 757–62. doi:10.3402/tellusa.v63i4.15857.

Diro, G. T., A. M. Tompkins, and X. Bi, 2012: Dynamicaldownscaling of ECMWF ensemble seasonalforecasts over EastAfricawith RegCM3. J. Geophys. Res., 117, D16103, doi:10.1029/2011JD016997

Doblas-Reyes, F. J., Coelho, C. A. D. S., & Stephenson, D. B. (2008). How much does simplification of probability forecasts reduce forecast quality?. Meteorological Applications, 15(1), 155-162.

Ferro, C. A. T. (2014). Fair scores for ensemble forecasts. Quarterly Journal of the Royal Meteorological Society, 140(683), 1917-1923.

García-Díez, Markel, Jesús Fernández, and Robert Vautard. "An RCM Multi-Physics Ensemble over Europe: Multi-Variable Evaluation to Avoid Error Compensation." Climate Dynamics 45, no. 11-12 (2015): 3141-56. doi:10.1007/s00382-015-2529-x.

Giorgi, F., Coppola, E., Solmon, F., Mariotti, L. and others (2012). RegCM4: model description and preliminary tests over multiple CORDEX domains. Clim Res 52:7-29, doi:10.3354/cr01018

EUPORIAS

Graham, R. J., Yun W. T., Kim J., Kumar A. and others (2011) Long-range forecasting and the Global Framework for Climate Services. Clim. Res. 47:47-55. doi:10.3354/cr00963

Funk, C.C., Peterson, P., Landsfeld, M., Pedreros, D.H., Verdin, J.P., Shukla, S., Husak, G., Rowland, J.D., Harrison, L., Hoell, A., and Michaelsen, J.2015, Sci. Data., 2:150066. doi: 10.1038/sdata.2015.66.

Hansen, J. W., Mason, S. J., Sun, L. and Tall A. 2011: Review of seasonal climate forecasting for agriculture in sub-saharan Africa. Experimental Agriculture, 47, pp 205-240. doi:10.1017/s0014479710000876.

Harris, I., Jones, P.D., Osborn, T.J. and Lister, D.H. (2014), Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. Int. J. Climatol., 34: 623–642. doi: 10.1002/joc.3711

Hazeleger, W., and Coauthors, 2010: EC-earth: a seamless earth-system prediction approach in action. Bulletin of the American Meteorological Society, 91, 1357-1363

Hunke EC, Lipscomb WH. 2010. 'CICE: The sea ice model documentation and software user's manual, version 4.1', Technical report LA-CC-06-012. Los Alamos National Laboratory: Los Alamos, NM.

Jolliffe, I. T. and Stephenson, D. B. 2003. Forecast Verification: A Practitioner's Guide in Atmospheric Science, Wiley, NY

Katragkou, E., García-Díez, M., Vautard, R., Sobolowski, S., Zanis, P., Alexandri, G., Cardoso, R. M., Colette, A., Fernandez, J., Gobiet, A., Goergen, K., Karacostas, T., Knist, S., Mayer, S., Soares, P. M. M., Pytharoulis, I., Tegoulias, I., Tsikerdekis, A., and Jacob, D.: Regional climate hindcast simulations within EURO-CORDEX: evaluation of a WRF multi-physics ensemble, Geosci. Model Dev., 8, 603-618, doi:10.5194/gmd-8-603-2015, 2015.

Legates, D.R. and C.J. Willmott (1990). Mean seasonal and spatial variability in global surface air temperature. Theoretical and Applied Climatology, 41, 11-21.

MacLachlan, C., Arribas, A., Peterson, K. A., Maidens, A., Fereday, D., Scaife, A. A., Gordon, M., Vellinga, M., Williams, A., Comer, R. E., Camp, J., Xavier, P. and Madec, G. (2015), Global Seasonal forecast system version 5 (GloSea5): a high-resolution seasonal forecast system. Q.J.R. Meteorol. Soc., 141: 1072–1084. doi: 10.1002/qj.2396

Madec G. 2008. NEMO Ocean Engine, Note du Pole de Modelisation.Institut Pierre-Simon Laplace (IPSL): Paris.

Maidment, R., D. Grimes, R.P.Allan, E. Tarnavsky, M. Stringer, T. Hewison, R. Roebeling and E. Black (2014) The 30 year TAMSAT African Rainfall Climatology And Time series (TARCAT) data set Journal of Geophysical Research DOI: 10.1002/2014JD021927

Megann, A., Storkey, D., Aksenov, Y., Alderson, S., Calvert, D., Graham, T., Hyder, P., Siddorn, J., and Sinha, B.: GO5.0: the joint NERC-Met Office NEMO global ocean model for use in coupled and forced applications, Geosci. Model Dev., 7, 1069–1092, doi: 10.5194/gmd-7-1069-2014, 2014.

Molteni F., Stockdale T., Balmaseda M., Balsamo G., Buizza R., Ferranti L., Magnusson L., Mogensen K., Palmer T. N., Vitart F. The new ECMWF seasonal forecast system (System 4). ECMWF Technical Memorandum; 2011, 656. Available at: http://www.ecmwf.int/publications/library/do/references/show?id=90277

Müller, W. A., Appenzeller, C., Doblas-Reyes, F. J., &Liniger, M. A. (2005). A debiased ranked probability skill score to evaluate probabilistic ensemble forecasts with small ensemble sizes. Journal of Climate, 18(10), 1513-1523.

Nicholas S. Novella, Wassila M. Thiaw, 2012: African Rainfall Climatology Version 2 for Famine Early Warning Systems. Journal of Applied Meteorology and Climatology, 52, 588–606, doi: 10.1175/JAMC-D-11-0238.1.

Ogallo, L., Bessemoulin, P., Ceron, J-P., Mason, S. and Connor, S.J. 2008: Adapting to climate variability and change: the Climate Outlook Forum process. WMO Bulletin, Vol. 57, No.2, pp.93-102.

Rae, J. G. L., Hewitt, H. T., Keen, A. B., Ridley, J. K., West, A. E., Harris, C. M., Hunke, E. C., and Walters, D. N.: Development of the Global Sea Ice 6.0 CICE configuration for the Met Office Global Coupled model, Geosci. Model Dev., 8, 2221-2230, doi:10.5194/gmd-8-2221-2015, 2015.

Rayner, N. A.; Parker, D. E.; Horton, E. B.; Folland, C. K.; Alexander, L. V.; Rowell, D. P.; Kent, E. C.; Kaplan, A. (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century J. Geophys. Res. v. 108, No. D14, 4407 10.1029/2002JD002670

Rienecker, M.M., M.J. Suarez, R. Gelaro, R. Todling, J. Bacmeister, E. Liu, M.G. Bosilovich, S.D. Schubert, L. Takacs, G.-K. Kim, S. Bloom, J. Chen, D. Collins, A. Conaty, A. da Silva, et al. (2011), MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. J. Climate, 24, 3624-3648, doi:10.1175/JCLI-D-11-00015.1

Rockel, B., Will, A., and A. Hense, 2008, Regional climate modeling with COSMO-CLM (CCLM), Meteorologische Zeitschrift, 17 (4), 347-348, doi: 10.1127/0941-2948/2008/0309

Schneider, Udo; Becker, Andreas; Finger, Peter; Meyer-Christoffer, Anja; Rudolf, Bruno; Ziese, Markus (2015): GPCC Full Data Reanalysis Version 7.0 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data. DOI: 10.5676/DWD_GPCC/FD_M_V7_050

Skamarock, William C., Joseph B. Klemp, Jimy Dudhia, David O. Gill, Dale M. Barker, Wei Wang, and Jordan G. Powers. "A Description of the Advanced Research WRF Version 3." NCAR/TN-475+STR, 2008.

Strandberg, G., Bärring, L., Hansson, U., Jansson, C., Jones, C., Kjellström, E., Kolax, M., Kupiainen, M., Nikulin, G., Samuelsson, P., Ullerstig, A., and S. Wang (2015), CORDEX scenarios for Europe from the Rossby Centre regional climate model RCA4, Report Meteorology and Hydrology (RMK) No. 116, 2014, http://www.smhi.se/publikationer/publikationer/cordex-scenarios-for-europe-from-the-rossby-centre-regional-climate-model-rca4-1.90272

Walters, D. N., Brooks, M. E., Boutle, I. A., Melvin, T. R. O., Stratton, R. A., Wells, H., Williams, K. D., Wood, N., Bushell, A. C., Copsey, D., Earnshaw, P. E., Gross, M. S., Hardiman, S. C., Harris, C. M., Heming, J. T., Klingaman, N. P., Levine, R. C., Manners, J., Martin, G. M., Milton, S. F., Mittermaier, M. P., Morcrette, C. J., Riddick, T. C., Roberts, M. J., Sanchez, C., Selwood, P. M., Stirling, A.J., Smith, C., Tennant, W. J., Vosper, S. B., Vidale, P. L., Wilkinson, J. M., Wood, N., Woolnough, S. J., and Xavier, P. K.: The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations, Geosci. Model Dev. Discuss., in preparation, 2016.

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

Weigel, A. P., Liniger, M. A., &Appenzeller, C. (2008). Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts?. Quarterly Journal of the Royal Meteorological Society, 134(630), 241-260.

Weigel, A. P., Liniger, M. A., &Appenzeller, C. (2007). The discrete Brier and ranked probability skill scores. Monthly Weather Review, 135(1), 118-124.

Weisheimer, A., & Palmer, T. N. (2014).On the reliability of seasonal climate forecasts. Journal of The Royal Society Interface, 11(96), 20131162.

Wilks, D. S. (2011).Statistical methods in the atmospheric sciences (Vol. 100). Academic press.

Williams, K. D., Harris, C. M., Bodas-Salcedo, A., Camp, J., Comer, R. E., Copsey, D., Fereday, D., Graham, T., Hill, R., Hinton, T., Hyder, P., Ineson, S., Masato, G., Milton, S. F., Roberts, M. J., Rowell, D. P., Sanchez, C., Shelly, A., Sinha, B., Walters, D. N., West, A., Woollings, T., and Xavier, P. K.: The Met Office Global Coupled model 2.0 (GC2) configuration, Geosci. Model Dev., 8, 1509-1524, doi:10.5194/gmd-8-1509-2015, 2015.

4 Lessons Learnt

The main lesson learnt is that dynamical and statistical downscaling of a global hindcast over Ethiopia in June-September cannot improve the global hindcast in terms of predictability skill. At the same time dynamical downscaling reproduces the skill from its driving global system while statistical downscaling loses a part of it. Our recommendation for developing similar experiments is to focus on regions and seasons where global seasonal prediction systems have a higher skill.
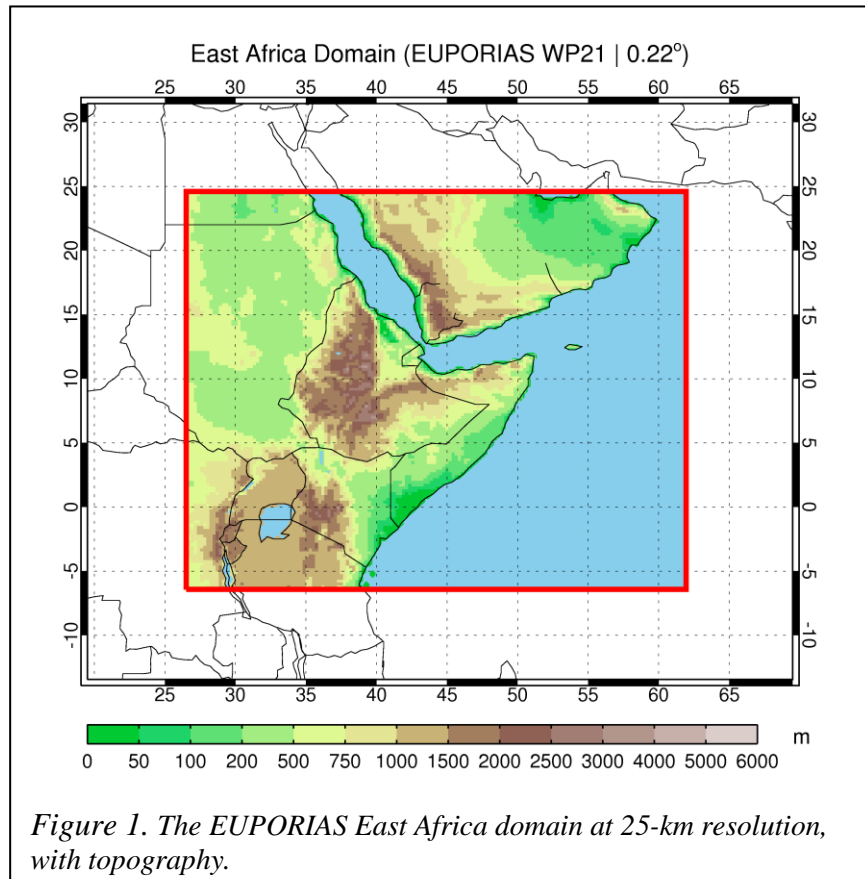
5 Links Built

The EC-EARTH seasonal hindcast from this deliverable was provided to the SPECS project and was used for dynamical downscaling over Europe in summer.
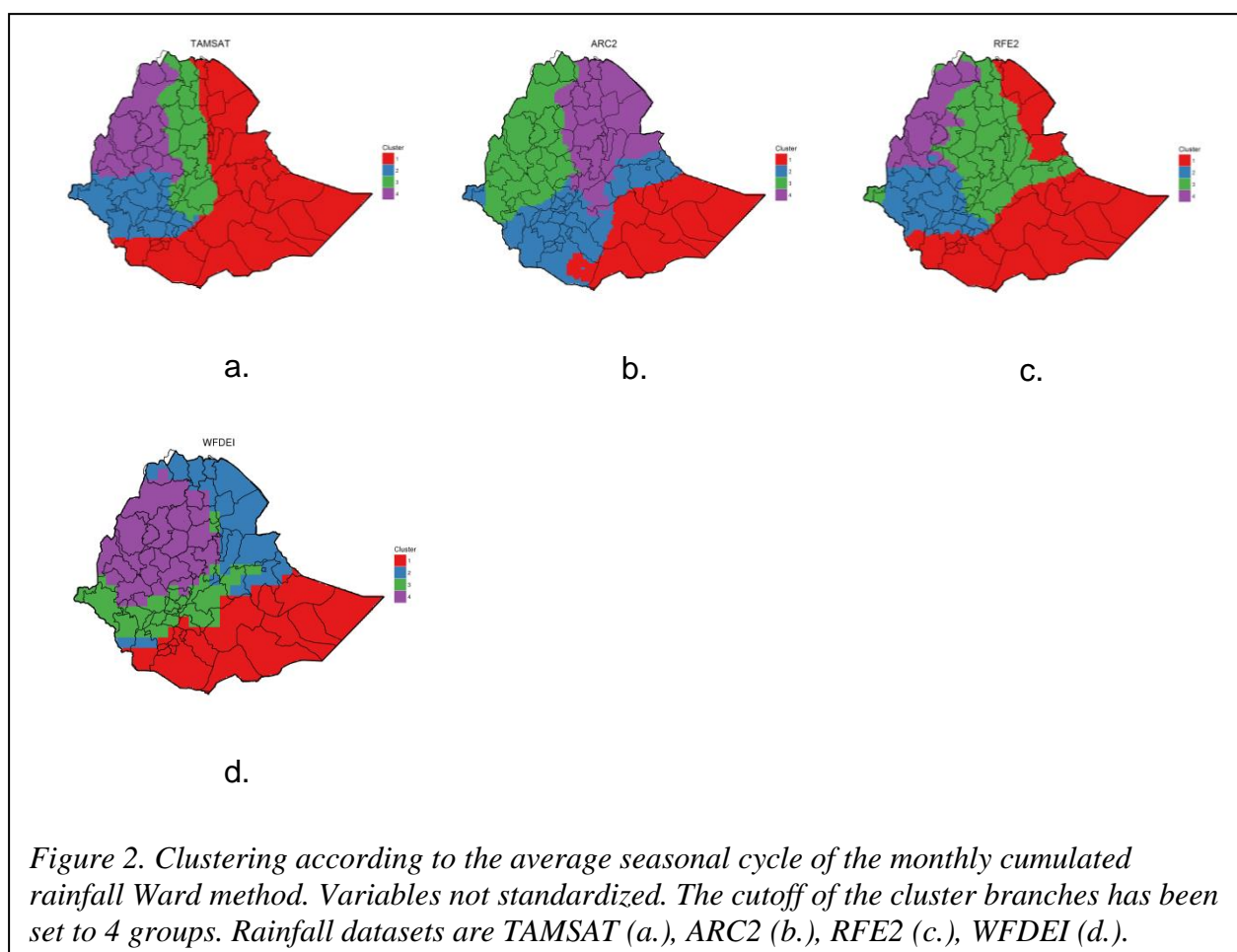
Dynamical and statistical downscaling results are used in WP2 for calculating climate information indices.

The downscaling tools provided by the downscaleR package were developed in collaboration with the SPECS project.

The downscaling framework used by UC (predictors and cross-validation framework) agree with that developed in the EU COST Action VALUE.
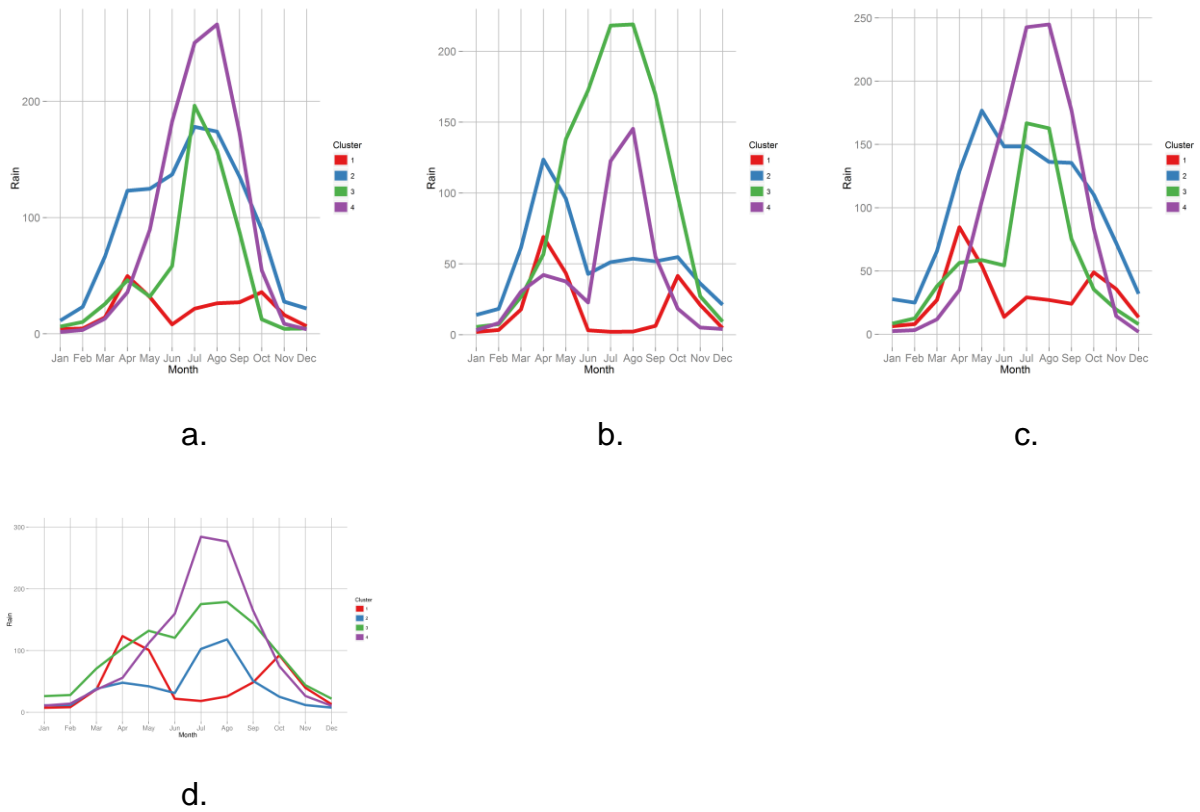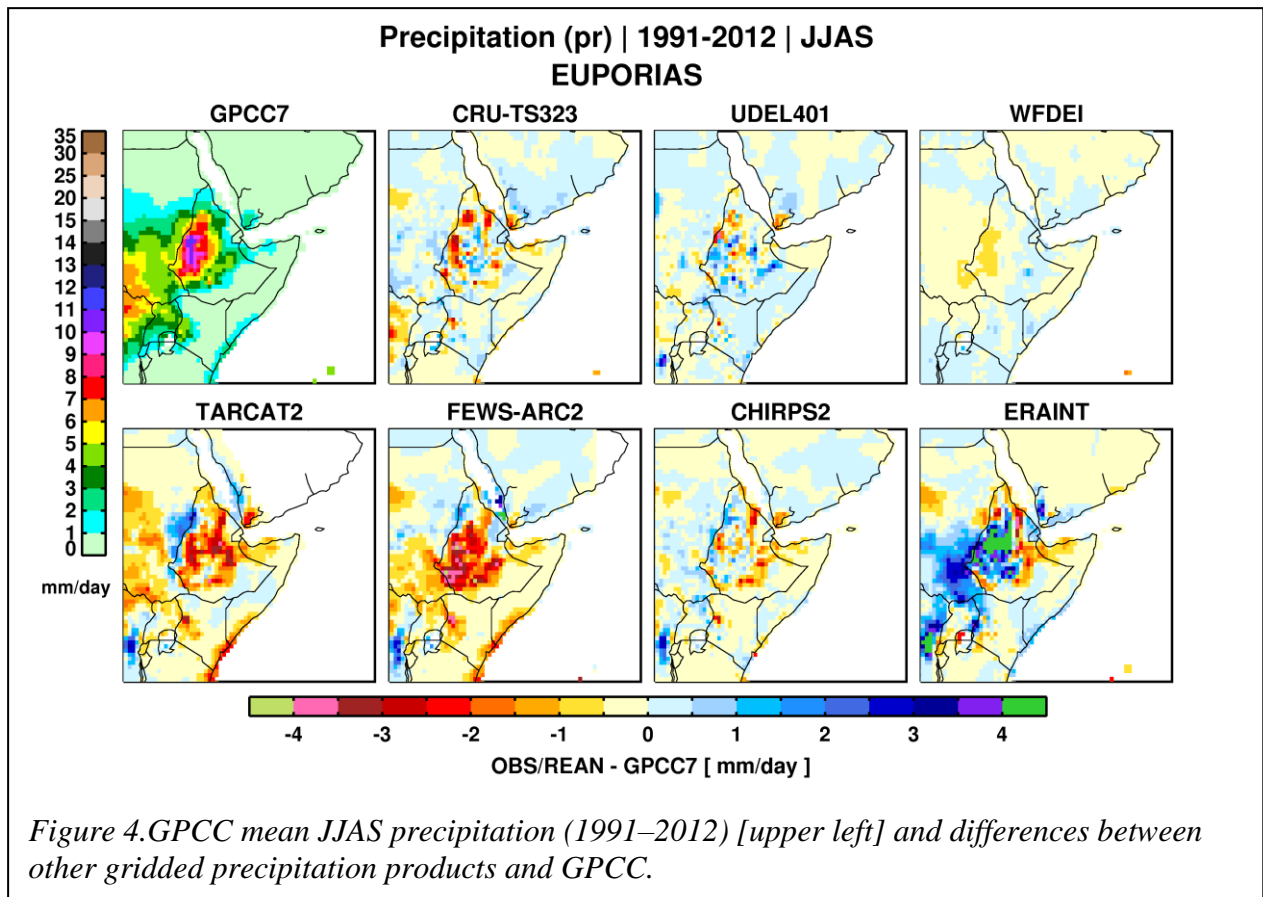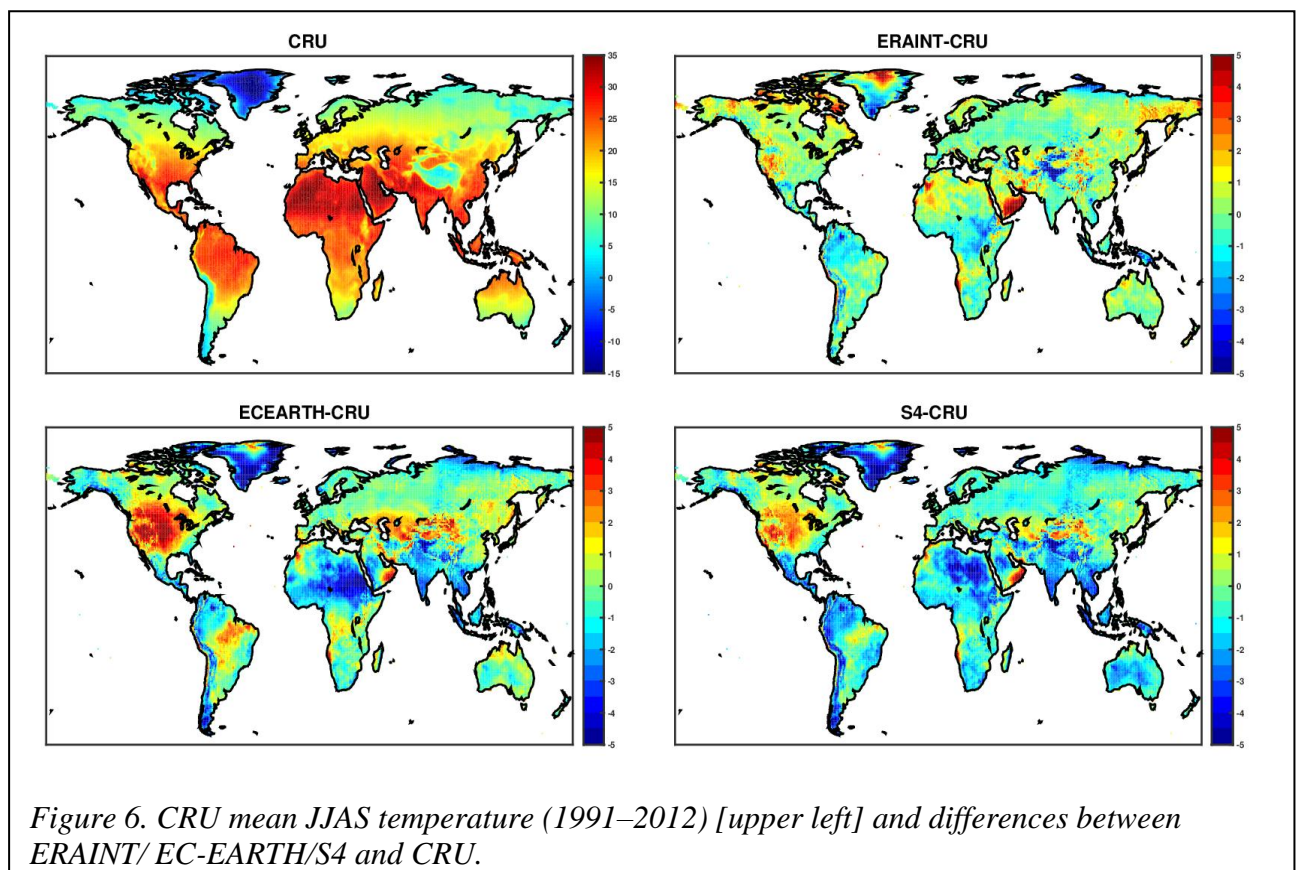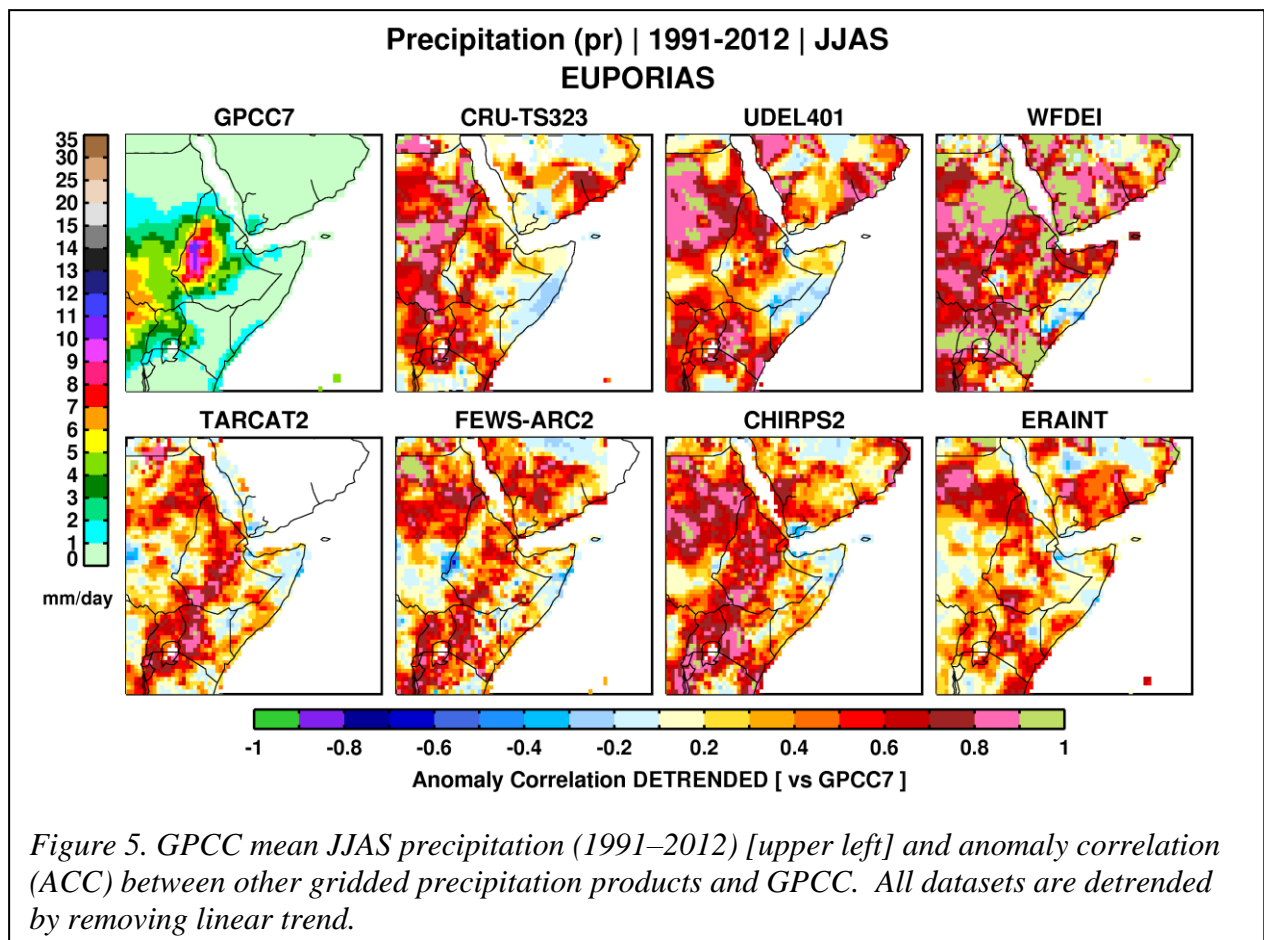
6 Figures

Figure 1. The EUPORIAS East Africa domain at 25-km resolution, with topography.

EUPORIAS

Figure 2. Clustering according to the average seasonal cycle of the monthly cumulated rainfall Ward method. Variables not standardized. The cutoff of the cluster branches has been set to 4 groups. Rainfall datasets are TAMSAT (a.), ARC2 (b.), RFE2 (c.), WFDEI (d.).

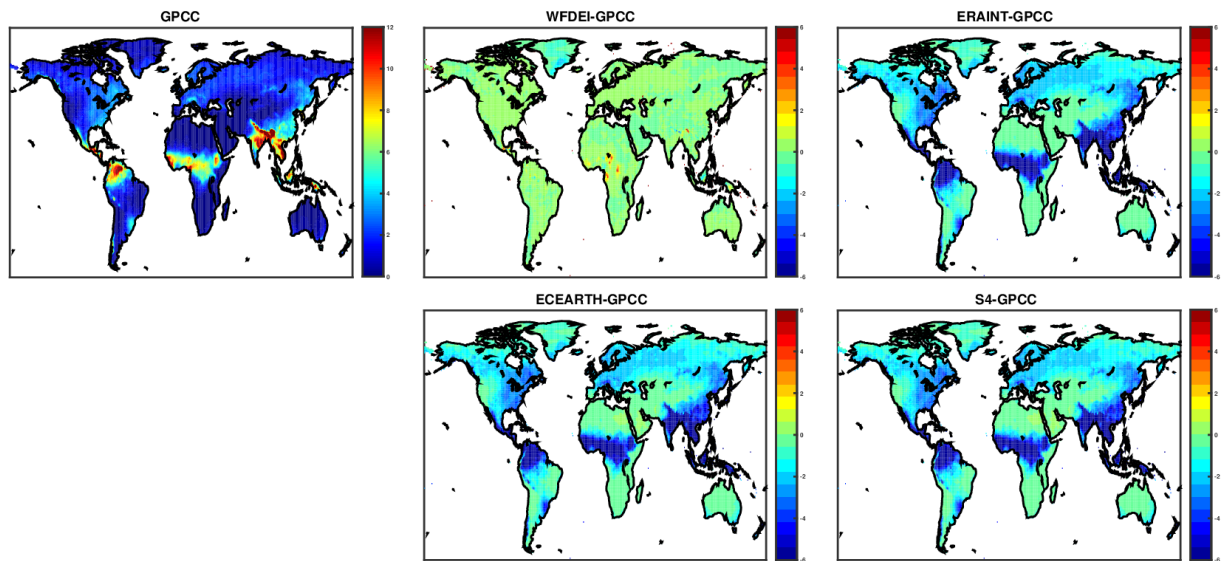*Figure 3. Average seasonal cycle of the monthly cumulated rainfall. The cutoff of the cluster branches has been set to 4 groups. Rainfall datasets are TAMSAT (a.), ARC2 (b.), RFE2 (c.), WFDEI (d.).*

EUPORIAS

Figure 4.GPCC mean JJAS precipitation (1991–2012) [upper left] and differences between other gridded precipitation products and GPCC.

EUPORIAS

Figure 5. GPCC mean JJAS precipitation (1991–2012) [upper left] and anomaly correlation (ACC) between other gridded precipitation products and GPCC. All datasets are detrended by removing linear trend.

Figure 6. CRU mean JJAS temperature (1991–2012) [upper left] and differences between ERAINT/ EC-EARTH/S4 and CRU.

Figure 7. GPCC mean JJAS precipitation (1991–2012) [upper left] and differences between WFDEI/ERA-Interim/ EC-EARTH/ and GPCC.
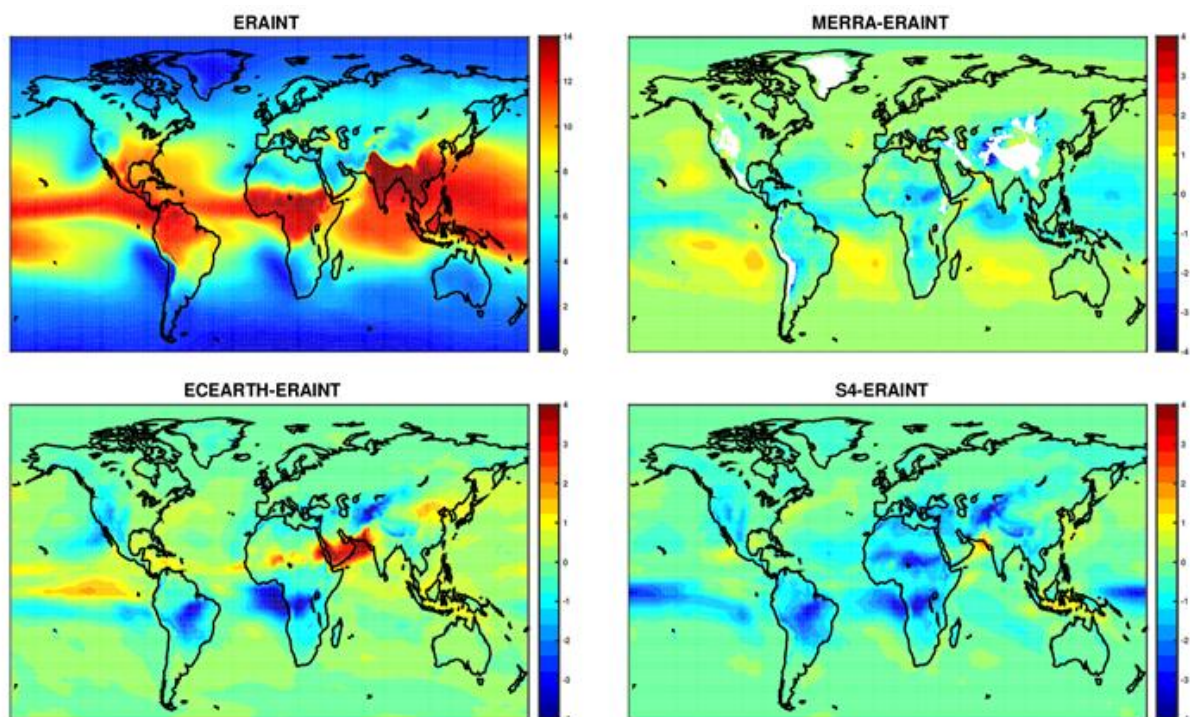
Figure 8. ERAINT mean JJAS specific humidity at 850 mb (1991–2012) [upper left] and differences between MERRA/EC-EARTH/S4 and ERAINT
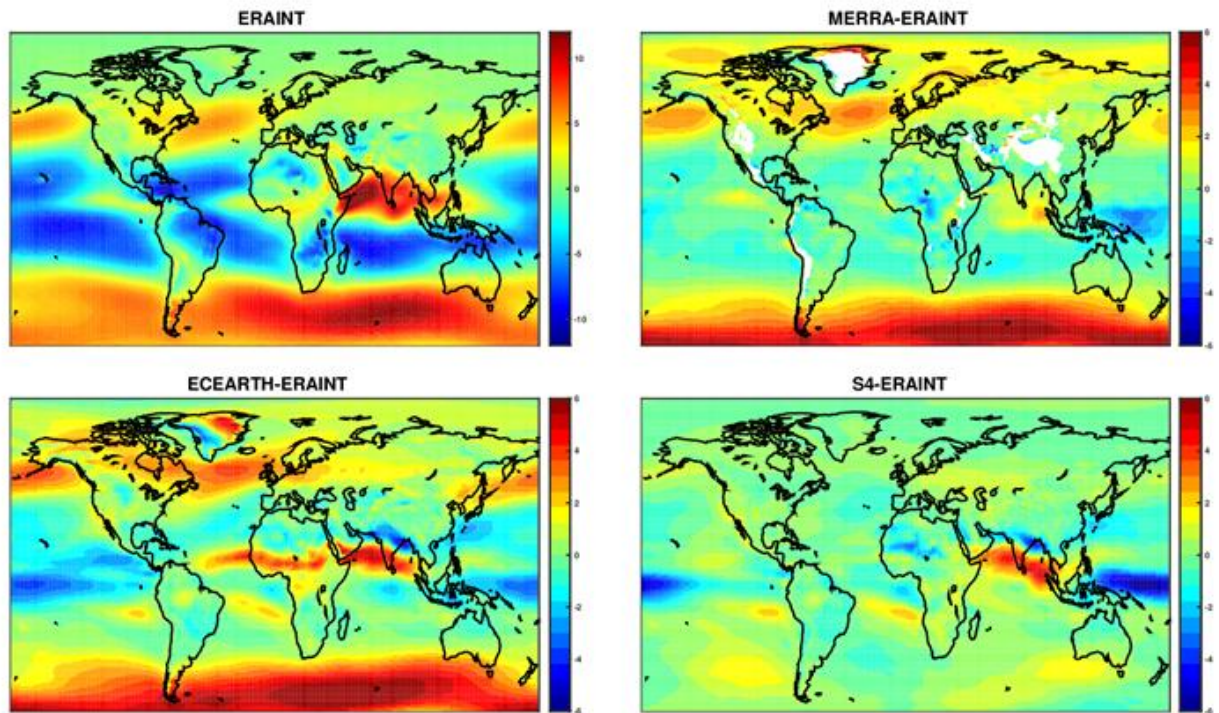
Figure 9. ERAINT mean JJAS zonal wind at 850 mb (1991–2012) [upper left] and differences between MERRA/EC-EARTH/S4 and ERAINT.
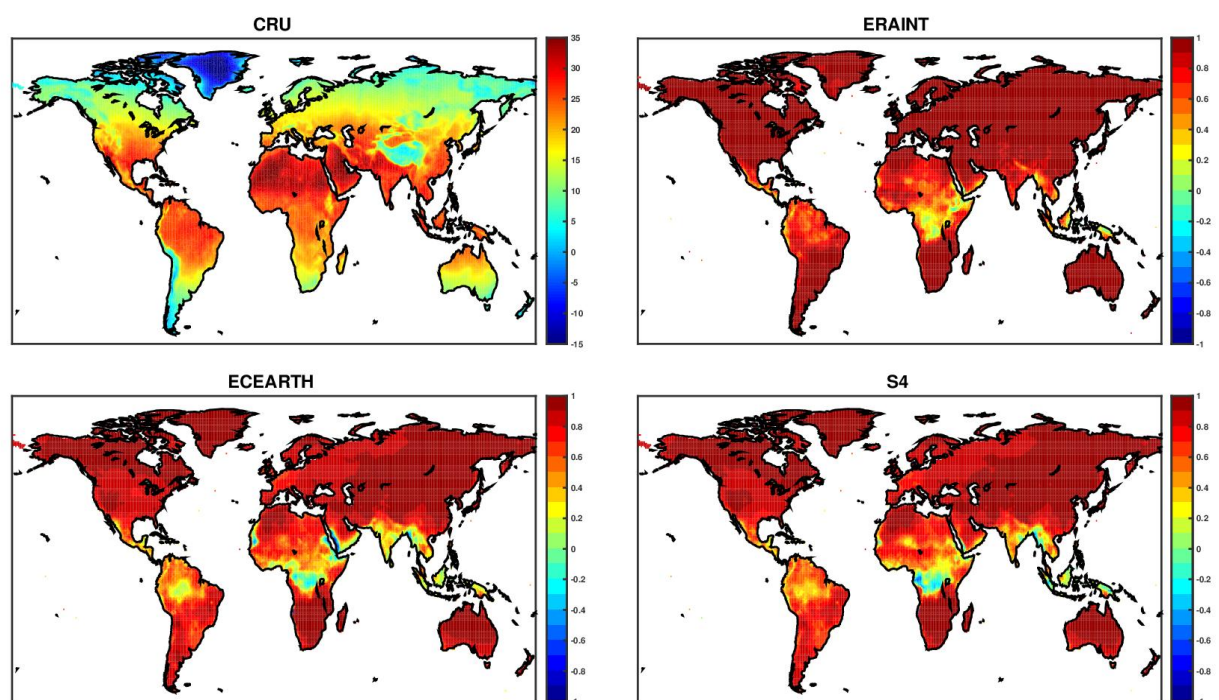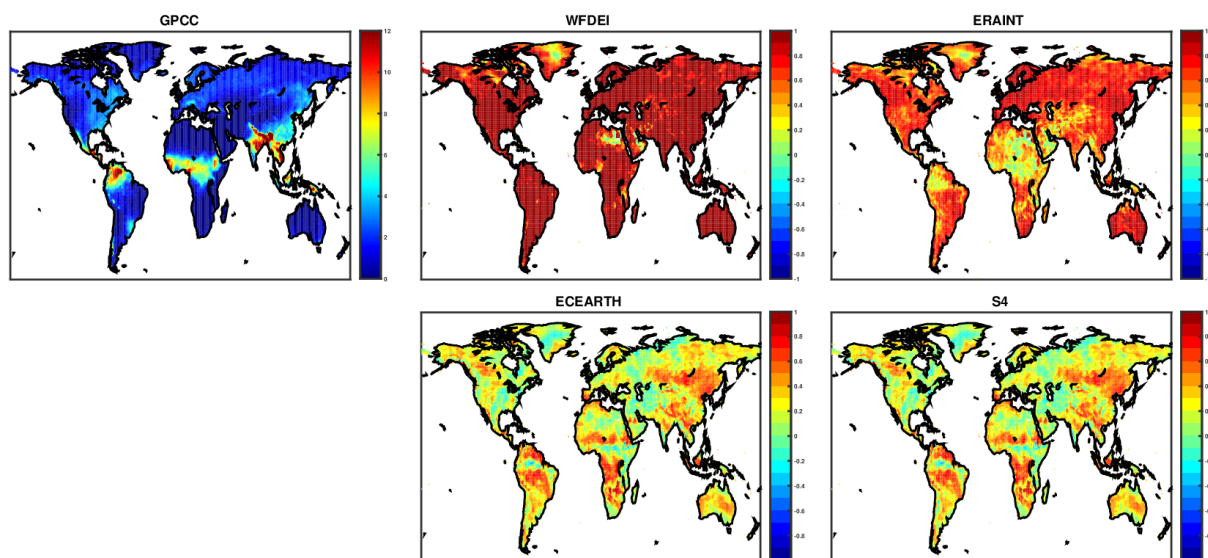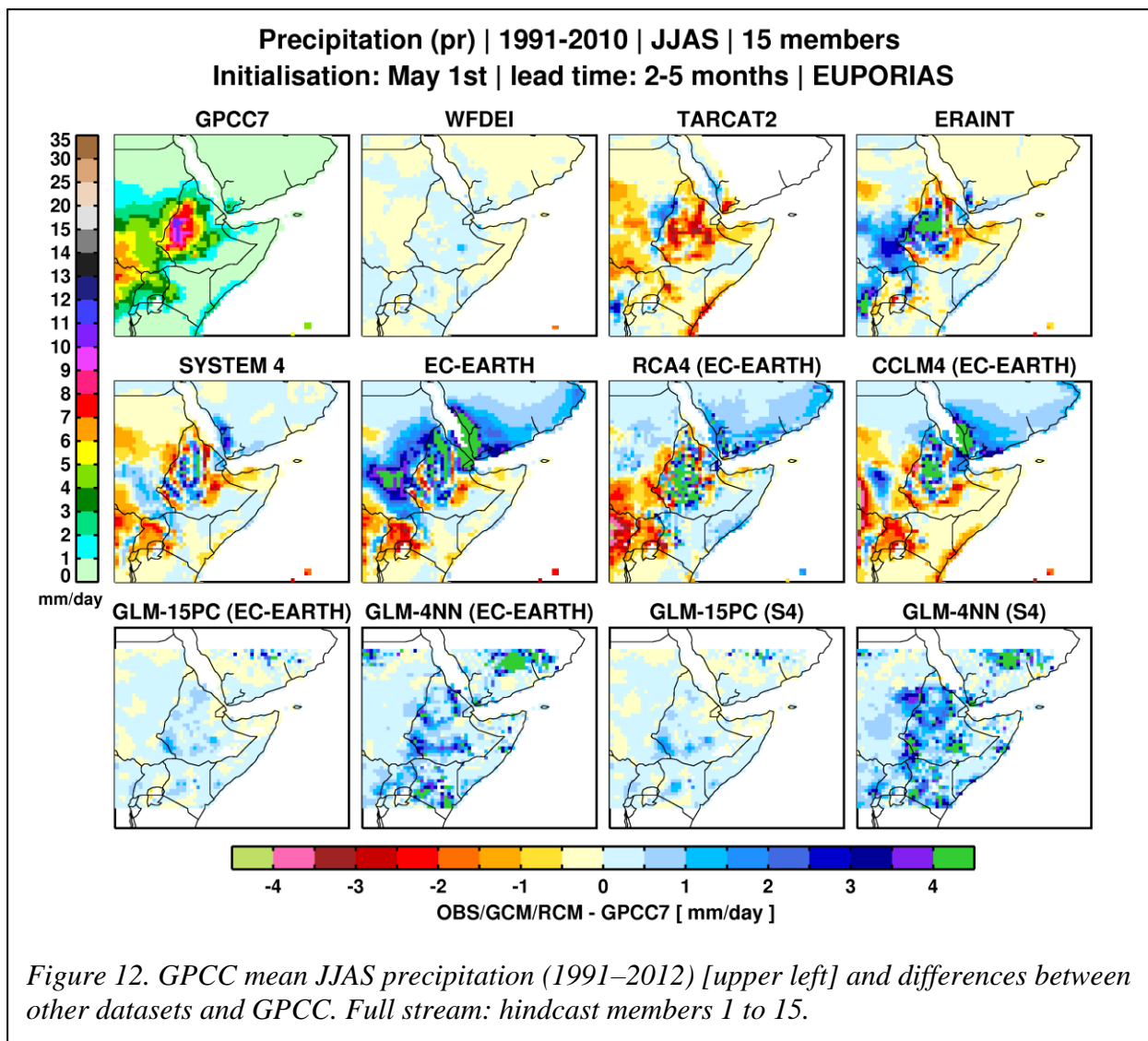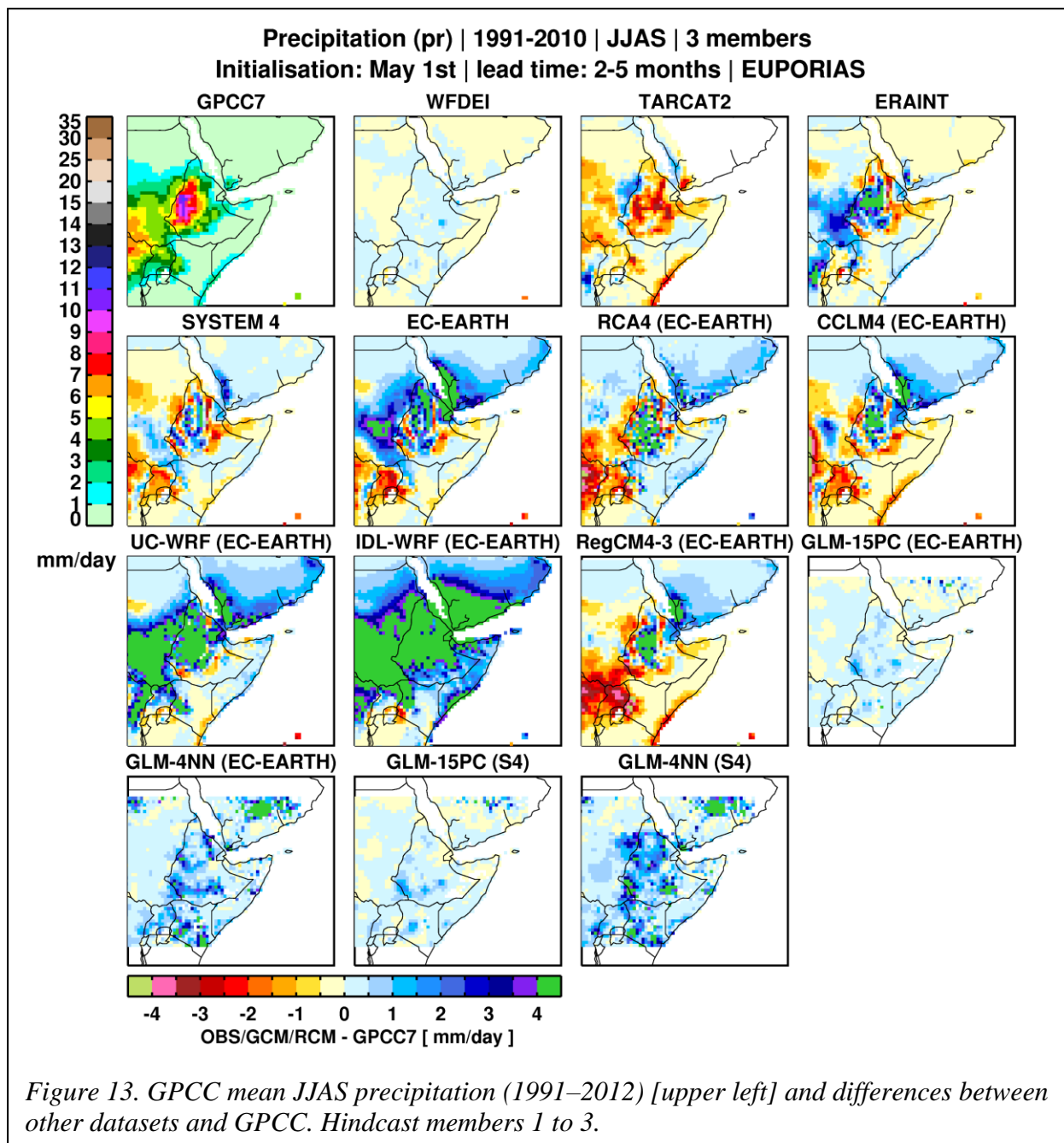

Figure 10. CRU mean JJAS temperature (1991–2012) [upper left] and anomaly correlation (ACC) between ERAINT/ EC-EARTH/S4 and CRU. All datasets are detrended by removing linear trend

EUPORIAS

Figure 11. GPCC mean JJAS precipitation (1991–2012) [upper left] and anomaly correlation (ACC) between ERAINT/ EC-EARTH/S4 and GPCC. All datasets are detrended by removing linear trend.

Figure 12. GPCC mean JJAS precipitation (1991–2012) [upper left] and differences between other datasets and GPCC. Full stream: hindcast members 1 to 15.

Figure 13. GPCC mean JJAS precipitation (1991–2012) [upper left] and differences between other datasets and GPCC. Hindcast members 1 to 3.
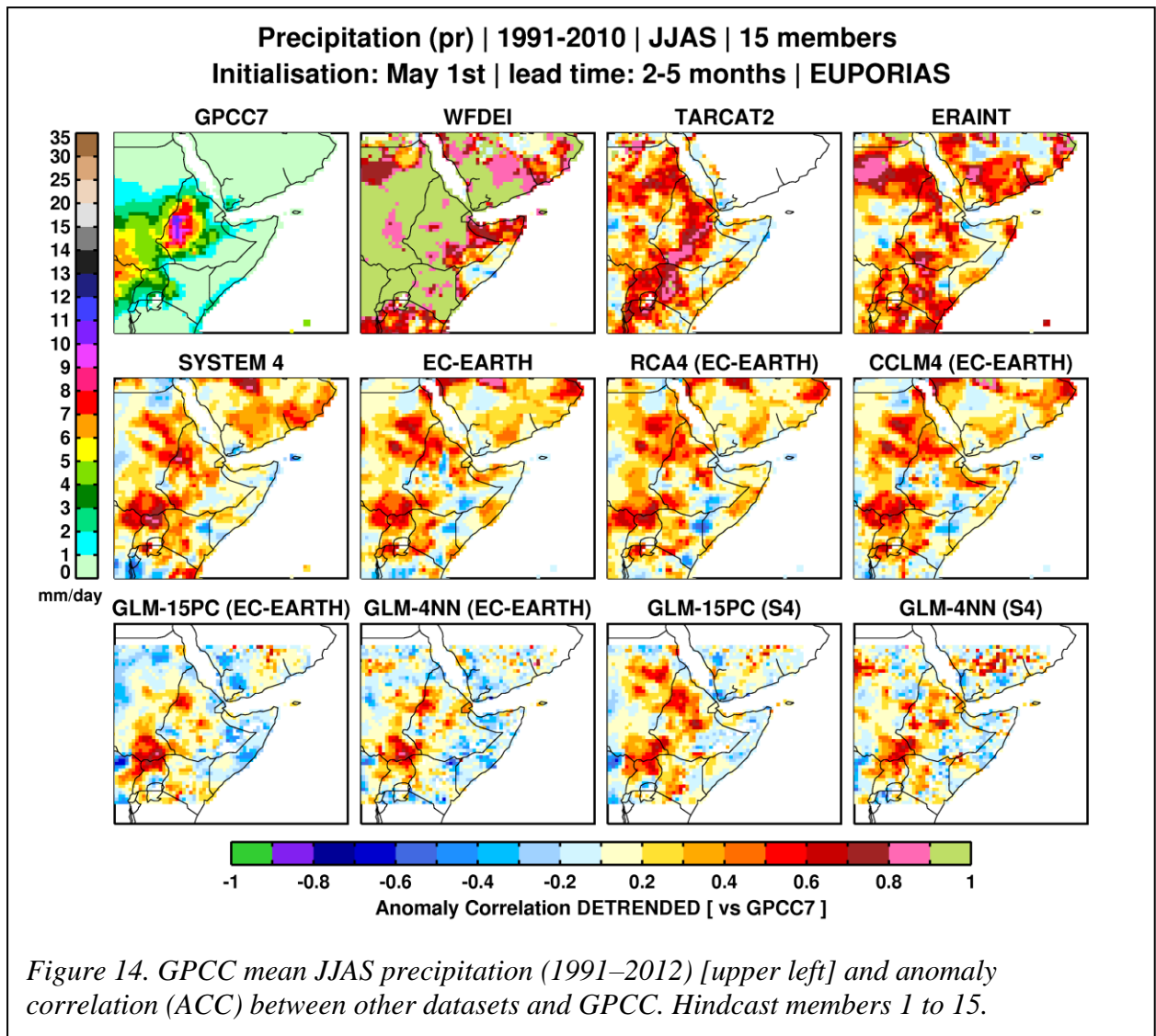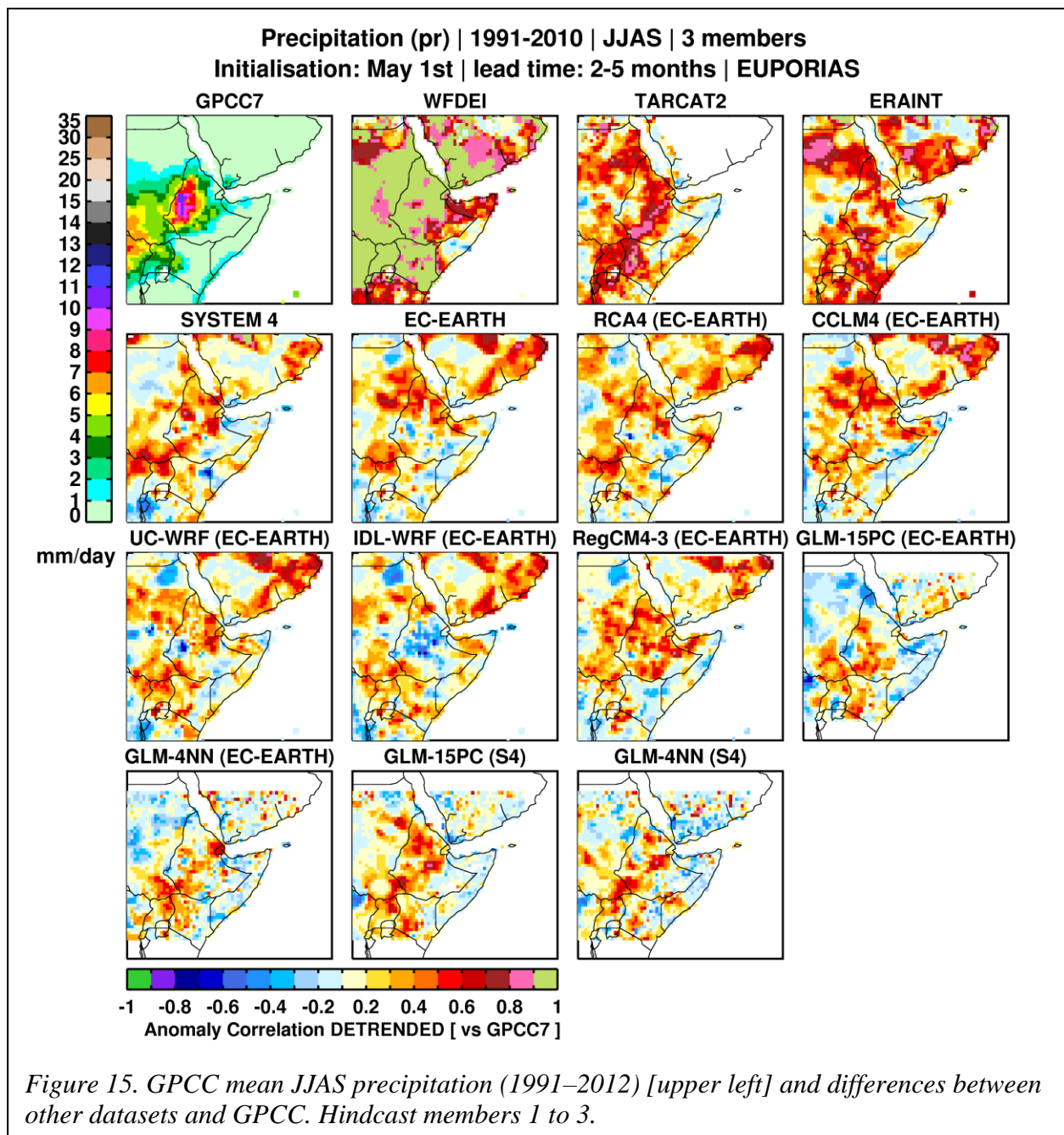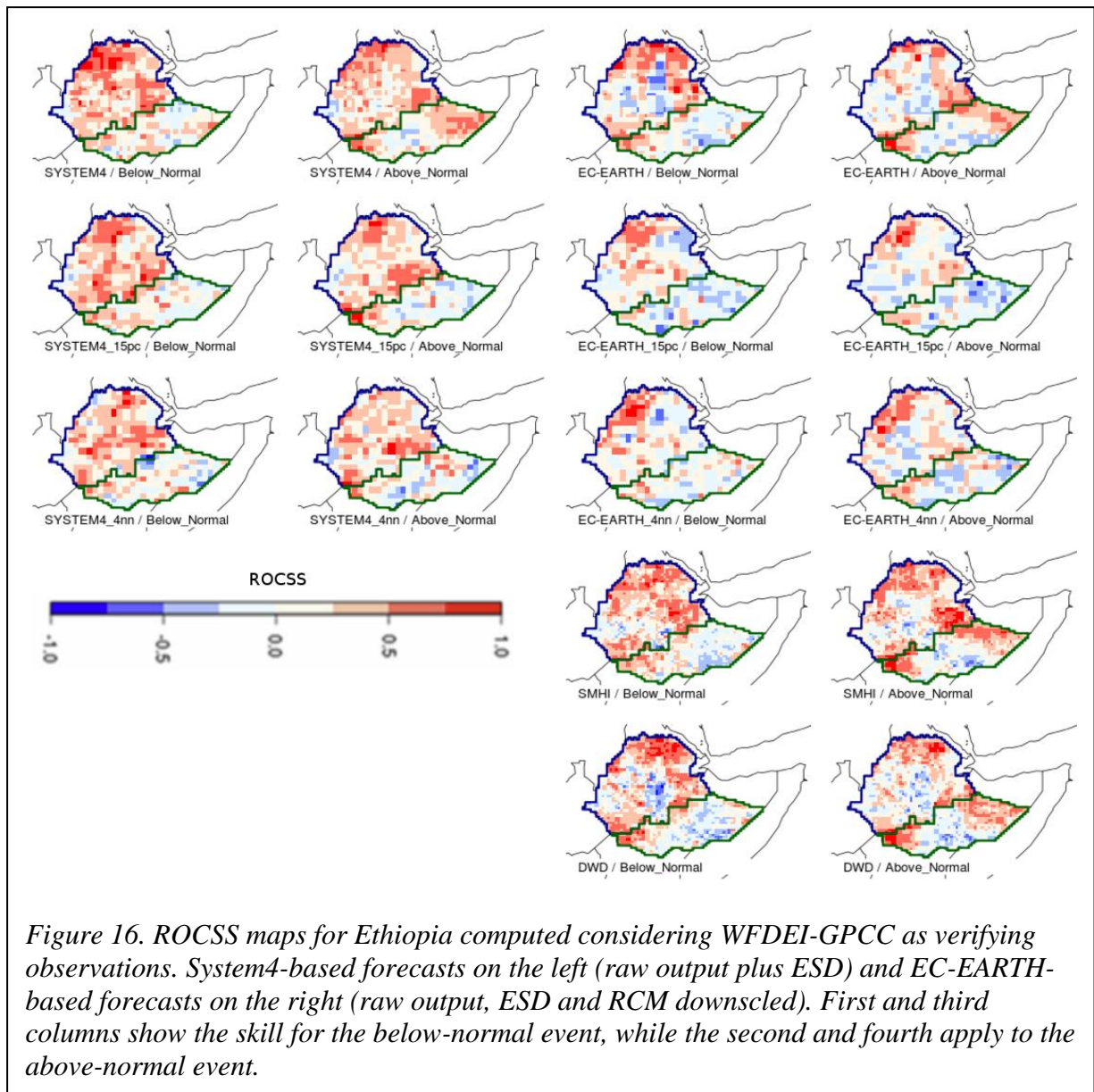
EUPORIAS

Figure 14. GPCC mean JJAS precipitation (1991–2012) [upper left] and anomaly correlation (ACC) between other datasets and GPCC. Hindcast members 1 to 15.
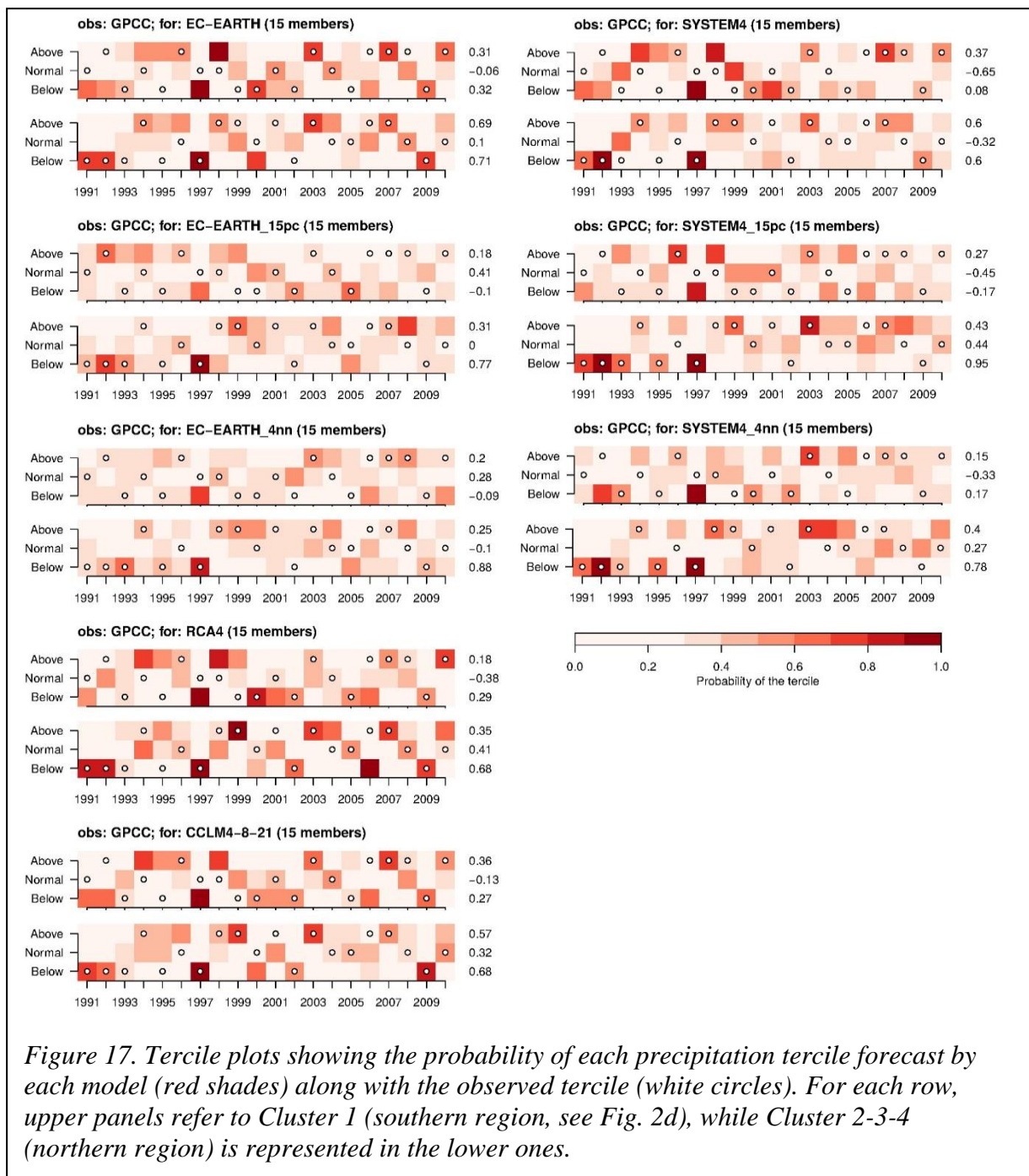
EUPORIAS

Figure 15. GPCC mean JJAS precipitation (1991–2012) [upper left] and differences between other datasets and GPCC. Hindcast members 1 to 3.

Figure 16. ROCSS maps for Ethiopia computed considering WFDEI-GPCC as verifying observations. System4-based forecasts on the left (raw output plus ESD) and EC-EARTH-based forecasts on the right (raw output, ESD and RCM downscled). First and third columns show the skill for the below-normal event, while the second and fourth apply to the above-normal event.
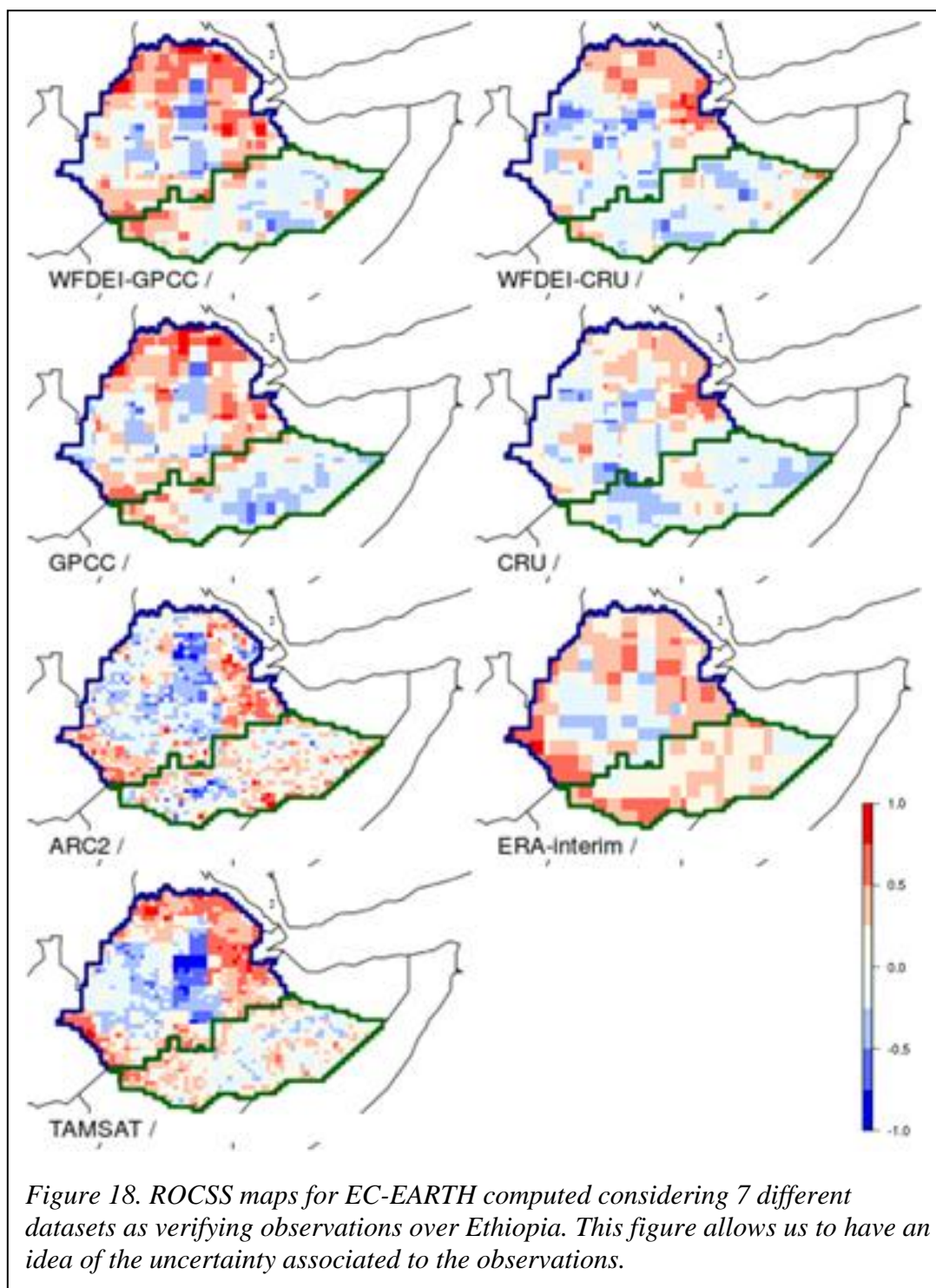
Figure 17. Tercile plots showing the probability of each precipitation tercile forecast by each model (red shades) along with the observed tercile (white circles). For each row, upper panels refer to Cluster 1 (southern region, see Fig. 2d), while Cluster 2-3-4 (northern region) is represented in the lower ones.

Figure 18. ROCSS maps for EC-EARTH computed considering 7 different datasets as verifying observations over Ethiopia. This figure allows us to have an idea of the uncertainty associated to the observations.
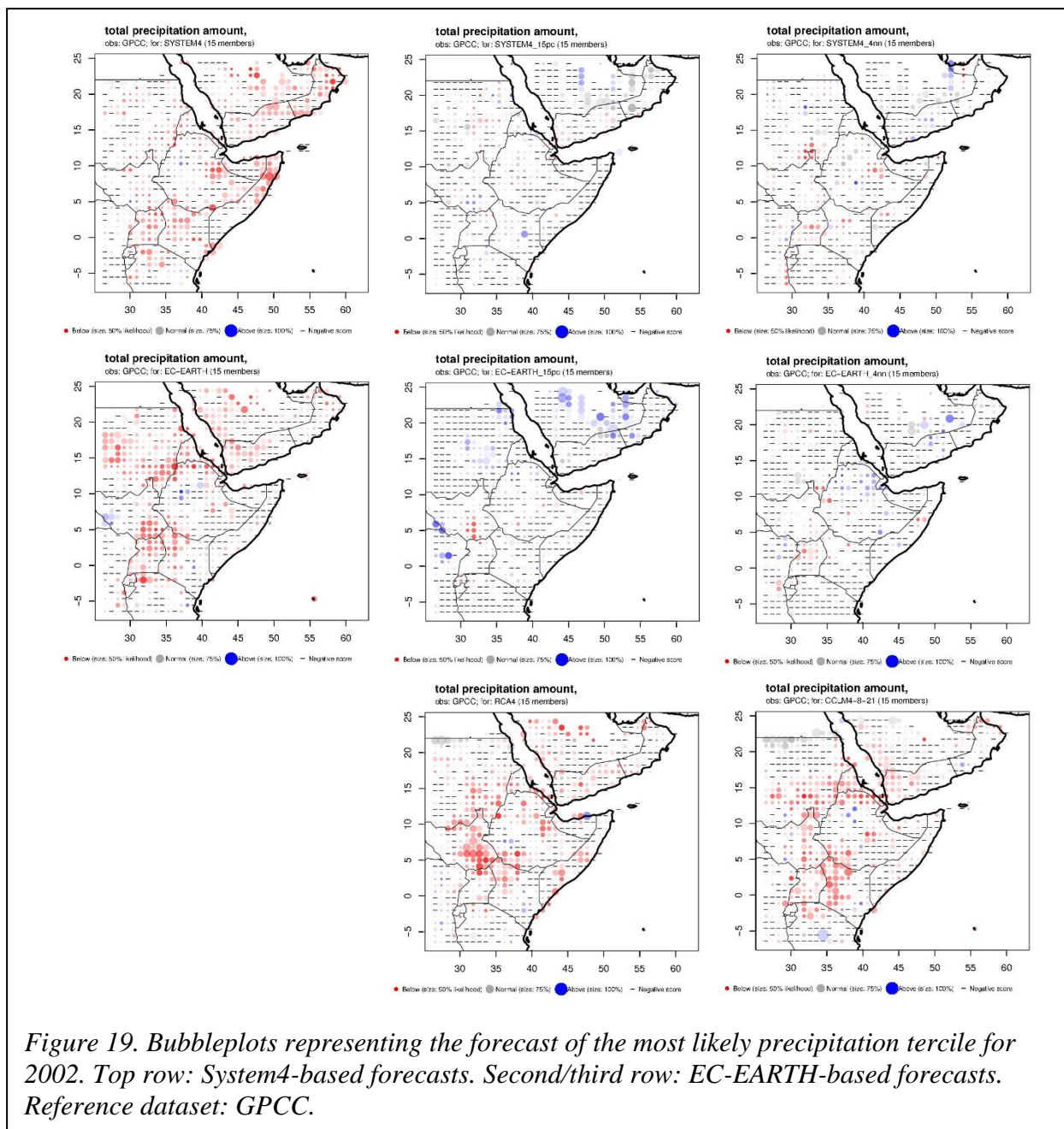
Figure 19. Bubbleplots representing the forecast of the most likely precipitation tercile for 2002. Top row: System4-based forecasts. Second/third row: EC-EARTH-based forecasts. Reference dataset: GPCC.
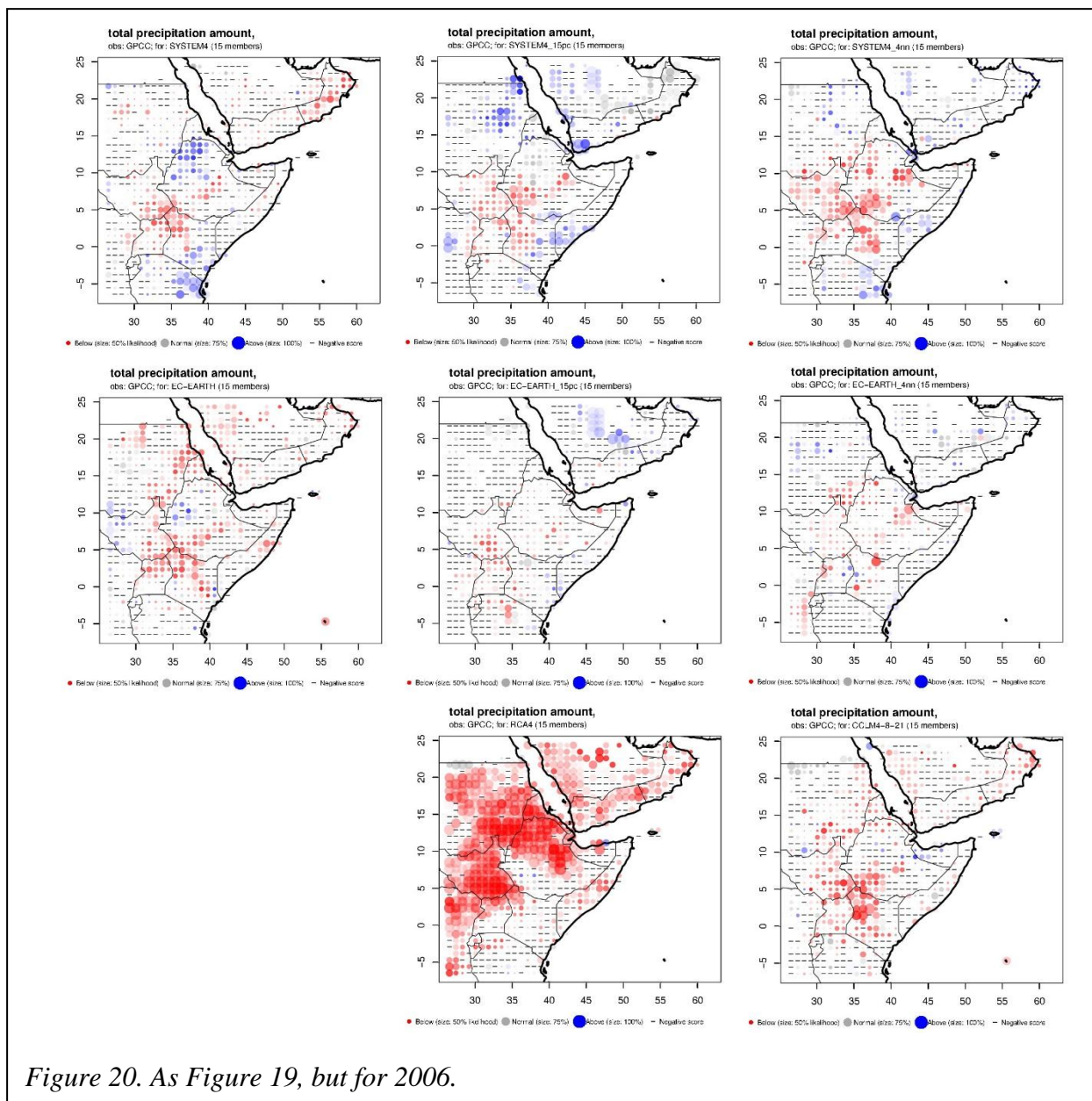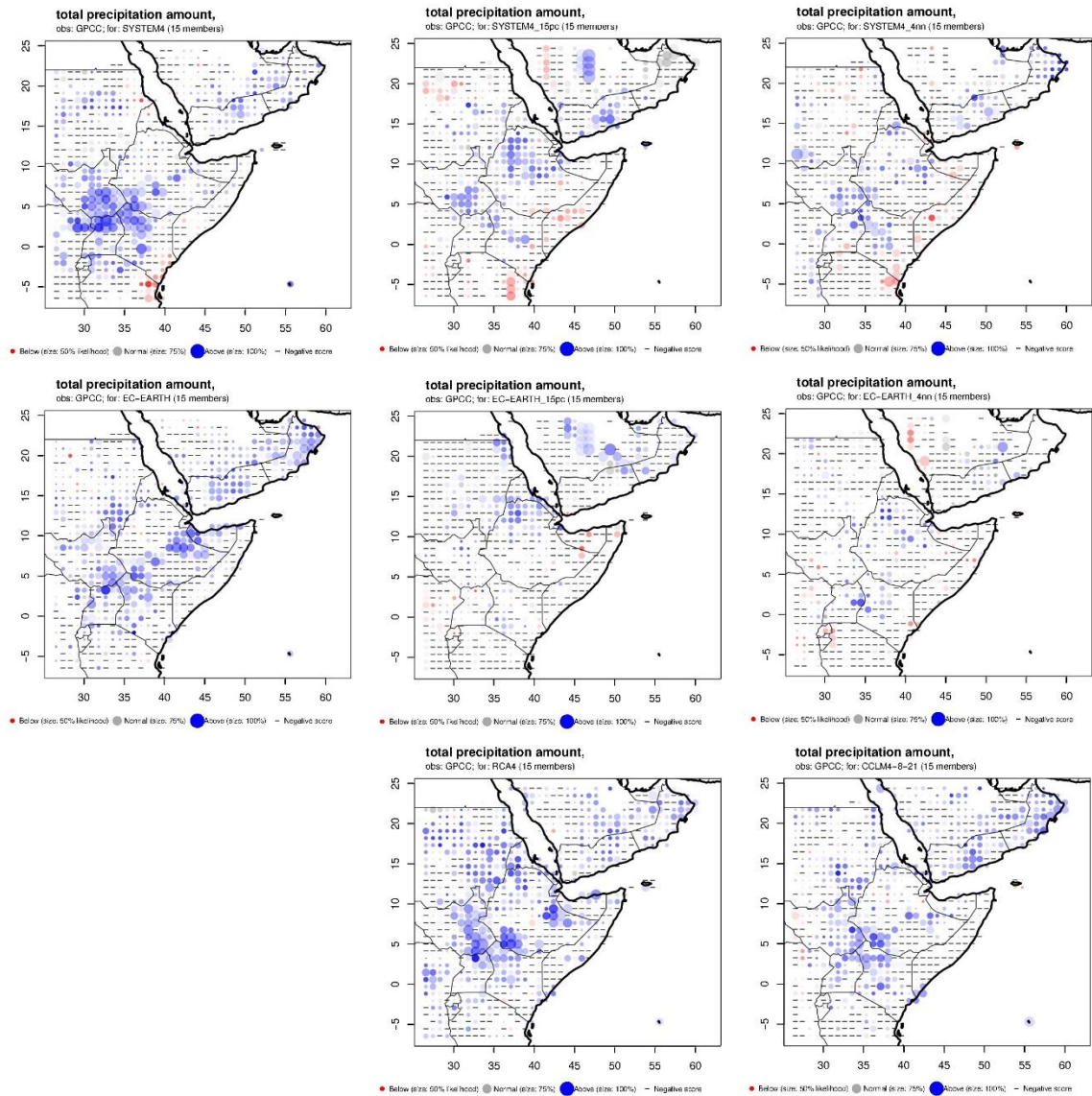
Figure 20. As Figure 19, but for 2006.

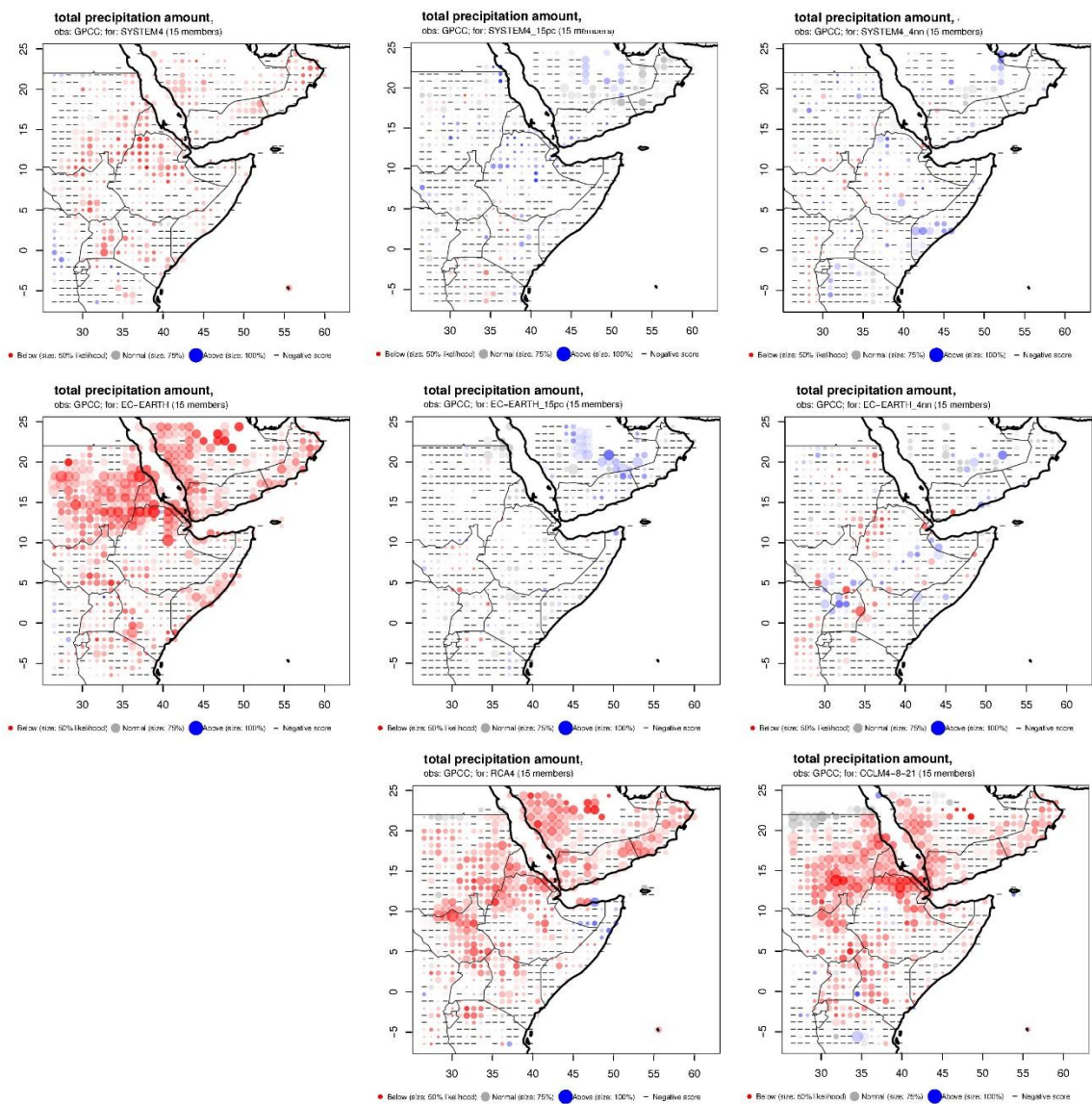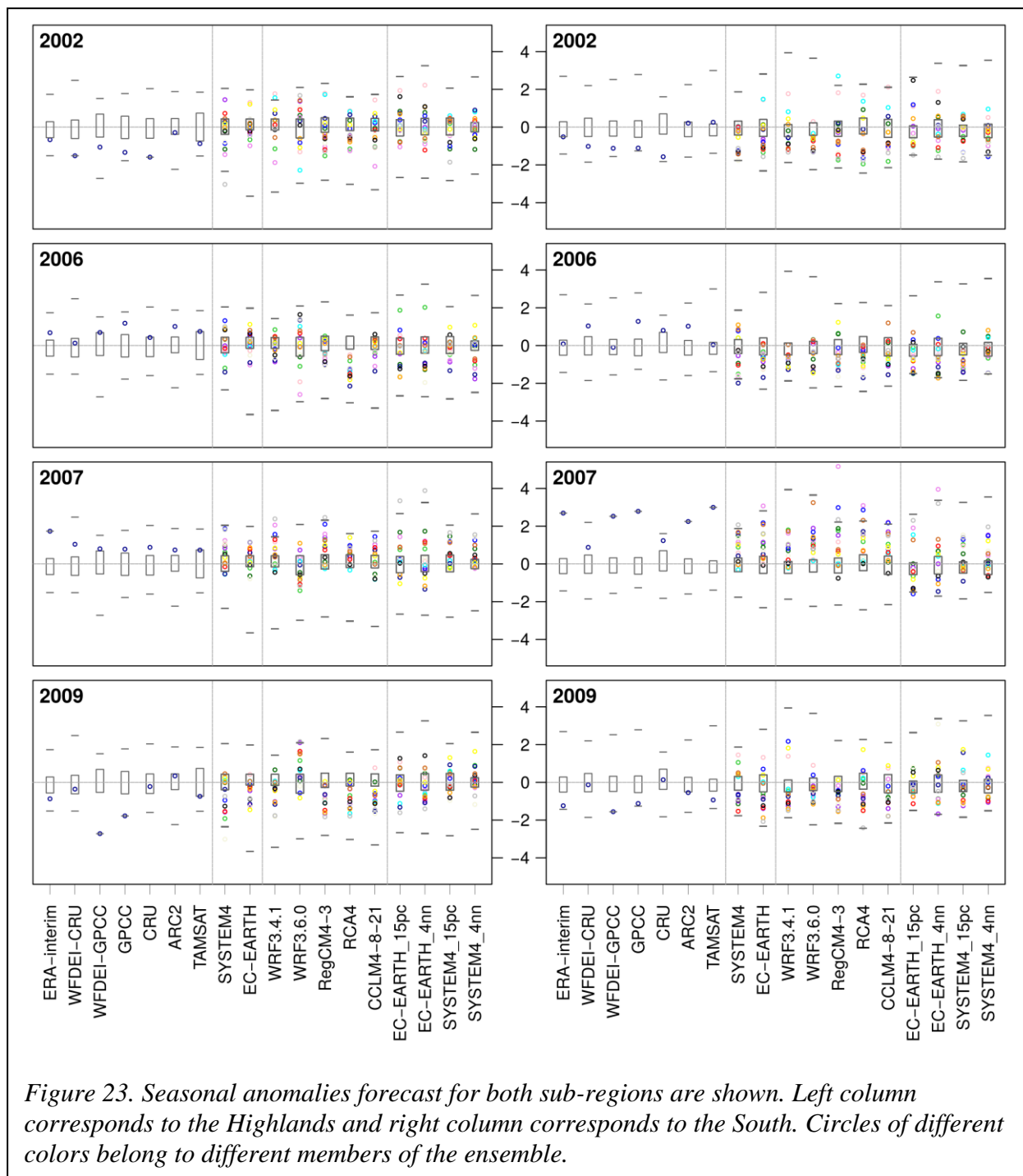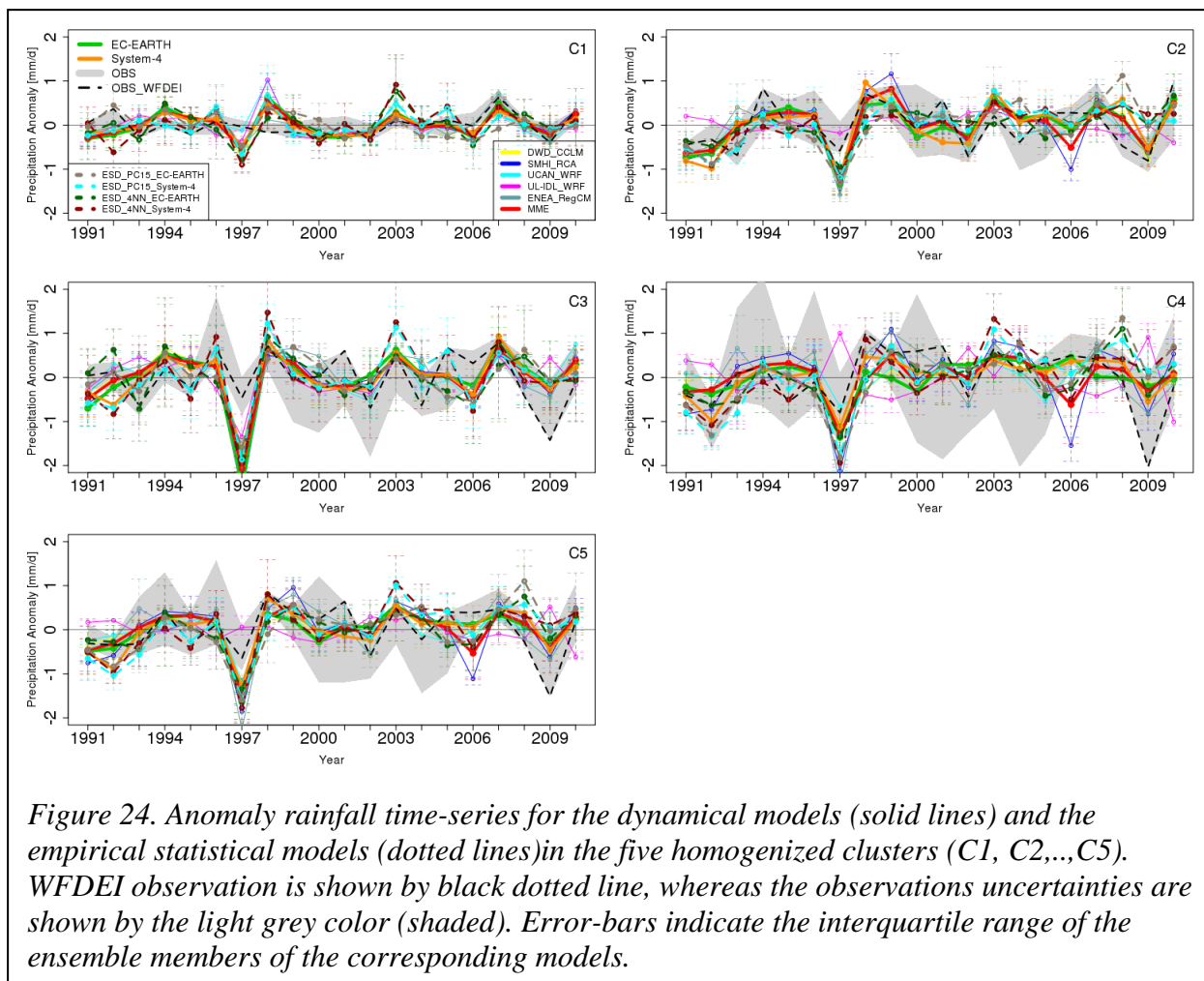*Figure 21. As Figure 19, but for 2007.*

Figure 22. As Figure 19, but for 2009.

Figure 23. Seasonal anomalies forecast for both sub-regions are shown. Left column corresponds to the Highlands and right column corresponds to the South. Circles of different colors belong to different members of the ensemble.

Figure 24. Anomaly rainfall time-series for the dynamical models (solid lines) and the empirical statistical models (dotted lines)in the five homogenized clusters (C1, C2,..,C5). WFDEI observation is shown by black dotted line, whereas the observations uncertainties are shown by the light grey color (shaded). Error-bars indicate the interquartile range of the ensemble members of the corresponding models.

Figure 25. Anomaly correlation coefficients for the statistical and dynamical models with respect to WFDEI in the five clusters.

Figure 26. Fair BSS for the lower rain tercile for the dynamical and statistical models over Ethiopia.

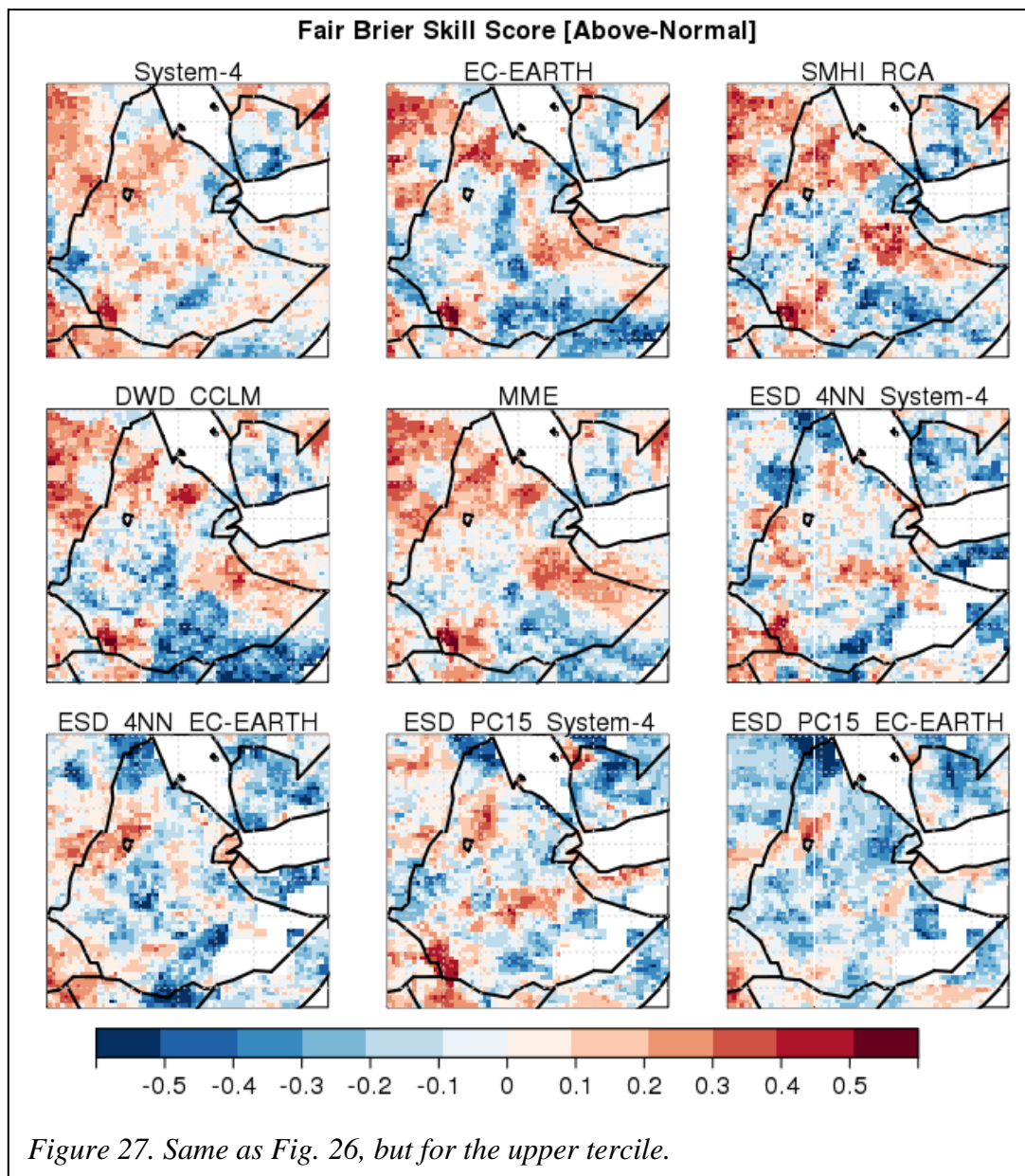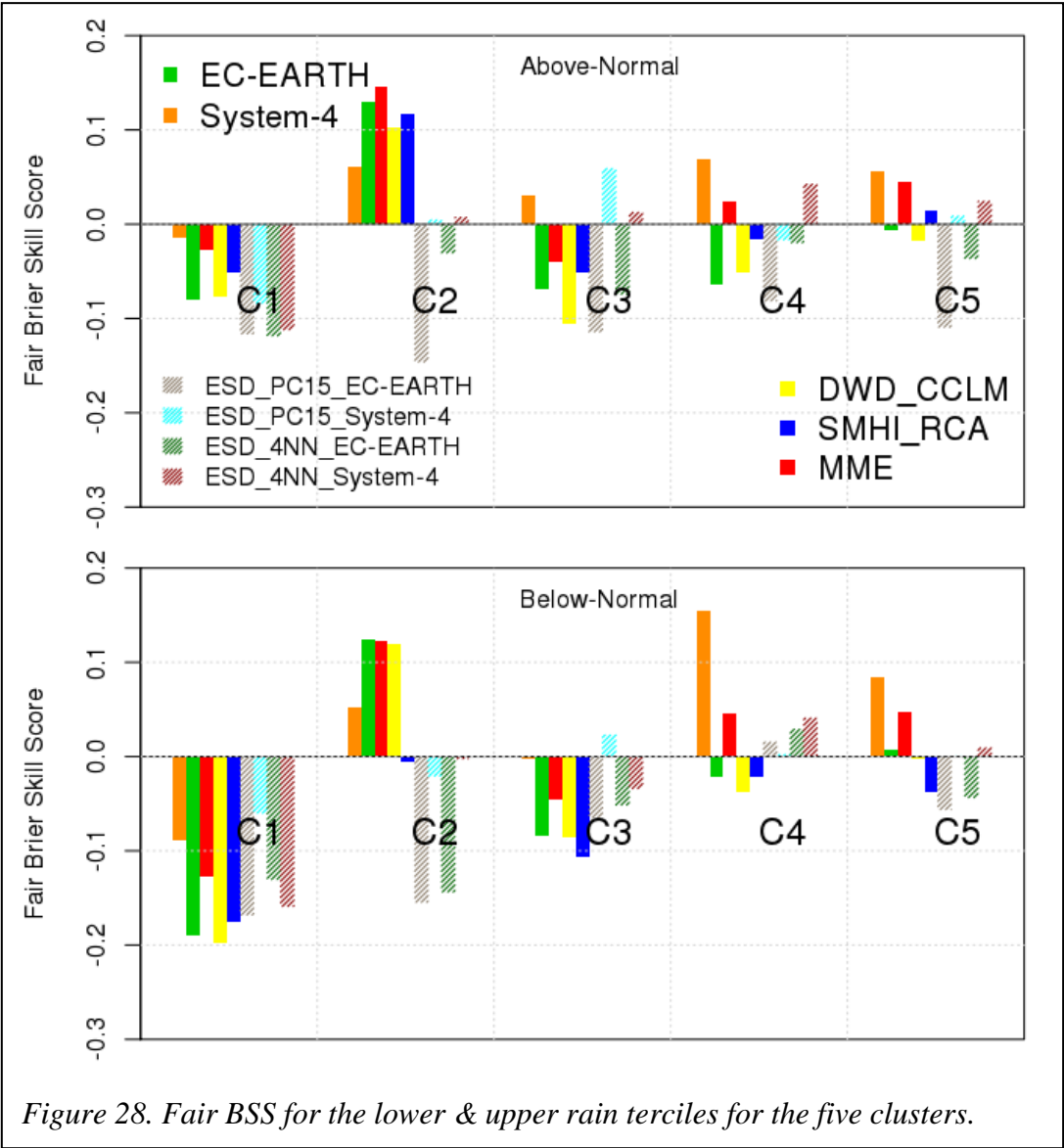*Figure 27. Same as Fig. 26, but for the upper tercile.*

Figure 28. Fair BSS for the lower & upper rain terciles for the five clusters.

Figure 29. Reliability diagram for the lower rain tercile for the five clusters.

Figure 30. Same as Fig. 29, but for the upper tercile.