Hydrology and
Earth System
Sciences
Discussions

Open Access

# Climate model validation and selection for hydrological applications in representative Mediterranean catchments

R. Deidda[1,3], M. Marrocu[2], G. Caroletti[1,3,6], G. Pusceddu[2], A. Langousis[7], V. Lucarini[1,5], M. Puliga[1,3], and A. Speranza[1,4]

[1]CINFAI, Consorzio Interuniversitario Nazionale per la Fisica delle Atmosfere e delle Idrosfere, Camerino (MC), Italy
[2]CRS4, Centro di Ricerca, Sviluppo e Studi Superiori in Sardegna, Loc. Piscina Manna, Edificio 1, 09010 Pula (CA), Italy
[3]Dipartimento di Ingegneria Civile, Ambientale e Architettura, Università di Cagliari, Piazza d'Armi, 09123 Cagliari, Italy
[4]Dipartimento di Fisica, Università di Camerino, Via Madonna delle Carceri 9, Camerino (MC), Italy
[5]Klimacampus, Institute of Meteorology, University of Hamburg, Germany
[6]Geophysical Institute, University of Bergen, Norway
[7]Department of Civil Engineering, University of Patras, Patras, Greece

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄◄ | ►►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Representative Mediterranean catchments**

R. Deidda et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◀ | ▶

◀ | ▶

Back | Close

Full Screen / Esc

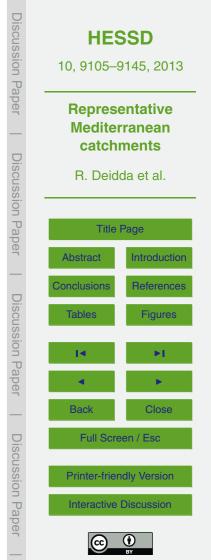Printer-friendly Version

Interactive Discussion

## Abstract

This paper discusses the relative performance of several climate models in providing reliable forcing for hydrological modeling in six representative catchments in the Mediterranean region. We consider 14 Regional Climate Models (RCMs), from the EU-FP6 ENSEMBLES project, run for the A1B emission scenario on a common 0.22-degree (about 24 km) rotated grid over Europe and the Mediterranean. In the validation period (1951 to 2010) we consider daily precipitation and surface temperatures from the E-OBS dataset, available from the ENSEMBLES project and the data providers in the ECA&D project. Our primary objective is to rank the 14 RCMs for each catchment and select the four best performing ones to use as common forcing for hydrological models in the six Mediterranean basins considered in the EU-FP7 CLIMB project. Using a common suite of 4 RCMs for all studied catchments reduces the (epistemic) uncertainty when evaluating trends and climate change impacts in the XXI century. We present and discuss the validation setting, as well as the obtained results and, to some detail, the difficulties we experienced when processing the data. In doing so we also provide useful information and hint for an audience of researchers not directly involved in climate modeling, but interested in the use of climate model outputs for hydrological modeling and, more in general, climate change impact studies in the Mediterranean.

## 1 Introduction

Climate Models (CMs) are numerical tools to simulate the past, present and future of Earth's climate. Hence, evaluating the accuracy of CMs is a crucial scientific and applicative objective, not only for the role of models in reconstructing the past and projecting the future state of the planet, but also because of their increasing relevance in the process of policymaking. For the latter purpose, it is necessary to summarize and evaluate the results originating from an increasing number of Global Climate Models (GCMs), providing climate projections over the whole planet. A common practice is

to build multi-model ensembles and study their statistics (mainly ensemble mean and spread). Note, however, that a multi-model ensemble is not statistically homogeneous (i.e. formed by statistically equivalent realizations of a process) and, therefore, using its mean to approximate the truth and its standard deviation to describe the uncertainty of the outputs, could be highly misleading (Lucarini, 2008; Annan et al., 2011).

In general GCMs are suited to provide large-scale climate predictions, not directly relevant to hydrological evaluations at a river basin level, which can be of interest for local policymaking. In order to refine GCM outputs, the most common approach is to use statistical and dynamical downscaling tools. Regional Climate Models (RCMs) are high-resolution dynamical models that take advantage of detailed representations of natural processes at high spatial resolutions capable of resolving complex topographies and land-sea contrast. However, they are run on a limited domain and thus require boundary conditions from a driving GCM (e.g., Giorgi and Mearns, 1999; Wang et al., 2004; Rummukainen, 2010). Thus, (i) the GCM-nested nature of regional climate modeling implies that RCM climate reconstructions and projections can critically depend on the driving GCM (e.g., Christensen et al., 1997; Takle et al., 1999; Lucarini et al., 2007); and (ii) precipitation and other atmospheric quantities, like temperature, although useful in improving climate modeling by highlighting and explaining differences in GCM parameterization and representation of climate features, they can hardly be considered climate state variables (Lucarini, 2008). The definition of reasonable climate scenarios has been an issue of major research efforts by the international scientific community (e.g., Lucarini, 2002; Fowler et al., 2007; Räisänen, 2007; Moss et al., 2010).
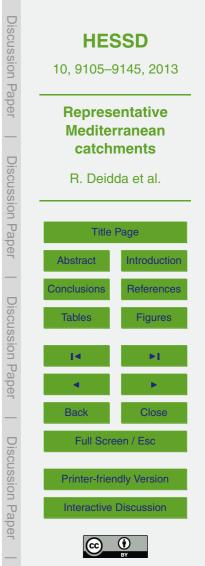
A key role in CMs is played by the atmospheric part of the hydrological cycle, not only because of its strong impact on the energy of atmospheric perturbations, but also because of the contribution of hydrometeors to human activities and the evolution of the environment. These contributions range from space-time availability of water resources and land policy, to extreme events like mudslides, avalanches, flash floods and droughts (e.g., Becker and Grünewald, 2003; Roe, 2005; Tsanis et al., 2011; Koutroulis et al., 2013; Muerth et al., 2013). The significant impact of the hydrological cycle to

human communities and the environment is also reflected by the number of studies focusing on the use of CM results for hydrological evaluations and assessments (e.g., Sulis et al., 2011, 2012; van Pelt et al., 2012; Guyennon et al., 2013; Cane et al., 2013; Velázquez et al., 2013). The relevance of this subject has also led many investigations towards ranking and validating CM performances based on hydrological measures.

Intercomparison studies have shown that no particular model is best for all variables and/or regions (e.g., Lambert and Boer, 2001; Gleckler et al., 2008). Most intercomparison and validation studies focus on evaluating hydrologically relevant parameters like temperature, precipitation, and surface pressure (e.g., Perkins et al., 2007; Giorgi and Mearns, 2002). Murphy et al. (2004) evaluated the skill of a 53-model ensemble in simulating 32 variables (from precipitation to cloud cover and upper-level pressures) to determine a Climate Prediction Index (CPI) that could provide an overall model weighting. Based on Murphy et al. (2004) CPI, Wilby and Harris (2006) evaluated climate models used in hydrological applications to create an impact-relevant CPI. The latter was based on the average bias of effective summer rainfall, which was found to be the most important predictor of annual low flows in the basin studied. Perkins et al. (2007) ranked 14 climate models based on their skill to reproduce simultaneously the probability density functions of observed precipitation, and maximum and minimum temperatures over 12 regions in Australia. Using results from the Coupled Model Intercomparison Project (CMIP3), Gleckler et al. (2008) ranked climate models by averaging the relative errors over 26 variables (precipitation, zonal and meridional winds at the surface and different pressure levels, 2 m temperature and humidity, top-of-the-atmosphere radiation fields, total cloud cover etc.). They also showed that defining a single index of model performance can be misleading, since it shades a more complex picture of the relative merits of different models. Johnson and Sharma (2009) derived the VCS (Variable Convergence Score) skill score to compare the relative performance of a total of 21 model runs from nine GCMs and two different emission scenarios in Australia, to their ensemble mean. They applied the VCS score to eight different variables and found that pressure, temperature, and humidity received the highest scores.

**Representative Mediterranean catchments**

R. Deidda et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Nevertheless, it is worth mentioning that CM results are currently tested only against observational (past) data, and the choice of the observables of interest is crucial for determining robust metrics able to audit effectively the models (Lucarini, 2008; Wilby, 2010). Unfortunately, data are often of nonuniform quality and quantity, due to e.g., the non stationarity and non homogeneity caused by changes in the network density, instrumentation, temporal sampling, and data collection strategies over time.

Recently, the detailed investigation of the behavior of CMs has been greatly facilitated by research initiatives aiming at providing open access outputs of simulations through projects like PRUDENCE (for RCMs; http://prudence.dmi.dk/), PCMDI/CMIP3 (GCMs included in the IPCC4AR; http://www-pcmdi.llnl.gov), ENSEMBLES (for RCMs; http://ensembles-eu.metoffice.com) and STARDEX (for RCMs; http://www.cru.uea.ac.uk/projects/stardex/).

In the framework of ENSEMBLES project, there has been an effort to produce values for some of the hydrologically relevant variables (i.e. precipitation, temperature and sea level pressure), on a regular data grid, based on objective interpolation of the observational network. This initiative has continued as a part of the URO4M (EU-FP7) project, which made the observed data fields (E-OBS) available on different grids for the 1950–2011 time-frame (Haylock et al., 2008; van den Besselaar et al., 2011). Recently, the E-OBS fields were newly released on a rotated grid consistent with that used by ENSEMBLES RCMs over western Europe. Although limited (due to the non-uniform spatial density of the observations used to produce the gridded data), the E-OBS fields constitute a reference for evaluating the performance of different CMs in the European and Mediterranean areas; from a technical point of view, they are built for direct comparisons with ENSEMBLES RCM outputs.

Following these recent initiatives, the European Union has funded the Climate Induced Changes on the Hydrology of Mediterranean Basins project (CLIMB; http://www.climb-fp7.eu), with the aim of producing a future-scenario assessment of climate change for significant hydrological basins of the Mediterranean (Ludwig et al., 2010), including: the Noce and Riumannu river basins in Italy, the Thau coastal lagoon

in France, Izmit bay in the Kokaeli region of Turkey, the Chiba river basin in Tunisia, and the Gaza aquifer in Palestine. The Mediterranean lands constitute an especially interesting area for hydrological investigation by climate scientists, given the high risk predicted by climate scenarios, and the pronounced susceptibility to droughts, extreme
5  flooding, salinization of coastal aquifers and desertification, as a consequence of the expected reduction of yearly precipitation and increase of the mean annual temperature.

The general goal of CLIMB project is to reduce the uncertainty of the process of assessing climate change impacts in the considered catchments. A major source of
10  uncertainty is certainly related to the wide scattering of climate model outputs. That said, our investigation focuses on reducing the uncertainty of climate model forcing, by intercomparing the performances of different RCM results from the ENSEMBLES project, and selecting a common subset of 4 models to drive hydrological model runs in the catchments. More precisely, this paper uses the newly released E-OBS fields,
15  to: a) evaluate the performance of ENSEMBLES RCMs in dealing with hydrologically relevant parameters in six Mediterranean catchments, and b) provide validated data to be used for hydrological modeling in successive steps of the CLIMB project.

Section 2 introduces CLIMB project in the context of the hydrological basins of interest, and Sect. 3 provides detailed information on the RCM datasets used. Section 4
20  describes the methods applied to audit ENSEMBLES past climate simulations, and Sect. 5 presents the obtained results, setting them in the context of previous research. Finally, Sect. 6 summarizes the main conclusions of this study.

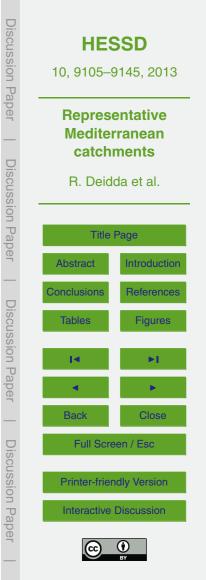## 2   The CLIMB project and the target catchments

As noted in the Introduction, results presented in this paper were produced in the
25  framework of the CLIMB project, aiming at selecting the most accurate ENSEMBLES Regional Climate Models (RCMs) to drive hydrological model runs in six (6) significant Mediterranean catchments. Figure 1 shows the location of the catchments, and Ta-

ble 1 summarizes their main characteristics. The areas of the catchments range from 250 to 3500 km$^2$. Since the horizontal resolution of all ENSEMBLES RCM outputs is approximately 24 km, all catchments can be embedded within a 4 × 4 stencil of model grid-points. From Table 1 one sees that the catchments differ in terms of their over-
all climatic characteristics, ranging from semi-arid (Gaza), Mediterranean (Chiba, Riu Mannu, Thau and Kokaeli), to humid continental (Noce) and, thus, they can be considered representative of the Mediterranean area. For a given climate model, the skill in reproducing accurately the local climatic features can be quite different, and the relative (and absolute) skill can vary considerably within the selected ensemble of climate
model results (see Sect. 5). Since our goal is to account, as much as possible, for all uncertainties related to the use of different climate models in different catchments, the validation phase described in this paper aims at selecting a common subset of four (4) CMs to drive hydrological model simulations in the considered catchments. In selecting the four best performing GCM-RCM combinations, we considered the additional con-
straint of maintaining at least two different RCMs nested in the same GCM, and two different GCMs forcing the same RCM. While the added constraint does not guarantee selection of the four best overall performing GCM-RCM combinations in each individual catchment, it allows for diversity of the selected model results in a uniform setting for whole catchments. In a separate effort, not described in this manuscript, we studied
the downscaled hydrologically relevant fields of the selected GCM-RCM combinations (as discussed in Sect. 4). Those fields account for the small scale variability associated with local topographic features and orographic constraints, crucial for hydrological modeling. While the obtained results will form the subject of an upcoming communication, for completeness, in Sect. 4 we summarize those findings crucial in understanding
the constraints imposed by our validation setting.

## 3 Climate models and reference dataset

The intercomparison and validation of CM results were performed for a subset of 14 Regional Climate Models (RCMs) from the ENSEMBLES project, run for the A1B emission scenario at 0.22 degree resolution.
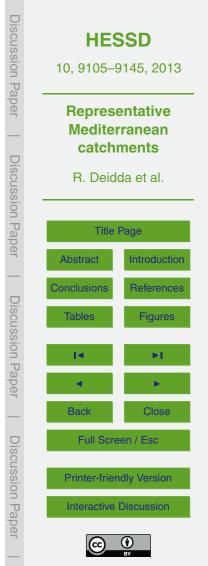
The choice of ENSEMBLES RCMs is particularly appealing due to the available standardizations: (i) all simulations were run on a common rotated grid of 0.22 degrees (this corresponds to a grid resolution of approximately 24 km at mid-latitudes), assuring an almost perfect overlap of common grid-points; (ii) almost all models cover the 150-yr time-frame from 1951 to 2100 at a daily level. The ENSEMBLES high resolution RCM runs are based on a standard portfolio of climate scenarios described in the Fourth Assessment Report of the International Panel for Climate Change (IPCC4AR; see for instance Solomon et al., 2007). The most complete dataset is given for the A1B scenario, which is considered as the most realistic.

In ENSEMBLES high resolution runs, each RCM is nested into a larger-scale field. The latter may originate from different GCMs, leading to different GCM-RCM combinations. For simplicity, we defined an acronym for each GCM and RCM considered in this study, listed in Tables 2 and 3, respectively. In all figures, we use symbols to display results from different RCMs, and colours to indicate runs forced by different GCMs. Figure 2 summarizes the combinations of symbols and colours used to display results from the 14 GCM-RCM combinations considered in this study.

Following the PCMDI/CMIP3 initiative, the ENSEMBLES project stimulated and guided several climatic centres towards standardization of model grids and outputs and, thus, promoted synergies across different research areas and interdisciplinary efforts. Nevertheless, when pre-processing the ENSEMBLES RCM outputs (i.e. before the validation phase), we experienced some minor discrepancies, requiring ad hoc treatments. Although all issues were manageable, they are worth mentioning, especially for scientists who are (or foresee) using these outputs to run hydrological models and perform impact studies:

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

– For most GCM-RCM combinations, the available time-frames cover the period from 1 January 1951 to 31 December 2100. However, some model outputs exhibit missing data on the last days of 2100, whereas for other models the missing values start at the end of 2099. In addition, two model simulations (i.e. BCM-HIR, HCS-HIR) stop in year 2050, while the BCM-RCA model run starts in 1961.

– Models HCH-RCA, HCS-CLM, HCS-HRM, HCL-HRM, HCH-HRM, HCS-HIR, and HCL-RCA use a simplified calendar with 12 months of 30 days each (i.e. 360 days per year), whereas the remainder use a standard Gregorian calendar with 365 days per regular year, and 366 days in leap years. Additionally, among these last models, some do not account for the leap year exception in 2100. We also detected some missing or incomplete data. More precisely, for some models the values in the last days of the simulation period are simply set to zero, rather than to an unambiguous default flag for missing values. While this is not an issue when working with temperature fields expressed in K, it may create problems when considering precipitation fields (expressed in $Kg\,m^{-2}\,s^{-1}$), since it is not apparent how to distinguish between missing data and zero precipitation (this is the case, e.g., for the last 390 days of data in the HCL-RCA simulation).

– For some models (see below), dry conditions are indicated by a very small positive or negative constant value, $P_{min}$, whereas the sea level elevation is set to a constant value, $z_{sea}$, different from zero. More precisely, $z_{sea} \approx 0.046\,m$ for HCH-RCA; $z_{sea} \approx 0.300\,m$ for BCM-HIR and HCS-HIR; $z_{sea} \approx 0.732\,m$ and $P_{min} \approx -9.0 \times 10^{-8}\,Kg\,m^{-2}\,s^{-1}$ for ECH-RMO; $z_{sea} \approx -0.002\,m$ and $P_{min} \approx 1.7 \times 10^{-18}\,Kg\,m^{-2}\,s^{-1}$ for ARP-HIR and ECH-HIR; $z_{sea} \approx -0.321\,m$ and $P_{min} \approx -1.5 \times 10^{-11}\,Kg\,m^{-2}\,s^{-1}$ for BCM-RCA, ECH-RCA and HCL-RCA. Also, for the last three models, missing temperature data at the end of the simulation period are indicated by a minimum temperature of 0 K, whereas HCL-HRM simulations exhibit some temperature values on the order of $10^{25}$ K. While the origin of the aforementioned discrepancies in the data cannot be easily identified (e.g. numerical errors, spurious effects of

model parameterizations, routines used to create the netcdf files in the ENSEM-BLES archive, etc.), one should properly treat them before using CM outputs to perform climate change impact studies. For example, unless properly identified, a minimum precipitation threshold may bias rainfall statistics (e.g. the annual frac-
⁵ tion of dry periods) and, from a practical point of view, influence hydrological and meteorological analysis (e.g. drought analysis).

For each considered catchment, the selected set of climate model data was validated using the CRU E-OBS dataset from the ENSEMBLES EU-FP6 project, made available by the ECA&D project (http://www.ecad.eu) and hosted by the Climate Research Unit
¹⁰ (CRU) of the Hadley Centre. E-OBS datafiles are gridded observational datasets of daily precipitation and temperature, developed on the basis of a European network of high-quality historical measurements. In particular, we used version 5.0 of the E-OBS dataset that covers the period from 1 January 1950 to 30 June 2011, and is available at four different grid resolutions: 0.25 and 0.5 degree regular latitude-longitude grids, and
¹⁵ 0.22 and 0.44 degree rotated pole grid. In our analysis, we use the rotated grid at 0.22 degree resolution, which matches exactly the grid of ENSEMBLES RCMs. Having an almost perfect point-to-point correspondence between ENSEMBLES RCM results and E-OBS reference data, greatly simplifies intercomparison, validation and calibration activities, since no interpolation or re-gridding of the data is needed.
²⁰ There are several additional reasons why the E-OBS dataset is considered to be the best available source for temperature and precipitation estimates in the considered catchments to pursue model validation: (1) E-OBS data have been obtained through kriging interpolation, which belongs to the class of best linear unbiased estimators (BLUE); (2) the original data (i.e. prior to interpolation) have been properly corrected to
²⁵ minimize biases introduced by local effects and orography; (3) the 95 % confidence intervals of the obtained estimates are also distributed, shedding light on the accuracy of the calculated areal averages, and (4) the surface elevation fields used for E-OBS interpolations are available as well. The latter can be used to assess which counterpart of the observed differences between ENSEMBLES RCM and E-OBS climatologies can

be attributed to different orographic representations and, also, to remove biases introduced by elevation differences in the E-OBS and corresponding RCM grid points. For example, to account for different surface elevation models used by ENSEMBLES RCMs and E-OBS, and before calculating CM performance metrics, we used the corresponding model orographies and a monthly constant lapse rate to translate surface temperatures at different elevations to those observed at sea level, as discussed in Sect. 4 below. For a more detailed description of the E-OBS dataset, the reader is referred to Haylock et al. (2008).

## 4  Metrics for validation

Several studies (see Introduction) have focused on metrics to assess the accuracy of climate model results. Note, however, that performance metrics should depend on the specific use of climate data. In this study, we focus on providing reliable climatic forcing for hydrological applications at a river basin level. For this purpose, one needs to downscale the CM fields at resolutions suitable to run hydrological models and assess climate change impacts. In what follows, we summarize the main aspects of the downscaling procedure in order to make more clear our validation setting. A detailed description of the downscaling procedures, together with obtained results, will form the subjects of an upcoming communication.

Precipitation downscaling was performed using the multifractal approach described in Deidda (2000) and Badas et al. (2006), starting from areal averages of daily precipitation. The latter were obtained by averaging rainfall values over a 4 × 4 stencil of ENSEMBLES grid-points centred in each catchment, covering an approximate area of 100 km × 100 km. This particular choice allowed to embed all catchments inside equally sized spatial domains. The size of the latter is within the range of space-time scale invariance of rainfall indicated by several studies (Schertzer and Lovejoy, 1987; Tessier et al., 1993; Perica and Foufoula-Georgiou, 1996; Venugopal et al., 1999; Deidda, 2000; Kundu and Bell, 2003; Gebremichael and Krajewski, 2004; Deidda et al.,
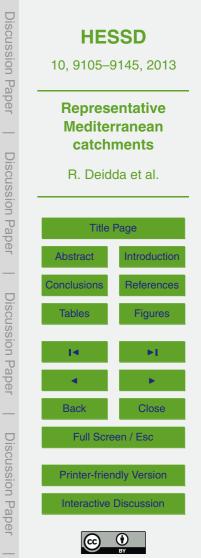
2004, 2006; Gebremichael et al., 2006; Badas et al., 2006; Veneziano et al., 2006; Veneziano and Langousis, 2005, 2010) and, thus, it can be used to define the integral volume of rainfall to be downscaled to higher resolutions of a few square kilometers. That said, the validation metrics of ENSEMBLES RCM simulations versus E-OBS data are calculated based on areal rainfall averages over a regular 4 × 4 grid-point stencil.

Temperature fields from ENSEMBLES RCMs were downscaled based on the procedure described in Liston and Elder (2006), which combines a spatial interpolation scheme (Barnes, 1964, 1973) with orographic corrections. Although, in the case of temperature, downscaling starts from the ENSEMBLES resolution (24 km × 24 km), we decided to adopt the same validation setting as that for precipitation (i.e. calculate areal averages of daily temperatures over a regular 4 × 4 grid-point stencil). Since temperature fields are particularly sensitive to elevation, in order to make homogeneous comparisons, we first reduced surface temperatures at different elevations to sea level, and then calculated areal averages of daily temperatures over a regular 4 × 4 grid-point stencil centered in each catchment. For the former, we used a standard monthly lapse rate for the Northern Hemisphere (Kunkel, 1989) and the corresponding ENSEMBLES model and E-OBS orographies. This appears to be a reasonable choice since, after reduction to sea level, the temperature field becomes quite smooth.

In summary, all metrics defined below are used to validate ENSEMBLES RCM simulations vs E-OBS data, using spatial averages of temperature and precipitation over a 4 × 4 grid-point stencil. This choice allows to better study climatic forcing at a common catchment scale. An additional advantage is that spatial averaging smooths and filters out some local biases present in E-OBS fields. These originate from the low density of observations, which cannot efficiently capture orographic effects on precipitation and temperature and, also, cannot account for the high spatial variability of daily rainfall.

It is worth mentioning that E-OBS dataset is based on observations obtained from a network of land-based stations. Hence, all E-OBS data over sea were masked and set to default missing values. Nevertheless, missing values were also found at some grid points over land, where the network density is low. An analysis of the data density

over time showed that there are E-OBS grid points over land where the availability of interpolated data depends on the period studied. This is due to changes in the observation network between 1950 and 2011. To account for this issue, when calculating areal averages of daily values over the 4 × 4 grid-point stencil, we maintained those points with less than 6 yr of missing data (i.e. 10 % of the 60 yr validation period from 1951 to 2010). To that extent, all validation metrics below were calculated using those grid points available for both E-OBS data and ENSEMBLES RCM simulations.

Following the discussion above, let us define $X_m(s, y)$ to be the monthly spatial average of variable $X$ (i.e. $X = T, P$ for monthly temperatures and rainfall intensities, respectively) over an area of approximately $100 \times 100$ km (i.e. a 4 × 4 stencil of ENSEMBLES climate model outputs), produced by climate model $m$ ($m = 1, \cdots, 14$) for month $s$ ($s = 1, \cdots, 12$) in year $y$ ($y = 1951, 1952, \cdots, 2100$). According to this notation, let us also denote by $m = 0$ the index for E-OBS. The $s$-th monthly mean and standard deviation of $X_m(s, y)$ over a $N_y$ yr climatological period starting in year $y_0$ are given by:

$$\mu_m^X(s) = \frac{1}{N_y} \sum_{y=y_0}^{y_0+N_y-1} X_m(s, y) \tag{1}$$

$$\sigma_m^X(s) = \sqrt{\frac{1}{N_y - 1} \sum_{y=y_0}^{y_0+N_y-1} \left[ X_m(s, y) - \mu_m^X(s) \right]^2} \tag{2}$$

The time-window for validation is set to 60 yr from 1951 to 2010, entirely covered by E-OBS data. Validation of climate model outputs over that period requires comparing specific statistics of $X_m$ ($m = 1, \cdots, 14$) to those of E-OBS (i.e. $X_{m=0}$). To do so we introduce average error measures for the absolute differences between statistics of the observed ($m = 0$) and modelled ($m = 1, \cdots, 14$) time series. Setting $y_0 = 1951$ and

HESSD

10, 9105–9145, 2013

Representative Mediterranean catchments

R. Deidda et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

|◄ | ►|

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

$N_y = 60$ in Eqs. (1) and (2), such error measures for the monthly climatological means and standard deviations of $X_m$, over the 60 yr period 1951–2010 are defined as:

$$E\mu_m^X = \frac{1}{12} \sum_{s=1}^{12} |\mu_m^X(s) - \mu_0^X(s)| \tag{3}$$

$$E\sigma_m^X = \frac{1}{12} \sum_{s=1}^{12} |\sigma_m^X(s) - \sigma_0^X(s)| \tag{4}$$

In addition, errors in the marginal distribution of $X_m$ can be quantified by averaging the absolute differences between the quantiles $x_m(\alpha_i)$ of the observed ($m = 0$) and simulated ($m = 1, \cdots, 14$) series at different probability levels $\alpha_i$ ($i = 1, \cdots, n$):

$$Eq_m^X = \frac{1}{n} \sum_{i=1}^{n} |x_m(\alpha_i) - x_0(\alpha_i)| \tag{5}$$

The above defined error metrics provide information on the reliability of a single variable, whereas for subsequent hydrological modeling we need to identify those models that perform best for a specific set of variables. As a minimum requirement for hydrological modeling, we seek for CMs that provide reliable estimates of precipitation and temperature: precipitation is the main source of water in the catchment, whereas temperature controls evaporation and evapotranspiration processes. Thus, in order to compare the performances of different models $m$ ($m = 1, \ldots, 14$) in reproducing, simultaneously, the statistics of the observed precipitation and temperature fields, we introduce the following dimensionless measures:

$$\epsilon\mu_m = \frac{1}{2} \frac{E\mu_m^P}{\sum_{k=1}^{M} E\mu_k^P} + \frac{1}{2} \frac{E\mu_m^T}{\sum_{k=1}^{M} E\mu_k^T} \tag{6}$$

$$\epsilon\sigma_m = \frac{1}{2}\frac{E\,\sigma_m^P}{\sum_{k=1}^{M}E\,\sigma_k^P} + \frac{1}{2}\frac{E\,\sigma_m^T}{\sum_{k=1}^{M}E\,\sigma_k^T} \qquad (7)$$

where $M = 14$ is the number of models.

Equations (6) and (7) can be used to assess the relative performance of different models in reproducing the monthly mean and standard deviation of the observed temperature and precipitation series.

## 5 Results and discussion

### 5.1 Seasonal distribution of precipitation and temperature

When assessing models' skills, it is important to examine their ability to reproduce the annual averages and seasonal variability of precipitation and surface temperature. These two variables specify the climate type according to the Köppen–Geiger climate classification system. Thus, a first skill measure is the ability of the models to reproduce the local climatic characteristics of specific basins.

Figures 3 and 4 show the mean monthly precipitation and temperature for each catchment, calculated using Eq. (1), for the $N_y = 60\,\text{yr}$ verification period (1951–2010). For the E-OBS observational reference, we use a dotted black line, whereas ENSEM-BLES RCM results are plotted using the reference introduced in Fig. 2. Based on the E-OBS seasonal variation of precipitation and temperature, Thau, Riu Mannu, Kokaeli and Chiba are characterized by a Mediterranean climate, with precipitation maxima in winter, and minima in summer. Gaza, although exhibiting a seasonal cycle of precipitation similar to the previous basins, is classified as semi-arid due to its low annual precipitation and high mean annual temperature. Noce, instead, has a humid continental climate, receiving regular precipitation during the year, with a single maximum during summer. In essence, although based on a sparse network, E-OBS data repro-

duce quite reasonably the expected seasonal climatological patterns in the considered catchments.

In comparing model results on precipitation with E-OBS observational data, Fig. 3 clearly shows significant discrepancies between ENSEMBLES RCMs and E-OBS climatologies, both in terms of magnitude and, in some cases, the observed seasonal cycle. In almost all catchments, several RCMs produce higher annual precipitation than that observed. More in detail: for Thau, Riu Mannu and Noce, 11 out of 14 models simulate higher precipitation amounts for at least ten months of the year. The same problem is detected for 9 RCMs in Kokaeli and 8 RCMs in Chiba catchments. An additional observation is that, in all catchments except Gaza, RCMs exhibit a larger relative error for high precipitation amounts in summer months. This positive bias is amplified in catchments with Mediterranean climate, where RCM precipitation can be up to ten times higher than the observed one. In Gaza, on the other hand, models are typically biased towards lower values of precipitation, with the exception of a handful of models, which are instead biased towards larger precipitation amounts. Note, also, that one model, BCM-RCA, produces unrealistic results.

A catchment where models' skills are problematic is Riu Mannu: while E-OBS indicates almost no precipitation during the summer months from June-August, some models (HCL-RCA, BCM-HIR, BCM-RCA, HCH-HRM) display relatively high amounts of summer precipitation. Although the skill of RCMs in reproducing seasonal precipitation is generally weaker during summer, especially for the drier southern and eastern European regions (Frei et al., 2006; Maraun et al., 2010), this finding is still surprising since several studies indicate that RCMs tend to underestimate precipitation during summer (e.g., Jacob et al., 2007).

Figure 3 also shows that, despite the aforementioned biases, almost all models predict a reliable seasonal precipitation cycle for Thau, Riu Mannu, Kokaeli, Chiba and Gaza catchments. On the contrary, in Noce catchment, RCMs fail to capture the seasonal cycle of precipitation: instead of a single maximum during summer, most models display a bimodal behaviour, with one maximum before the summer season and one

Title Page

Abstract          Introduction

Conclusions          References

Tables          Figures

◀          ▶

◀          ▶

Back          Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

directly following it. This pattern indicates a humid subtropical climate, typical of most lowlands in Adriatic Italy, south of the Noce region. A possible reason is that, at 24 km resolution, RCMs are limited in resolving small-scale features necessary to correctly reproduce orographic precipitation. Moreover, the observed differences between CM

5 results and E-OBS observational data might stem from the irregular and sparse network of E-OBS stations, especially in view of the fact that most stations are located at low altitudes.

Unsurprisingly, RCMs tend to perform better in modeling surface temperature. Figure 4 shows that all RCMs reproduce a reliable yearly cycle of monthly climatological

10 temperatures (previously reduced to sea level, see Sect. 4). Nevertheless, when compared to E-OBS, several RCMs exhibit significant biases, often larger than 5 K, while the inter-model spread can be as large as 10 K (especially during summer). More in detail: for Noce, HCL-RCA exhibits a negative bias larger than 5 K during winter; for Kokaeli, ECH-HIR exhibits a positive bias of about 4 K during winter; whereas for Chiba,

15 HCL-HRM overestimates temperature during summer by 5 K. However, given the good representation of the seasonal temperature cycle, these discrepancies can be properly reduced using simple bias-correction techniques.

## 5.2 Selection of best performing ENSEMBLES models

Error metrics introduced in Eqs. (3) and (4) were calculated for each ENSEMBLES

20 RCM for the 60 yr verification period from 1951 to 2011, and then plotted for each catchment in a separate scatterplot (Fig. 5). One sees that, in all catchments except Noce, most models (i.e. 11 out of 14 models in Thau basin; 13 models in Riu Mannu and Kocaeli catchments, and all 14 models in Gaza and Chiba) exhibit errors lower than 1.5 mm d$^{-1}$ in their monthly means, $E\mu^P$, and lower than 1 mm d$^{-1}$ in their monthly

25 standard deviations, $E\sigma^P$. For Noce basin, the corresponding errors are significantly larger.

Concerning the temperature error metrics $E\mu^T$ and $E\sigma^T$ in Fig. 6, one notes that the error in the monthly means is much larger than the one in the standard deviations:

for almost all models and most catchments (i.e. all models in Thau, Riu Mannu and Chiba basins, 13 out of 14 models in Kokaeli catchment, and 12 out of 14 models in Noce and Gaza), $E\mu^T$ and $E\sigma^T$ are smaller than 3 K and 0.5 K, respectively. Again, one concludes that Noce is the catchment where ENSEMBLES RCMs tend to perform worse even for temperature: only 4 out of 14 models exhibit errors smaller than 1.5 K in $E\mu^T$. Note that in Gaza (the second less performing catchment), 8 out of 14 RCMs exhibit errors smaller than 1.5 K in $E\mu^T$. Concerning Noce basin, it is worth noting that one cannot easily conclude to what extent the calculated temperature and precipitation errors originate from limitations of the observation network at high elevations in the Alps.

Figures 5 and 6 also show that model performances can vary significantly from one catchment to another. For instance, HCS-CLM is one of the two worst performers for precipitation in Thau, but it performs best in Noce (see Fig. 5). For all models, precipitation errors in Gaza are quite small (see Fig. 5), while temperature errors are significant (see Fig. 6). This means that, as expected, among all considered RCMs it is not possible to identify a subset of models that performs best for both variables in all catchments. An option is to base model selection on dimensionless metrics capable of weighting the errors in the variables of interest, even when the choice is driven by additional constraints as discussed in Sect. 2. With this aim, we introduced the dimensionless error metrics in Eqs. (6) and (7) to account for both precipitation and temperature RCM performances in reproducing E-OBS observational data. Results are presented in Fig. 7, where one sees that the selection of HCH-RCA, ECH-RCA, ECH-RMO, ECH-REM models (marked with an additional black circle) are the best choices for Thau, Riumannu, Kokaeli and Chiba catchments, under the additional constraint of maintaining two different RCMs nested in the same GCM, and two different GCMs forcing the same RCM. For Noce and Gaza, where other choices would have been slightly better, the selected 4 models still display good performances.

In order to check the general behaviour of the probability distributions of the simulated precipitation and temperature fields against E-OBS, we used Eq. (5) to compute

**Representative Mediterranean catchments**

R. Deidda et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

|◄ | ►|

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version
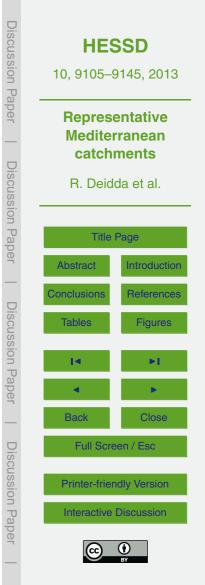
Interactive Discussion

the mean absolute errors in the quantiles at 100 uniformly spaced probability levels. The results are presented in the scatterplots of Fig. 8. One sees that the 4 selected models display reasonable performances in all considered catchments, thus confirming the selection. To assist the reader in identifying the 4 selected models in all figures,

5  we have drawn thicker the corresponding lines in Figs. 3 and 4, and added black circles in all scatterplots (Figs. 5–8).

Last but not least, it is interesting to analyze and intercompare the variability of the mean annual precipitation and temperature over the five 30 yr climatological periods between 1951 and 2100. In Fig. 9, where results for precipitation are presented, one

10  clearly observes a very large variability in the simulations of the 14 models, with some models predicting mean annual precipitation three times higher than others. One also gets an idea of how drastically the behaviour of a single model can change in different catchments. These findings support the need for extensive analyses, as that presented here, before proceeding with hydrological modeling in specific catchments. For exam-

15  ple, HEC-HIR model gives the largest annual precipitation in Thau and Noce, greatly overestimating E-OBS, while the same model in Riumannu and Gaza gives a very small annual precipitation, greatly underestimating E-OBS. The five 30 yr climatologies for the 4 selected models are shown in Fig. 9 inside vertical rectangles. Again, one visually observes that the selection of the 4 models is a reasonable compromise for all

20  catchments.

An additional observation one makes is that, for each catchment, the variability in the 30 yr climatological periods for a single model is much smaller than the variability among different models. Figure 10 shows similar results for temperatures, where the variability among different models, although still larger than the variability among the

25  30 yr climatological periods for a single model, is of comparable magnitude.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

## 6   Conclusions

Validation of climate models is typically performed using observational data, by study-ing the skills of different models in reproducing climate features in the study area. One major task is the choice of such observations. This study focuses on providing reliable climatic forcing for hydrological applications at a river basin level and, therefore, pre-cipitation and surface temperature were chosen as verification variables since: (i) they are used to specify the climate of a region in several climate classification systems, like the Köppen-Geiger one, (ii) they represent a minimum requirement for hydrologi-cal modeling, being respectively the main source of water in the catchments and the main control parameter for evaporation and evapotranspiration, and (iii) precipitation and temperature observation networks are the most dense and readily available ones (in contrast to, say, radiation or evapotranspiration measurements).

Another basic problem of model validation is that of selecting appropriate metrics to weight the relative influence of different variables. Since we needed to compare model performances in simultaneously reproducing the statistics of temperature and precipitation fields, we introduced dimensionless normalized metrics.

RCMs have been used in two major scientific projects, PRUDENCE and ENSEM-BLES, to produce future climate projections for EU. The corresponding model simula-tions have been studied and validated in several recent papers (i.e., Christensen and Christensen, 2007, for PRUDENCE; and Lorenz and Jacob, 2010, for ENSEMBLES). While most validation studies examined model performances by focusing on medium-to-large scale areas (e.g. Christensen et al., 2010, studied ENSEMBLES RCM results by dividing Europe in several large areas), RCMs were found to only partially reproduce climate patterns in Europe (Jacob et al., 2007; Christensen et al., 2010).

Our study suggests that, when interest is at relatively small spatial scales associated with hydrological catchments, as it is the case of CLIMB project, validation of CM results should be conducted at a single-basin level, rather than at macro-regional scales. In this case, it is necessary to check models' skills in reproducing prescribed observations

at specific river basins, since averaging over quite large areas might bias the assessment. For example, for Riu Mannu, Thau and Chiba catchments, model performances can vary significantly (see Sect. 5), even though these catchments are included in the same large-scale area in Christensen and Christensen (2007) study.

In this work we validated RCM results at scales suitable to run hydrological models and conduct climate impact studies for representative Mediterranean catchments. We found that the performance of a single RCM in reproducing observational data can change significantly in different river basins. This finding highlights the need for extensive analyses of climate model outputs before proceeding with hydrological modeling in specific catchments.

Another important finding is that, at least for temperature and precipitation studied at a river basin level, the variability in the 30 yr climatological periods for a single model is much smaller than the variability among different models, as Figs. 9 and 10 clearly show. We also stressed that the validation process in complex terrains, as is the case of the Alps (Noce catchment), may be significantly affected by weaknesses of model grids and the representativeness of the observational network. Actually, it can be problematic to interpret the differences between model outputs and observations, since they may originate from a combination of issues: (a) the model grid is too coarse (e.g., the Noce case, located in the Alps, where the maximum elevation considered by the models is approximately 2500 m, quite lower than real orography); and (b) the observational network is too sparse to provide a proper basis for models' validation.

Projects like ENSEMBLES and PRUDENCE stimulated and guided several climatic centres towards standardization of procedures, model grids and outputs and, thus, promoted synergies across different research areas and interdisciplinary efforts. Nevertheless, our study indicates that errors and inconsistencies are still present, suggesting basin-specific pre-processing of CM outputs before proceeding with hydrological modeling and climate impact assessments.

To what concerns hydrological applications, we note that while our validation-based model selection provides a reasonable indication of the four best performing models

in the considered catchments, some model deficiencies still need to be addressed. Specifically, by applying bias- and quantile-correction techniques, we were able to reduce the differences between observed and modelled probability distributions. In addition, by accounting for the effects of orography on precipitation, temperature and other variables, and making proper use of downscaling tools, we reproduced local climate attributes and, to some extent, the observed small scale variability. These results, as well as hydrological modeling projections in the considered catchments (currently addressed by several working groups of the CLIMB project) will form the subjects of forthcoming communications.

## References

Annan, J. D., Hargreaves, J. C., and Tachiiri, K.: On the observational assessment of climate model performance, Geophys. Res. Lett., 38, L24702, doi:10.1029/2011GL049812, 2011. 9108

Badas, M. G., Deidda, R., and Piga, E.: Modulation of homogeneous space-time rainfall cascades to account for orographic influences, Nat. Hazards Earth Syst. Sci., 6, 427–437, doi:10.5194/nhess-6-427-2006, 2006. 9116, 9117

Barnes, S. L.: A technique for maximazing details in numerical weather map analysis, J. Appl. Meteorol., 3, 396–409, 1964. 9117

Barnes, S. L.: Mesoscale objective analysis using weighted timesetime observations, NOAA Tech. Memo. ERL NSSL-62, Tech. rep., National Severe Storms Laboratory, Norman, OK, 1973. 9117

Becker, A. and Grünewald, U.: Flood risk in Central Europe, Science, 300, 1099, doi:10.1126/science.1083624, 2003. 9108

Cane, D., Barbarino, S., Renier, L. A., and Ronchi, C.: Regional climate models downscaling in the Alpine area with multimodel superensemble, Hydrol. Earth Syst. Sci., 17, 2017–2028, doi:10.5194/hess-17-2017-2013, 2013. 9109

Christensen, J. H. and Christensen, O. B.: A summary of the PRUDENCE model projections of changes in European climate by the end of this century, Clim. Change, 81, 7–30, doi:10.1007/s10584-006-9210-7, 2007. 9126

Christensen, J. H., Machenhauer, B., Jones, R., Schär, C., Ruti, P., Castro, M., and Visconti, G.: Validation of present-day regional climate simulations over Europe: LAM simulations with observed boundary conditions, Clim. Dynam., 13, 489–506, doi:10.1007/s003820050178, 1997. 9108

Deidda, R.: Rainfall downscaling in a space-time multifractal framework, Water Resour. Res., 36, 1779–1794, 2000. 9116

Deidda, R., Badas, M. G., and Piga, E.: Space-time scaling in high-intensity Tropical Ocean Global Atmosphere Coupled Ocean-Atmosphere Response Experiment (TOGA-COARE) storms, Water Resour. Res., 40, W02506, doi:10.1029/2003WR002574, 2004. 9116

Deidda, R., Badas, M. G., and Piga, E.: Space-time multifractality of remotely sensed rainfall fields, J. Hydrol., 322, 2–13, doi:10.1016/j.jhydrol.2005.02.036, 2006. 9117

Fowler, H., Blenkinsop, S., and Tebaldi, C.: Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling, Int. J. Climatol., 27, 1547–1578, 2007. 9108

Frei, C., Schöll, R., Fukutome, S., Schmidli, J., and Vidale, P. L.: Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models, J. Geophys. Res., 111, D06105, doi:10.1029/2005JD005965, 2006. 9121

Gebremichael, M. and Krajewski, W.: Assessment of the statistical characterization of small-scale rainfall variability from radar: analysis of TRMM ground validation datasets, J. Appl. Meteorol., 43, 1180–1199, 2004. 9116

Gebremichael, M., Over, T., and Krajewski, W.: Comparison of the scaling characteristics of rainfall derived from space-based and ground-based radar observations, J. Hydrometeorol., 7, 1277–1294, 2006. 9117

Giorgi, F. and Mearns, L. O.: Introduction to special section: regional climate modeling revisited, J. Geophys. Res., 104, 6335–6352, doi:10.1029/98JD02072, 1999. 9108

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◀ | ▶

◀ | ▶

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Giorgi, F. and Mearns, L.: Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the Reliability Ensemble Averaging (REA) Method, J. Climate, 15, 1141–1158, 2002. 9109

Gleckler, P. J., Taylor, K. E., and Doutriaux, C.: Performance metrics for climate models, J. Geophys. Res., 113, D06104, doi:10.1029/2007JD008972, 2008. 9109

Guyennon, N., Romano, E., Portoghese, I., Salerno, F., Calmanti, S., Petrangeli, A. B., Tartari, G., and Copetti, D.: Benefits from using combined dynamical-statistical downscaling approaches – lessons from a case study in the Mediterranean region, Hydrol. Earth Syst. Sci., 17, 705–720, doi:10.5194/hess-17-705-2013, 2013. 9109

Haylock, M. R., Hofstra, N., Klein-Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006, J. Geophys. Res., 113, D20119, doi:10.1029/2008JD010201, 2008. 9110, 9116

Jacob, D., Baerring, L., Christensen, O. B., Christensen, J. H., de Castro, M., Deque, M., Giorgi, F., Hagemann, S., Hirschi, M., Jones, R., Kjellstroem, E., Lenderink, G., Rockel, B., Sanchez, E., Schaer, C., Seneviratne, S. I., Somot, S., van Ulden, A., and van den Hurk, B.: An inter-comparison of regional climate models for Europe: model performance in present-day climate, Clim. Change, 81, 31–52, doi:10.1007/s10584-006-9213-4, 2007. 9121

Johnson, F. and Sharma, A.: Measurement of GCM Skill in Predicting Variables Relevant for Hydroclimatological Assessments, J. Climate, 22, 4373–4382, doi:10.1175/2009JCLI2681.1, 2009. 9109

Koutroulis, A., Tsanis, I., Daliakopoulos, I., and Jacob, D.: Impact of climate change on water resources status: a case study for Crete Island, Greece, J. Hydrol., 479, 146–158, 2013. 9108

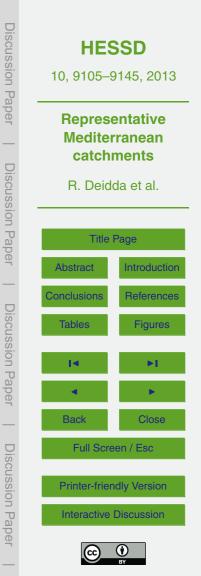Kundu, P. K. and Bell, T. L.: A stochastic model of space-time variability of mesoscale rainfall: Statistics of spatial averages, Water Resour. Res., 39, 1328, doi:10.1029/2002WR001802, 2003. 9116

Kunkel, K. E.: Simple procedures for extrapolation of humidity variables in the mountainous western United States, J. Climate, 2, 656–669, 1989. 9117

Lambert, S. J. and Boer, G. J.: CMIP1 evaluation and intercomparison of coupled climate models, Clim. Dynam., 17, 83–106, 2001. 9109

Liston, G. E. and Elder, K.: A Meteorological Distribution System for High-Resolution Terrestrial ModModel (MicroMet), J. Hydrometeorol., 7, 217–234, 2006. 9117

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

|◄ | ►|

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Lucarini, V.: Towards a definition of climate science, Int. J. Environ. Pollut., 18, 413–422, 2002. 9108

Lucarini, V.: Validation of climate models, in: Encyclopaedia of Global Warming and Climate Change, edited by: Philander, G., 1053–1057, SAGE, Thousand Oaks, USA, 2008. 9108, 9110

Lucarini, V., Danihlik, R., Kriegerova, I., and Speranza, A.: Does the Danube exist? Versions of reality given by various regional climate models and climatological data sets, J. Geophys. Res., 112, D13103, doi:10.1029/2006JD008360, 2007. 9108

Ludwig, R., Soddu, A., Duttmann, R., Baghdadi, N., S., B., Deidda, R., Marrocu, M., Strunz, G., Wendland, F., Engin, G., Paniconi, C., Prettenthaler, F., Lajeunesse, I., Afifi, S., Cassiani, G., Bellin, A., Mabrouk, B., Bach, H., and Ammerl, T.: Climate-induced changes on the hydrology of mediterranean basins – a research concept to reduce uncertainty and quantify risk, Fresen. Environ. Bull., 19, 2379–2384, 2010. 9110

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user, Rev. Geophys., 48, RG3003, doi:10.1029/2009RG000314, 2010. 9121

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T., Meehl, G. A., Mitchell, J. F. B., Nakicenovic, N., Riahi, K., Smith, S. J., Stouffer, R. J., Thomson, A. M., Weyant, J. P., and Wilbanks, T. J.: The next generation of scenarios forc limate change research and assessment, Nature, 463, 747–756, 2010. 9108

Muerth, M. J., Gauvin St-Denis, B., Ricard, S., Velázquez, J. A., Schmid, J., Minville, M., Caya, D., Chaumont, D., Ludwig, R., and Turcotte, R.: On the need for bias correction in regional climate scenarios to assess climate change impacts on river runoff, Hydrol. Earth Syst. Sci., 17, 1189–1204, doi:10.5194/hess-17-1189-2013, 2013. 9108

Murphy, J. M., Sexton, D. M. H., Barnett, D., Jones, G. S., Webb, M. J., Collins, M., and Stainforth, D. A.: Quantification of modelling uncertainties in a large ensemble of climate change simulations, Nature, 430, 768–772, 2004. 9109

Perica, S. and Foufoula-Georgiou, E.: Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions, J. Geophys. Res., 101, 26347–26361, 1996. 9116

Perkins, S., Pitman, A., Holbrook, N., and McAneney, J.: Evaluation of the AR4 climate models simulated daily maximum temperature, minimum temperature and precipitation over Australia using probability density functions, J. Climate, 20, 4356–4376, 2007. 9109

Räisänen, J.: How reliable are climate models?, Tellus A, 59, 2–29, doi:10.1111/j.1600-0870.2006.00211.x, 2007. 9108

Roe, G.: Orographic precipitation, Annu. Rev. Comput. Sci., 33, 645–671, doi:10.1146/annurev.earth.33.092203.122541, 2005. 9108

Rummukainen, M.: State-of-the-art with regional climate models, WIREs Climate Change, 1, 82–96, doi:10.1002/wcc.8, 2010. 9108

Schertzer, D. and Lovejoy, S.: Physical modeling and analysis of rain and clouds by anysotropic scaling of multiplicative processes, J. Geophys. Res., 92, 9693–9714, 1987. 9116

Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L.: Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK and New York, NY, USA, 2007. 9113

Sulis, M., Paniconi, C., Rivard, C., Harvey, R., and Chaumont, D.: Assessment of climate change impacts at the catchment scale with a detailed hydrological model of surface-subsurface interactions and comparison with a land surface model, Water Resour. Res., 47, W01513, doi:10.1029/2010WR009167, 2011. 9109

Sulis, M., Paniconi, C., Marrocu, M., Huard, D., and Chaumont, D.: Hydrologic response to multimodel climate output using a physically based model of groundwater/surface water interactions, Water Resour. Res., 48, W12510, doi:10.1029/2012WR012304, 2012. 9109

Takle, E. S., Gutowski, W., Arritt, R. W., Pan, Z., Anderson, C. J., da Silva, R. S., Caya, D., Chen, S.-C., Giorgi, F., Christensen, J., Hong, S.-Y., Juang, H.-M. H., Katzfey, J., Lapenta, W. M., Laprise, R., Lopez, P., Liston, G. E., McGregor, J., Pielke, A., and Roads, J. O.: Project to Intercompare Regional Climate Simulations (PIRCS): description and initial results, J. Geophys. Res., 104, 19443–19461, doi:10.1029/1999JD900352, 1999. 9108

Tessier, Y., Lovejoy, S., and Schertzer, D.: Universal multifractals: theory and observations for rain and clouds, J. Appl. Meteorol., 32, 223–250, 1993. 9116

Tsanis, I., Koutroulis, A., Daliakopoulos, I., and Jacob, D.: Severe climate-induced water shortage and extremes in Crete, Clim. Change, 106, 667–677, 2011. 9108

van den Besselaar, E., Haylock, M., Klein-Tank, A., and van der Schrier, G.: A European daily high-resolution observational gridded data set of sea level pressure, J. Geophys. Res., 116, D11110, doi:10.1029/2010JD015468, 2011. 9110

van Pelt, S. C., Beersma, J. J., Buishand, T. A., van den Hurk, B. J. J. M., and Kabat, P.: Future changes in extreme precipitation in the Rhine basin based on global and regional climate model simulations, Hydrol. Earth Syst. Sci., 16, 4517–4530, doi:10.5194/hess-16-4517-2012, 2012. 9109

Velázquez, J. A., Schmid, J., Ricard, S., Muerth, M. J., Gauvin St-Denis, B., Minville, M., Chaumont, D., Caya, D., Ludwig, R., and Turcotte, R.: An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources, Hydrol. Earth Syst. Sci., 17, 565–578, doi:10.5194/hess-17-565-2013, 2013. 9109

Veneziano, D. and Langousis, A.: The areal reduction factor: a multifractal analysis, Water Resour. Res., 41, W07008, doi:10.1029/2004WR003765, 2005. 9117

Veneziano, D. and Langousis, A.: Scaling and Fractals in Hydrology, in: Advances in Data-based Approaches for Hydrologic Modeling and Forecasting, edited by: Sivakumar, B. and Berndtsson, R., 107–242, World Scientific, Singapore, 2010. 9117

Veneziano, D., Langousis, A., and Furcolo, P.: Multifractality and Rainfall Extremes: a Review, Water Resour. Res., 42, W06D15, doi:10.1029/2003WR002574, 2006. 9117

Venugopal, V., Foufoula-Georgiou, E., and Sapozhnikov, V.: Evidence of dynamic scaling in space-time rainfall, J. Geophys. Res., 104, 31599–31610, 1999. 9116

Wang, Y., Leung, L. R., McGregor, J. L., Lee, D. K., Wang, W. C., Ding, Y., and Kimura, F.: Regional climate modeling: progress, challenges and prospects, J. Meteorol. Soc. Jpn., 82, 1599–1628, 2004. 9108

Wilby, R. L.: Evaluating climate model outputs for hydrological applications, Hydrolog. Sci. J., 55, 1090–1093, 2010. 9110

Wilby, R. L. and Harris, I.: A framework for assessing uncertainties in climate change impacts: low-flow scenarios for the River Thames, UK, Water Resour. Res., 42, W02419, doi:10.1029/2005WR004065, 2006. 9109

**Table 1.** Main topographic and 60 yr (1951–2100) climatological characteristics of the considered catchments: area ($S$), mean elevation ($\overline{z}$), mean annual precipitation ($\overline{P}$) and sea level temperature ($\overline{T}$), and minimum and maximum values of the monthly averages of precipitation and sea level temperatures.

| | Catchment (Country) | $S$ (km$^2$) | $\overline{z}$ (m) | $\overline{P}$ (mm yr$^{-1}$) | $\overline{T}$ (K) | min/max $P$ (mm month$^{-1}$) | min/max $T$ (K) |
|---|---|---|---|---|---|---|---|
| 1 | Thau (France) | 250 | 150 | 616 | 288.4 | 20/87 | 280.3/297.3 |
| 2 | Riu Mannu (Italy) | 472 | 300 | 466 | 290.8 | 3/72 | 283.2/299.7 |
| 3 | Noce (Italy) | 1367 | 1600 | 851 | 288.6 | 40/99 | 275.7/299.7 |
| 4 | Kocaeli (Turkey) | 3505 | 400 | 545 | 288.3 | 13/88 | 278.7/298.2 |
| 5 | Gaza (Palestine) | 365 | 50 | 217 | 294.6 | 0/56 | 286.6/301.6 |
| 6 | Chiba (Tunisia) | 286 | 200 | 377 | 292.1 | 2/60 | 285.1/300.5 |

**Table 2.** Acronyms of the Global Climate Models (GCMs) used as drivers of ENSEMBLES Regional Climate Models (RCMs) considered in this study.

| Acronym | Climatological center and model |
| --- | --- |
| HCH | Hadley Centre for Climate Prediction, Met Office, UK HadCM3 Model (high sensitivity) |
| HCS | Hadley Centre for Climate Prediction, Met Office, UK HadCM3 Model (standard sensitivity) |
| HCL | Hadley Centre for Climate Prediction, Met Office, UK HadCM3 Model (low sensitivity) |
| ARP | Meteo-France, Centre National de Recherches Meteorologiques, France CM3 Model Arpege |
| ECH | Max Planck Institute for Meteorology, Germany ECHAM5 / MPI OM |
| BCM | Bjerknes Centre for Climate Research, Norway BCM2.0 Model |

**Table 3.** Acronyms of the Regional Climate Models (RCMs) considered in this study.

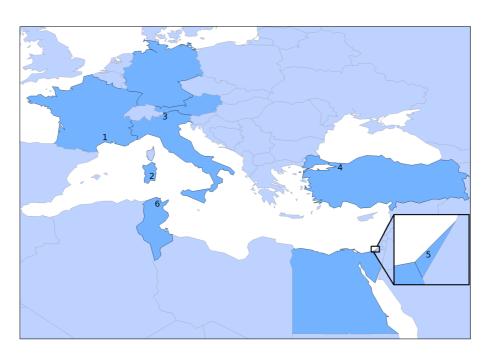| Acronym | Climatological center and model |
|---------|--------------------------------|
| RCA | Swedish Meteorological and Hydrological Institue (SMHI), Sweden RCA Model |
| HIR | Danish Meteorological Institute (DMI), Denmark HIRHAM5 Model |
| CLM | Federal Institute of Technology of Zurich (ETHZ), Switzerland CLM Model |
| HRM | Hadley Center for Climate Prediction, Met Office, UK HadRM3Q3 Model |
| RMO | Koninklijk Nederlands Meteorologisch Instituut (KNMI), Netherlands RACMO2 Model |
| REM | Max Planck Institute for Meteorology, Hamburg, Germany REMO Model |

**Fig. 1.** Locations of representative Mediterranean catchments considered in this study (see also Table 1): 1 – Thau (France), 2 – Riu Mannu (Italy), 3 – Noce (Italy), 4 – Kocaeli (Turkey), 5 – Gaza (Palestine), 6 – Chiba (Tunisia).

| GCM / RCM | HadleyC, Std. | HadleyC, Low | HadleyC, High | ECHam 5 | ARPege | BCM |
|---|---|---|---|---|---|---|
| HadRM ● | HCS-HRM ● | HCL-HRM ● | HCH-HRM ● | | | |
| REMo ◆ | | | | ECH-REM ◆ | | |
| HIRham ■ | HCS-HIR ■ | | | ECH-HIR ■ | ARP-HIR ■ | BCM-HIR ■ |
| CLM ▲ | HCS-CLM ▲ | | | | | |
| RacMO ▼ | | | | ECH-RMO ▼ | | |
| RCA 3 | | HCL-RCA ► | HCH-RCA ► | ECH-RCA ► | | BCM-RCA ► |

**Fig. 2.** Combinations of Global Climate Models (GCMs) and Regional Climate Models (RCMs) considered in this study. In all figures, we use the same color (symbol) to refer to the same GCM (RCM). Model acronyms are introduced in Tables 2 and 3.
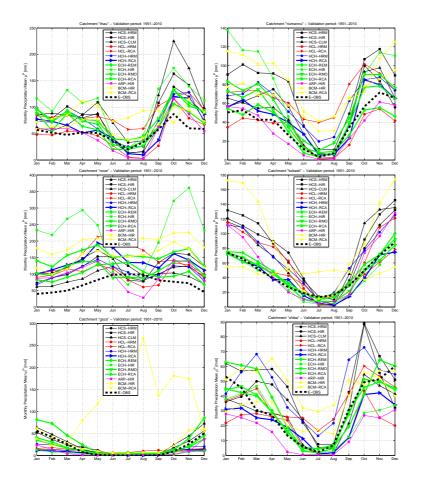
**Fig. 3.** Mean monthly precipitation (mm) over the 60 yr verification period (1951–2100) for the 14 ENSEMBLES models and E-OBS. Each subplot displays areal averages over a 4 × 4 grid-point stencil centered in each catchment.
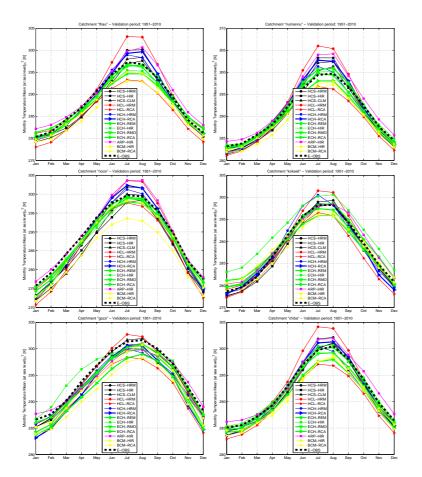
Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

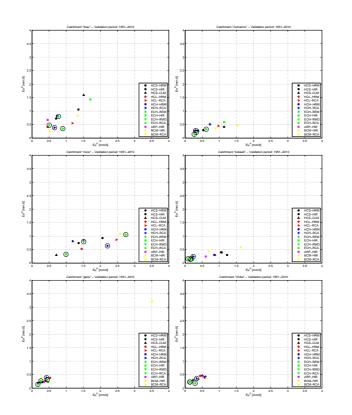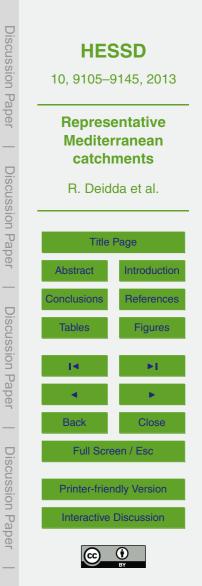**Fig. 4.** Same as Fig. 3, but for mean monthly temperatures (K) reduced to sea level.

**Fig. 5.** Scatterplot of errors in the mean and standard deviation of monthly precipitation (mm d$^{-1}$) over the 60 yr verification period (1951–2100), computed using Eqs. (3) and (4) for each of the 14 ENSEMBLES models with respect to E-OBS observational reference. Each subplot displays areal averages over a 4 × 4 grid-point stencil centered in each catchment.
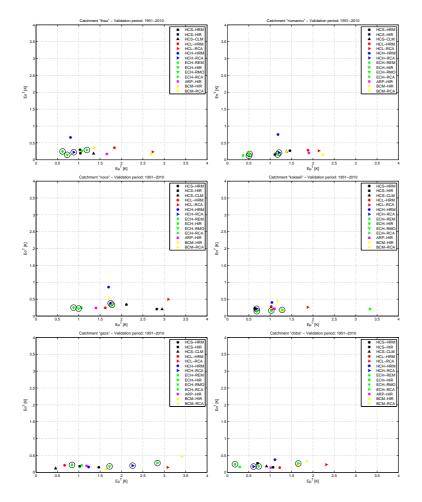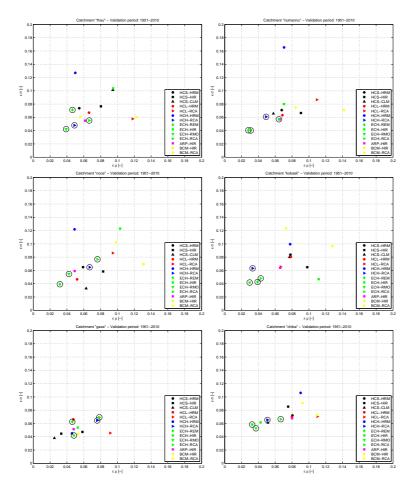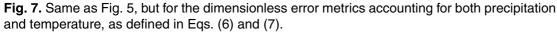
Title Page

Abstract    Introduction

Conclusions    References

Tables    Figures

|◄    ►|

◄    ►

Back    Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Fig. 6.** Same as Fig. 5, but for temperatures (K) reduced to sea level.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Fig. 7.** Same as Fig. 5, but for the dimensionless error metrics accounting for both precipitation and temperature, as defined in Eqs. (6) and (7).

Title Page

Abstract　　Introduction

Conclusions　　References

Tables　　Figures

◀　　▶

◁　　▷

Back　　Close

Full Screen / Esc

Printer-friendly Version
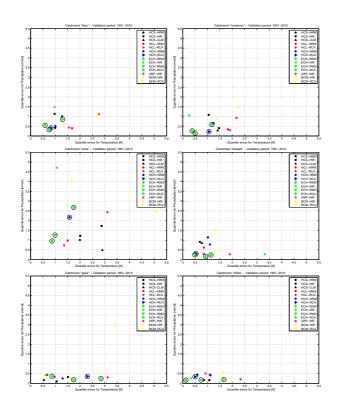
Interactive Discussion

**Fig. 8.** Scatterplots of mean absolute errors in the quantiles of daily precipitation and temperature distributions, computed using Eq. (8) for each of the 14 ENSEMBLES models with respect to E-OBS observational reference. Quantiles are calculated at 100 uniformly spaced probability levels. Each subplot displays areal averages over a 4 × 4 grid-point stencil centered in each catchment.
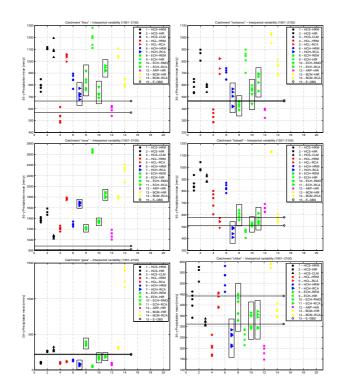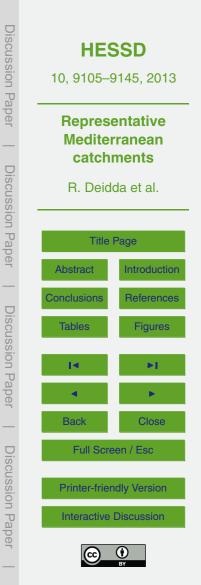
Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Fig. 9.** Variability of the mean annual precipitation (mm yr$^{-1}$) over the five 30 yr climatological periods between 1951 and 2100. Models HCS-HIR and BCM-HIR stop in year 2050. The reference values of E-OBS climatologies in the two 30 yr periods between 1951–2010 are indicated with empty circles and horizontal lines. Each subplot displays results based on areal averages over a 4 × 4 grid-point stencil centered in each catchment.
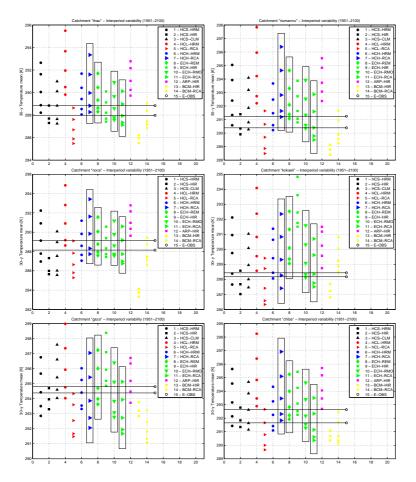
**Fig. 10.** Same as Fig. 9, but for temperatures (K) reduced to sea level.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

|◄ | ►|

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion