

Grid-based performance evaluation of GCM-RCM combinations for rainfall reproduction

Ali Danandeh Mehr¹  · Ercan Kahya¹

Received: 21 April 2015 / Accepted: 16 February 2016 / Published online: 1 March 2016
© Springer-Verlag Wien 2016

Abstract Prior to hydrological assessment of climate change at catchment scale, an applied methodology is necessary to evaluate the performance of climate models available for a given catchment. This study presents a grid-based performance evaluation approach as well as an intercomparison framework to evaluate the uncertainty of climate models for rainfall reproduction. For this purpose, we used outputs of two general circulation models (GCMs), namely ECHAM5 and CCSM3, downscaled by a regional climate model (RCM), namely RegCM3, over ten small to mid-size catchments in Rize Province, Turkey. To this end, five rainfall-borne climatic statistics are computed from the outputs of ECHAM5-RegCM3 and CCSM3-RegCM3 combinations in order to compare with those of observations in the province for the reference period 1961–1990. Performance of each combination is tested by means of scatter diagram, bias, mean absolute bias, root mean squared error, and model performance index (MPI) measures. Our results indicated that ECHAM5-RegCM3 overestimates the total monthly rainfall observations whereas CCSM3-RegCM3 tends to underestimate. In terms of maximum monthly and annual maximum rainfall reproduction, ECHAM5-RegCM3 shows higher performance than CCSM3-RegCM3, particularly in the coastland areas. In contrast, CCSM3-RegCM3 outperforms ECHAM5-RegCM3 in reproducing the number of rainy days, especially in the inland

areas. The results also revealed that if a GCM-RCM combination performs well for a portion (statistic) of a catchment, it is not necessarily appropriate for the other portions (statistics). Moreover, the MPI measure demonstrated the superiority of ECHAM5-RegCM3 to CCSM3-RegCM3 up to 33 % excellence for annual rainfall reproduction in Rize Province.

1 Introduction

According to the Intergovernmental Panel on Climate Change (IPCC), climate change is “a long-term, typically decades or longer, change in the state of the climate that can be identified by changes in the mean and/or the variability of its properties”. It may be due to natural internal processes, external forcing, or to persistent anthropogenic changes in the composition of the atmosphere or land (IPCC 2007). Based upon significant effect of climate change on water resources (Özdoğan 2011; Grossi et al. 2013; Shen et al. 2014) and flood events (Kundzewicz et al. 2013), general circulation models (GCMs) are increasingly used as main sources of information required to hydrological impact studies specifically in the regions vulnerable to climate change.

GCMs are main tools to generate climate change projections. Progress in GCMs during the recent decades is vast. Their validation with observations for the twentieth century has indicated that they are able to simulate the basic features of the climate of the Earth (Şen 2013). However, owing to the inherent uncertainty of GCMs (Foley 2010), they still suffer from spatial and temporal differences in their climate outputs, particularly in precipitation (Varis et al. 2004; Kundzewicz et al. 2008; Şen 2013). In addition, GCMs commonly have relatively low spatial resolution and therefore unable to resolve significant catchment-scale features, including topography and land use, as needed in hydrologic modelling and

✉ Ali Danandeh Mehr
danandeh@itu.edu.tr

Ercan Kahya
kahyae@itu.edu.tr

¹ Civil Engineering Department, Istanbul Technical University, Maslak, 34469 Istanbul, Turkey

impact assessment (Moradkhani et al. 2010; Najafi et al. 2011a; Bozkurt et al. 2012). One way to make informed decision associated with a specific region is to obtain fine-scale regional information (i.e. downscaling) using one or more regional climate models (RCMs) that takes boundary conditions from different GCMs. Then either a multi-model approach (e.g. Van der Linden and Mitchell 2009; Najafi et al. 2011b) or an intercomparison strategy (e.g. Vaze et al. 2011; Bozkurt et al. 2012) can be used to find more appropriate combination(s). In both, a catchment-scale performance evaluation of GCM-driven RCM outputs at a certain period of time, commonly referred as reference period, would be crucial to realize inevitable biases of the ensemble or the best GCM-RCM combination.

In addition to intercomparison studies (e.g. Hofstadter and Bidegain 1997; Errasti et al. 2011; Barfus and Bernhofer 2014; Anandhi and Nanjundiah 2015), there are number of researches about the performance of GCM-RCM combinations (e.g. Jones and Reid 2001; Xuejie et al. 2002; Harvey and Wigley 2003; Jacob et al. 2007; Smith and Chandler 2010; Bozkurt et al. 2012; Roosmalen et al. 2010; Fu et al. 2013; Deidda et al. 2013). Xuejie et al. (2002), using a GCM-RCM combination model, CSIRO-driven RegCM2, evaluated the performance of the combination for present-day reproducing and future change of different measures of extreme rainfall and temperature events in China. The authors reported that the CSIRO-RegCM2 combination can reproduce extreme events well in the country. Nieto and Rodríguez-Puebla (2006) compared the observational and simulated precipitation variations over the Iberian Peninsula by using empirical orthogonal function and spectral analyses. Their results pointed out that the Geophysical Fluid Dynamics Laboratory Coupled Model (version 2) resulted in the best correlation with the observed annual rainfall cycle. Perkins et al. (2007) evaluated the performance of the IPCC's fourth assessment report models over Australia in terms of simulating daily maximum temperature, daily minimum temperature, and rainfall probability density functions (PDF). The MIROC-M, CSIRO, and ECHO-G models were recommended as more skilful ones for use in climate change impact assessments over the country. In a similar study, Maxino et al. (2008) reported that the CSIRO, IPSL, and MIROC-M are relatively well models to capture the PDFs of the maximum and minimum temperatures and rainfall over the Murray-Darling Basin in Australia. Fu et al. (2013) developed a score-based multi-criteria method to assess and rank the regional performance of 25 GCMs for the southeastern Australia. The results indicated that two GCMs including CGCM3.1-T47 developed by the Canadian Centre for Climatic Modelling and Analysis and INM:CM30 developed by the Institute for Numerical Mathematics in Russia are the two best models, whilst the Beijing Climatic Centre CM1, Bjerknæs Centre for Climatic Research BCM20, and Atmospheric Research

CSIRO:MK35 resulted in relatively crude representations of monthly rainfall climate.

There are also some regional climate change impact assessments and intercomparison studies in Turkey (e.g. Önoel and Semazzi 2009; Ozkul 2009; Önoel et al. 2014; Bozkurt and Sen 2013). To the best of our knowledge, the only performance evaluation study for the country (so far) has been carried out by Bozkurt et al. (2012). Comparing three GCM-RCM combinations (i.e. ECHAM5-RegCM3, CCSM3-RegCM3, and HadCM3-RegCM3) with those of different sets of observations at eastern Mediterranean–Black Sea region for the period 1960–1990, the authors recommended that all of the combinations can be considered for climate change impact assessment in the study region.

Considering spatial resolution degree of current GCMs or even GCM-RCM combinations, none of the abovementioned results obtained from either intercomparison or performance evaluation studies are directly applicable for catchment-scale floodplain engineering works. Additionally, most of the studies compare climate model outputs with reanalysis data that might not be reliable enough for performance evaluation at catchment scale, specifically in the ungauged or poorly gauged catchments. The reason behind this is that not all reanalysis data are created by observations of the catchment of interest. Some data types, such as precipitation (depending on the reanalysis) are obtained by running (presumably newer) GCM or numerical weather prediction models. Addressing these issues, an applied grid-based GCM-RCM performance evaluation method as well as an intercomparison study between two GCM-RCM combinations are presented in this study. The results will be used for selecting the more reliable climate model outputs as boundary conditions for an ongoing catchment-scale climate change impact assessment project FRA-PFC Rize: “Flood Risk Assessment: Present and Future Conditions in Rize Province” in Turkey. In addition, in the present study, we provide a grid-based overall performance analysis, whereas most of the previous intercomparison studies did not discuss the overall performance of climate models. Although some catchment-scale hydrological impact studies have been carried out in recent years (e.g. Graham et al. 2007a, 2007b; Bastola et al. 2011; Mondal and Mujumdar 2012; Halmstad et al. 2013; Demirel et al. 2013), to the best of our knowledge, there is only a few studies in the relevant literature focused on the performance evaluation of GCM-RCM combinations in catchment scale (e.g. van Vliet et al. 2012 and Deidda et al. 2013). Since such studies make use of small areas to compare the performance of several GCM-RCMs for their ability to reproduce hydrological variables (such as rainfall, temperature, and others), there is still need for more studies to establish a reference for hydrological applications in general. Information about the implemented models, methods, and observational data together with the obtained results has been described in the following sections.

2 Models description

Our analysis is restricted to the two GCMs (i.e. ECHAM5 and CCSM3) integrated with an RCM namely RegCM3 due to the fact that they were only available models for the study region providing daily outputs at the time of analysis, and therefore, the ongoing climate change impact assessment project on the study area is based on projections obtained by these combinations. Since the results are planned to be used as a basis for (i) selecting more reliable GCM-RCM combination, (ii) implementing/developing an appropriate bias correction method, and (iii) subsequent projected extreme rainfall modeling (Danandeh and Kahya, submitted), five rainfall-borne statistics (are explained in Section 5), derived from GCM-RCM outputs have been considered for performance evaluation in this study.

RegCM3, an upgraded version of RegCM2 developed by the International Centre for Theoretical Physics in Italy, is an RCM with sigma-pressure vertical coordinate that comprises a radiative transfer package, a non-local boundary layer scheme, an ocean surface flux parameterization, an explicit moisture scheme, and a large-scale cloud and precipitation scheme with several options for cumulus convection scheme for the model simulations (Bozkurt et al. 2012). Detailed descriptions of physical parameterizations of RegCM3 can be found in Pal et al. (2007).

ECHAM5 is a coupled atmosphere-ocean GCM developed at the Max Planck Institute for Meteorology, Germany. The model is used in a number of configurations which differ in the vertical extent of the atmosphere as well as the relevant processes. ECHAM5 is described in detail in Roeckner et al. (2003). CCSM3 is the third generation of the Community Climate System Models that has recently been released to the climate community. CCSM3 is a coupled GCM with components representing the atmosphere, ocean, sea ice, and land surface connected by a flux coupler. The model was designed to produce realistic simulations over a wide range of spatial resolutions, enabling inexpensive simulations lasting several millennia or detailed studies of continental-scale dynamics, variability, and climate change (Collins et al. 2006). The code, documentation, input datasets, and model simulations are freely available from the model website. Detailed descriptions of numerical parameterizations and physics of CCSM3 can be found in Collins et al. (2006).

In order to downscale the outputs of ECHAM5 and CCSM3 over Turkey, the relevant simulations were performed continuously from 1 January 1960 to 31 December 1990 at Istanbul Technical University through the project of the United Nations Development Program entitled “Enhancing the Capacity of Turkey to Adapt to Climate Change”. Detailed information about the project can be found in Bozkurt et al. (2012) and Önoğlu et al. (2014). The reference period of GCM-RCM outputs used in this study has been

selected identical to the present-day performance evaluation period (i.e. 1960–1990) implemented in the project. Considering the adjustment period (spin-up), we removed the first year outputs from the simulations’ archive before computing climate statistics.

3 Rainfall climatology of the study area

The study area is Rize Province, located in eastern Black Sea region between the Pontic Mountains and the Black Sea in Turkey, which is considered as the “wettest” corner of the country (Fig. 1). The province with a total area approximately around 3900 km² encompasses ten small to mid-size catchments. According to Koppen-Geiger climate classification, it has a borderline oceanic/humid subtropical climate similar to most of the eastern Black Sea region, with warm summers and cool winters. Annual precipitation averaging in the province is around 2500 mm with a maximum rate commonly in late autumn (October to December). Rainfall climatology (i.e. long-term mean) of the study area at reference period 1961–1990 is investigated in two portions as inland and coastland areas by mainly considering the elevation differences from sea level that depict two primary precipitation regimes. With respect to the well-documented significant effect of coastline on rainfall distribution over the coastal area of the eastern Black Sea region (Eris and Agiralioglu 2009), such a separation provides more accurate climatology for each portion due to elimination of averaging error. A digital elevation model including the borders of the study region and utilized rain gauges locations to derive the province climatology is presented in Fig. 1.

According to the figure, coastland is classified as the portion where the elevation is below 1500 m and inland corresponds to the areas where the elevation is greater than 1500 m, mainly constituting the headwaters of the catchments. As shown in the figure, precipitation data were obtained from ten rain gauges. Half of them are operated by the Turkish State Meteorological Service (TSMS) and the remaining stations are operated by the Turkish State Hydraulic Works known as DSI in Turkey. DSI provides only monthly records. Thus, we used the relevant data only for estimating the monthly rainfall climatology of the study area. Homogeneity of the annual rainfall records in the eastern Black Sea region for the period of 1960–2005 has been already proved by Eris and Agiralioglu (2009). Rainfall climatology of each gauge for the reference period is presented in Table 1 that indicates notable annual rainfall difference between gauges located in the coastland and inland portions. In contrast to the generally accepted idea of the increase of precipitation with height, the orographic effects cannot be seen for the study area. It means that mean annual precipitation decrease with distance from sea. Such a pattern was already shown for entire eastern

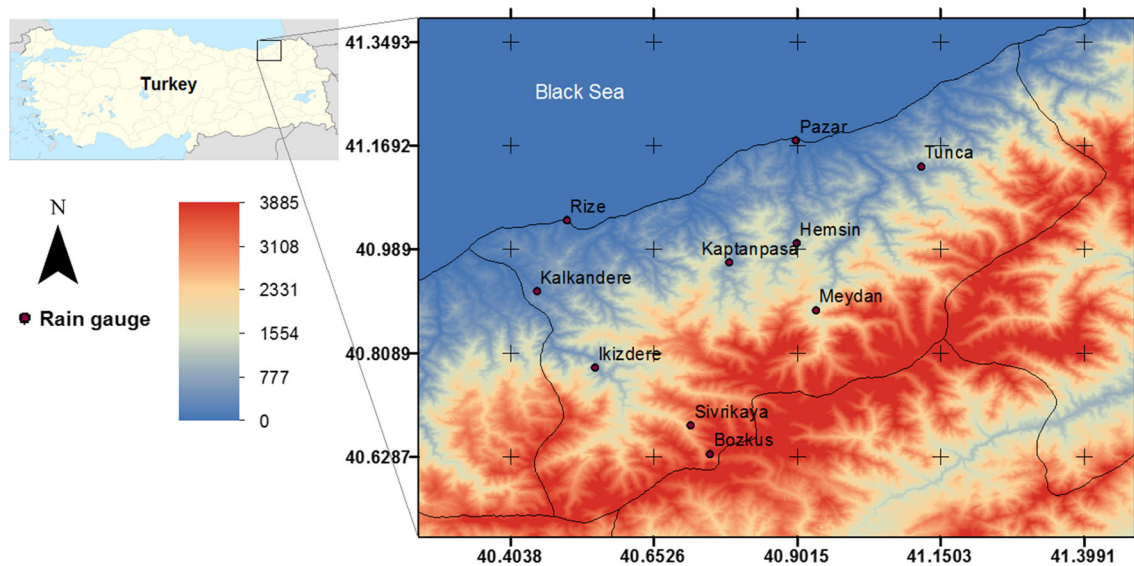


Fig. 1 Locations of the study area, Rize Province, and rain gauges used in the study

Black Sea region by Eris and Agiralioğlu (2009). The authors investigated the precipitation records of 33 rain gauge stations (46-year period) across the entire eastern Black Sea region and demonstrated that mean annual precipitation decreases with elevation as well as distance from the sea.

In order to estimate areal distribution of rainfall climatology, two different averaging methods including arithmetic mean and isohyet lines have been used in this study, and the results have been presented in Table 2. We utilized Kriging optimal interpolation method (Kitanidis 1997) to generate isohyet maps. For instance, isohyets produced for autumn season have been presented in Fig. 2. Five selected RegCM3 grids, providing good spatial coverage of the province, having the resolution of 27 km are also shown in the figure. In this study, the results of isohyet line method have been considered to compare with those of climate model outputs at each grid (i.e. grid-based performance evaluation). In addition, with respect to the total area of grids overlaid on each portion, the

area weighted average method has also been applied to portion-based performance evaluation.

4 Evaluation criteria

Our review in the literature demonstrated that there is no agreement in the scientific community on how to determine the best-performing climate models. It is up to the researchers which performance measures are used. In the present study, precipitation realization performance of the GCM-RCM combinations is analysed based on five measures comprising (i) scatterplots of simulation/observation, (ii) bias (simulation–observation), (iii) mean absolute bias (MAB), (iv) root mean squared error (RMSE), and (v) model performance index (MPI). The first four measures are applied only to individual climatic statistics whereas the last one (i.e. MPI) is applied to represent the overall performance of the model. Concepts and

Table 1 Monthly and annual rainfall climatology (mm) of utilized rain gauges

Rain gauge	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year
Rize	223	170	147	101	102	125	131	189	218	274	254	236	2170
Kalkandere	174	162	135	118	123	151	168	194	191	279	219	193	2107
Pazar	182	142	111	81	83	152	142	164	210	270	235	211	1983
Tunca	196	163	120	106	121	140	131	151	190	246	188	198	1950
Kaptanpasa	134	133	96	82	92	140	108	92	138	188	150	175	1528
Hemsin	147	99	87	83	79	138	90	110	153	170	148	170	1474
Meydan	121	79	73	81	112	99	79	70	88	117	123	132	1174
Ikizdere	107	75	62	76	84	93	54	72	86	140	135	135	1119
Sivirkaya	51	57	66	119	157	114	90	76	70	134	84	64	1082
Bozkus	20	30	34	53	94	80	57	39	28	54	38	26	553

Table 2 Climatology of Rize Province (total monthly precipitation, mm)

Area	Method	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year
Coastland	Isohyet	155	123	100	73	77	128	112	130	160	206	184	178	1625
	Arithmetic	181	145	118	97	99	144	130	156	191	247	204	204	1918
Inland	Isohyet	132	97	83	92	112	114	93	99	113	157	129	130	1351
	Arithmetic	68	61	60	85	120	95	68	61	62	103	87	80	950
Province	Isohyet	141	108	90	84	98	119	101	111	132	177	151	149	1461
	Arithmetic	134	110	94	92	108	123	104	117	137	187	156	153	1514

mathematical expressions of the scatter diagram, bias, MAB, and RMSE are well documented in a number of model performance studies (e.g. Danandeh Mehr et al. 2013; Liu et al. 2014; Demirel and Moradkhani 2015). In order to avoid repetition, only the MPI measure is therefore described in this section. The MPI, firstly introduced by Reichler and Kim (2008), is an overall performance evaluation index, which measures the reliability of a model based on the scaled-normalized error variance (I^2) of a broad range of climate variables (v).

$$I_i^2 = e_{vi}^2 / e_{vi}^2 \tag{1}$$

$$e_{vi}^2 = \sum_{j=1}^J \left(w_i (M_{vij} - O_{vj}) / \sigma_{vj}^2 \right) \tag{2}$$

where e_{vi}^2 is normalized error variance of i th model; e_{vi}^2 is the average error found in a reference ensemble of models; J is the number of grid cells in the catchment; M_{vij} and O_{vj} are the modelled and corresponding observed climatology, respectively; σ^2 is the inter-annual variance from the observations; w_i is proper weight needed for area and mass averaging model; subscripts i and j are the indices for each model and grid points, respectively; and the overbar indicates averaging.

To rank climate models with respect to overall climatic statistics, the overall MPI is derived as the mean over the I^2 of all climate variables. This is based on the assumption that a

reliable model should represent several components of the climate system of whole catchment.

$$MPI = I_i^2 \tag{3}$$

The MPI varies around one with values greater than one for underperforming models and values less than one for more accurate models. The expression $M-O$ in Eq. (2) indicates the error between simulation and observation. The smaller the error, the smaller the index and the better the model. More details about MPI can be found in Reichler and Kim (2008).

5 Results and discussion

In this section, ability of ECHAM5-RegCM3 and CCSM3-RegCM3 to reproduce five climatic statistics at each RCM grid as well as two different lands across Rize Province are investigated. The climatic statistics include the following: (i) total monthly rainfall (i.e. sum of long-term mean of daily rainfall at each month; TMR), (ii) monthly maximum rainfall (i.e. the maximum value among long-term mean of daily rainfalls at each month; MMR), (iii) total annual rainfall (i.e. sum of long-term mean of daily rainfall; TAR), (iv) annual maximum rainfall (i.e. the maximum of daily rainfalls at each year; AMR), and (v) number of rainy days (NRD). The choice of each grid as performance evaluation zone is dictated by the

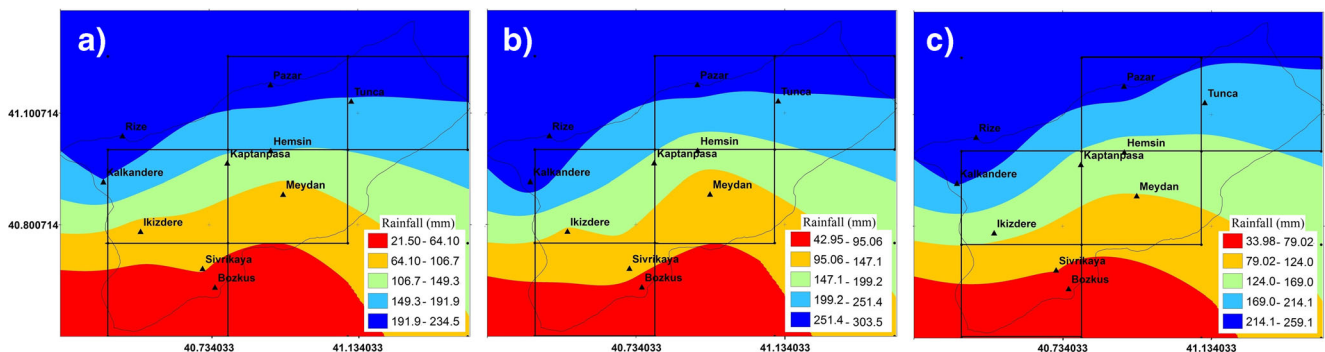


Fig. 2 Examples of long-term mean of total monthly rainfall over Rize Province: **a** September, **b** October, and **c** November

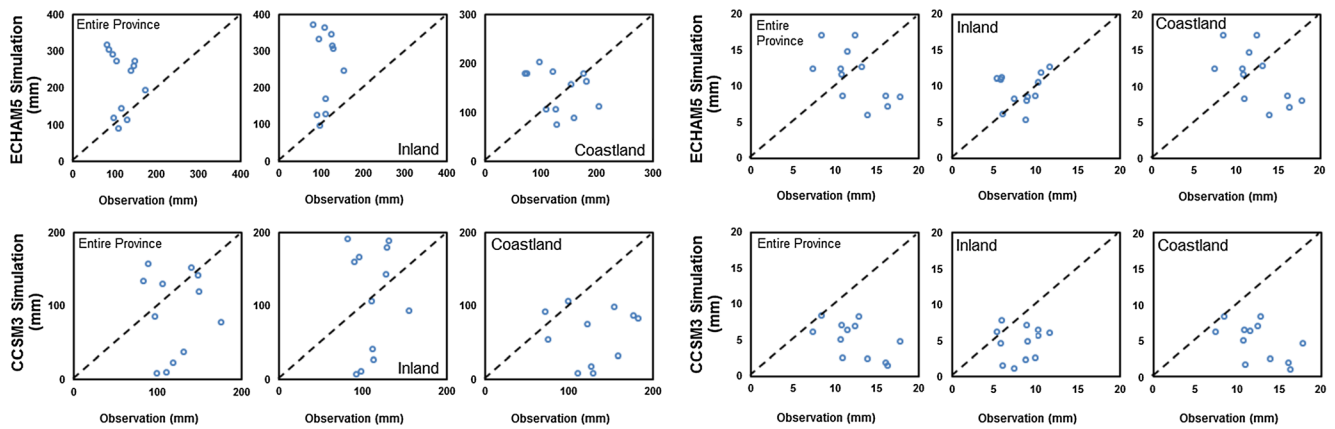


Fig. 3 Scatterplots of total monthly rainfall (*left*) and monthly maximum rainfall (*right*) at each portion and the entire Rize Province for the period 1961–1990

type of observational data sets available at each grid. It is worth mentioning that simulation results at the reference period 1961–1990 composed of 10,957 days ($365 \times 30 + 7$ leap days) at each grid.

5.1 Monthly performance

It is conventional to evaluate simulation/prediction models first by scatterplots of simulation/observation, so that we will be able to realize the uncertainty features of simulation results. The scatterplots of the observed and modelled TMR and MMR data for the two portions and the entire province are provided in Fig. 3. Significant differences between the results of different GCM-RCM combinations are clear in the figure. According to the Fig. 3 (left), generally, ECHAM5 tends to

overestimate the TMR for the entire province whilst CCSM3 yields slightly dryer condition. In terms of MMR reproduction associated with each portion, it can be indicated that ECHAM5 resulted in higher performance than CCSM3, particularly for the inland portion (see Fig. 3 (right)). To evaluate the efficiency of the GCM-RCM models more thoroughly, variations of the observed and modelled TMR and MMR values at each month for each portion are depicted in Fig. 4. Differing largely from each other, both GCMs show the same annual trends, which are more or less consistent with the observed climate trend. In contrast with ECHAM5, the CCSM3 underestimates the TMR amount in seasons other than winter and early spring. With respect to the MMR simulation results (see Fig. 4 (left)), both GCM-RCM combinations reproduce substantial biases in the coastland portion. Relatively good

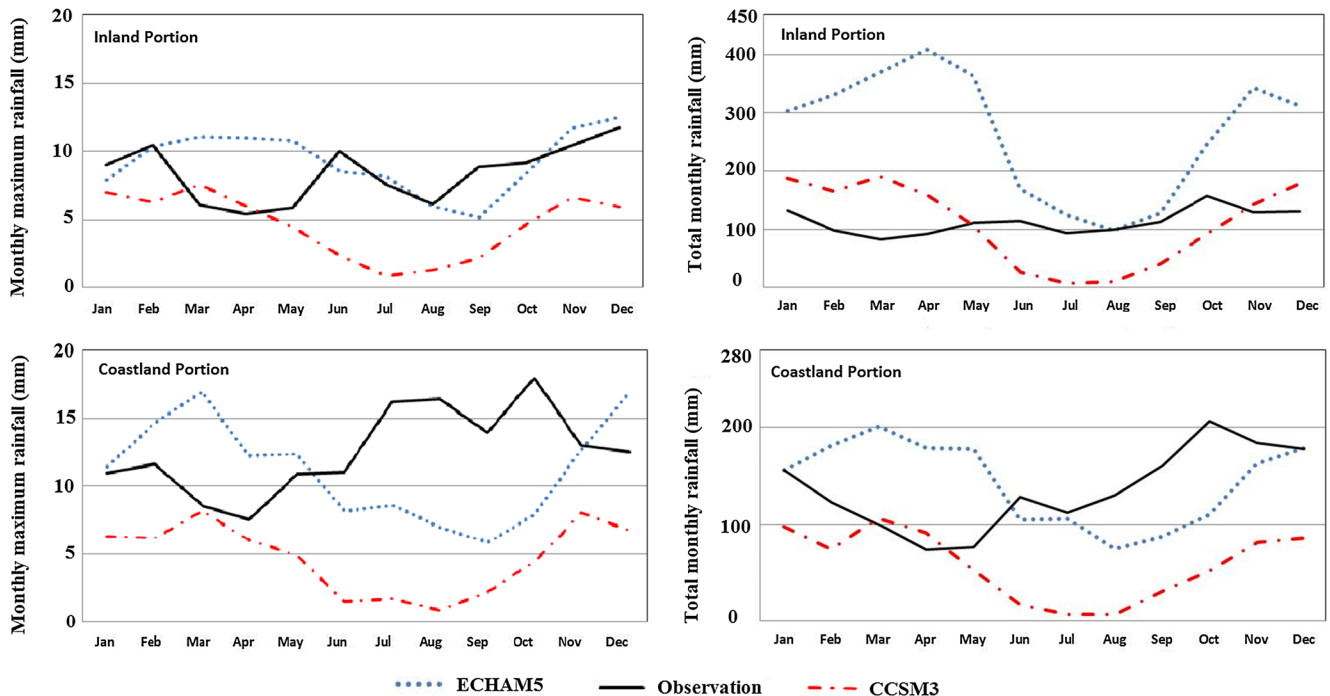


Fig. 4 Total monthly rainfall (*top*) and monthly maximum rainfall (*bottom*) climatology of Rize Province for the period 1961–1990

Table 3 Portion-based bias, MAB, and RMSE of GCMs driven RCM simulations (mm)

Statistic	Model	Portion	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	MAB	RMSE
TMR	ECHAM5	Inland	172	234	288	317	251	55	31	-2.0	14	89	215	182	154	188
		Coastland	0.1	58	101	105	101	-23	-6.0	-55	-73	-95	-21	1.0	53	67
	CCSM3	Inland	56	68	108	-68	-6.0	-88	-87	-89	-73	-64	14	48	64	70
		Coastland	-58	-49	6.0	18	-24	-112	-105	-123	-130	-154	-102	-92	81	93
MMR	ECHAM5	Inland	-1.1	-0.1	5.1	5.6	4.9	-1.5	0.6	-0.2	-3.7	-0.7	1.2	0.8	2.1	2.9
		Coastland	0.5	3.0	8.4	4.7	1.5	-2.9	-7.6	-9.4	-8.1	-10	-0.4	4.4	5.1	6.1
	CCSM3	Inland	-2.1	-4.2	1.6	0.6	-1.4	-7.6	-6.7	-4.8	-6.7	-4.5	-3.8	-5.8	4.1	4.7
		Coastland	-4.6	-5.5	-0.4	-1.5	-6.0	-9.5	-14.5	-15.6	-11.8	-13.5	-5.0	-5.8	7.8	9.2

agreement with reference climatology is only provided by ECHAM5 in the inland portion. The coastland area in eastern Black Sea frequently receive more rainfall in the second half-year (July–November), whilst both ECHAM5 and CCSM3 simulated the opposite pattern of approximately first half-year rainfall (December–May) greater than the second half-year. Apart from February, March, and April, ECHAM5 outperforms CCSM3 at whole year. It can be concluded that the major uncertainty in maximum monthly simulation for the province belongs to the coastland portion, which may originate from the local effect of the Black Sea that is not well simulated by the models.

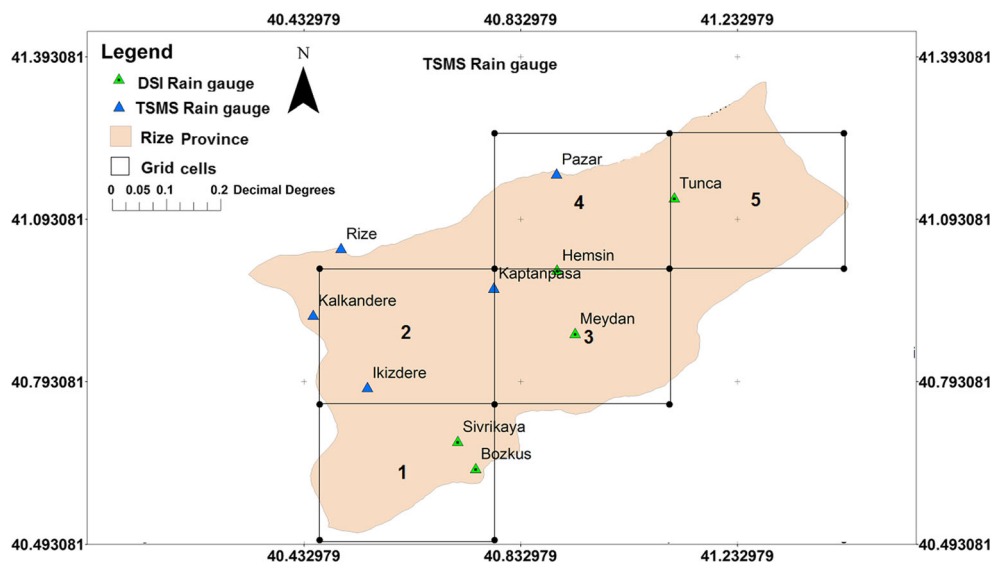
In order to provide profound discussion on simulation results, Table 3 presents quantitative errors of climate models for the both TMR and MMR statistics at each portion. Considering TMR, the highest bias value comes from ECHAM5 reproduction at the inland portion, whereas it interestingly reproduces the lowest bias for the coastland estimation. The spatial inconsistency of a climate model in fine resolution reveals the fact that GSM-RCM intercomparison studies should be carried out in catchment scale if the results are expected to be applied in hydrological design. The RMSE

values of ECHAM5 and CCSM3 for the TMR estimation vary from 67 to 188 and 70 to 93 mm, respectively. According to the mean observed monthly rainfall for the entire province being equals to 122 mm, the results indicated up to 154 and 76 % uncertainty for ECHAM5 and CCSM3, respectively. Regarding the mean of observed MMR (equal to 9.0 and 13 mm at the inland and coastland portions, respectively, see Fig. 4), the best simulation performance for reproducing MMR over the province belongs to ECHAM5 with 32 and 47 % uncertainty at inland and coastland portions, respectively. Corresponding uncertainty values for CCSM3 results are 52 and 71 %, respectively.

5.2 Annual performance

In order to identify uncertainty features of annual rainfall reproductions, three different statistics including TAR, AMR, and NRD are considered in this study. Owing to types of the data sets available in TSMS and DSI rain gauges (see Section 3), the annual performance is examined in grid-based level where at least one TSMS rain gauge is located in the grid of interest. In grid-based analysing, expected results

Fig. 5 Distribution of rain gauges Turkish State Meteorological Service (TSMS) and Turkish State Hydraulic Works (DSI) and the location of five RegCM3 grid boxes surrounding the Rize Province



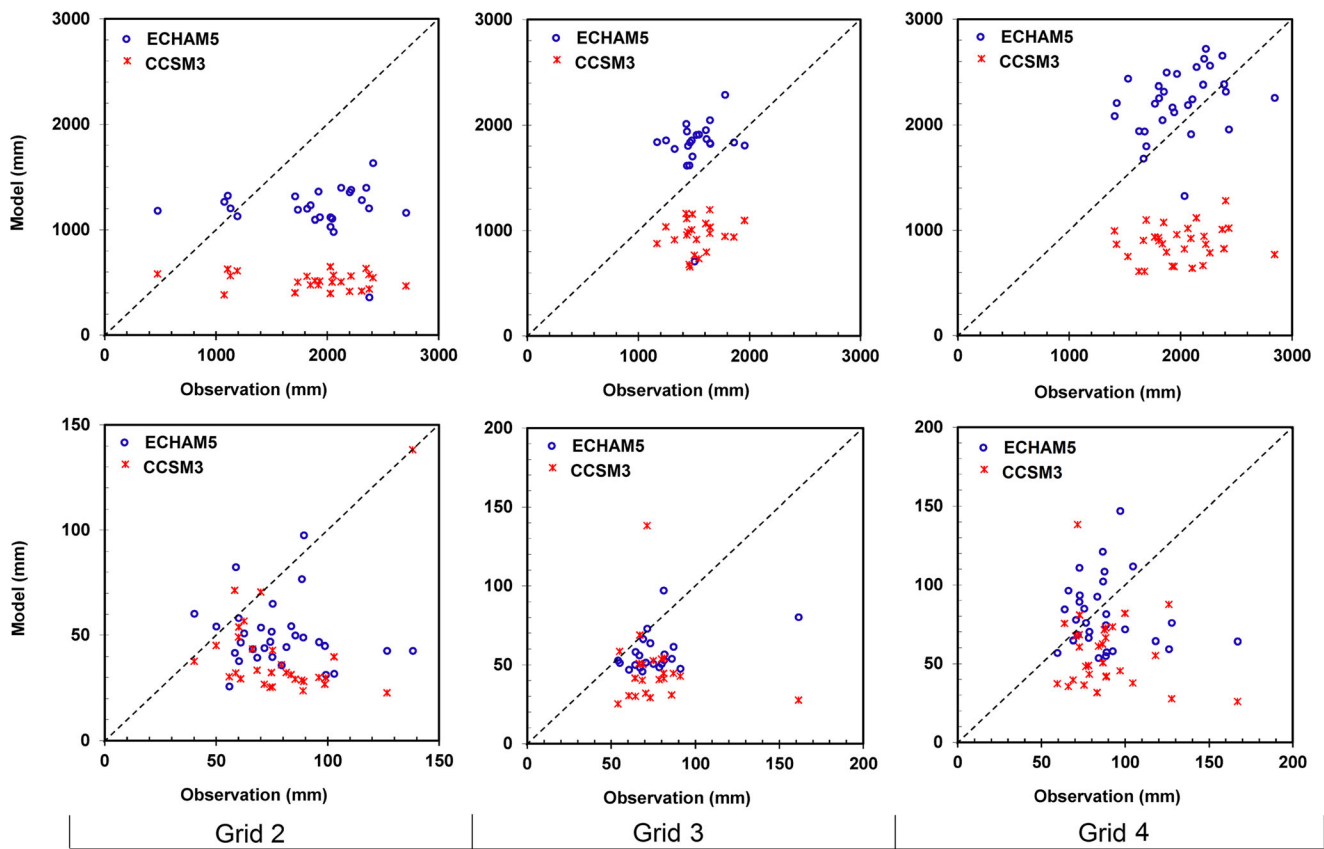


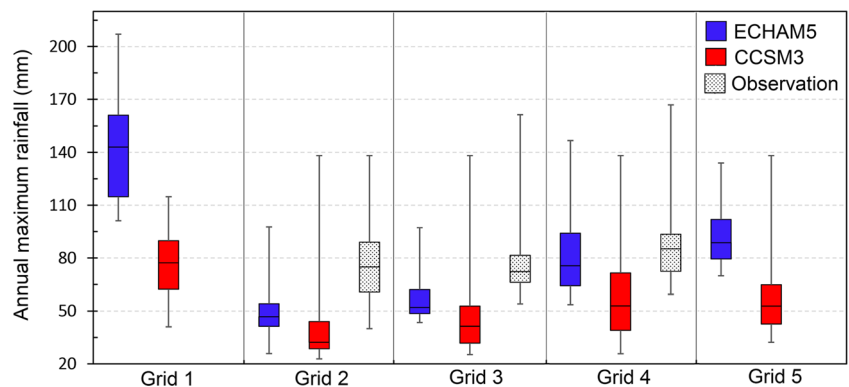
Fig. 6 Scatterplots of modelled and observational total annual rainfall (*top*) and annual maximum rainfall (*bottom*). Plots were given for the RegCM3 grid possessing at least one TSMS rain gauge

are more reliable for different part of the province; however, one can develop portion-based deduction using the mean of performance of grids laying at each portion. In the present study, we derived overall result based upon both arithmetic mean and MPI calculated at grid level. Areal distribution of TSMS and DSI rain gauges across the province as well as grids identification number adopted for the study have been displayed in Fig. 5.

Thirty-year scatterplots of the TAR and AMR for grids 2 and 4 as well as 23-year scatterplot for grids 3 (due to 7-year recording gap in Kaptanpasa station) have been presented in Fig. 6. Because of the lack of daily precipitation

records at grids 1 and 5, the comparison was only performed for grids 2 through 4. Highly scattered simulations of CCSM3-RegCM3 particularly in grids 2 and 4, which roughly represent the coastland portion, indicate the inability of the combination to simulate reliable TAR and AMR specifically in coastland portion. According to the figure, ECHAM5-RegCM3 provides wetter condition than CCSM3-RegCM3 in all grids. Dense scatter of ECHAM5-RegCM3 outputs around the perfect model diagonal line in all grids implies superiority of ECHAM5 to CCSM3 for both total annual and annual maximum rainfall simulation at Rize Province. Thus, projected AMR series

Fig. 7 Box and whisker plots of modelled and observational annual maximum rainfall values at the RegCM3 grids for the period 1961–1990



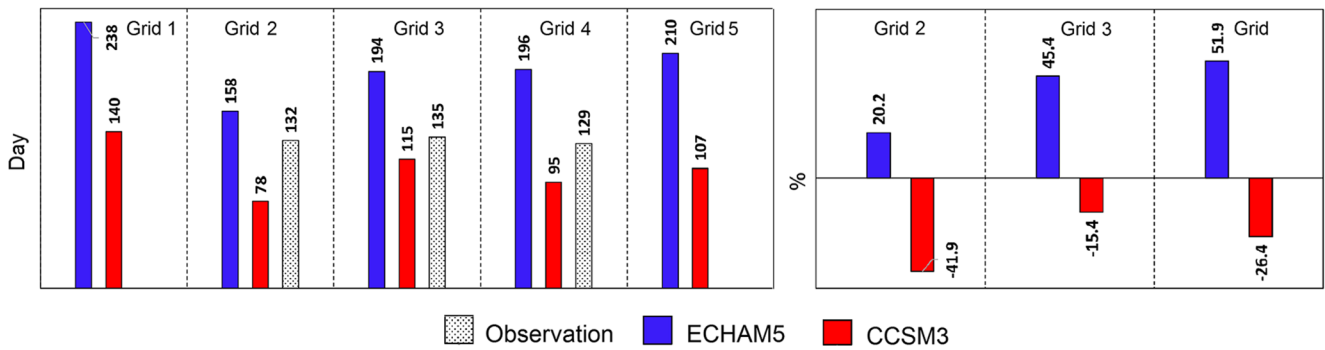


Fig. 8 Number of rainy days at each RegCM3 grids (left) and percentage of simulation error (right). The percentage errors were given for the grid possessing at least one TSMS rain gauge

produced by ECHAM5-RegCM3 may be more reliable to consider in the FRA-PFC Rize project.

Owing to the importance of projected AMR series on extreme events, efficiency of the climate models is considered to supplementary evaluation using boxplot of the observed and modelled AMR series at reference period (see Fig. 7). The horizontal line inside the boxes represents the median and the upper and lower vertical lines, whiskers, indicate the 75 and 25th percentiles, respectively. The figure depicts a positive spatial trend in both minimum and maximum values over the province, which indicates an increasing rainfall pattern from the southwest to northeast of the study region. Both ECHAM5 and CCSM3 more or less follow this pattern in all grids, except in grid 1. The figure shows considerable differences between modelled and observational extremes. Only in grid 2, CCSM3 reproduced maximum value of AMR series (equals to 139 mm) and its event year (i.e. 1979) identical to those of observational data series interestingly. Considering the medians, the figure also implies that ECHAM5 reproduces closer values to the observed median than those of CCSM3.

As another annual climatic statistic for grid-based inter-comparison, Fig. 8 compares the observed and simulated

long-term average of NRD, defined here as days of rainfall in excess of 1 mm. The figure demonstrates that the annual NRD in the entire province falls in the reasonably consistent range 129–135 days. With respect to ECHAM5 and CCSM3 reproductions, the rainy days are within the wide range 158–196 and 78–115 days, respectively. The figure also indicates that ECHAM5 tends to overestimate the observational NRD for the entire province in the range 20–52 %. By contrast, CCSM3 underestimates in the range of 15–42 %. It is obvious from the figure that the lowest percentage error (–15.4 %) of rainy days estimation belongs to CCSM3 and the highest one (51.9 %) belongs to ECHAM5.

The annual simulation results at each grid have been numerically compared in Table 4. As qualitatively discussed, the tabulated results verify the superiority of ECHAM5 to CCSM3 in terms of simulating both TAR and AMR. The averaged RMSE values of ECHAM5 and CCSM3 for TAR estimation are 578 and 1081 mm, respectively. Considering mean observed annual rainfall at grids 1–3 equals to 1800 mm (not given in the table), the results indicated up to 32 and 60 % uncertainty for ECHAM5 and CCSM3, respectively. In terms of simulating AMR, uncertainty value of ECHAM5 increases up to 42 % whereas corresponding value at CCSM3 remains constant around 60 %. With respect to the averaged error measures computed for rainy days simulation, it can be pointed out that CCSM3 provides more agreement than ECHAM5 with observed rainy days climatology.

In addition to the implementation of simple arithmetic mean of grid errors (Table 4, column 7), the MPI measure was also used in this study (Table 5) to investigate the overall

Table 4 Grid-based MAB and RMSE of GCMs driven RCM simulations

Statistic	Model	Measure	Grid 2	Grid 3	Grid 4	Average
Total annual (mm)	ECHAM5	MAB	776	367	371	505
		RMSE	885	413	434	578
	CCSM3	MAB	1315	617	1110	1014
		RMSE	1466	616	1181	1081
Maximum annual (mm)	ECHAM5	MAB	32	30	25	29
		RMSE	42	27	34	34
	CCSM3	MAB	42	37	35	38
		RMSE	54	43	45	47
Rainy days	ECHAM5	MAB	31	61	68	53
		RMSE	34	63	71	56
	CCSM3	MAB	52	20	34	35
		RMSE	54	24	37	38

Table 5 Overall comparison of the climate models

Index	Statistic (v)	ECHAM5	CCSM3
e^2	TAR	12	13.4
	AMR	3.1	9.1
F^2	TAR	1.3	1.4
	AMR	0.3	1.0
MPI		0.8	1.2

performance of each model regarding the annual statistics (i.e. TAR and AMR). Here, we scaled the normalized error variance by the average error found in ensemble of GCM-RCM combinations used in this study. According to the relative difference between MPI values given in the table, ECHAM5-RegCM3 is 33 % more accurate than CCSM3-RegCM3 for annual rainfall reproduction. Referring to the average errors given in Table 4, one might claim that the superiority of the ECHAM5 implied from Table 5 is perhaps not so surprising. For this purpose, it should be mentioned that the averaging of grid-based errors to judge about the entire province is limited to statistics with same dimension whereas the MPI does not suffer from this drawback.

6 Conclusions

Catchment-scale performance evaluation of GCM-RCM combinations is required for selecting a proper set of GCM-RCM combinations to run hydrological models in small catchments. This important issue was addressed in this study to investigate precipitation realization performance of two GCM-RCM combinations, ECHAM5-RegCM3 and CCSM3-RegCM3, over ten small to mid-size catchments situated in Rize Province, northeastern part of Turkey. To this end, different daily, monthly, and annual precipitation statistics as well as two performance evaluation frameworks, grid- and portion-based, are implemented, which are of interest to a wide audience of hydrologists and practitioners.

The results show that both ECHAM5-RegCM3 and CCSM3-RegCM3 reproduced the monthly rainfall cycle rather weak. In general, the former tends to overestimate total monthly rainfall whilst the latter yields slightly dryer condition in the entire province. Consequently, CCSM3-RegCM3 underestimates observational total annual rainfall. For monthly and annual maximum rainfall statistics, ECHAM5-RegCM3 resulted in higher performance than CCSM3-RegCM3, particularly in the inland portion. Relating to the number of rainy days in different parts of the study area, ECHAM5-RegCM3 tends to overestimate the relatively constant rainy days of the entire province. By contrast, CCSM3-RegCM3 underestimates the statistic in the range 20–50 days. Since the major uncertainty of both models appeared in the coastland portion, implementation of model projections in this portion evidently requires special cares such as a bias correction procedure.

The evaluation performed here showed that no GCM-RCM combination reproduces satisfactory results for the entire province. When a combination performs well for a portion/grid of study region, it is not necessarily a well combination for the other portions/grids. Similarly, the evaluation for different variables showed that no combination was best/worst in all variables even though all of the variables are from a same

family (i.e. rainfall-born statistics). In other words, if a combination performs well for reproducing a specific statistic, it is not necessarily a well combination to reproduce the other statistics. These conclusions are consistent with those of reported by Kjellström et al. (2010) at regional-scale evaluation of GCM-RCM combinations. Therefore, it is suggested that when the interest is impact assessment at even relatively small areas, as it is the case of FRA-PFC Rize project, validation of GCM-RCM outputs should be conducted at a single-grid level, rather than at regional scales. In this case, it is also necessary to check models' performance in reproducing different prescribed observations in order to make a fair judgment among models available for the catchment of interests.

Acknowledgments This work is part of a research project supported by the Scientific and Technological Research Council of Turkey (TUBITAK) under Grant Number 112Y214. The authors thank the project research team: D. Z. Şeker, M. Özger, D. Bozkurt, O. Şen, and H. Erdem for their contribution in the study.

References

- Anandhi A, Nanjundiah RS (2015) Performance evaluation of AR4 climate models in simulating daily precipitation over the Indian region using skill scores. *Theor Appl Climatol* 119:551–566
- Barfus K, Bemhofer C (2014) Assessment of GCM performances for the Arabian Peninsula, Brazil, and Ukraine and indications of regional climate change. *Environ Earth Sci*. doi:10.1007/s12665-014-3147-3
- Bastola S, Murphy C, Sweeney J (2011) The role of hydrological modelling uncertainties in climate change impact assessments of Irish river catchments. *Adv Water Resour* 34(5):562–576
- Bozkurt D, Sen OL (2013) Climate change impacts in the Euphrates–Tigris Basin based on different model and scenario simulations. *J Hydrol* 480:149–161
- Bozkurt D, Turuncoglu U, Sen OL, Onol B, Dalfes HN (2012) Downscaled simulations of the ECHAM5, CCSM3 and HadCM3 global models for the eastern Mediterranean-Black Sea region: evaluation of the reference period. *Clim Dyn* 39:207–225
- Collins WD, Bitz CM, Blackmon ML, et al. (2006) The Community Climate System Model Version 3 (CCSM3). *J Clim* 19:2122–2143
- Danandeh Mehr A, Kahya E, Olyae E (2013) Streamflow prediction using linear genetic programming in comparison with a neuro-wavelet technique. *J Hydrol* 505:240–249
- Deidda R, Marrocu M, Caroletti G, et al. (2013) Regional climate models' performance in representing precipitation and temperature over selected Mediterranean areas. *Hydrol Earth Syst Sci* 17:5041–5059
- Demirel MC, Moradkhani H (2015) Assessing the impact of CMIP5 climate multi-modeling on estimating the precipitation seasonality and timing. *Clim Chang*. doi:10.1007/s10584-015-1559-z
- Demirel MC, Booj MJ, Hoekstra AY (2013) Impacts of climate change on the seasonality of low flows in 134 catchments in the river Rhine basin using an ensemble of bias-corrected regional climate simulations. *Hydrol Earth Syst Sci* 7(10):4241–4257
- Eris E, Agiralioğlu N (2009) Effect of coastline configuration on precipitation distribution in coastal zones. *Hydrol Process* 23(25):3610–3618
- Errasti I, Ezcurra A, Sáenz J, Ibarra-Berastegi G (2011) Validation of IPCC AR4 models over the Iberian Peninsula. *Theor Appl Climatol* 103(1–2):61–79

- Foley AM (2010) Uncertainty in regional climate modelling: A review. *Prog Phys Geogr* 34(5):647–670
- Fu G, Liu Z, Charles SP, Xu Z, Yao Z (2013) A score-based method for assessing the performance of GCMs: a case study of south-eastern Australia. *J Geophys Res - Atmos* 118:4154–4167
- Graham LP, Andréasson J, Carlsson B (2007a) Assessing climate change impacts on hydrology from an ensemble of regional climate models, model scales and linking methods—a case study on the Lule River basin. *Clim Chang* 81:293–307
- Graham LP, Hagemann S, Jaun S, Beniston M (2007b) On interpreting hydrological change from regional climate models. *Clim Chang* 81: 97–122
- Grossi G, Caronna P, Ranzi R (2013) Hydrologic vulnerability to climate change of the Mandrone glacier (Adamello-Presanella group, Italian Alps). *Adv Water Resour* 55:190–203
- Halmstad A, Najafi MR, Moradkhani H (2013) Analysis of precipitation extremes with the assessment of regional climate models over the Willamette River Basin, USA. *Hydrol Process* 27:2579–2590
- Harvey LDD, Wigley TML (2003) Characterizing and comparing control-run variability of eight coupled AOGCMs and of observations part 1: temperature. *Clim Dyn* 21:619–646
- Hofstadter R, Bidegain M (1997) Performance of general circulation models in south-eastern South America. *Clim Res* 9:101–110
- IPCC (2007) Intergovernmental Panel on Climate Change fourth assessment report on scientific aspects of climate change for researchers, students, and policymakers.
- Jacob D, Bärring L, Christensen OB, et al. (2007) An inter-comparison of regional climate models for Europe: model performance in present-day climate. *Clim Chang* 81:31–52
- Jones PD, Reid PA (2001) Assessing future changes in extreme precipitation over Britain using regional climate model integrations. *Int J Climatol* 21:1337–1356
- Kitanidis PK (1997) Introduction to geostatistics: applications to hydrogeology. Cambridge University Press, 249 pp.
- Kjellström E, Boberg F, Castro M, Christensen HJ, Nikulin G, Sánchez E (2010) Daily and monthly temperature and precipitation statistics as performance indicators for regional climate models. *Clim Res* 44: 135–150
- Kundzewicz ZW, Kanae S, Seneviratne SI, et al. (2013) Flood risk and climate change: global and regional perspectives. *Hydrol Sci J* 59(1):1–28
- Kundzewicz ZW, Mata LJ, Arnell NW, et al. (2008) The implications of projected climate change for freshwater resources and their management. *Hydrol Sci J* 53(1):3–10
- Liu Z, Mehran A, Phillips TJ, AghaKouchak A (2014) Seasonal and regional biases in CMIP5 precipitation simulations. *Clim Res* 60: 35–50
- Maxino CC, McAvaney BJ, Pitman AJ, Perkins SE (2008) Ranking the AR4 climate models over the Murray-Darling Basin using simulated maximum temperature, minimum temperature and precipitation. *Int J Climatol* 28:1097–1112
- Mondal A, Mujumdar PP (2012) On the basin-scale detection and attribution of human-induced climate change in monsoon precipitation and streamflow. *Water Resour Res* 48:W10520. doi:10.1029/2011WR011468
- Moradkhani H, Baird RG, Wherry S (2010) Impact of climate change on floodplain mapping and hydrologic ecotones. *J Hydrol (Amsterdam)* 395:264–278
- Najafi MR, Moradkhani H, Piechota TC (2011a) Statistical downscaling of precipitation using machine learning with optimal predictor selection. *J Hydrol Eng* 16(8):650–664
- Najafi MR, Moradkhani H, Piechota TC (2011b) Ensemble streamflow prediction: climate signal weighting vs. climate forecast system reanalysis. *J Hydrol* 442–443:105–116
- Nieto S, Rodríguez-Puebla C (2006) Comparison of precipitation from observed data and general circulation models over the Iberian Peninsula. *J Climate* 19(17):4254–4275
- Önol B, Semazzi FHM (2009) Regionalization of climate change simulations over Eastern Mediterranean. *J Clim* 22(8):1944–1961
- Önol B, Bozkurt D, Turuncoglu UU, et al. (2014) Evaluation of the twenty-first century RCM simulations driven by multiple GCMs over the Eastern Mediterranean–Black Sea region. *Clim Dyn* 42: 1949–1965
- Özdoğan M (2011) Climate change impacts on snow water availability in the Euphrates–Tigris basin. *Hydrol Earth Syst Sci* 15:2789–2803
- Ozkul S (2009) Assessment of climate change effects in Aegean river basins: the case of Gediz and Buyuk Menderes Basins. *Clim Change* 97:253–283
- Pal JS, Giorgi F, Bi X, et al. (2007) Regional climate modeling for the developing world: the ICTP RegCM3 and RegCM3. *Bull Am Meteorol Soc* 88(9):1395–1409
- Perkins SE, Pitman AJ, Holbrook NJ, McAneney J (2007) Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *J Clim* 20:4356–4376
- Reichler T, Kim J (2008) How well do coupled models simulate today's climate. *Bull Am Meteorol Soc* 89:303–311
- Roeckner E, Bäuml G, Bonaventura L, et al. (2003) The atmospheric general circulation model ECHAM5. Part I: model description. *Max Planck Inst Meteorol Rep* 349:127
- Roosmalen LV, Christensen JH, Butts MB, et al. (2010) An intercomparison of regional climate model data for hydrological impact studies in Denmark. *J Hydrol* 380:406–419
- Şen ÖL (2013) A holistic view of climate change and its impacts in Turkey. Report. Istanbul Policy Centre, Sabanci University, Istanbul
- Shen Y, Oki T, Kanae S, et al. (2014) Projection of future world water resources under SRES scenarios: an integrated assessment. *Hydrol Sci J* 59(10):1775–1793
- Smith I, Chandler E (2010) Refining rainfall projections for the Murray Darling Basin of south-east Australia—the effect of sampling model results based on performance. *Clim Chang* 102:377–393
- Van Der Linden P, Mitchell JFB (Eds) (2009) ENSEMBLES: climate change and its impacts: summary of research and results from the ENSEMBLES Project. Met Office Hadley Centre, Fitz Roy Road, Exeter EX1 3 PB, UK, 160.
- Van Vliet MTH, Blenkinsop S, Burton A, et al. (2012) A multi-model ensemble of downscaled spatial climate change scenarios for the Dommel catchment, Western Europe. *Clim Chang* 111:249–277
- Varis O, Kajander T, Lemmälä R (2004) Climate and water: from climate models to water resources management and vice versa. *Clim Chang* 66:321–344
- Vaze J, Teng J, Chiew FHS (2011) Assessment of GCM simulations of annual and seasonal rainfall and daily rainfall distribution across south-east Australia. *Hydrol Process*. doi:10.1002/hyp.7916
- Xuejie G, Zongci Z, Giorgi F (2002) Changes of extreme events in regional climate simulations over East Asia. *Adv Atmos Sci* 19(5): 927–942