Assessment of uncertainties in the hydro-climatic modeling chain over heterogeneous landscapes

Dissertation zur Erlangung des Doktorgrades an der Fakultät für Geowissenschaften der Ludwig-Maximilians-Universität München

vorgelegt von

David Gampe, München

eingereicht am 29. Mai 2018, München

Supervisor: Prof. Dr. Ralf Ludwig, Department of Geography, Ludwig-Maximilians-Universität, Munich

2nd Supervisor: Prof. Dr. Karsten Schulz, Institute of Water Management, Hydrology and Hydraulic Engineering, Universität für Bodenkultur, Vienna

Tag der mündlichen Prüfung: 19.10.2018

Je me souviens.

Acknowledgments

After a very exciting time, this thesis finally comes to an end. This, sometimes hard and frustrating, but extremely instructive, positively challenging and most of the time also very enjoyable journey would not have been possible without the support of numerous people along the way whom I want to express my sincere gratitude to.

First and foremost, I want to thank my supervisor Prof. Dr. Ralf Ludwig for giving me the opportunity to start this PhD project and the confidence he put in me throughout the years to also accomplish it. His passion for science was inspiring and motivating and I benefited a great deal from his experience and deep understanding of climate change and hydrology. His kindness and the endless support he gave me while at the same time giving me the freedom to find my own path and pursue new ideas cannot be considered a given and deserve special thanks. I am also extremely grateful for the opportunity he gave me to accompany him on the field trips to Québec and hope to have the chance to enjoy his *Kaiserschmarrn aux bleuets* in Chicoutimi again sometime in the future.

Special thanks are to Prof. Dr. Karsten Schulz and Prof. Dr. Harald Kunstmann for being my mentors in the course of the MICMoR graduate program, for the time they invested and their interest to follow my progress as well as the fruitful discussions we had in our meetings. In this line, I also want to thank the MICMoR research school for giving me the opportunity to broaden my horizon through various courses and for the financial support to attend international conferences. I want to particularly thank Dr. Bärbel Elija Bleher for her guidance and also her patience with delayed reports. To all the MICMoR fellows for the interesting discussions and the fun times we had in the yearly retreats, which I truly enjoyed.

I also want to thank all the colleagues and friends from the 3rd and 2nd floor, who made working at the institute really enjoyable. Particular thanks to the GLOBAQUA office: Verena Huber García, Dr. Swen Meyer, Janja Vrzel and Lyndsy Schultz for being such great office roommates, the extremely nice atmosphere we had and for never complaining about me cursing and shouting at my computer. To Ben Müller and Florian Willkofer for their patience and willingness to discuss R-Figures, posters and papers, helping me with coding issues and all the fun board game nights we had. To Vera Erfurth, Gudrun Lampart and Karin Drexler for making my life a lot easier and taking care of contracts and bureaucracy behind the scenes. I also wish to express my gratitude to Dr. Wilfried Hagg for the possibility to join his group and getting involved in teaching as well as the freedom he gave me to finalize my thesis.

A very warm thank you belongs also to the entire team of Rapid Maxvorstadt for all the fun times we had and the many, many defeats we suffered together and still had a blast. Many hits & hugs to Augsburg Rolling Thunder and a huge thank you for all the bruises and for introducing me to a whole new world that I didn't know existed. I also wish to thank the "Stadioncrew" for our field trips to the stadium every other Saturday!

Many thanks belong to the most important people in my life: my family and friends who filled my life with joy and helped me through the inevitable thesis-related crises that occurred along the way. Finally, I especially want to thank Isabella for the numerous sneak previews we saw, her support, love and for always being there for me over the past years.

Summary

Following recent climate projections, the Mediterranean region is likely to experience a severe reduction in precipitation and increased temperatures and is thus likely to be exposed to increased drought risk under future climate. This will increase the pressure on water resources in regions that are already suffering from water scarcity under today's climate. To assess the impact of climate change on aquatic ecosystems and potentially derive adaptation strategies, hydrological models form a great tool. These models are driven by climate scenarios provided through General Circulation or Global Climate Model (GCMs), which are then dynamically downscaled by Regional Climate Models (RCMs), under different Representative Concentration Pathways (RCPs) and form the so-called hydro-climatic or hydroclimatological modeling chain. Despite all improvements and increased process understanding and their implementation in climate models, the hydro-climatic modeling chain is still prone to numerous sources of uncertainty. The detection, evaluation, quantification and potentially also reduction of the related uncertainties is of great importance to allow for robust estimates of climate change impacts and derive potential adaptation measures. This thesis assesses the contribution of each of the uncertainty sources in selected, data scarce river basins in the Mediterranean region and over complex terrain. The main research statements and findings are briefly summarized in the following.

Observational uncertainty depicts a considerable source of uncertainty in the hydro-climatic modeling chain and shall thus not be neglected. Gridded data sets are applied in various ways in the hydro-climatic modeling chain. They serve as a reference to evaluate, select and bias correct RCMs and are applied for the calibration and validation of hydrological models. Available reference data sets show large differences, especially for precipitation and over complex terrain. As precipitation plays a key role in the water balance it is considered the most important variable in the hydro-climatic modeling chain. Therefore, available data sets for precipitation were analyzed over a mountainous catchment. This study revealed large differences in the representation of the general climatology, while the deviations are even more pronounced for indicators such as consecutive dry and wet days and heavy or extreme precipitation. Bias correction of RCMs with an ensemble of reference data sets revealed large differences in the historical and future representation of precipitation while having only a minor impact on RCM selection and climate change signals.

The involved uncertainty sources in climate projections should be assessed on the catchment scale. Uncertainties in climate projections arise from the GCMs, RCMs and the RCPs. Knowledge on the relative contribution of each source of uncertainty is crucial for model selection and potential model improvement. As in previous studies a variance decomposition approach was applied to retrieve the relative contribution of each uncertainty source to the overall uncertainty. In contrary to previous studies the assessment is performed on the catchment scale, where the climate data is usually applied in climate change impact assessment studies, rather than on the continental scale and is based on the most recent climate projections. The results confirm previous findings with the GCMs dominating the

uncertainty for precipitation and temperature. However, scenario uncertainty originating from the RCP is higher than in previous studies for both, temperature and precipitation. As expected, the choice of the RCM is most relevant for surface runoff and evapotranspiration as these variables are highly dependent on the surface and land use scheme implemented in the corresponding RCM.

The direct usage of RCM projections allows for an estimation of future water scarcity in data scarce areas. Apart from data scarcity, the setup, calibration and validation of hydrological models is usually a time consuming task. Today's RCMs provide simulations of the variables of the water balance at a 0.11° (~ 12 km) grid, allowing for a first-guess estimation of future changes on precipitation, evapotranspiration and surface runoff. The RCM ensemble applied shows a strong decrease in precipitation and surface runoff over the selected river basins around the Mediterranean. As these basins are already exposed to water scarcity under today's climate, this indicates increased pressure on the water resources under future conditions. A high agreement among the RCMs in the direction and magnitude of change indicates robustness in the climate projections.

Remote sensing allows to increase the robustness of a hydrological model in a data scarce area. The distributed and physically based hydrological Water Flow and Balance Simulation Model (WaSiM) was applied in the Gaza Strip, Palestine, to derive the impacts of climate change on the water balance. This area is, as many Mediterranean catchments, characterized by data scarcity which hampers the setup and validation of hydrological models. A simple remote sensing approach that links surface brightness temperature and vegetation coverage to derive evapotranspiration patterns was applied to improve the parameterization of the hydrological model and to better represent irrigated areas. The approach showed great potential to increase model robustness especially under data scarcity. An R-package was compiled in the course of this thesis to allow for an efficient processing of a series of remote sensing images. The model results reveal increased drought risk under future climate and highlight the need to define adaptation strategies to avoid a further over-exploitation of the ground water aquifer that serves as the main freshwater source in the area.

The results of this thesis are presented in four peer-reviewed scientific publications and address different uncertainty sources in the hydro-climatic modeling chain. The thesis highlights the potential of remote sensing to increase the robustness of hydrological models in data scarce areas. An assessment of the dominant uncertainty sources confirms the importance of GCMs for model selection if only a subset of models is included in hydrological studies. The importance of observational uncertainty of precipitation is of special focus in this thesis and shall not be neglected in future studies. Rather, an ensemble approach similar to climate models should be considered.

Zusammenfassung

Aktuellen Klimaprojektionen zufolge wird der Mittelmeerraum im Zuge des Klimawandels sehr wahrscheinlich verringerten Niederschlägen, stark erhöhten Temperaturen und einem daraus resultierenden erhöhten Dürrerisiko ausgesetzt sein. All dies wird den Druck auf die Wasserressourcen gerade in solchen Gebieten erhöhen, welche bereits unter den heutigen klimatischen Bedingungen unter Wasserknappheit leiden. Bei der Abschätzung und Quantifizierung der Auswirkungen des Klimawandels auf aquatische Ökosysteme und der Erstellung entsprechender Anpassungsstrategien, spielen hydrologische Modelle eine zentrale Rolle. Diese Modelle werden mit Klimaprojektionen aus Globalen Klimamodellen (GCMs), welche mit Hilfe Regionaler Klimamodelle (RCMs) dynamisch auf eine höhere Auflösung skaliert werden, für verschiedene Strahlungsantriebe (RCPs) im Rahmen der hydro-klimatischen Modellkette angetrieben. Trotz intensiver Verbesserung der Klimamodelle und erhöhten Prozessverständnisses ist die hydro-klimatische Modellkette nach wie vor mit sehr vielen Unsicherheiten behaftet. Um robuste Modellergebnisse zu ermöglichen, ist eine genaue Abschätzung, Evaluierung und Quantifizierung der Unsicherheiten innerhalb der Modellkette unabdingbar und daher von hoher Relevanz. Die vorliegende Doktorarbeit zielt darauf ab, den Beitrag der Unsicherheiten jeder einzelnen Komponente der hydro-klimatischen Modellkette zur Gesamtunsicherheit abzuschätzen und zu quantifizieren. Hierbei sind der oftmals datenarme Mittelmeer- und der Alpenraum von speziellem Interesse und werden anhand ausgewählter Untersuchungsgebiete analysiert. Im Folgenden werden die wesentlichen Hypothesen und die zugehörigen Resultate dieser Arbeit kurz vorgestellt.

Die Unsicherheiten der Referenzdatensätzen leisten einen nicht unerheblichen Beitrag zur Gesamtunsicherheit und sollten daher im Zuge der hydro-klimatischen Modellkette nicht vernachlässigt bzw. ignoriert werden. Flächenverteilte Referenzdatensätze sind im Rahmen der hydro-klimatischen Modellkette an mehreren Stellen von großer Bedeutung; Zum einen werden sie für die Evaluierung und Korrektur der RCMs benötigt, zum anderen sind sie bei der Kalibierung und Validierung hydrologischer Modelle bedeutend. Die verfügbaren Referenzdatensätze, vor allem für Niederschlag, weisen jedoch große Unterschiede auf, welche über Regionen mit komplexer Topographie noch verstärkt werden. Da Niederschlag eine tragende Rolle in der Wasserhaushaltsgleichung einnimmt, stellt er die wichtigste Größe der hydro-klimatischen Modellkette dar. Aus diesen Gründen wurden im Rahmen dieser Arbeit verschiedene Niederschlagsreferenzdatensätze für ein alpines Einzugsgebiet untersucht. Die verwendeten Datensätze wiesen große Unterschiede hinsichtlich der Wiedergabe des Jahresganges auf, welche für Indikatoren wie die Anzahl der aufeinanderfolgenden Niederschlagstage sowie extreme Niederschlagsereignisse noch verstärkt wurden. Die Korrektur der RCMs mit einem Ensemble aus Referenzdatensätzen zeigte große Unterschiede in den absoluten Werten des Niederschlags auf, hatte jedoch nur geringen Einfluss auf die Auswahl geeigneter RCMs und die resultierenden Klimaänderungssignale.

Die Unsicherheiten in den Klimaprojektionen müssen auf der Einzugsgebietsebene be-

trachtet und analysiert werden. Unsicherheiten in den Klimaprojektionen können auf die drei wesentlichen Komponenten GCM, RCM und RCP zurückgeführt werden. Wie bereits in vorherigen Veröffentlichungen, wurde der Ansatz der Varianzzerlegung angewandt, um den relativen Beitrag jeder der genannten Komponenten zur Gesamtvarianz zu ermitteln. Im Gegensatz zu den vorherigen Studien, wurde dieser Ansatz jedoch auf der Einzugsgebietsebene angewandt, da dies gewöhnlich auch die typische räumliche Skala ist, auf welcher die Daten verwendet werden. Die Ergebnisse unterstützen vorherige Studien darin, dass die Unsicherheit durch die GCMs die anderen Komponenten überlagert. Allerdings ist der Beitrag der RCMs und der RCPs deutlich höher als in vorherigen Untersuchungen. Dies trifft sowohl auf Temperatur, als auch auf Niederschlag zu. Für die Variablen Evapotranspiration und Oberflächenabfluss ist der Beitrag der RCMs dominant, was auf die unterschiedlichen Parametrisierungen von Landnutzung und -oberfläche in den RCMs zurückzuführen ist.

Die direkte Verwendung von RCM Projektionen ermöglicht es, Abschätzungen über eine künftige Entwicklung der Wasserknappheit selbst in datenarmen Gebieten zu treffen. Unabhängig von der Datenverfügbarkeit ist das Aufsetzen, Kalibrieren und Validieren eines hydrologischen Modelles in der Regel sehr zeit- und arbeitsintensiv. Heutige RCMs ermöglichen es jedoch, eine erste Abschätzung der Wasserbilanz unter zukünftigem Klima abzuleiten, da diese alle Komponenten der Wasserhaushaltsgleichung, sprich Niederschlag, Evapotranspiration und Oberflächenabfluss, auf einem $0.11°$ (~ 12 km) Raster bereitstellen. Das hier angewendete RCM-Ensemble zeigt für den Großteil der betrachteten Flusseinzugsgebiete eine starke Abnahme in Niederschlag, Evapotranspiration und Oberflächenabfluss auf. Da die betroffenen Einzugsgebiete bereits unter heutigen klimatischen Bedingungen von Wasserknappheit betroffen sind, wird dies sehr wahrscheinlich den Druck auf die verfügbaren Wasserressourcen erhöhen und eine fortschreitende Übernutzung selbiger zur Folge haben. Eine hohe Übereinstimmung sowohl in der projizierten Richtung der Veränderung, als auch der absoluten Werte, weist auf eine relativ hohe Robustheit der Projektionen dahingehend hin.

Mithilfe der Fernerkundung kann die Robustheit hydrologischer Modelle speziell in datenarmen Regionen erhöht werden. Das flächenverteilte Wasserhaushalts-Simulations-Modell (WaSiM) wurde im Gazastreifen, Palästina, angewendet um die Auswirkungen des Klimawandels auf die Wasserressourcen ableiten zu können. Wie viele andere Gebiete im Mittelmeerraum ist auch der Gazastreifen durch eine schlechte Datenverfügbarkeit gekennzeichnet, welche die Anwendung und Validierung hydrologischer Modelle erschwert und einschränkt. Ein relativ einfacher Fernerkundungsansatz wurde angewendet um Verdunstungsmuster abzuschätzen, welche eingesetzt wurden um bewässerte landwirtschaftliche Flächen im Einzugsgebiet zu ermitteln und im Modell zu parametrisieren. Um eine Vielzahl von Satellitenbildern effizient prozessieren zu können, wurde hierfür im Rahmen dieser Doktorarbeit ein R-Paket kompiliert. Aufgrund der Übertragbarkeit und der einfachen Prozessierung, stellt dieser Ansatz eine vielversprechende Alternative dar um die Robustheit hydrologischer Modelle, speziell in datenarmen Regionen, zu erhöhen. Die Modellergebnisse zeigen ein erhöhtes Dürrerisiko unter künftigen Klimabedingungen auf. Dieses wird

den Druck auf die Wasserressourcen des Grundwasseraquifers sehr wahrscheinlich erhöhen, wodurch die Erstellung und die Umsetzung sinnvoller Anpassungsstrategien unumgänglich wird, um die landwirtschaftliche Produktivität auch unter künftigen Bedingungen aufrecht erhalten zu können.

Die Ergebnisse dieser Doktorarbeit sind in vier wissenschaftlichen Veröffentlichungen zusammengefasst, wobei jede dieser Veröffentlichungen eine andere Unsicherheitskomponente der hydro-klimatischen Modellkette beleuchtet. Diese Doktorarbeit unterstreicht das Potential der Fernerkundung um die Robustheit hydrologischer Modelle unter Datenknappheit zu erhöhen. Da GCMs den stärksten Einfluss auf die Gesamtunsicherheiten haben, sollte ihnen bei der Wahl der Klimaprojektionen besondere Aufmerksamkeit geschenkt werden. Des Weiteren wurde die Bedeutung der Referenzunsicherheiten hervorgehoben, welche auf in zukünftigen Studien verstärkt berücksichtigt werden muss. Es wird daher die generelle Empfehlung ausgesprochen, auch für die Referenzdatensätze, analog zur Verwendung von Klimaprojektionen, ein Ensemble aus mehreren Datensätzen zu verwenden.

Contents

List of Figures

Abbreviations

ANOVA Analysis of Variance

AOGCM Atmosphere–Ocean Global Climate Models

AR Assessment Report

BC Bias Correction

CLIMB Climate Induced Changes on the Hydrology of Mediterranean Basins

CMIP Coupled Model Intercomparison Project

EURO-CORDEX Coordinated Downscaling Experiment - European Domain

ECMWF European Centre for Medium-Range Weather Forecasts

ESM Earth System Models

GCM General Circulation Model / Global Climate Model

GLOBAQUA Managing the effects of multiple stressors on aquatic ecosystems with water scarcity

IPCC Intergovernmental Panel on Climate Change

MME Mutli Model Ensemble

NDVI Normalized Difference Vegetation Index

NOAA National Oceanic and Atmospheric Administration

NWP Numerical Weather Prediction

QM Quantile Mapping

RCF Representative Climate Futures

RCM Regional Climate Model

RCP Representative Concentration Pathway

SRES Special Report on Emissions Scenarios

WaSiM Water Flow and Balance Simulation Model

WCRP World Climate Research Programme

1 Introduction

1.1 Projected climate change impacts on the Mediterranean region

Climate change depicts one of the greatest challenges for humanity in the 21st century (Karl and Trenberth, 2003) and will likely have severe impacts on the natural and human systems. Among others, climate change is driving biodiversity loss (Rockström et al., 2009), challenges food security (Schmidhuber and Tubiello, 2007), will likely increase flood risk and flood probability (Arnell and Gosling, 2016) on the global scale and affects the tourism sector (Scott et al., 2012). Due to the longevity of atmospheric CO_2 climate change will lead to irreversible increased global temperatures and changes in precipitation (Solomon et al., 2009).

Information about possible future climate changes under global warming is becoming more and more an integral part of developing suitable adaptation strategies at global, national, regional and local levels. Such information is based on long-term climate simulations generated by climate models under a number of future emission scenarios. These climate projections provide a range of possible developments of future climate and resulting changes in variables such as precipitation and surface air temperature and provide the basis for hydrological modeling with focus on climate change impacts. In addition to the different realizations created by the scenarios, climate model simulations are characterized by differences in their model structure and parameterization schemes, e.g. of cloud microphysics in convective clouds (Gettelman and Morrison, 2015). Different models thus generate different realizations of climate variables causing uncertainties in the projections. Following previous studies in the context of climate science (Déqué et al., 2007; Di Luca et al., 2012; García-Díez et al., 2015; Prein and Gobiet, 2017), uncertainty is defined as spread of climate or hydrological models, as well as reference data sets in this thesis. The two terms *spread* and *uncertainty* are used interchangeably. *Robustness* in contrary is given in case of a high model or data set agreement, i.e. small spread, and can be interpreted as low uncertainty. In addition to a small spread, the agreement in the direction of change for multiple climate models of a given variable is considered as robustness in the climate change signal.

Projections of future climate show a high robustness for mean air temperature as the models generally agree in a positive change and show significant increases for most areas around the globe also for different climate model generations (Knutti et al., 2013). In contrary to air temperature projections for precipitation are more complex as they are highly dependent on other variables, linked to topography and potentially strongly influenced by land use (changes) on the regional scale (Barkhordarian et al., 2013). The uncertainty in precipitation projections therefore is generally larger than for air temperature (Beniston et al., 2007; Knutti and Sedláček, 2013). However, climate projections are characterized by strong regional differences such as the arctic amplification (Screen and Simmonds, 2010) and therefore need to be assessed on a regional scale.

The Mediterranean region can be considered as one of the hot-spots of climate change, characterized by a strong increase in air temperature and severe decrease in precipitation of up to 50% and a resulting rise in drought frequency of up to 40% (Christensen et al., 2007). Projections for consecutive dry days and a decrease of total precipitation are consistent throughout most models and depict increased water stress under future climate (Sillmann et al., 2013b). The frequency, intensity and duration of heat waves are projected to increase and the drought period in the Mediterranean is projected to expand temporally with earlier onsets and later end dates (Beniston et al., 2007). There is a general tendency towards increased and intensified extremes under future climate with change rates exceeding the general long-term climate change signal (Fischer et al., 2013).

The assessment of the impacts of climate change on aquatic ecosystems and water resources is of great importance and hydrological models are an essential tool for water management in this context (Devia et al., 2015). Hydrological models provide a simplified representation of the complex real world system and can be categorized by e.g. their type, complexity, dynamic and spatial scale (Moradkhani and Sorooshian, 2009). As the Mediterranean region is already prone to seasonal water scarcity and projections indicate an intensification of these phenomena it is of high importance to assess future impacts on water resources especially in this region.

In the course of the project CLIMB ('Climate Induced Changes on the Hydrology of Mediterranean Basins'), funded under the European Union's Seventh Programme, seven hydrological basins in the Mediterranean were examined. Hydrological modeling served as the primary tool to estimate the impacts of climate change on the water resources in these study areas. In this project, novel geophysical field monitoring techniques, remote sensing tools and hydrological models were applied to increase the process understanding and provide the basis for water management (Ludwig et al., 2010). Hydrological modeling using the physically based Water Flow and Balance Simulation Model (WaSiM) (Schulla and Jasper, 2007) revealed increased pressure on the water resources in spring in a catchment in Sardinia, Italy under changing climate (Meyer et al., 2016) that translates to reduced agricultural productivity for the future period (Bird et al., 2016). The application of WaSiM in another region revealed an increased drought risk due to climate change and demonstrated the need for adapted irrigation practices in the Gaza Strip in Palestine (Gampe et al., 2016a).

As aquatic ecosystems are highly complex an integrative assessment is of extreme importance. The project GLOBAQUA (Managing the effects of multiple stressors on aquatic ecosystems with water scarcity) examines river basins in the Mediterranean under multiple pressures (Navarro-Ortega et al., 2015). In this context climate change has proven to be one of the potential stressors on the water balance with decreased precipitation and negative changes in surface runoff for four river basins of the Ebro (Spain), Adige (Italy), Evrotas (Greece) and the Sava (Eastern Europe) (Gampe et al., 2016b). These projected reductions in runoff lead, together with socio-economic changes, to a decrease in water quality e.g. for the Ebro River (Herrero et al., 2018).

As climate change impact assessment studies are prone to uncertainties, challenges remain to communicate the results to decision makers and eventually implement necessary adaptation strategies (Maslin and Austin, 2012). It is therefore of high importance to identify and quantify these sources of uncertainty to continuously improve the understanding and the representation of processes, such as the parameterization of convection, and ultimately improve climate projections.

This cumulative thesis has been carried out within the two mentioned projects, CLIMB & GLOBAQUA, to determine and quantify the contribution of related uncertainty sources included in most of the climate change impact studies. In the following the typical hydro-climatic modeling chain is introduced in section 1.2. An overview on which sources of uncertainty can be identified and how these are currently addressed and included in the literature is provided in section 1.3 followed by the formulation of the aims and goals of this thesis and how the presented publications add to the discussion in section 1.4. Four publications emerged in the course of this thesis and are presented in section 2. A conclusion of the main findings and an outlook on potential future work in this research domain is presented in section 3.

1.2 The typical hydro-climatic modeling chain

The typical hydro-climatic modeling chain is shown in Figure 1 (Muerth et al., 2013; Teutschbein and Seibert, 2010; Xu et al., 2005). Climate forcing is typically provided through climate models, either directly from General Circulation or Global Climate Models (GCMs) or dynamically downscaled through Regional Climate Models (RCMs). As due to computational constraints it is usually not feasible to simulate the entire ensemble of available climate scenarios, a sub-selection of RCMs is chosen (Wilcke and Bärring, 2016) and then often post-processed through e.g. bias correction (BC) or a further interpolation to a finer grid (Velázquez et al., 2013). The so-derived climate model ensemble then provides the required spatio-temporal resolution to drive the selected hydrological model(s).

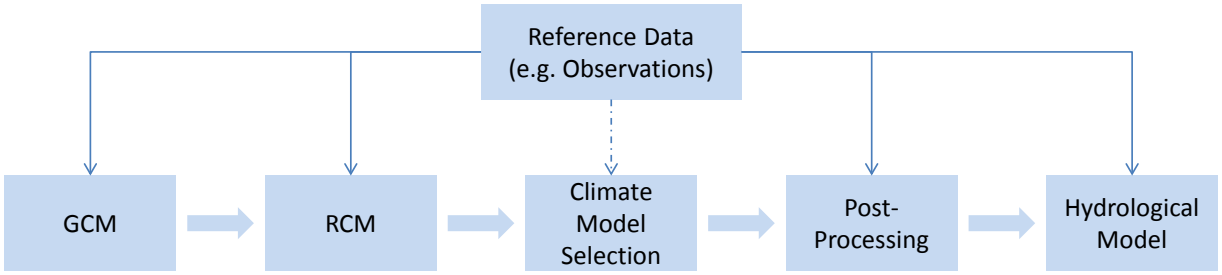


Figure 1: Typical hydro-climatic modeling chain as usually applied in climate change impact studies (Muerth et al., 2013; Teutschbein and Seibert, 2010; Xu et al., 2005). As reference data sets play a crucial role in most steps of these chain they are included as well.

(GCMs) can be considered the primary and irreplaceable tool for the generation of future climate (Giorgi and Gutowski, 2016; Knutti et al., 2013); These are made available through

e.g. the Coupled Model Intercomparison Project (CMIP) coordinated by the World Climate Research Programme (WCRP), where they have been systematically defined and harmonized regarding their simulation horizon, simulated scenarios, experiments included and are centrally stored and provided to the scientific community (Meehl et al., 2007).From the previous to the most recent fifth phase of the intercomparison projects (CMIP5) the number of models stored, their spatial resolution and their complexity increased. Already in the previous phase (CMIP3) integrated Atmosphere-Ocean Global Climate Models (AOGCMs), which include interactions between atmosphere, ocean, land and sea ice, became the standard. In CMIP5 some models are now coupled to biogeochemical compartments allowing for a detailed representation of carbon fluxes and thus the carbon cycle and are categorized as Earth System Models (ESMs) allowing for an interactive computation of emission concentrations however are even more computational demanding (Taylor et al., 2012). As for this thesis does not focus on carbon fluxes or atmosphere-land-ocean interactions, the term GCM will be used for all climate models simulating the global climate in the following without differentiating AOGCMs and ESMs. Today, GCMs simulate the globes climate at a spatial grid of 1.25 - 2.5° with differences in their number of vertical layers and process parameterization (Collins et al., 2011b; Sillmann et al., 2013b; Voldoire et al., 2013).

Projection of future climate in CMIP5 simulations are performed after model year 2005 and forced with the Representative Concentration Pathways (RCP) presented by Van Vuuren et al. (2011). These RCPs replace the emissions scenarios of previous climate model generations, as presented in the Special Report on Emissions Scenarios (SRES) for the 4^{th} Assessment Report (AR4) of the Intergovernmental Panel on Climate Change (IPCC) (Nakicenovic et al., 2000) used in CMIP3. For the SRES scenarios the climate models responded to the increased CO_2 individually while in case of the RCPs they are driven by a forced response (Taylor et al., 2012). The highest forcing is constructed in RCP8.5, resulting in an increased radiative forcing throughout the 21st century and finally reaching 8.5 Wm^{-2} by the end of the century, while the RCP4.5 can be considered an intermediate scenario with an increased radiative forcing of 4.5 Wm^{-2} by the end of the 21st century (Taylor et al., 2012). While two additional RCPs (2.6 & 6.0) are constructed, most simulations are available with the previous two radiative forcings. Simulations with available CMIP5 models results in a global temperature change of around 2°C for RCP4.5 and more than 4°C under RCP8.5 (Knutti and Sedláček, 2013) as presented in Figure 2.

The spatial representation of current GCMs is too coarse to capture local forcings and thus not capable of providing an accurate representation of local climate, especially for extremes. These effects get elevated over complex terrains and for areas that are dominated by small scale weather patterns (Giorgi et al., 2001). However, small scale climate change signals and projections are essential for users and decision makers to estimate climate change impacts and for the development of feasible adaptation strategies. Therefore, regionalization or downscaling techniques have been developed and established over the past decades (Giorgi et al., 2009). These techniques can be grouped in statistical or dynamical downscaling. In case of statistical downscaling, the large scale climatic state is derived from GCMs

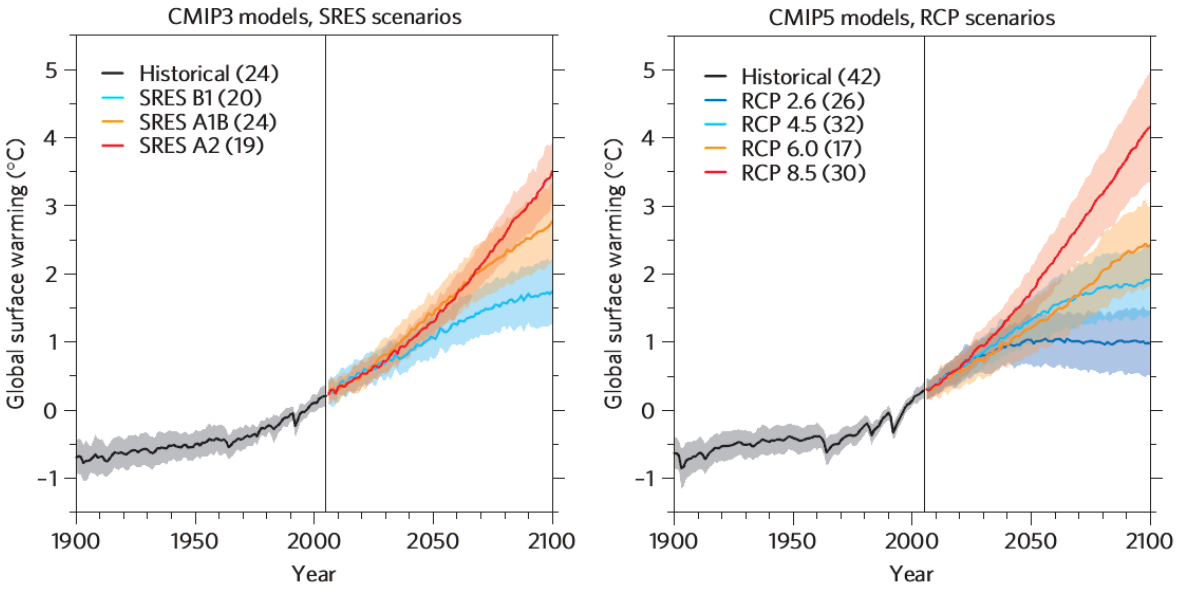


Figure 2: Projected global air temperature change relative to the period 1986 - 2005 for the SRES and RCP scenarios based on CMIP3 and CMIP5 simulations (adapted from Knutti and Sedláček, 2013). The number in brackets refers to the number of models available.

and applied to a statistical model. This establishes a statistical relationship between the large scale GCM fields (predictors) and local variables (predictands) usually observations from station data (Themeßl et al., 2011). Statistical downscaling is often combined with error removal or bias correction (BC) of climate model data to derive reliable high resolution input data for impact models (Wood et al., 2004). While statistical downscaling is a computational efficient alternative usually based on accepted statistical methods it is not based on physically consistent processes and requires very long and reliable observational data series with a high network density to allow for robust performance.

These constraints can be overcome through dynamical downscaling techniques, which however are computational demanding and thus generally limited in the number of available scenarios (Fowler et al., 2007). The usual procedure in dynamical downscaling is the one-way nesting approach where the large scale forcing provided by either a GCM or reanalysis data serves as boundary conditions for the RCM run on a smaller domain with higher spatial resolution (Giorgi and Gutowski, 2015). This process is usually only performed one-way, meaning there is generally no feedback from the RCM to the boundary conditions provided e.g. through GCMs. As such two-way nesting is even more computational demanding, studies are so far mostly carried out on even smaller domain sizes (e.g. Bowden et al., 2012) however, great efforts are currently undertaken to provide large sets of fully coupled RCMs for e.g. the Mediterranean region (Ruti et al., 2016). To allow for model comparison, provide a benchmark for evaluation and resulting model improvement and for a largest possible uncertainty assessment (Giorgi et al., 2009), an overarching framework was created for dynamical downscaling through the Coordinated Downscaling Experiment that provides systematic downscaling for various domains, e.g. over Europe

(EURO-CORDEX). For EURO-CORDEX several modeling groups from various institutes contributed to provide a large range of RCM simulations at 0.44° and 0.11° resolution (Jacob et al., 2014). These downscaled climate scenarios consequently have a better representation of local features, such as topography, due to their higher resolution (Castro et al., 2005). Several studies showed the added value of RCM simulations over GCM results in various areas and for a multitude of variables especially over complex terrain (Di Luca et al., 2012; Feser et al., 2011; Rummukainen, 2016). A direct comparison between the two downscaling techniques however is challenging (Casanueva et al., 2016) and the performance of both approaches is highly dependent on season, variable, location and the metrics applied for evaluation (Yoon et al., 2012). However, statistical downscaling requires a time series that captures the full range of possible events, as the representation of extremes is potentially underrepresented in the observations, resulting in artifacts and inconsistencies when future extremes are of interest (Vrac et al., 2012). Despite the inconsistencies in temporal variability, dynamical downscaling also allows for better representation of spatial variability and potential changes therein (Flaounas et al., 2013). For these reasons and for comparability of different river basin studies as well as the availability of many RCM scenarios through the EURO-CORDEX initiative, dynamical downscaling is preferred over statistical downscaling in this thesis. Nevertheless, in case a high resolution and long time series of observations is available, a combination of both approaches, i.e. statistical downscaling of RCM simulations, shows great potential (e.g. Chen et al., 2012).

In most climate change impact assessment studies it is usually not feasible to consider a large number of climate projections for computational reasons and related challenges to cope with large data sets (Mendlik and Gobiet, 2016). Therefore, the available RCM ensemble is usually reduced and only a few simulations are considered. However, the selection of the 'right' simulations to include is usually not a trivial choice. Model selection is often based on the availability of models (e.g. Hagemann et al., 2013) or based on internal agreement upon the partners involved within the corresponding project (e.g. Muerth et al., 2013). More advanced alternatives are based on sensitivity to climate forcing (Flint and Flint, 2012) or the concept of Representative Climate Futures (RCF) as proposed by Whetton et al. (2012) and similarly by e.g. Kienzle et al. (2012). Here the climate change signals of available climate projections are included to derive scenarios based on temperature and precipitation changes such as the RCF *warm & dry*. These rather simple clustering approaches can also be extended to include multiple additional variables, such as shortwave downward radiation, relative humidity etc., to derive a subset of climate models to cover the entire spread of the available ensemble by means of clustering the climate change signals and a scoring approach (Wilcke and Bärring, 2016). Following the approach presented by Masson and Knutti (2011), the selection can also be based on model independence to maximize diversity and minimize model dependency as further elaborated by Mendlik and Gobiet (2016). Despite these approaches, a performance based model selection is still widely applied (Biemans et al., 2013; Deidda et al., 2013; Kotlarski et al., 2017; Pierce et al., 2009). Here, the climate models are selected based on validity, i.e. their ability to reproduce the historical climatology as

provided through reference data sets. This approach is based on the assumption that models that are capable of reproducing the historical climate are more credible and their validity will remain also for the future horizon which can be considered a prerequisite for the models' ability to project future climate (Kotlarski et al., 2014).

Despite the increased resolution and the added value of RCMs over the coarser GCM simulations (Di Luca et al., 2012; Feser et al., 2011; Rummukainen, 2016) and their general capability of reproducing the most important regional climatic features and a mostly good representation of regional climate (Maraun and Widmann, 2015), RCMs are still prone to important biases (Kotlarski et al., 2014). If general circulation patterns for example are slightly misplaced in the GCM the nested RCM is neither expected nor able to correct these features and resulting climatology on the regional scale can be largely biased (Giorgi and Gutowski, 2015). To assess model biases or model errors, RCM or GCM simulations are generally compared to regional reference data sets. Recent studies assessed biases in the climatology of current RCMs (e.g. García-Díez et al., 2015; Kotlarski et al., 2014), as well as for extremes (e.g. Nikulin et al., 2011). For climate change impact assessment studies a good representation of the climatology but also of daily variations and extremes are of high importance, however, RCMs are usually not capable of reproducing these on the regional scale. Bias correction (BC) or bias adjustment is, despite being controversially discussed (Ehret et al., 2012), thus often a necessary step (Dosio, 2016; Muerth et al., 2013; Piani et al., 2010; Teutschbein and Seibert, 2013). Various techniques to adjust model biases have emerged over the past years. Among these distribution based scaling approaches showed great potential for BC of precipitation (Chen et al., 2013; Themeßl et al., 2011) and temperature (Räisänen and Räty, 2013). Among these, Quantile Mapping (QM) approaches are widely used to adjust model biases (Addor et al., 2016; Pierce et al., 2015) based on the distributions of historical data. This correction approach fits the distribution of precipitation in the RCMs to the distribution (quantiles) of the reference data set. However, BC approaches usually assume stationarity, i.e. the observed bias will remain constant in future projections, which might lead to over- or underestimated changes in precipitation and temperature (Bellprat et al., 2013; Christensen et al., 2008). Additionally, the issue of inter-variable dependencies remains (Wilcke et al., 2013) as typically each variable (e.g. precipitation) is treated separately. Statistical bias correction methods, such as QM, are also only effective if the systematic bias is not due to a misrepresentation of the general climatology in RCMs introduced e.g. through location errors (Maraun and Widmann, 2015).

As presented in Figure 1, reference data sets play a key role in the hydro-climatic modeling chain as they are involved in almost all components, either for the evaluation of climate models, for their selection, for the application of a successful BC and lastly for the calibration and simulation of the reference period with hydrological models. Although several meteorological variables, e.g. air temperature, relative humidity, wind speed and incoming shortwave downward radiation, are important and required for hydrological applications, this thesis will focus mostly on precipitation as it is key variable in the hydrological cy-

cle and one of the most widely used climate variables in other studies (Schneider et al., 2016; Daly et al., 2017). The data sets introduced in the following are therefore mentioned with a focus on precipitation, although many of these also contain multiple other variables. Despite local observations from station data there is a growing number of gridded data sets available. These data sets differ in their domain size, spatial and temporal resolution, and originate from different sources and methods and usually a trade-off between spatial resolution and domain size must be made. Gridded data sets are either reanalysis products, derived through remote sensing or constructed through interpolation of station observations. Reanalysis data sets are constructed similar to forecasts produced by numerical weather prediction (NWP). In contrary to NWP the observations are reanalyzed and the model is evaluated against observations at every time step and adjusted if needed. Through this procedure a three-dimensional state of the atmosphere is constructed and provides various variables in a continuous way (Bengtsson et al., 2004). These data sets are usually global products and provide information on various variables over a large domain and cover a large time period however, they usually lack the high spatial resolution required for regional or catchment scale studies. Widely used reanalysis data sets are the Twentieth Century Reanalysis data set provided through the National Oceanic and Atmospheric Administration (NOAA) (Compo et al., 2011) and the ERA-Interim data set constructed at the European Centre for Medium-Range Weather Forecasts (ECMWF) (Dee et al., 2011). Remotely sensed data sets for precipitation are commonly derived from passive microwave information (Joseph et al., 2009) or combined with information from infrared sensors (Ashouri et al., 2015). High-resolution data sets based on station observations are usually only available on country level (e.g. over Spain by Herrera et al., 2012) or cover a specific geographical region e.g. the Alps (Isotta et al., 2014). However, coarser data sets based on station data are available at the continental scale, e.g. over Europe, such as the well-known E-OBS data set (Haylock et al., 2008).

The final component of the hydro-climatic modeling chain is the hydrological model itself. Available models differ in their complexity, dynamic and spatial scale (Moradkhani and Sorooshian, 2009) and can be lumped, conceptual or physically based and fully distributed. The choice of the model is dependent on the desired application but also driven by the availability of data. Simple models usually only require a few parameters and include e.g. relatively simple evapotranspiration schemes physically based models require more input parameters and apply a more complex evapotranspiration scheme (Velázquez et al., 2013). The hydrological model applied in this study, WaSiM, allows for a flexible spatial resolution and variable setup regarding the complexity of the model. In the study carried out within this thesis, WaSiM (Schulla and Jasper, 2007) was run at a 100 m grid with the Penman-Monteith equation to calculate potential evapotranspiration (Gampe et al., 2016a). Depending on the study area, the modular structure allows to include or exclude modules such as snow and irrigation which would require additional input data and parameters.

1.3 State-of-the-art: uncertainty sources and assessment in hydrological climate change impact studies

As presented in Figure 1, various components are involved in the hydro-climatic modeling chain each introducing additional uncertainty to the resulting climate change impacts. Uncertainties in the climate scenarios can attributed to three main sources: scenario, global and regional model uncertainty (Déqué et al., 2007, 2012). Scenario uncertainty originates from the emission scenario and the climate models' response to these for the SRES scenarios or the radiative forcing directly introduced through the more recent RCPs. Global model uncertainty originates from the various GCMs providing different boundary conditions for RCMs. The latter are responsible for regional model uncertainty as they resolve the regional climatology based on the provided boundary conditions. Furthermore, internal variability in the climate system represented in the GCMs and RCMs and contributes to the overall uncertainty as well as the region of interest (Giorgi et al., 2009). In addition to the uncertainties in the climate scenarios, observational uncertainty introduced by the reference data sets included in the study, adds another layer of uncertainty to the modeling chain. Finally, the choice of the hydrological model and the parameterization of the model add additional uncertainty. Precise knowledge of uncertainties in climate projections and their quantification is essential to increase the credibility of the projections and increase the confidence in the data sets to derive robust adaption strategies (Foley, 2010). Additionally, precise information on the major uncertainty sources is crucial to improve climate projections and their usage in impact studies and provides valuable information for e.g. model selection if only a subset of available climate models is considered.

A common way to quantify the contribution of various uncertainty sources included in climate projections is an assessment of the explained variance by means of an analysis of variance (ANOVA) as presented by von Storch and Zwiers (1999). The approach is based on a decomposition of the overall variance of the ensemble for a given variable, typically the climate change signal, which can be assumed as a sum of the individual variances arising from the different sources of uncertainty. Following the ANOVA approach, the variance of an ensemble of climate models V can be written as sum of the variance originating from the GCMs G, the RCMs R and the RCP S and their respective interaction terms as:

$$V = G + R + S + GR + GS + RS + GRS$$

A decomposition of the variance then allows to estimate the contribution of each uncertainty source to the overall variance of the ensemble. This approach does thus not allow for an absolute assessment of uncertainties but rather estimates the relative contribution. The ANOVA approach was applied in various studies to quantify the uncertainty in climate projections (Déqué et al., 2007, 2012; Ferro, 2004; Prein et al., 2011). A detailed description of the method and how it was implemented in this thesis can be found in Gampe et al. (2016b) and Gampe et al. (2018).

1.3.1 Observational uncertainty

As described in section 1.2 available reference data sets originate from various sources such as reanalysis, station observations or remote sensing. Although all data sets consider and assimilate observations at some point, they cannot be characterized as observations hence the term 'observational uncertainty' is somewhat misleading. Within this thesis 'observational uncertainty' generally refers to the uncertainty originating from the reference data sets, regardless of being observational or just observationally constrained similar to Sillmann et al. (2013a). Previous studies revealed a large spread in available reference data sets that can be of similar magnitude as the CMIP5 spread depending on the applied indices, variable and region (Palazzi et al., 2013). Similar to climate models, there is generally a higher agreement between available reference data sets on temperature than on precipitation. This is due to the higher spatial variability of precipitation e.g. through orographic effects or convective events, but also originates from measurement errors especially through wind effects maximized for solid precipitation (Rasmussen et al., 2012). As gridded observational data sets are constructed by interpolation of station data, the choice of the interpolation method can have additional impact on the resulting precipitation in a particular grid cell (Chen et al., 2017). These differences increase with lower station density, which is both typical and especially relevant for higher elevations (Prein and Gobiet, 2017).

Sillmann et al. (2013a) highlighted considerable differences in reanalysis data sets also when compared to gridded observations especially when extreme indices are evaluated. As presented by Henn et al. (2018), the uncertainties in gridded precipitation data sets significantly increase over complex topography with differences in annual precipitation sums of 200mm for the Western United States. However, the related biases and uncertainties might be even larger as indicated by a relatively simple comparison of precipitation and streamflow. A considerable contribution to this topic was the study carried out by Prein and Gobiet (2017) that included several observational data sets and demonstrated the added values of high resolution data sets. Their results highlight the need to consider multiple data sets, possibly from different sources, for a robust evaluation of climate models and a correction of model biases. To this point climate model evaluation is usually based on a singular reference data set (e.g. Deidda et al., 2013; Frei et al., 2006; Kotlarski et al., 2014) which itself might include a considerable bias as demonstrated by Prein and Gobiet (2017). While other methods for model selection have been developed, the credibility of climate models can only be assessed by comparing to climate observations and thus is of high importance (Giorgi et al., 2009). Alexander and Arblaster (2017) evaluated recent CMIP5 simulations over Australia comparing one high and one coarse resolution data set as reference. They revealed differences in bias detection especially for extreme indices. A great contribution to the field was made by Kotlarski et al. (2017) who evaluated observational uncertainty on the pan-European scale focusing on various regions typically used to analyze and evaluate climate models. They revealed considerable biases of up to 2° C for temperature and differences of 70 mm for precipitation over Europe. Major difference

can be found for wet day frequency and extreme precipitation. Depending on the region this also has considerable influence on model selection, especially where high resolution data sets are available. Gómez-Navarro et al. (2012) found large differences for model scoring and ranking over Spain using three reference data sets for minimum and maximum temperature as well as precipitation. Sunyer et al. (2013) showed that the differences in observational data sets are mainly not due to differences in their spatial resolution. They revealed large differences for bias detection over Denmark especially for extreme precipitation where the data sets did not even agree in the sign of the bias. Herger et al. (2018) introduced observational uncertainty in the model selection process. These studies highlight the need to consider multiple reference data sets not only for model evaluation and selection, but also for BC. Up to now many studies were published focusing on BC inter comparison based on various methods to adjust model bias (e.g. Chen et al., 2013; Teutschbein and Seibert, 2012). However, most BC studies only consider a single data set (e.g. Addor et al., 2016; Jeon et al., 2016) rather than multiple data sets and thus neglect observational uncertainty. Despite all these differences, observational uncertainty was generally lower than model uncertainty from climate models (Kotlarski et al., 2017) and the magnitude of GCM bias exceeded the observational uncertainty (Alexander and Arblaster, 2017). Nevertheless it is well known at this point that observational uncertainty contributes considerably to the overall uncertainty and thus shall not be neglected (Bellprat et al., 2012; Kotlarski et al., 2017; Prein and Gobiet, 2017).

1.3.2 Uncertainties of climate projections

Despite the structural harmonization in the CMIPs, the individual GCMs differ considerably in their representation of processes, model structure and have different physical parameterization (Kay et al., 2015). Recent CMIP5 models for example have similar and uniform anthropocentric emissions, while natural emissions, e.g. originating from soil nitrogen or climate-sensitive lightning, differ considerably (Lamarque et al., 2013). This results in large differences regarding their representation of various processes and causes, among other reasons, different projections for future climate. In addition, the CMIP model ensemble is a so called ensemble-of-opportunity as the contribution of a modeling group to such a project is largely dependent on non-scientific reasons, such as the availability of a climate model or the required financial and computational resources to fulfill the simulations, resulting in a model sampling that is neither random nor systematic (Tebaldi and Knutti, 2007) and thus challenging the interpretation of such model ensembles (Kay et al., 2015). It is nevertheless important to consider more than one of these simulations. The preferable use of such multi model ensembles (MMEs) over a single model realization approach to assess climate change impacts is well established (Masson and Knutti, 2011; Mendlik and Gobiet, 2016). These MMEs largely capture model uncertainty from the RCMs, global model uncertainty from the GCMs as well as the underlying physical uncertainties and sample internal variability (Collins et al., 2011a), i.e. the natural variability of the climate system in the absence of external forcing due to the chaotic nature of the system (Deser et al., 2012). While MMEs

are considered to capture most of the uncertainty involved and their application became a standard procedure, the quantification of the uncertainty sources or their relative impact are usually not explicitly addressed. The magnitude of the relative contribution of each source of uncertainty to the overall uncertainty however is dependent on the assessed variable, such as precipitation or air temperature (Déqué et al., 2007; Gampe et al., 2016b) and the considered time horizon.

Scenario uncertainty depicts a major source of uncertainty that introduces further uncertainty not fully represented by the model spread due to the unpredictability of human activity and resulting future emissions that influence the climate system (Foley, 2010). This is especially challenging as uncertainty originating from other sources generally arise from model imperfections or different parameterization schemes and can potentially be reduced (Giorgi et al., 2009), however, due to the complexity of the climate system parameterizing and neglecting some processes will always be unavoidable. As the scenarios, both SRES and RCP, follow a certain path until the year 2100, the contribution of scenario uncertainty is highly dependent on the considered projection horizon as only little uncertainty is introduced for the near future. Prein et al. (2011) showed that the relative contribution of scenario uncertainty is negligible for the near future up to 2050 while it contributes with up to 35% to the overall uncertainty by the end of the 21st century for the variables air temperature, geopotential height and specific humidity. Déqué et al. (2007) showed a significant contribution of scenario uncertainty to the overall uncertainty only for summer air temperature. The results of Hawkins and Sutton (2009) reveal the large regional and temporal differences of various uncertainty sources. Their findings confirm that scenario uncertainty is negligible in the near future while being the major source of uncertainty by the end of the 21st century for global mean air temperature, as presented in Figure 3. For precipitation in contrary, the contribution is negligible for most areas of the world also in the distant future (Hawkins and Sutton, 2010). The higher relative contribution of scenario uncertainty to temperature projections compared to other climate variables is generally well established (Déqué et al., 2007; Fowler et al., 2007). Two potential strategies for climate change impact assessment studies emerge from the unpredictable nature of scenarios based on human activity. A common approach follows the MME strategy and considers multiple available scenarios in the respective impact study to consider a large range of potential scenarios (e.g. Prudhomme et al., 2014; Schewe et al., 2014). Alternatively, the selection of scenarios can be based on feasibility thus exclude apparently unrealistic emission scenarios such as the low emission scenario RCP 2.6 from the ensemble (Mora et al., 2013; Sanford et al., 2014). Following this approach recent studies focus on RCP8.5 as it can be considered a business-as-usual scenario based on recent emissions (Gerstengarbe et al., 2015; Kay et al., 2015; Maure et al., 2018). With focus on decision makers and resulting adaptation strategies, a selection of optimistic and pessimistic scenarios, thus potentially consider the attitude and impact of the decision maker, can also be valuable (Giuliani and Castelletti, 2016).

Hawkins and Sutton (2009) showed that model uncertainty in temperature projections

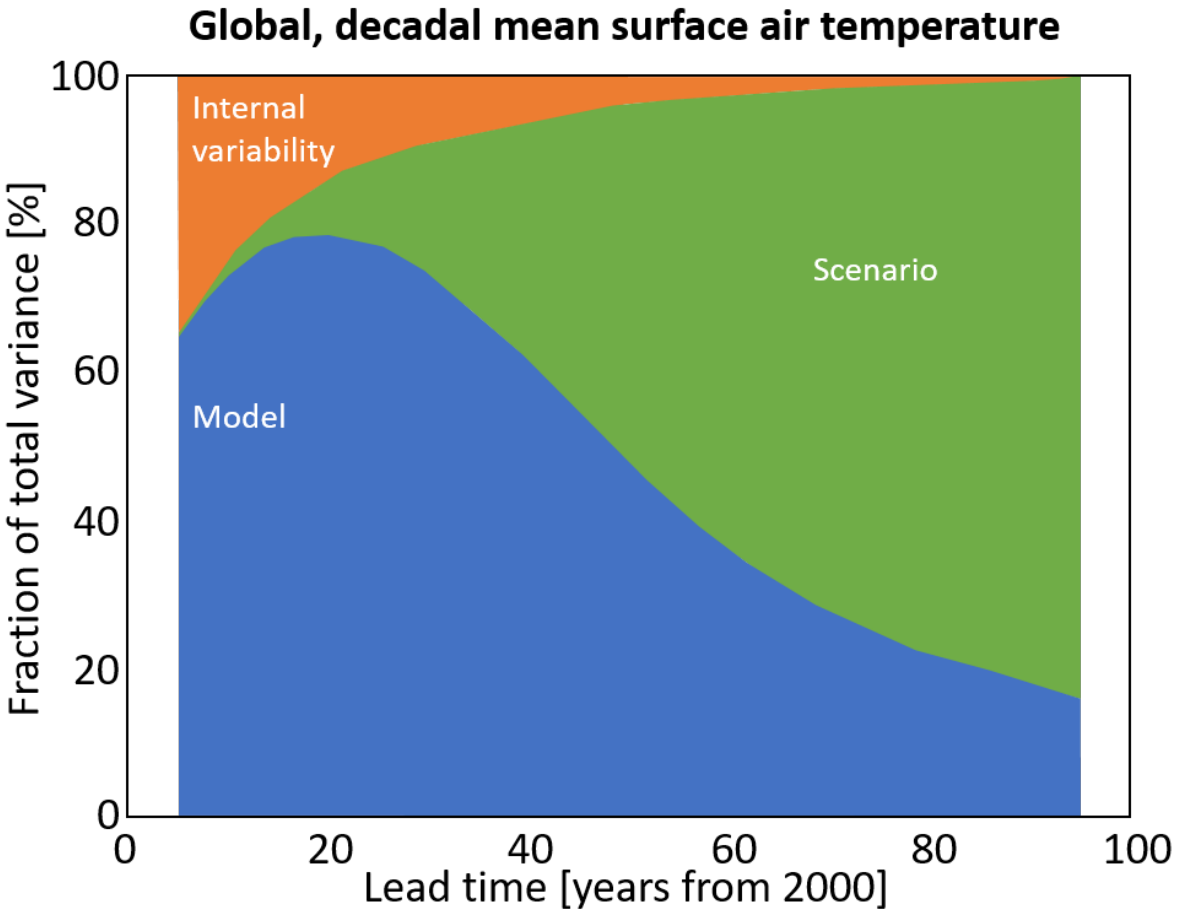


Figure 3: The fraction of total variance in global, decadal mean air temperature projections explained depending on lead time (adapted from Hawkins and Sutton, 2009).

from GCMs is dominant for mid-century, i.e. the 2040 horizon, while it is significantly reduced by the end of the century as scenario uncertainty becomes more important. These results were confirmed in the follow-up study (Hawkins and Sutton, 2010) where model uncertainty from the GCMs dominates clearly until mid-century for air temperature. For precipitation, model uncertainty is the dominant source throughout the entire 21st century (Hawkins and Sutton, 2010). The findings presented by Prein et al. (2011) confirm previous studies with model uncertainty being the dominant source of uncertainty for precipitation and temperature for both, mid-century and end-of-the-century projections. For air temperature, model uncertainty contributes with 70-85% to the overall uncertainty for mid-century projections and, with the slight exception of the winter season, with over 50% for the end of the century. For precipitation, 70-90% of the uncertainty are introduced through model uncertainty with no significant difference in the projection horizon. The results by Hawkins and Sutton (2009, 2010) and Prein et al. (2011) are however based on the CMIP3 GCM ensemble and do not account for uncertainty introduced through dynamical downscaling. Déqué et al. (2007) concluded that uncertainty originating from GCMs generally contributes most to the overall ensemble spread over Europe at the end of the 21st century for precipitation and air temperature. With the exception of summer precipi-

tation, a similar analysis with a different ensemble of available RCMs resulted in a similar contribution of GCM and RCM to the overall uncertainty (Déqué et al., 2012). Fowler et al. (2007) conclude that the contribution of RCMs is higher for precipitation than for temperature, however the main source of uncertainty is due to the structure and different physical parameterization schemes of the driving GCMs. Similar results were presented by Giorgi and Gutowski (2016) over Europe with the GCM dominating the uncertainty for both, air temperature and precipitation. Only summer precipitation, which is likely due to convective events and the corresponding parameterization schemes in the RCMs, makes an exception here. As shown by Graham et al. (2007), this translates directly to hydrological quantities where the GCMs depict the largest source of uncertainty for future runoff projections followed by RCM and scenario. These results indicate the importance of GCMs for ensembles of climate models and model selection. Déqué et al. (2007) conclude that the number of GCMs should at least match the number of RCMs for an efficient analysis of uncertainty quantification.

In near future projections the internal variability of the climate system plays an important role for both, precipitation and temperature (Hawkins and Sutton, 2010). For global air temperature projections, internal variability explains up to 40% of the total uncertainty in the near future while being negligible for the end of the century as presented by Hawkins and Sutton (2009) and shown in Figure 3. Following Prein et al. (2011) internal variability accounts for around 10% for air temperature and slightly more for precipitation projections even at the end of the century. Stronger than for the other uncertainty components, there is a high seasonal variability in the contribution of internal variability especially for precipitation. However, the assessment of internal variability in the mentioned studies is based on ensembles-of-opportunity and thus derived from statistical measures as presented by Hawkins and Sutton (2010). Reliable assessment of internal variability of the climate became a widely researched issue over the past years. A perturbed physics ensemble, e.g. the same model with small perturbation in the initial conditions and chaotic evolution of climate, is needed to fully quantify the role of natural variability (Deser et al., 2012; Gutowski Jr et al., 2016), although bootstrapping shows promising results to derive a robust estimation (Addor and Fischer, 2015). Internal variability, often considered as 'noise' in climate models (Deser et al., 2012), plays a crucial role in the time of emergence of the climate change signal, i.e. when the signal can be clearly separated from the noise (Hawkins and Sutton, 2012). As presented by Kay et al. (2015) a large ensemble produces a spread similar to CMIP5 simulations purely based on internal variability. The contribution of internal variability to the overall uncertainty is furthermore characterized by strong regional differences as shown by Hawkins and Sutton (2009).

1.3.3 Uncertainties of hydrological models

Despite the uncertainties in the related meteorological forcing discussed in the previous sections, uncertainty in hydrological models arises mainly from parameter uncertainty and due

to model structure (Liu and Gupta, 2007). The latter can be addressed by applying several hydrological models of different complexity levels with the same forcing. Velázquez et al. (2013) applied an ensemble of four hydrological models of different complexity over Québec and Bavaria. They concluded a rather low level of model uncertainty for high flows however with high regional differences, higher uncertainty for mean flows and even more pronounced impact on low flows. Schewe et al. (2014) analyzed water scarcity under future climate and revealed, despite large regional differences, that the spread of an ensemble of five global hydrological models is of similar or even greater magnitude than the spread introduced through GCMs. Similar results were obtained by Wada et al. (2013) who concluded that the choice of the global hydrological model outweighs the uncertainty introduced through GCM. Hattermann et al. (2018) showed that the contribution of the hydrological models to the overall uncertainty is highly dependent on the hydrological regime of the river and varies seasonal, however, the uncertainty introduced by the GCM usually obscures the other uncertainty sources.

Parameter uncertainty is often altered by the calibration procedure. The impact of the selection of the period used for calibration was analyzed in the study presented by Seiller et al. (2012) where an ensemble of lumped models was calibrated and evaluated under periods with different climate conditions. The authors found strong differences and large uncertainties if only one model is selected, however a robust result for a larger number of models as model sensitivity balances the different responses. Similar results were obtained by Brigode et al. (2013) for 89 catchments in France, who attributed the lack of robustness, i.e. the dependence of the model parameters on the calibration period, as the major source of uncertainty. In the study presented by Harder and Pomeroy (2014) the parameter uncertainty of different empirical schemes to separate solid and liquid precipitation was analyzed, revealing substantial uncertainties arising from non-physical parameters of 160mm difference in snow water equivalent resulting in changing snow cover duration of up to 26 days and thus also on discharge from snow melt. A similar study over Québec revealed large uncertainties due to different snow modules and evapotranspiration schemes of different models (Seiller and Anctil, 2014). Muerth et al. (2013) showed that, e.g. for high flows, model uncertainty dominates over uncertainty arising from BC as the corresponding processes are more dependent on model structure than input data. Uncertainty introduced by natural variability generally dominated uncertainty in the study presented by Seiller and Anctil (2014).

The issue of parameter uncertainty increases if the study area lacks the required data for model parameterization. The issue of data scarcity challenges hydrological modeling in many catchments in the Mediterranean (Cudennec et al., 2007; Gunkel et al., 2015). As elaborated in various studies, remote sensing data can be included in the calibration or parameterization process to reduce uncertainties in hydrological model results (Karimi and Bastiaanssen, 2015).Silvestro et al. (2015) used land surface temperature and soil moisture information based on satellite imagery to calibrate a hydrological model for two Italian catchments to increase model robustness. Brown and Pervez (2014) used satellite images

to derive the Normalized Difference Vegetation Index (NDVI) to assess and parameterize irrigated areas over the United States. Remote sensing data is also applied to calibrate and validate snow cover (e.g. Shrestha et al., 2014) as well as soil moisture and evapotranspiration for improved model calibration (Kunnath-Poovakka et al., 2016) and validation (Meyer et al., 2016).

1.4 Aims and goals of this thesis

According to recent climate projections the Mediterranean region is likely to experience severe impacts on water resources. It is of high importance to provide robust estimates of future climate change impacts especially in areas suffering from water scarcity already under the current climate. This can be obtained through hydrological modeling, which is however often hampered by data scarcity especially in the Mediterranean. The thesis provides robust estimates on water resources in the Mediterranean as well as the quantification of the related uncertainties in selected river basins and was carried out within the projects CLIMB and GLOBAQUA, both funded under the European Union's Seventh Programme. The different steps in the uncertainty cascade involved in hydrological climate change impact studies as presented in Figure 4 are addressed. The cumulative thesis is compiled of four scientific publications and the relevant research questions address three main research areas and will be presented in the following.

1.4.1 The role of observational uncertainty in the hydro-climatic modeling chain

Although observational uncertainty is more and more included in climate change and impact studies and some studies highlighted the importance of observational uncertainty, it is still underrepresented and widely neglected. Precipitation is considered to be the most important variable for hydrological climate change impact assessment studies (Giorgi and Gutowski, 2015). As observational uncertainty is considerably larger for precipitation than for temperature, there is an emphasis on the assessment of precipitation data sets. Similar to the uncertainty assessment of climate projections, studies of an extensive assessment of observational uncertainty focused on larger domains rather than hydrological catchments. It is thus one of the key questions how available precipitation data sets differ on the catchment scale, especially when compared to interpolated station data. One goal of this thesis is to provide insights in the performance of available reference grids with respect to climatology and hydrologically relevant indicators such as consecutive dry and wet days. As available gridded products differ considerably in their spatial resolution it is important to assess and evaluate if coarse data sets perform worse and thus can or should be discarded from the analysis. It is expected that observational uncertainty contributes considerably to the overall uncertainty, especially over complex terrain and that the selection of a reference data

set should be carried out thoroughly and not be an arbitrary choice as in most hydrological modeling studies. Based on this, two research question emerge and can be formulated as:

Q1: How well do gridded precipitation products perform on the catchment scale especially over complex topography?

Q2: What is the relative contribution of observational uncertainty to the overall uncertainty in the hydro-climatic modeling chain?

It is expected that observational uncertainty has a significant impact on climate model evaluation, bias correction and model selection. However, the issue of observational uncertainty is so far neglected in the literature and studies mostly focus on a comparison of bias correction methods. One of the key questions to answer is hence to evaluate the impact of observational uncertainty on bias correction and projected future precipitation. Additionally, as model selection is still widely based on validity, it is important to determine how the choice of the reference grid influences the evaluation and selection of RCMs and thus changes the future projection envelope. This consequently leads to the third research question that covers observational uncertainty:

Q3: What is the influence of observational uncertainty on post-processing, i.e. bias correction, climate model evaluation, model selection and the resulting future projections?

1.4.2 Assessment of uncertainties in the recent EURO-CORDEX climate projections

Climate projections include a range of uncertainty originating from different sources. Previous studies assessed the relative contribution of GCM, RCM and scenario uncertainty usually for precipitation and temperature. As the potential of direct usage of RCM data to assess future water scarcity is explored, it is also important to assess the contribution of uncertainty sources for all variables of the water balance. In addition, previous studies assessed the relative contribution to the overall variance on the regional scale, while the application is usually carried out on a finer scale, e.g. hydrological catchments. It is therefore important to quantify the related uncertainties also on the catchment scale as the role of the RCMs might be underestimated in previous studies due to a smoothing of errors over a larger area. Additionally, all studies mentioned in the previous sections have been carried out using the CMIP3 projections that are based on the SRES emission scenarios. As the most recent CMIP5 projections are based on the RCPs they introduce radiative forcing externally instead of model-dependent calculation based on the provided emissions. Thus the relative contribution of the RCPs but also GCMs and RCMs to the overall uncertainty has to be evaluated again to allow for model generation comparison and support model selection. Following this, the fourth research question of this thesis emerges:

Q4: What is the relative contribution of each of the uncertainty sources involved in recent regional climate projections on the catchment scale and for all the variables of the water balance in the downscaled CMIP5 climate projections?

1.4.3 Climate change impact assessment under data scarcity

Hydrological modeling, especially in the Mediterranean region, is often hampered by the lack of available data for parameterization, calibration and validation. It is however of high relevance to assess water resources under future climate especially in areas already suffering from water scarcity. However, issues remain in how hydrological models can provide robust results in these areas. The great potential of remote sensing information for hydrological modeling has been illustrated in the previous section, however, the processing of several remote sensing images can be time consuming and inefficient. It is therefore crucial to provide a computational efficient, affordable method to include remote sensing information in the hydrological modeling chain that allows to increase the robustness of the model results without requiring additional input data. Based on this the following research question can be formulated as:

Q5: How can remote sensing data be applied to improve the robustness of a hydrological model while being computational efficient?

As the setup, calibration and validation of hydrological models is a rather complex task, there is also a need explore alternative ways to estimate future changes on the hydrological quantities. Recent RCMs include various schemes of the land surface and land use (Kotlarski et al., 2014) thus also provide information on evapotranspiration and surface runoff. This allows to assess changes in the water balance and provide estimates on future water scarcity in selected Mediterranean river basins. This leads to the following research question:

Q6: Is the Mediterranean region likely to experience increased pressure on the water balance and thus likely to be exposed to increased water scarcity under future climate?

2 Scientific Publications

This cumulative thesis is comprised of four peer-reviewed scientific publications. Three of these have been published in peer-reviewed international journals an additional paper has been submitted to a peer-reviewed journal and is currently under review. In the following the publications are listed with the publishing journal and the journals impact factor according to the 2016 Thomson Reuters Journal Citation Report. The papers are not presented in chronological order, but rather follow the order of the research questions presented in the previous section and the hydro-climatic uncertainty cascade as shown in Figure 4. Additionally, the contribution of each author to the corresponding publication is provided. Paper I evaluates observational uncertainty in an Alpine catchment by comparing an ensemble of available gridded precipitation data sets based on hydrological relevant indicators. Paper II applies the selected reference data sets to bias correct the RCM ensemble and evaluates the impact of observational uncertainty for post-processing, i.e. bias correction, and model selection. Paper III addresses and quantifies the uncertainties in recent climate projections provided through the EURO-CORDEX initiative. In addition, future water scarcity in selected Mediterranean river basins are estimated. Paper IV depicts an example for the typical hydro-climatic-modeling chain, as a subset of available RCMs is selected, bias corrected and then used to drive a physically based hydrological model. This study covers uncertainty originating from the hydrological model and highlights the benefit of remote sensing to derive robust model results.

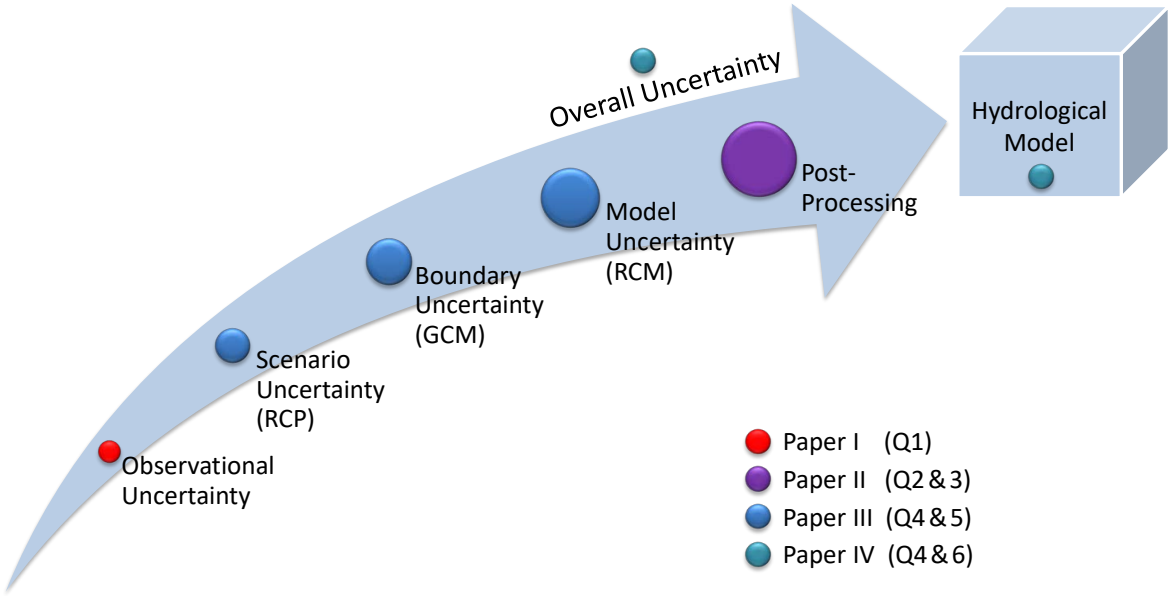


Figure 4: Uncertainty sources included in the climate forcing of hydrological models in the typical hydro-climatic-modeling chain. Width of the arrow symbolizes the accumulated uncertainty, colored circles indicate the corresponding papers which address the contribution of these sources. The addressed research questions are given in brackets.

2.1 Paper I: Evaluation of Gridded Precipitation Data Products for Hydrological Applications in Complex Topography. Hydrology

Gampe, D.,& Ludwig, R. (2017). *Evaluation of Gridded Precipitation Data Products for Hydrological Applications in Complex Topography.* Hydrology, 4(4), 53.

Paper I focuses on observational uncertainty in available gridded precipitation data sets that could be applied to drive a hydrological model over complex terrain. The catchment assessed is the Adige river basin located in Northern Italy that was also addressed in Paper II & III. This catchment was selected as it is offers a relatively dense network of precipitation gauges and a robust assessment of future water quantities is of high importance due to intensively irrigated agriculture and hydro power generation in the catchment. A high resolution gridded precipitation data set was constructed from station measurements using the interpolation tool provided within the hydrological model WaSiM, as this is typically how the meteorological forcing for a hydrological model is constructed if station data are selected. This grid serves as a reference grid to evaluate an ensemble of available reanalysis and observational as well as one remote sensing data set using several indicators of high relevance for hydrological modeling such as consecutive dry and wet days. The coefficient of variation is then applied to evaluate contribution of spatial resolution to the overall spread of the ensemble. This paper adds to the discussion on observational uncertainty on the catchment scale with the inclusion of multi-source data sets and serves as an example for a tailored analysis for hydrological climate change impact assessment studies. This paper addresses the following research question:

Q1: How well do gridded precipitation products perform on the catchment scale especially over complex topography?

Author's contributions: The study was carried out by D. Gampe, supervised by R. Ludwig. Both authors contributed to the concept of this paper, while the first author implemented the calculations and analyses and prepared the first version of the manuscript.
Status: published
Journal: Hydrology (MDPI)
Impact factor: NA (Journal indexed in GeoRef and Scopus [from Vol. 5; 2018])

hydrology

MDPI

Article

Evaluation of Gridded Precipitation Data Products for Hydrological Applications in Complex Topography

David Gampe * and Ralf Ludwig

Department of Geography, Ludwig-Maximilians-Universität München (LMU), 80333 Munich, Germany; r.ludwig@lmu.de

* Correspondence: david.gampe@lmu.de; Tel.: +49-(0)-89-2180-6648

Received: 6 October 2017; Accepted: 11 November 2017; Published: 16 November 2017

Abstract: Accurate spatial and temporal representation of precipitation is of utmost importance for hydrological applications. Uncertainties in available data sets increase with spatial resolution due to small-scale processes over complex terrain. As previous studies revealed high regional differences in the performance of gridded precipitation data sets, it is important to assess the related uncertainties at the catchment scale, where these data sets are typically applied, e.g., for hydrological modeling. In this study, the uncertainty of eight gridded precipitation data sets from various sources is investigated over an alpine catchment. A high resolution reference data set is constructed from station data and applied to quantify the contribution of spatial resolution to the overall uncertainty. While the results demonstrate that the data sets reasonably capture inter-annual variability, they show large seasonal differences. These increase for daily indicators assessing dry and wet spells as well as heavy precipitation. Although the higher resolution data sets, independent of their source, show a better agreement, the coarser data sets showed great potential especially in the representation of the overall climatology. To bridge the gaps in data scarce areas and to overcome the issues with observational data sets (e.g., undercatch and station density) it is important to include a variety of data sets and select an ensemble for a robust representation of catchment precipitation. However, the study highlights the importance of a thorough assessment and a careful selection of the data sets, which should be tailored to the desired application.

Keywords: reference precipitation; precipitation comparison; gridded data sets; hydrological impact; reanalysis data sets; remote sensing data sets; observations; meteorology; uncertainty

1. Introduction

Precipitation plays an important role in the hydrological cycle and is one of the most widely used climate variables [1,2]. While there is a clear link between the amount, intensity, and distribution of precipitation to various processes in the ecosystem [3], this relation is nonlinear, and heavy precipitation does not necessarily result in high river discharge [4,5]. Additionally, the link of heavy precipitation and flood occurrence is challenging to assess and quantify [6]. Nevertheless, accurate assessment of precipitation is of utmost importance as it provides the meteorological input for hydrological and other impact models and studies. As precipitation varies greatly in space and time [7,8], gridded precipitation information in high temporal and spatial resolution is required. The need for such high-resolution data sets is crucial and not only limited to hydrological applications, but many other fields, such as evaluation of the performance of climate simulations and possible detection and adjustment of model biases. The need for high resolution is increased over areas with complex topography with high spatial variability [9].

Available gridded precipitation datasets differ in their domain size, spatial and temporal resolution, and originate from different sources and methods. Usually, a trade-off between spatial

resolution and domain size must be made. While global and continental data sets, provided e.g., through the National Oceanic & Atmospheric Administration (NOAA) [10], the European Centre for Medium-Range Weather Forecasts (ECMWF) [11] and the German Weather Service (DWD) [12], provide information on precipitation over a large domain and cover a large time period, they lack the high spatial resolution required for regional or catchment scale studies. High-resolution data sets are usually only available on a country level [13] or cover a specific geographical region [14]. These data sets are either reanalysis products [10,11,15–17], derived through remote sensing [18,19] or interpolated station observations [12,14,20]. Most of these data sets use station data either directly, or assimilate observed precipitation at some stage. The density of the included stations differs greatly between, but also within the data sets, restricting the effective resolution and spatial consistency [21]. Recent efforts were made to merge different data sets to achieve a better representation and combining the advantages of each data set [22], as well as collecting sub-daily station observations [23] to increase the temporal resolutions. However, observations are prone to severe undercatch of precipitation, which is amplified in case of solid precipitation and over mountainous areas [21,24,25]. Therefore, there is merit in including data sets from a variety of sources for a robust estimation of reference precipitation.

Various studies to assess the performance of these data sets were carried out in recent years, revealing considerable differences between the products. Most of these studies focus either on global, continental, or regional performance assessment often aiming in the evaluation of Regional Climate Model (RCM) performance [2,26]. Regional studies assess the benefit of higher resolution and consequently a better topographic representation, the issues of station density and the representation of regional weather phenomena [21]. Detailed source specific assessment showed limitations in the representation of several weather patterns for reanalysis data sets [27] and the restriction of satellite information under specific cloud cover conditions [15,28]. Gridded precipitation data products are very often applied as meteorological input for hydrological modeling exercises in data scarce areas to bridge the data gap at the catchment or smaller regional scale [29,30]. It is therefore important to also assess the performance of the data set at smaller scales over hydrological basins. Some studies detect the deviations on the catchment scale; however they focus on reanalysis [27] or selected regional data sets [31]. Additional efforts were carried out in the past focusing on 3-hourly and daily error analysis for satellite precipitation products over alpine catchments. They revealed large errors especially in the summer months and showed the difficulties to adequately address errors in satellite products with respect to the station density of in situ measurements [32]. Other studies identified the impact of different satellite precipitation products on the reproduction of flood events in Northern Italy. These studies showed both the potential and advantages of these data products, but also their limitations [33,34].

However, there is still a need to evaluate long-term uncertainties and differences in existing precipitation products on the catchment scale, and exploiting as many of the available products as possible from different sources. It is important to assess not only the climatology, but also the daily extremes. This study contributes to recent findings by including an ensemble of eight available gridded precipitation data sets derived from observations and through reanalysis or remote sensing. These are evaluated against a high-resolution reference grid derived from local observations over a river basin for the period 1989–2008, with a focus on hydrological implications. Therefore, the comprehensive assessment does not only focus on the general climatology, but also on dry and wet spells, as well as extreme precipitation on the catchment scale. The study area is the catchment of the Adige in located in the southern part of the Alps in Northern Italy, covering an area of 12,100 km^2. This catchment is selected due to a relatively high network of observation stations, and the challenging orography for the comparison with the ensemble of precipitation data sets. Due to the complex topography, the uncertainty introduced by the spatial resolution of the data sets was of special interest in this study. The overall purpose of this study was motivated from the application point of view, rather than atmospheric assessment of regional patterns. It is clear, that the outcomes of this study are highly

regionally dependent and cannot be transferred to other catchments, but will give insights in local performance and uncertainties of precipitation data sets, especially over an alpine catchment.

2. Study Area

The focus of this study was the catchment of the Adige located in Northern Italy. The catchment, shown in Figure 1, covers an area of 12,100 km^2. The Adige River stretches over a length of 409 km from the Southern Alps to the Adriatic Sea, passing three Italian provinces.

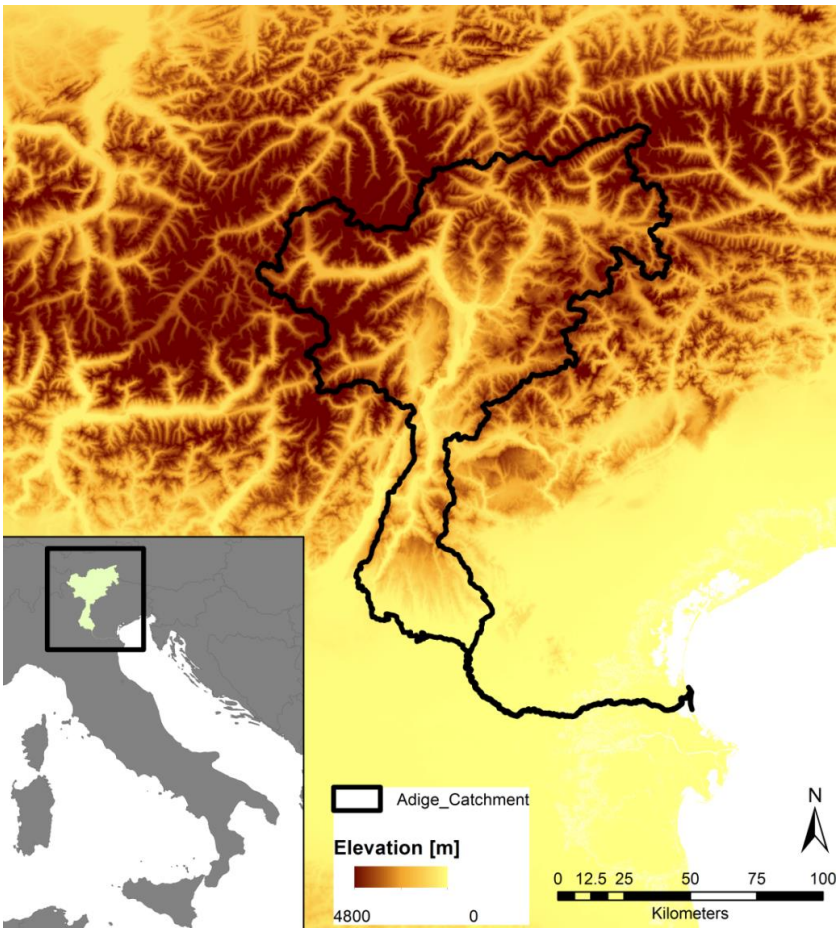


Figure 1. Location of the Adige catchment and corresponding elevation.

The area is characterized by a humid climate with annual mean precipitation ranging from 500 mm in the inner-Alpine dry areas, such as the Venosta valley [14,35], to 1600 mm. Precipitation distribution shows a pronounced summer peak and relatively dry winters [36]. An overview of the observed mean annual precipitation at the available stations is presented in Figure 2. The area is characterized by strong elevation differences ranging from sea level to 3865 m, as presented in Figure 1. The hydrology is characterized by snow and glacier melting in the spring months, and intense precipitation events in summer [36]. Discharge is used for irrigation of intensive agriculture - mainly fruit trees - and through various touristic activities. Additionally, 30 major reservoirs used for power generation are operated in the area, increasing pressure on ecosystems due to altered hydrology [37].

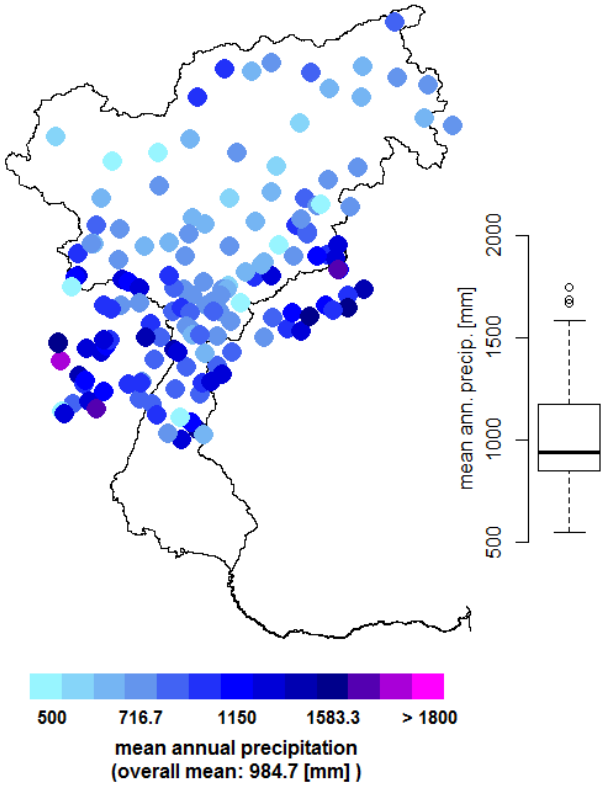


Figure 2. Mean annual precipitation sums for the observation stations.

3. Data Sets and Methods

3.1. Construction of the High-Resolution Reference Data Set

Daily precipitation data for the period 1989–2008 derived from gauge measurements was available over the catchment. The station network for the study area covers a total of 153 stations as presented in Figure 2; however the temporal and orographic coverage varies considerably. Apart from the temporal coverage, the spatial distribution also varies strongly throughout the area, as shown in Figure 2. Most stations are located in the central to southern part of the catchment while especially the north-western part is sparsely covered. The southernmost part of the catchment is only covered by far away stations, as local stations were not available.

As the study aims in evaluating gridded precipitation products, the meteorological interpolation module of the hydrological model Water Flow and Balance Simulation Model (WaSiM, [38]) was chosen for the construction of the precipitation grid with a spatial resolution of 1 km and daily time step. This procedure was selected as it is computationally efficient, and furthermore, the resulting data set will serve as meteorological forcing of this hydrological model in consecutive studies. Additionally, the module 'regional superposition' allows for overlaying and combining multiple interpolation methods over various areas to consider topographic effects and specifications of the catchment. In the following the selected method is briefly described, whereas for a detailed insight, it is referred to in the technical report [38].

The precipitation at each grid cell was calculated based on the four nearest stations in respective quadrants, ideally on a regular grid. If no station was available in the respective quadrant, the scheme automatically switched to Inverse Distance Weighing (IDW), which is also widely used for interpolation of meteorological variables [39–41], and can be helpful in areas with low station density [42]. For mountainous areas, elevation dependent regression was applied to account for topographic effects. Here, the inversion levels were calibrated to better represent the characteristics of the catchment. In this study, in a first step, bilinear interpolation and elevation dependent regression were applied separately

over the entire catchment. In a second step, the regional superposition module was applied to combine both approaches. In this way topographic effects could be captured, but were not overestimated, and the inner-Alpine dry areas were better captured compared to exclusive elevation-dependent regression. Additionally, the residuals for the regression were calculated for various sub regions to achieve better performance, and a better representation of the observed fields. The final grid was constructed for the period 1989–2008, which represents the maximum temporal coverage of the analyzed gridded data products presented in Section 3.3. All analyses presented in this study refer to this time frame only. The grid will be referred to as ADG-1KGPR, representing the gridded precipitation over the Adige domain with a 1 km grid size.

3.2. Performance Validation of the Constructed Reference Grid

To assess the performance of the grid, a leave-one-out-cross-validation (LOOCV) approach was applied to allow for best-possible representation of the observed precipitation. Three main indicators were calculated to assess the performance of the interpolation method; The Percent Bias (PBIAS) assesses the tendency of the modeled data to exceed or underestimate the observed data in percent, hence the optimal value is zero. The Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE) are two widely used indicators when it comes to model performance assessment in meteorology and other fields [43,44]. Figure 3 shows the presented evaluation criteria at the precipitation gauges and their corresponding grid cells in the interpolated 1 km gridded result. Stations outside the catchment are left out of the analysis to avoid confusing results and just presented as empty circles in the figure. Figure 3a represents the MAE*, which is the MAE based on all dates for the period 1989–2008, whereas Figure 3b shows the MAE, which is based only on days with precipitation occurrence. Therefore MAE showed higher values as MAE*, as the absolute error at days with no or only low precipitation was relatively small. Figure 3c shows the RMSE, and in Figure 3d, the corresponding PBIAS is shown. Both mean RMSE and MAE show similar or better results as comparable studies over various study areas [40,42,44–46].

Figure 4 shows the monthly distribution of the four presented criteria. As expected, the course of the RMSE (Figure 4a), MAE* (Figure 4c), and MAE (Figure 4d) followed the distribution of monthly precipitation, with the maximum occurring in the summer and fall months resulting from convective and strong precipitation events. However, the highest spread and overestimation of PBIAS (Figure 4b) was visible for the winter months, which are likely also the months with the highest uncertainty in the observations, due to higher undercatch of snow compared to liquid precipitation [24,25].

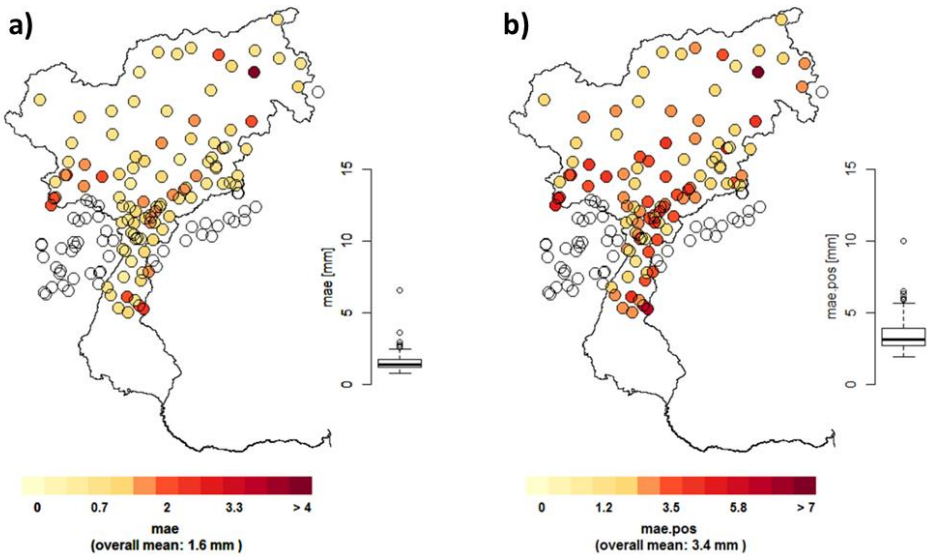


Figure 3. *Cont.*

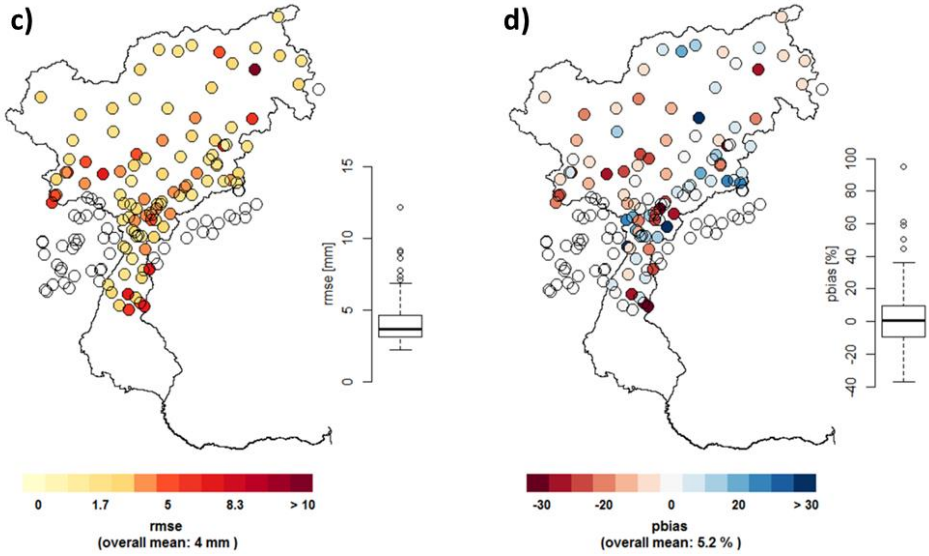


Figure 3. Evaluation of interpolation performance at various precipitation stations over the Adige catchment. Assessment through the indicators: MAE* (**a**); Mean Absolute Error (MAE) (**b**) and Root Mean Square Error (RMSE) (**c**); and Percent Bias (PBIAS) (**d**). Stations outside the catchment are masked.

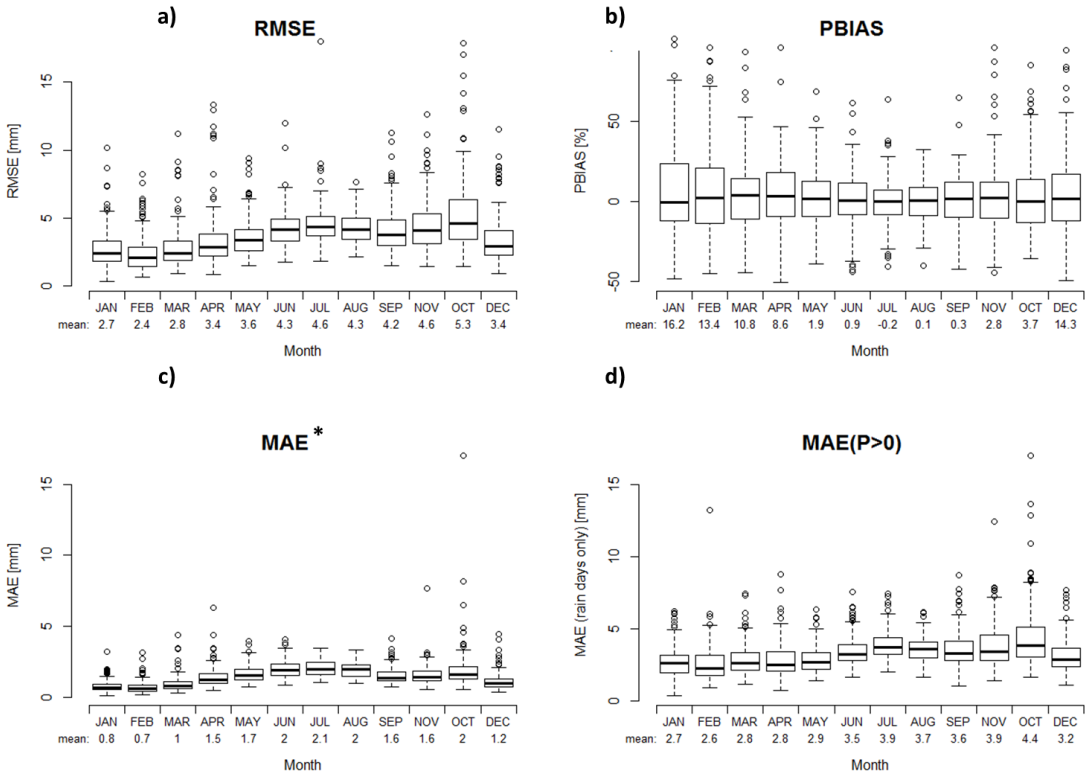


Figure 4. Evaluation of interpolation performance at monthly scale, boxplots result from mean results at each observation station over the Adige catchment. Four indicators presented: RMSE (**a**); PBIAS (**b**); MAE* (**c**); and MAE (**d**).

3.3. Available Precipitation Data Sets

In the following section, the available precipitation data sets included in this study are briefly introduced. Table 1 gives a short overview over the data sets and their horizontal resolution, as well as the covered time period. The main criteria for the selection of data sets were daily resolution and

a temporal coverage, at least for the 20-year period (1989–2008). For this reason, any monthly data sets or daily data sets with a shorter temporal coverage, or a starting date late than 1989, were not included here. The starting and end date were determined to find the maximum overlap for the selected data sets. The horizontal resolution of the included gridded products varied from a 5 km grid to 1° × 1° (~125 km) spacing. Coarser grids, such as the global 20th Century Reanalysis [10] and the NCEP/NCAR Reanalysis [16] data set, were not included due to the relatively small catchment area and the respective coverage of these global data sets. Data sets with a shorter time period available due to a later starting date, such as the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis [47], the Climate Prediction Center (CPC) Morphing Technique (CMORPH) [48] and the Global Precipitation Measurement Mission (GPM) [49], were also not included in this study. This led to a selection of eight data sets, as presented in Table 1, which are either interpolated products from station observations, derived from satellite or reanalysis data sets. The data products are shortly introduced in the following, while for detailed information it is referred to the respective publications.

Three observational data sets are included in this study: the Alpine Gridded Data Set (EURO4m-APGD), derived through the EURO4m project [14], the E-OBS data set in version 11 provided through the European Climate Assessment and Dataset (ECA&D), and derived within the ENSEMBLES project [20], as well as the Full Data Daily Product from the Global Precipitation Climatology Centre (GPCC-FDD) provided through the DWD [12]. The EURO4m-APGD was derived from numerous observation stations from various national observatories, and covers the greater Alpine area for the time period 1971–2008. With a 5 km grid size, this data set represents the highest resolution observational data product available for this area so far. The E-OBS data set is widely used in various studies and based on selected observation stations that provide a long data record and ensured quality control. Therefore, less stations are included compared to the EURO4m-APGD [50]; however, the E-OBS data set covers the time period 1951–2015 (and updates are ongoing), with a daily resolution on a 25 km grid over Europe [20]. GPCC-FDD provides precipitation information from station data on a 1.0° grid for the time period 1988–2013. While GPCC offers a multitude of precipitation data sets, the FDD is recommended for daily analysis with focus on extremes and was therefore selected for this study. The GPCC-FDD includes similar stations compared to E-OBS for the interpolation, as indicated by the corresponding variable in the data set.

Four reanalysis products are included in this study, which are available at various resolutions. While most reanalysis data sets assimilate precipitation in some way, are not affected by system changes, temporal and spatial coverage of the observations and undercatch are problematic [51]. The reanalysis data set with the highest horizontal resolution is the MESAN downscaling product on a ~5 km grid. MESAN is derived through the downscaling of HIRLAM 0.2° forecasts, and assimilates observed precipitation from rain gauge data [15]. Additionally, the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), provided through the National Aeronautics and Space Administration (NASA) which shows considerable improvements over Version 1 [52] with an improved assimilation scheme, was included. In MERRA-2, observed precipitation was used as forcing for the parameterization of the land surface. The data set offers global coverage and starts in 1980, while updates are still ongoing [53], with a horizontal resolution of ~50 km. The two coarsest reanalysis data sets in this study are provided through the European Centre for Medium-Range Weather Forecasts (ECMWF). First, the ERA-Interim global reanalysis at a resolution of approximately 80 km (0.75°) and second, the Reanalysis for the 20th Century (ERA-20C) [54], were chosen. ERA-Interim assimilates various observations from multiple sources in a 12 hr cycle [11], while ERA-20C only assimilates surface and mean sea level pressures and marine winds [55]. Similar to [21] the coarse data sets are not expected to perfectly reproduce the daily precipitation over a complex medium sized catchment. However ERA-Interim is widely used as meteorological forcing for hydrological models in data scarce regions, making this an important data set to include. Especially for the long time period covered, it is also worth including the ERA-20C, since this data set would allow simulations

starting in 1900. Additionally, it is interesting to see how this data set, which does not assimilate many observations, performs on the catchment scale.

Table 1. Overview of data sets included in this study with corresponding spatial resolution, temporal coverage, data set type or source respectively, and reference publication.

Data Set	Resolution	Temporal Coverage	Type/Source	Reference
MESAN	5 km	1989–2010	Downscaling/Reanalysis	[15]
EURO4m-APGD	5 km	1971–2008	Observations	[14]
E-OBS v. 11	25 km	1950–2015	Observations	[20]
PERSIANN-CDR	0.25° (~30 km)	1983–present	Multisatellite (infrared), corrected	[19]
MERRA-2	0.5° latitude × 0.625° longitude (~50 km)	1980–present	Reanalysis	[53]
ERA-Interim	0.75° (~80 km)	1979–present	Reanalysis	[11]
GPCC–FDD v1.0	1.0° (~100 km)	1988–2013	Observations	[12]
ERA-20C	125 km	1900–2010	Reanalysis	[54]

The only remotely sensed data product included was the bias-corrected version of the multi-satellite Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) specified as PERSIANN-CDR (Climate Data Record). PERSIANN uses infrared brightness temperature derived from geostationary satellite information to estimate the rainfall rate [19]. To reduce the biases in the so-derived precipitation but keep the spatial and temporal patterns at 0.25° resolution, the 2.5° monthly data set by the Global Precipitation Climatology Project (GPCP) [56] is applied for correction.

It is important to note that most data sets are not independent from each other, as most of them include the same station observations directly, or assimilate them in some way. This is however, a common problem in comparison studies, and cannot be avoided.

3.4. Remapping of the Reference Grid

When comparing spatial data sets, a common grid has to be defined to perform the analysis. The highest spatial resolution possible for a comparison is defined by the selected reference grid, and was thus 1 km in this study. This approach however favors higher resolution data sets, as the coarser grids are not able to resolve the fine-scale processes [3]. A second possibility is to perform the comparison at multiple defined resolutions that correspond to meaningful data sets, e.g., the highest resolution and a medium range resolution, as done in some studies [21,50]. Lastly, the comparison can be conducted on the coarser grid, which provides a fair comparison for the coarser grids; however small scale information are lost, which might not be desirable [3].

In this study, a combination of the first and the latter approaches was chosen. A conservative remapping of the reference grid to the resolution of each of the gridded data sets was conducted. The direct evaluation of each data set hence was carried out on the corresponding, coarser grid of each data set. All data sets were then disaggregated to the highest resolution of 1 km to account for the catchment delineations, the number of grid cells, and the correct areal weights for coarser grid cells, where only a small percentage of the grid cell might be within the catchment. Obviously, the coarser data sets had a lower effective resolution, and could not represent the topography at the detailed 1 km resolution. However, there was no additional source of uncertainty introduced through this disaggregation, but rather already captured within the uncertainties of the data sets. Nevertheless, the inter-comparison of the data sets was at a higher resolution and penalized the coarser data sets. This approach was chosen, as the performance should be evaluated on catchment scale with respect to hydrological modeling, which will likely be carried out at even higher resolution.

4. Results

In the following, the results of the comparison are presented, starting with a generic section on the climatological performance of mean precipitation at various time scales. The following subsection then introduces some indicators for a more detailed comparison based on daily precipitation. The uncertainty introduced by elevation only is evaluated in the last section.

4.1. Mean Precipitation

The mean annual precipitation over the Adige catchment for all data sets included in this study is presented in Figure 5, as well as the ensemble mean, derived from all data sets. There was a considerable spread in the catchment mean of 350 mm·y^{-1} identifiable in the ensemble, corresponding to 35% of the ensemble mean annual precipitation. The highest mean annual precipitation was found in ERA-Interim (1196 mm) and PERSIANN-CDR (1140 mm), while E-OBS (847 mm) and MESAN (868 mm) showed the lowest annual mean precipitation sums. The general inner-alpine dry valleys were present in the high resolution data sets EURO4m-APGD and MESAN, although the extent and the magnitude are overestimated in the latter. These dryer areas were also reproduced by the E-OBS data set; however, there was a strong North-South gradient with lower precipitation in the South. The coarser data sets were not able to reproduce these small scale dry areas, which was not surprising.

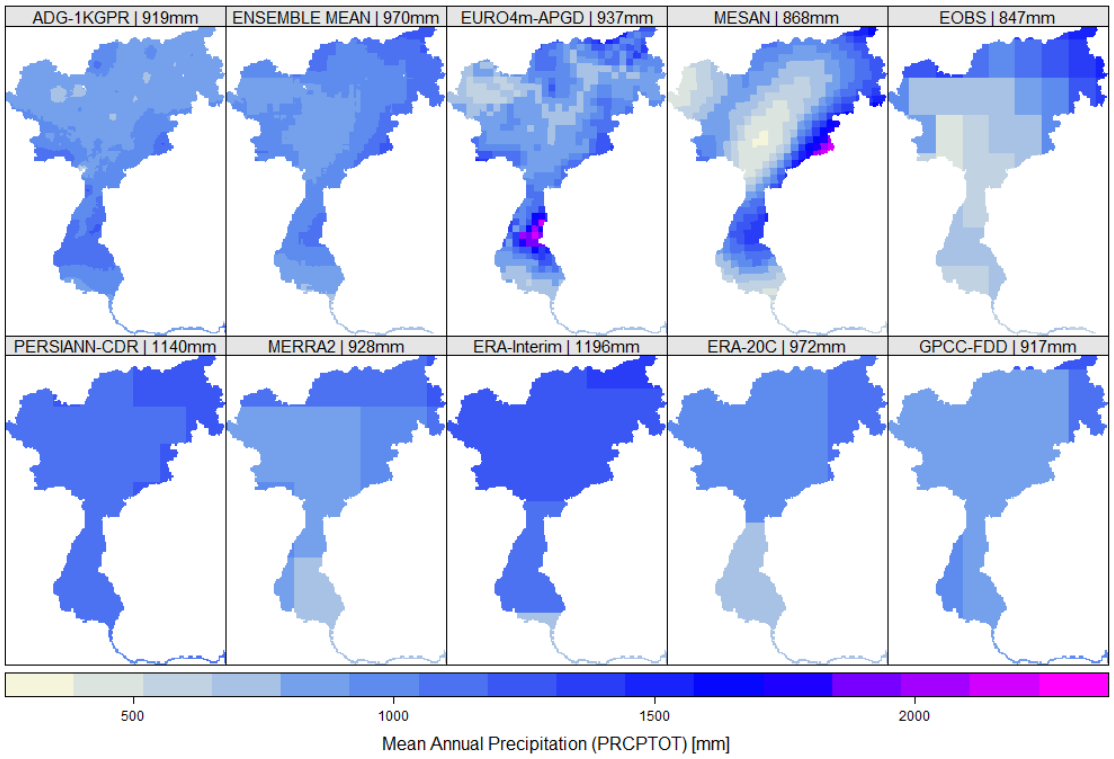


Figure 5. Mean annual precipitation over the catchment for the applied data sets including the ensemble mean; numbers in the panels correspond to the mean catchment precipitation.

To evaluate the temporal representation and variability of the mean catchment precipitation, Figure 6 shows Taylor diagrams [57] based on daily (a), monthly (b), and annual (c) catchment precipitation of the gridded data sets (circles) and the corresponding aggregated reference grid (triangles) with respect to the reference grid (grey triangle). Dashed lines display the correlation of the evaluated data sets compared to their corresponding reference grid, green arcs the RMSE, and blue, dashed arcs display the standard deviation of each grid. Best agreement, measured by correlation, for both, aggregated reference grids and the evaluated data sets, is given for monthly

data. All data sets captured the monthly precipitation reasonably well, with correlation coefficients between 0.8 and 0.95. As expected, daily correlations were considerably lower, with no data set showing correlation greater than 0.8, and overall lower correlations for annual precipitation (with the exception of EURO4m-APGD and MERRA-2). Inter-annual variability, here, simply the standard deviation of the annual precipitation, was represented very well in E-OBS, MESAN, ERA-20C and EURO4m-APGD, while the other data sets underestimated the variability considerably. The ranking for RMSE varied depending on the time scale; however, EURO4m-APGD, MERRA-2 and ERA-20C showed comparatively low errors for all three time scales, while MESAN showed higher errors for annual precipitation. The triangles in the Taylor Diagrams show the effect introduced purely by the spatial resolution and the different grids of the included data sets. They were all derived through aggregation of the ADG-1KGPR to the corresponding coarser grid. As expected, they all showed very high correlations (>0.95) and low RMSE compared to the reference grid at 1 km resolution.

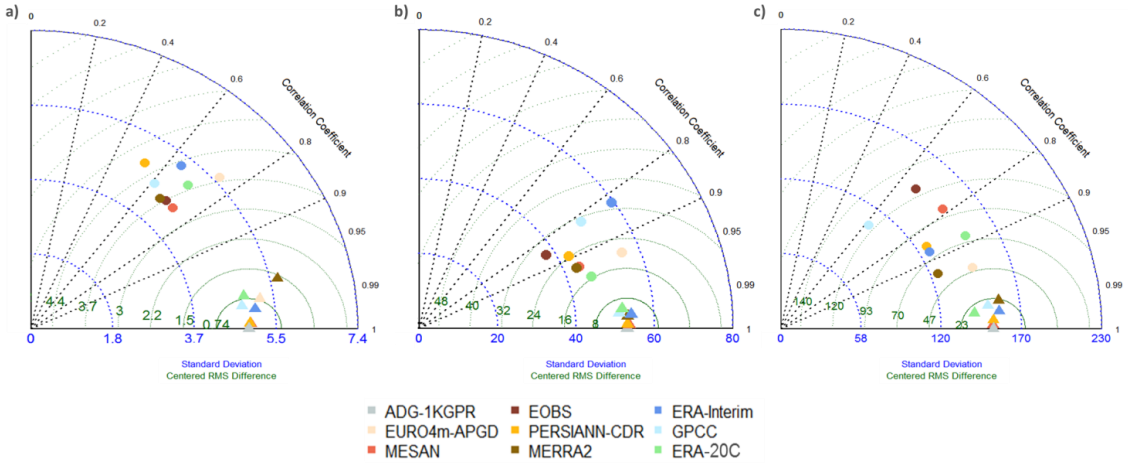


Figure 6. Taylor Diagram based on daily (**a**); monthly (**b**) and annual (**c**) catchment mean precipitation for the Adige catchment. Circles correspond to the gridded precipitation data sets, while triangles refer to aggregated reference grid.

Boxplots for mean monthly precipitation over the catchment are presented in Figure 7. In contrary to Figure 6, the spread within each data set comes from the spatial variability within the long-term monthly mean precipitation over the catchment, derived from grids similar to the annual precipitation sums presented in Figure 5. Darker grey boxplot (first in each month) correspond to the reference grid, colored boxplot to the evaluated data sets and the shaded grey boxplots to each corresponding aggregated reference grid. The latter represent the uncertainty that purely comes from the spatial resolution, as they are conservatively remapped as described in Section 3. Throughout all months, this results in less variability, hence shorter boxplots, with a tendency towards decreasing median precipitation for all months with increasing spatial resolution.

In general, the annual precipitation cycle was captured by all data sets; however, there were distinct differences for the individual data sets and months. The annual cycle was more pronounced in the EMCWF reanalysis products ERA-Interim (dark blue) and ERA-20C (green), and considerably less pronounced in PERSIANN-CDR (yellow). The latter showed an overestimation in winter precipitation of up to 100%. This was caused by the overestimation of precipitation due to cold clouds, where the algorithm for PERSIANN was less accurate [19,28] and was less pronounced in summer. ERA-20C and especially ERA-Interim considerably overestimated summer precipitation, also compared to other reanalysis data sets included in the study. This was important to point out, as precipitation undercatch and the underrepresentation of high-elevation gauges potentially led to an underestimation of precipitation in the observational data sets. Similar to the results presented in Figure 5, E-OBS (dark red) produced a larger spread for monthly precipitation throughout the year, resulting from very high precipitation in the northern, alpine part of the catchment, and very low precipitation in the

central and southern parts. This also led to an underestimation of summer precipitation in this data set. All data sets derived from observations agreed for most months, and represent similar magnitudes, which was not too surprising, as they were not independent from each other, and several stations were included in all these data sets. Additionally, the high-resolution data sets better represented the annual cycle and the precipitation magnitudes compared to the other data sets.

The presented differences in monthly precipitation would likely have strong implications for hydrological impact modeling. The presented overestimation in winter for PERSIANN-CDR would likely lead to greater snow pack accumulation, resulting in higher discharge from snow melt in spring. The overestimation in summer in the ERA-Interim and ERA-20C data sets might lead to higher soil moisture in summer, and consequently increased discharge and higher sensitivity to strong precipitation events with respect to flood events.

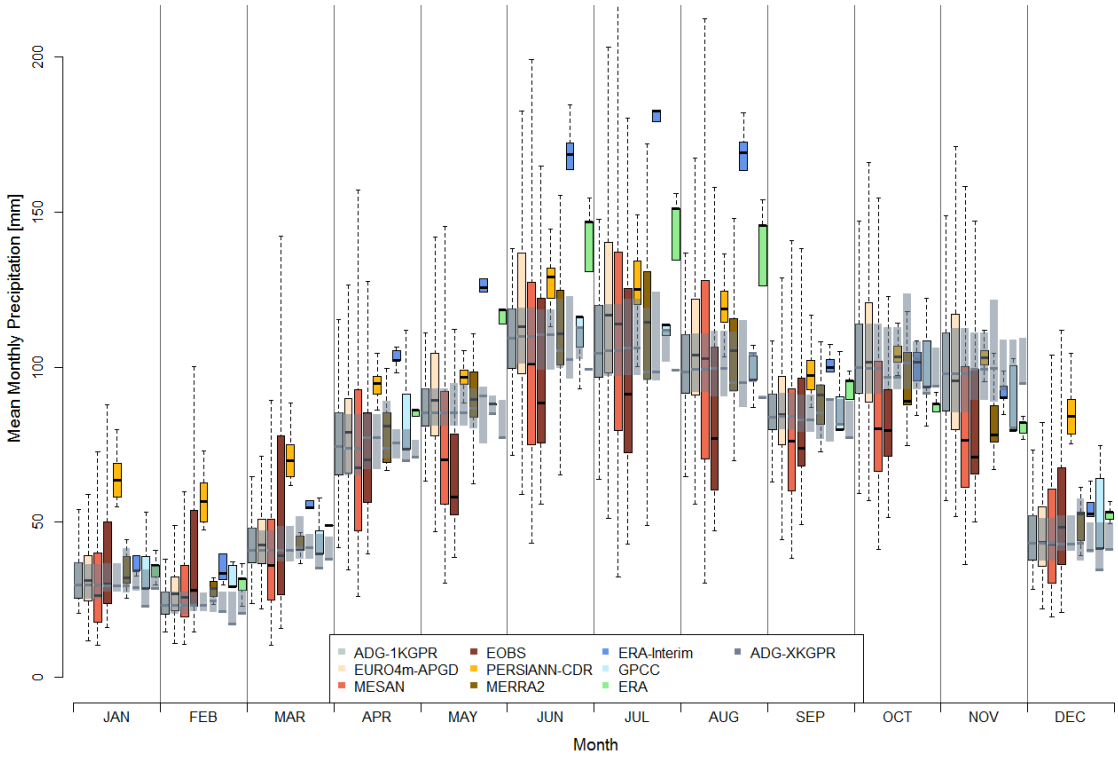


Figure 7. Monthly Precipitation Sums for the applied data sets. The grey boxplots correspond to the high resolution Water Flow and Balance Simulation Model (WaSiM) aggregation, while shaded boxes show the boxplots for the respective grids.

4.2. Comparison of Hydrologically Important Indicators

To further assess the differences in the data sets, three additional indicators addressing extreme precipitation were derived to estimate the implications for hydrology. These indicators were the maximum number of consecutive days with precipitation <1 mm (CDD, numbers of days), the maximum number of consecutive days with precipitation >1 mm (CWD, numbers of days) [58], and the contribution of heavy precipitation (>95th percentile) to the annual precipitation (R95pTOT, in %) [59]. These were selected, as they provided insights in the performance of the data sets for heavy precipitation, as well as the timing and distribution of precipitation and wet/dry days. All of these indicators are based on daily precipitation.

Figure 8 shows the deviations in CDD for the period 1989–2008 for all selected data sets compared to the corresponding remapped reference grid. The upper panel shows the absolute amount for the reference grid, as well as the ensemble mean. To indicate the spread and the uncertainty within

the ensemble, the standard deviation is presented in the top right. This reveals a larger spread, i.e., higher uncertainties, in the central part of the catchment, which is generally characterized by a higher number of consecutive dry days compared to the alpine northern area. The high resolution data sets tend to better represent CDD, whereas the coarser data sets considerably underestimate these. This can be at least partly attributed to the horizontal resolution and smoothed orography, as the aggregated reference grids are based on a 1 km grid and are not interpolated to the respective resolution. The highest underestimation can be found in PERSIANN-CDR (−65 days for the catchment mean), which can be attributed to the less accurate algorithm for cold clouds already mentioned, that overestimates precipitation on cloud covered days especially in winter. The strong underestimation in the ERA-Interim (−32 days) data can not only be attributed to horizontal resolution, as the coarser ERA-20C and GPCC-FDD show a better representation of CDD, but is rather caused by the general overestimation in the frequency of low precipitation days.

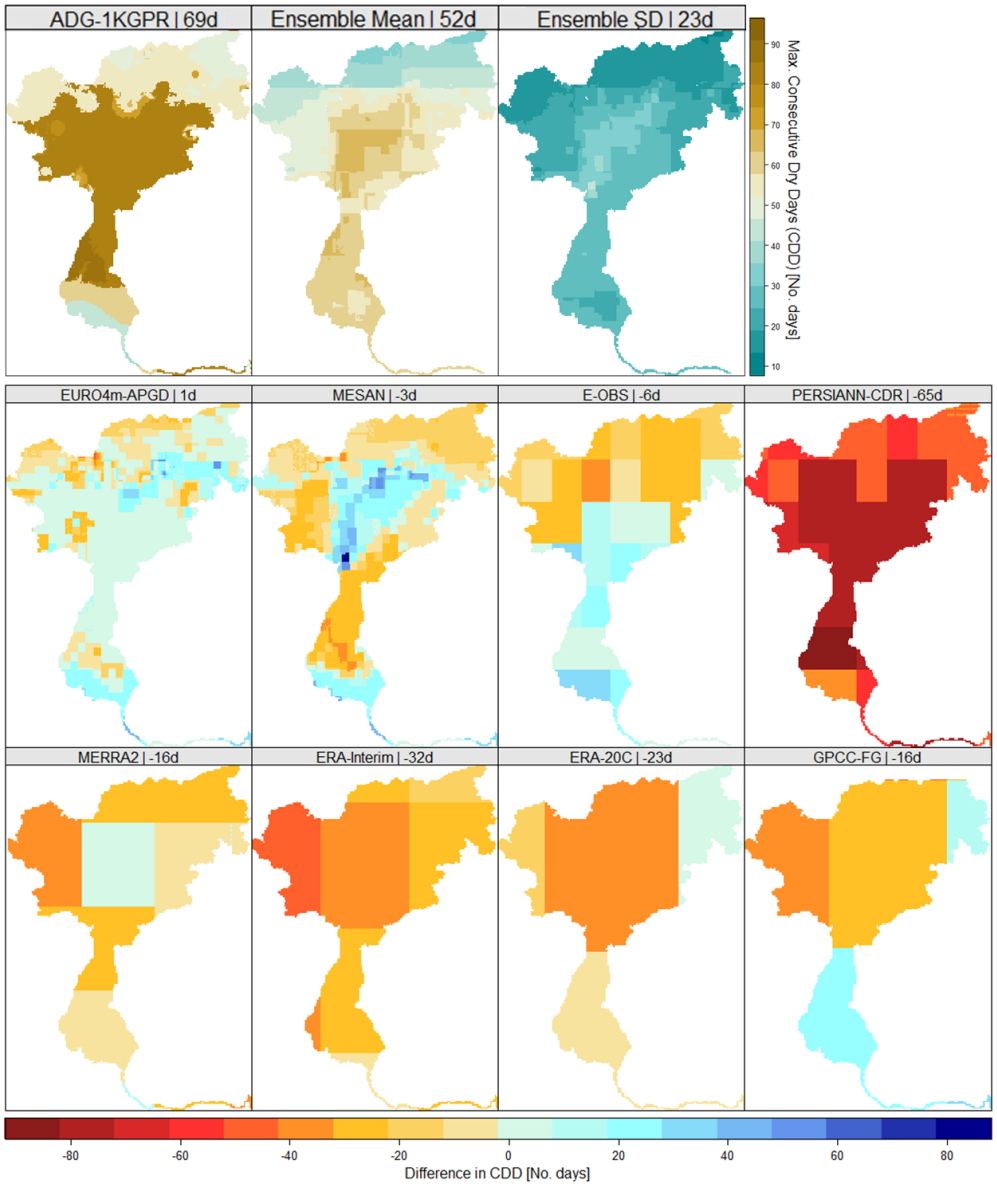


Figure 8. Maximum number of consecutive dry days (CDD) in ADG-1KGPR, the ensemble mean and corresponding standard deviation (**top row**) and the deviations from the corresponding reference grid for the eight gridded data sets (**middle & bottom row**).

This was confirmed through the maximum number of consecutive wet days (CWD), i.e., days with precipitation >1 mm, presented in Figure 9 in the same way as described for CDD. In general, there is higher agreement throughout the different data sets with the reference grid. The model spread, shown as standard deviation in the top right, was more homogenous than for CDD. Additionally, the data sets agree more, resulting in a smaller spread, i.e., standard deviation of seven compared to 23 days (for CDD). In contrary to CDD, there was no evidence that high resolution data sets perform better, as the coarsest grids, ERA-20C and GPCC-FDD both showed very low deviations from the corresponding reference grid. The highest deviations were found in the MERRA-2 and ERA-Interim reanalysis for the northern part of the catchment through orographic precipitation. These data sets overestimated the mean CWD by 11 days (MERRA-2) and by nine days (ERA-Interim). PERSIANN-CDR also showed a slight overestimation, mostly for the central area of the catchment. However, as for CDD, the data sets derived from observations showed less deviations to the reference data set compared to the reanalysis products, with the exception of the coarse ERA-20C.

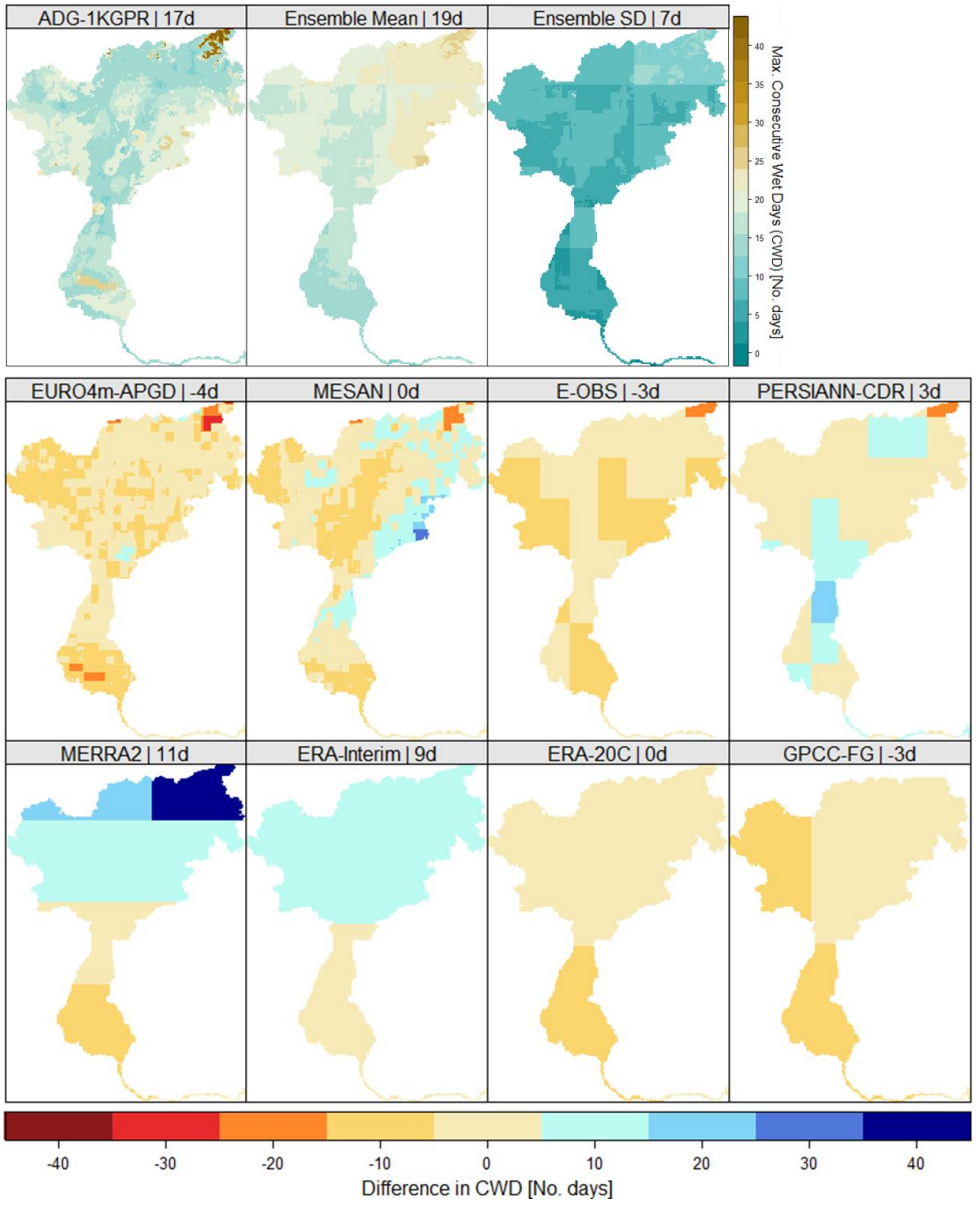


Figure 9. Same as Figure 8 but for consecutive wet days (CWD).

To assess the representation of heavy precipitation events in the data sets, the percent contribution of precipitation from heavy precipitation days, i.e., greater than the 95th percentile, to the total annual precipitation (R95pTOT) is presented in Figure 10. As expected, the coarser data sets tended to reproduce lower values, meaning a lower contribution of heavy precipitation days to the annual precipitation. The high resolution data sets again agreed best with the reference grid, although MESAN showed more homogenous patterns as the EURO4m-APGD and the reference grid, originating from the downscaling of the HIRLAM results. A strong underestimation was present for E-OBS, PERSIANN-CDR, ERA-Interim and ERA-20C. For ERA-Interim there was a known error reported by the ECMWF with the representation of convective events. The convective available potential energy (CAPE) was zero in the data set for the afternoon time steps where most convective events occur. Despite the coarse resolution, MERRA-2 represented the contribution of heavy precipitation well, though it missed out the north-eastern part of the area with a considerably lower share due to the horizontal resolution. In the GPCC data set R95pTOT, the northern part was better reproduced compared to the coarse EMCWF grids; however, the southern part showed lower R95pTOT. This originated from the very low station density in the GPCC-FDD data set in the southern part, with only one station per grid cell, according to the respective variable in the data set.

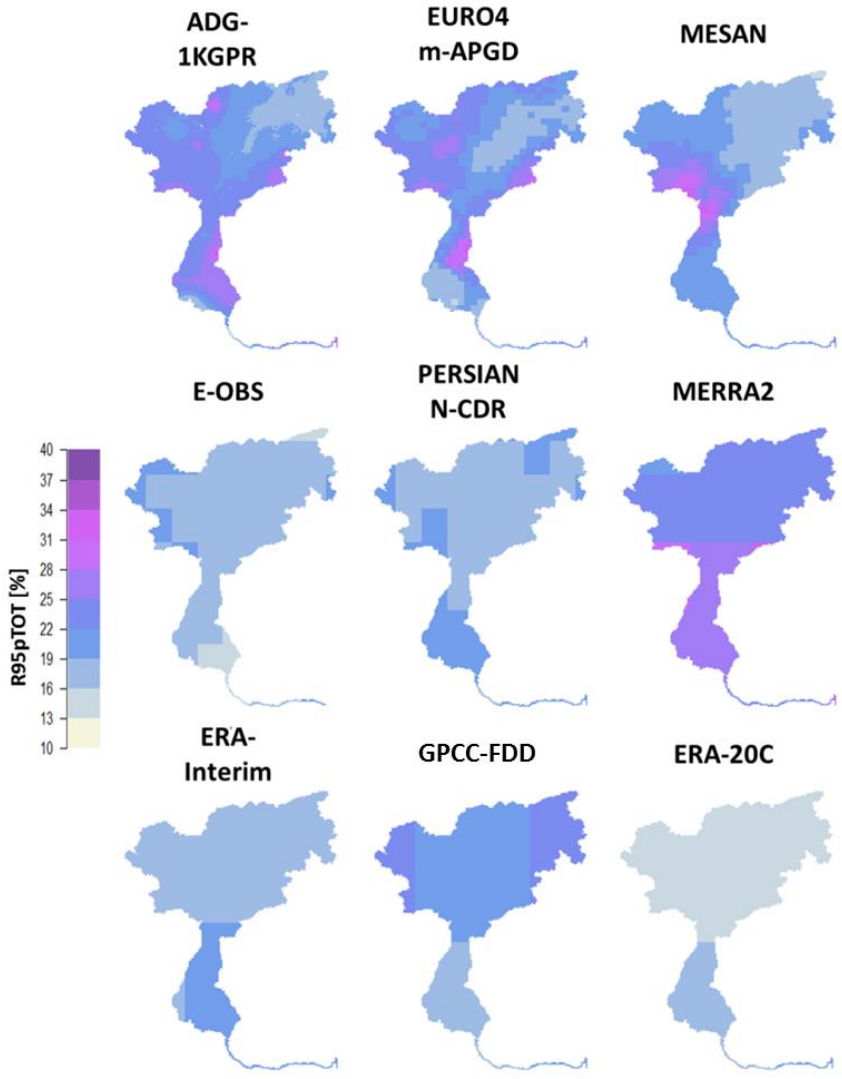


Figure 10. Contribution of heavy precipitation (>95th percentile) to annual precipitation in % (R95pTOT) for the reference grid and the selected data sets.

4.3. Uncertainties and Recommendation

To estimate the relative contribution of spatial resolution compared to the uncertainty introduced by the data sets themselves, the coefficient of variation (CV) is presented. CV is defined as the standard deviation normalized by the mean and given in percent. The data sets were remapped to the corresponding grids of all data sets included and the climatological monthly catchment precipitation was presented at each resolution. These were then compared to the CV derived from the data sets at their native resolution, which were grouped in observational (blue) and reanalysis (red) data sets. PERSIANN-CDR was excluded from this analysis, as it was the only remote sensing data set in the study. Figure 11 shows boxplots for the CV introduced through spatial resolution only, while the red and blue squares correspond to the mean CV in the reanalysis and observation data sets. The CV introduced through elevation only, was derived by remapping each data set to the grids of the other data sets, and calculating CV separately per data set. The variation shown as boxplots hence stems only from the effect of different resolutions and does not include uncertainty through the data sets themselves. The CV shown in red and blue in contrary are based on the variation derived from the data sets at their individual grids, which also includes uncertainty originating from the different resolutions. With the exception of January and March, the uncertainty introduced by the source of the data sets exceeded the contribution of the grid resolution. The relative variability from spatial resolution was higher for the winter months, where the observational data sets also showed greater variability as the reanalysis data sets.

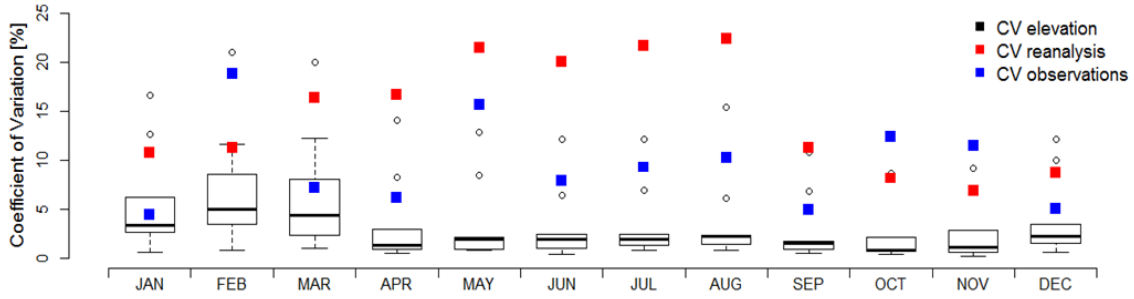


Figure 11. Coefficient of Variation (CV) with respect to the source of the data sets for observational grids (blue) and reanalysis data sets (red), compared to CV introduced only by elevation (boxplots, black).

As presented, there was a considerable spread within the available precipitation data sets. To cover most of the uncertainty and benefit from the strength of each selected data set, similar to climate change impact assessment studies, an ensemble approach is recommended to serve as a reference [50]. However, based on the presented indicators, three of the presented data sets were not able to represent the precipitation conditions in the catchment. This conclusion can be drawn based on the comparison to the reference grid, as well as with respect to the other data sets. As also suggested in previous studies [51] the corresponding data set were excluded from the ensemble in this study. This was not done to reduce the overall spread, but rather based on a thorough selection process, as the resulting differences in the data sets were not introduced to different methods or sources of the data sets, but rather bad performance and limited applicability of these data sets over the study area. Therefore, PERSIANN-CDR was excluded, due to the systematic overestimation of precipitation in the winter months, and the considerably underestimation of CDD. Similarly, ERA-Interim and ERA-20C were excluded, as they overestimated summer precipitation and produced too few consecutive dry days and were additionally, due to their coarse resolution, not able to perform well for heavy precipitation (R95pTOT). As the remaining data sets did not show considerable deviations from both the reference and the ensemble mean, they were included to derive the ensemble reference grid presented in Figure 12 for the presented indicators. The mean annual precipitation (Figure 12a) represented small-scale features such as the dry Venosta valley in the north-western part of the catchment, as well as the higher precipitation over the alpine north-eastern part. This was not as

pronounced in the reference data set, possibly through low station density, but identifiable in most other gridded data sets. R95pTOT, presented in Figure 12b, follows the patterns of the high resolution data sets, however, with less artifacts introduced through interpolation as for the reference grid, the stations were highly visible in the corresponding R95pTOT. However, the absolute range of R95pTOT was only about two-thirds compared to the reference grid. The maximum number of consecutive dry days (Figure 12c) followed the reference grid; the grid structure of the E-OBS grid was clearly visible, resulting in very sharp and unrealistic boundaries. These were less pronounced for consecutive wet days (Figure 12d) where the higher number of CWD in the north east was dominated by the reanalysis data sets, which showed a higher number of consecutive wet days in this area.

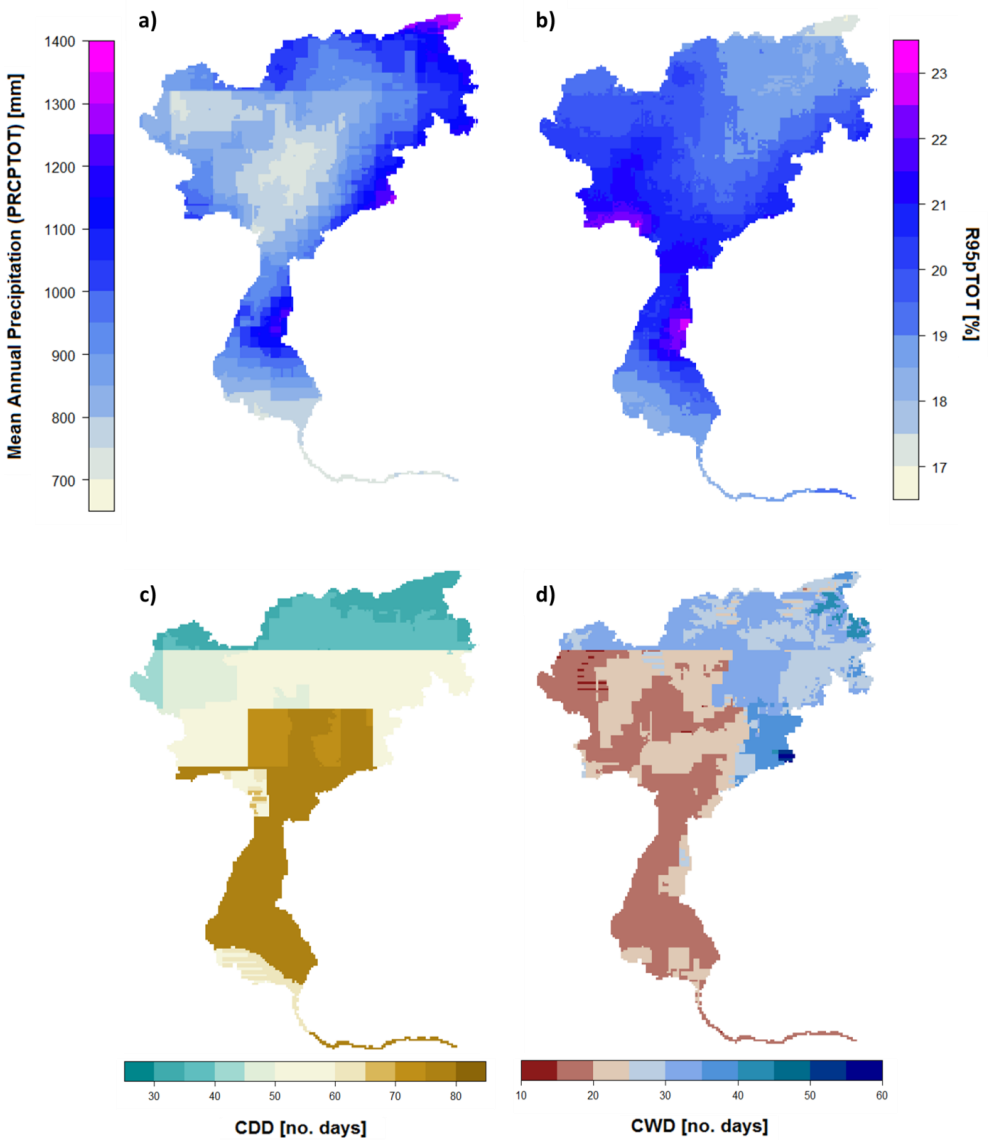


Figure 12. Indicators for the selected grids ensemble, (**a**) mean annual precipitation in mm; (**b**) R95pTOT in %; (**c**) CDD in number of days; and (**d**) CWD in number of days.

5. Discussion

The presented differences in available precipitation data products highlight the need to take into account the uncertainties related to finding the real reference precipitation. This becomes crucial when data sets are applied to calibrate and validate hydrological models, or are chosen to select a subset or bias correct climate models. Therefore, the uncertainties and differences within the ensemble of

available daily, gridded precipitation data products have to be evaluated on the catchment scale. In this study, the uncertainty of gridded precipitation data sets, from various sources and derived through different methods, was investigated over an alpine catchment. It has to be pointed out that findings over this area do not allow for a performance assessment of the presented data sets in general. Whereas the data sets show similar annual precipitation sums and a good representation of the inter-annual variability, i.e., standard deviation, they show clear differences on daily time scale. Although all data sets capture the annual cycle of precipitation, they show large differences in the magnitudes, up to 100% in some months for mean catchment precipitation. This confirms the results presented by previous studies over Europe [50] and the alpine region [21] also on a catchment scale. This study confirms the limitations of the satellite precipitation product PERSIANN-CDR, which led to a systematic overestimation of precipitation, especially in winter, likely due to non-raining clouds where the algorithm performs worse than for e.g., warm tropical convective systems [19]. This overestimation, especially in winter was, also identified over most of Europe in general [50]. Increased precipitation in winter will likely result in a larger snow pack that will influence the peaks in discharge for the spring months, due to snow melt. The slight overestimation of summer precipitation by PERSIANN-CDR confirms recent findings for the PERSIANN data product with respect to convective events and better agreement in fall, where precipitation is dominated by stratiform systems [32]. However, the bias-corrected CDR applied in this study improved the PERSIANN data set considerably, as the overestimation was not as pronounced in this study over the same region. Clear outliers in the summer months were the two reanalysis data sets from EMCWF, especially ERA-Interim, which confirms the findings in the studies mentioned. Similar to the results presented for the comparison with station data [60], there is a higher agreement between reanalysis data sets and observational data sets in winter. This can be confirmed also by the lower coefficients of variation in the winter months for both groups of data sets. Differences in the performance of the observational data sets is strongly linked to station density [2], which is extremely low in the E-OBS data set [21] and, especially for the southern part of the catchment, also for GPCC-FDD. This leads to underestimated precipitation for E-OBS in annual, as well as seasonal (summer), precipitation and an increased number of consecutive dry days for the southern parts.

One source of uncertainty in the different data sets is the spatial resolution. As expected, the high resolution data sets generally better represent the regional or local patterns for both mean precipitation, as well as for several indicators such as consecutive dry and wet days, and the contribution of heavy precipitation to the annual precipitation. Nevertheless, this is not dependent on the source of the data set, as higher resolution reanalysis show similar patterns to the reference grid, such as dry areas and areas where heavy precipitation has a higher share in the annual precipitation. However, the reanalysis data sets included in this study show less consecutive dry days and more consecutive wet days compared to the observational data sets. To evaluate the effect of different resolutions on the model spread, the coefficient of variation was selected to estimate the relative contribution of the spatial resolution. The spread originating from spatial resolution is considerably lower, and generally dominated by the model spread. This is crucial, as it highlights the applicability also of coarser data sets on the catchment scale, assuming that the data set is able to represent the general precipitation features of the area.

As this study focusses on the applicability of gridded precipitation products as meteorological input for hydrological models, it is important to assess not only their climatology, but also evaluate their performance on daily time steps. Therefore, three indicators were presented, focusing on potential impacts on low and high flows, as well as the occurrence of high precipitation events. The results presented in this study indicated strong differences in the number of consecutive dry (CDD) and wet days (CWD), and the contribution of heavy precipitation to the annual precipitation (R95pTOT). For the reanalysis data sets, the results highlight a higher contribution of low but steady precipitation events resulting in a lower R95pTOT, more CWD, and less CDD. These precipitation characteristics highly impact potential hydrological modeling, as steady moderate rainfall events infiltrates in the soil

and will be available for plants, whereas higher intensity events of the same total amount might lead to flooding, and to eventually lower soil moistures [61]. Additionally, there is a higher sensitivity of flood response to temporal than spatial variability [8]. The sensitivity for discharge response to the spatial resolution of the rainfall event over smaller catchments is low, and the impact only moderate over large catchments [62]. The differences in the length of the dry spells is particularly important for low flows [63] which are of interest, especially for the southern part of the Adige catchment, and the agricultural areas that are characterized by irrigation, and affected by seasonal water scarcity [37].

As observational data sets are prone to undercatch and measurement errors, which are amplified in winter [24,25], and the location and density of precipitation stations [21] it is obvious that these data sets have large uncertainties, especially in the higher elevations where station density is low, or no stations are present at all. To bridge this gap, (regional) reanalysis data sets show great potential [50], and should therefore be included in the analysis to find the best possible reference precipitation. Although the remote sensing data set did not perform acceptably over this catchment, it is still important not to neglect these data sets and include them in the selection process. Similar to climate impact assessment studies and confirming recent findings [50], an ensemble approach is presented and recommended in this study to derive a more robust precipitation data set. This will help to overcome the issues with observations, especially over complex topography, and to quantify and reduce the related uncertainties. Nevertheless, the recent study highlights the importance of detecting data sets that show high deviations on the catchment scale, and exclude them from the ensemble. This is not proposed to reduce the ensemble spread, but rather is based on a careful selection process. As presented, a variety of indicators should be evaluated, and selected specifically for the desired application, to identify data sets that are not capable of reproducing the given conditions in a specific area or catchment. The data sets excluded from the ensemble all showed clear mismatches with not only the reference grid, but also the rest of the ensemble, and could therefore be excluded from the ensemble. Additionally, a more dense station network is required, also in the higher elevations, to reduce the uncertainties in the observational data sets. This is especially crucial for the alpine areas in the northern parts of the catchment that are characterized by complex topography. It must be highlighted, that the selection of data sets is highly dependent on the region and hardly transferable to other regions and climate zones. The final ensemble mean now provides a robust basis for further activities such as hydrological modeling, climate model selection, and potentially for bias correction of climate model data.

Acknowledgments: This work has been supported by the European Communities 7th Framework Programme Funding under Grant agreement no. 603629-ENV-2013-6.2.1-Globaqua.

Author Contributions: The study was carried out by the first author as part of his PhD research, supervised by the second author. Both authors contributed to the concept and writing of this paper, with the first author preparing the first version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Schneider, U.; Ziese, M.; Meyer-Christoffer, A.; Finger, P.; Rustemeier, E.; Becker, A. The new portfolio of global precipitation data products of the Global Precipitation Climatology Centre suitable to assess and quantify the global water cycle and resources. *Proc. Int. Assoc. Hydrol. Sci.* **2016**, *374*, 29. [CrossRef]
2. Daly, C.; Slater, M.E.; Roberti, J.A.; Laseter, S.H.; Swift, L.W. High-resolution precipitation mapping in a mountainous watershed: Ground truth for evaluating uncertainty in a national precipitation dataset. *Int. J. Climatol.* **2017**, *37*, 124–137. [CrossRef]
3. Prein, A.F.; Gobiet, A.; Truhetz, H.; Keuler, K.; Goergen, K.; Teichmann, C.; Fox Maule, C.; van Mejgaard, E.; Déqué, M.; Nikulin, G.; et al. Precipitation in the EURO-CORDEX 0.11° and 0.44° simulations: High resolution, high benefits? *Clim. Dyn.* **2016**, *46*, 383–412. [CrossRef]

4. Ivancic, T.J.; Shaw, S.B. Examining why trends in very heavy precipitation should not be mistaken for trends in very high river discharge. *Clim. Chang.* **2015**, *133*, 681–693. [CrossRef]
5. Stephens, E.; Day, J.J.; Pappenberger, F.; Cloke, H. Precipitation and floodiness. *Geophys. Res. Lett.* **2015**, *42*, 316–323. [CrossRef]
6. Madsen, H.; Lawrence, D.; Lang, M.; Martinkova, M.; Kjeldsen, T.R. Review of trend analysis and climate change projections of extreme precipitation and floods in Europe. *J. Hydrol.* **2015**, *519*, 3634–3650. [CrossRef]
7. Bacchi, B.; Kottegoda, N.T. Identification and calibration of spatial correlation patterns of rainfall. *J. Hydrol.* **1995**, *165*, 311–348. [CrossRef]
8. Paschalis, A.; Fatichi, S.; Molnar, P.; Rimkus, S.; Burlando, P. On the effects of small scale space–time variability of rainfall on basin flood response. *J. Hydrol.* **2014**, *514*, 313–327. [CrossRef]
9. Hijmans, R.J.; Cameron, S.E.; Parra, J.L.; Jones, P.G.; Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **2005**, *25*, 1965–1978. [CrossRef]
10. Compo, G.P.; Whitaker, J.S.; Sardeshmukh, P.D.; Matsui, N.; Allan, R.J.; Yin, X.; Gleason, E., Jr.; Vose, R.S.; Rutledge, G.; Bessemoulin, P.; et al. The twentieth century reanalysis project. *Q. J. R. Meteorol. Soc.* **2011**, *137*, 1–28. [CrossRef]
11. Dee, D.; Uppala, S.M.; Simmons, A.J.; Berrisford, P.; Poli, P.; Kobayashi, S.; Andrae, U.; Balmaseda, M.A.; Balsamo, G.; Bauer, P.; et al. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **2011**, *137*, 553–597. [CrossRef]
12. Schamm, K.; Ziese, M.; Raykova, K.; Becker, A.; Finger, P.; Meyer-Christoffer, A.; Schneider, U. GPCC Full Data Daily Version 1.0: Daily Land-Surface Precipitation from Rain Gauges Built on GTS Based and Historic Data. Available online: https://rda.ucar.edu/datasets/ds497.0/ (accessed on 30 July 2017).
13. Herrera, S.; Gutiérrez, J.M.; Ancell, R.; Pons, M.R.; Frías, M.D.; Fernández, J. Development and analysis of a 50-year high-resolution daily gridded precipitation dataset over Spain (Spain02). *Int. J. Climatol.* **2012**, *32*, 74–85. [CrossRef]
14. Isotta, F.A.; Frei, C.; Weilguni, V.; Percec Tadic, M.; Lassègues, P.; Rudolf, B.; Pavan, V.; Cacciamani, C.; Antolini, G.; Ratto, S.M.; et al. The climate of daily precipitation in the Alps: Development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. *Int. J. Climatol.* **2014**, *34*, 1657–1675. [CrossRef]
15. Häggmark, L.; Ivarsson, I.; Gollvik, S.; Olofsson, O. Mesan, an operational mesoscale analysis system. *Tellus* **2000**, *52*, 2–20. [CrossRef]
16. Kalnay, E.; Kanamitsu, M.; Kistler, R.; Collins, W.; Deaven, D.; Gandin, L.; Iredell, M.; Saha, S.; White, G.; Woollen, J.; et al. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **1996**, *77*, 437–471. [CrossRef]
17. Rienecker, M.M.; Suarez, M.J.; Gelaro, R.; Todling, R.; Bacmeister, J.; Liu, E.; Bosilovich, M.G.; Schubert, S.D.; Takacs, L.; Kim, G.-K.; et al. MERRA: NASA's modern-era retrospective analysis for research and applications. *J. Clim.* **2011**, *24*, 3624–3648. [CrossRef]
18. Kummerow, C.; Barnes, W.; Kozu, T.; Shiue, J.; Simpson, J. The tropical rainfall measuring mission (TRMM) sensor package. *J. Atmos. Ocean. Technol.* **1998**, *15*, 809–817. [CrossRef]
19. Ashouri, H.; Hsu, K.L.; Sorooshian, S.; Braithwaite, D.K.; Knapp, K.R.; Cecil, L.D.; Nelson, B.R.; Prat, O.P. PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bull. Am. Meteorol. Soc.* **2015**, *96*, 69. [CrossRef]
20. Haylock, M.R.; Hofstra, N.; Klein Tank, A.M.G.; Klok, E.J.; Jones, P.D.; New, M. A European daily high resolution gridded data set of surface temperature and precipitation for 1950–2006. *J. Geophys. Res.* **2008**, *113*, D20. [CrossRef]
21. Isotta, F.A.; Vogel, R.; Frei, C. Evaluation of European regional reanalyses and downscalings for precipitation in the Alpine region. *Meteorologische Zeitschrift* **2015**, *24*, 15–37. [CrossRef]
22. Beck, H.E.; van Dijk, A.I.J.M.; Levizzani, V.; Schellekens, J.; Miralles, D.G.; Martens, B.; de Roo, A. MSWEP: 3-h 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 589–615. [CrossRef]
23. Blenkinsop, S.; Lewsis, E.; Chan, S.C.; Fowler, H.J. Quality-control of an hourly rainfall dataset and climatology of extremes for the UK. *Int. J. Climatol.* **2016**, *37*, 722–740. [CrossRef] [PubMed]
24. Goodison, B.E.; Louie, P.Y.T.; Yang, D. *WMO Solid Precipitation Measurement Intercomparison*; Final Report, WMO/TD-No. 872; World Meteorological Organization: Geneva, Switzerland, 1998.

25. Rasmussen, R.; Baker, B.; Kochendorfer, J.; Meyers, T.; Landolt, S.; Fischer, A.P.; Black, J.; Thériault, J.M.; Kucera, P.; Gochis, D.; et al. How well are we measuring snow: The NOAA/FAA/NCAR winter precipitation test bed. *Bull. Am. Meteorol. Soc.* **2012**, *93*, 811–829. [CrossRef]
26. Kotlarski, S.; Keuler, K.; Christensen, O.B.; Colette, A.; Déqué, M.; Gobiet, A.; Goergen, K.; Jacob, D.; Lüthi, D.; van Meijgaard, E.; et al. Regional climate modeling on European scales: A joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geosci. Model. Dev.* **2014**, *7*, 1297–1333. [CrossRef]
27. Bosilovich, M.G.; Chen, J.; Robertson, F.R.; Adler, R.F. Evaluation of global precipitation in reanalyses. *J. Appl. Meteorol. Climatol.* **2008**, *47*, 2279–2299. [CrossRef]
28. Kidd, C.; Bauer, P.; Turk, J.; Huffman, G.J.; Joyce, R.; Hsu, K.L.; Braithwaite, D. Intercomparison of high-resolution precipitation products over Northwest Europe. *J. Hydrometeorol.* **2012**, *13*, 67–83. [CrossRef]
29. Nkiaka, E.; Nawaz, N.R.; Lovett, J.C. Evaluating global reanalysis precipitation datasets with rain gauge measurements in the Sudano-Sahel region: Case study of the Logone catchment, Lake Chad Basin. *Meteorol. Appl.* **2017**, *24*, 9–18. [CrossRef]
30. Palazzi, E.; Hardenberg, J.; Provenzale, A. Precipitation in the Hindu-Kush Karakoram Himalaya: Observations and future scenarios. *J. Geophys. Res. Atmos.* **2013**, *118*, 85–100. [CrossRef]
31. Sikorska, A.E.; Seibert, J. Value of different precipitation data for flood prediction in an alpine catchment: A Bayesian approach. *J. Hydrol.* **2016**. [CrossRef]
32. Maggioni, V.; Nikolopoulos, E.I.; Anagnostou, E.N.; Borga, M. Modeling satellite precipitation errors over mountainous terrain: The influence of gauge density, seasonality, and temporal resolution. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 4130–4140. [CrossRef]
33. Mei, Y.; Anagnostou, E.N.; Nikolopoulos, E.I.; Borga, M. Error analysis of satellite precipitation products in mountainous basins. *J. Hydrometeorol.* **2014**, *15*, 1778–1793. [CrossRef]
34. Nikolopoulos, E.I.; Anagnostou, E.N.; Borga, M. Using high-resolution satellite rainfall products to simulate a major flash flood event in northern Italy. *J. Hydrometeorol.* **2013**, *14*, 171–185. [CrossRef]
35. Frei, C.; Schär, C. A precipitation climatology of the Alps from high-resolution rain-gauge observations. *Int. J. Climatol.* **1998**, *18*, 873–900. [CrossRef]
36. Chiogna, G.; Majone, B.; Cano Paoli, K.; Diamantini, E.; Stella, E.; Mallucci, S.; Lencioni, V.; Zandonai, F.; Bellin, A. A review of hydrological and chemical stressors in the Adige catchment and its ecological status. *Sci. Total Environ.* **2016**, *540*, 429–443. [CrossRef] [PubMed]
37. Navarro-Ortega, A.; Acuña, V.; Bellin, A.; Burek, P.; Cassiani, G.; Choukr-Allah, R.; Dolédec, S.; Elosegi, A.; Ferrari, F.; Ginebreda, A.; et al. Managing the effects of multiple stressors on aquatic ecosystems under water scarcity. The GLOBAQUA project. *Sci. Total Environ.* **2015**, *503*, 3–9. [CrossRef] [PubMed]
38. Schulla, J.; Jasper, K. *Model Description Wasim-Eth*; Technical Report; Institute for Atmospheric and Climate Science, Swiss Federal Institute of Technology: Zürich, Switzerland, 2007.
39. Camera, C.; Bruggeman, A.; Hadjinicolaou, P.; Pashiardis, S.; Lange, M.A. Evaluation of interpolation techniques for the creation of gridded daily precipitation (1×1 km^2); Cyprus, 1980–2010. *J. Geophys. Res. Atmos.* **2014**, *119*, 693–712. [CrossRef]
40. Di Luzio, M.; Johnson, G.; Daly, C.; Eischeid, J.K.; Arnold, J. Constructing retrospective gridded daily precipitation and temperature datasets for the conterminous united states. *Am. Meteorol. Soc.* **2008**, *47*, 475–497. [CrossRef]
41. Lu, G.Y.; Wong, D.W. An adaptive inverse-distance weighting spatial interpolation technique. *Comput. Geosci.* **2008**, *34*, 1044–1055. [CrossRef]
42. Hasenauer, H.; Merganicova, K.; Petritsch, R.; Pietsch, S.A.; Thornton, P.E. Validating daily climate interpolations over complex terrain in Austria. *Agric. For. Meteorol.* **2003**, *119*, 87–107. [CrossRef]
43. Chai, T.; Draxler, R.R. Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geosci. Model. Dev.* **2014**, *7*, 1247–1250. [CrossRef]
44. Hutchinson, M.F.; McKenney, D.W.; Lawrence, K.; Pedlar, J.H.; Hopkinson, R.F.; Milewska, E.; Papadopol, P. Development and testing of Canada-wide interpolated spatial models of daily minimum–maximum temperature and precipitation for 1961–2003. *Am. Meteorol. Soc.* **2009**, *48*, 725–741. [CrossRef]
45. Hunter, R.D.; Meentemeyer, R.K. Climatologically aided mapping of daily precipitation and temperature. *J. Appl. Meteorol.* **2005**, *44*, 1501–1510. [CrossRef]
46. Xia, Y.; Fabian, P.; Winterhalter, M.; Zhao, M. Forest climatology: Estimation and use of daily climatological data for Bavaria, Germany. *Agric. For. Meteorol.* **2001**, *106*, 87–103. [CrossRef]

47. Huffman, G.J.; Adler, R.F.; Bolvin, D.T.; Gu, G.; Nelkin, E.J.; Bowman, K.P.; Hong, Y.; Stocker, E.F.; Wolff, D.B. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multi-year, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.* **2007**, *8*, 38–55. [CrossRef]
48. Joyce, R.J.; Janowiak, J.E.; Arkin, P.A.; Xie, P. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeorol.* **2004**, *5*, 487–503. [CrossRef]
49. Hou, A.Y.; Kakar, R.K.; Neeck, S.; Azarbarzin, A.A.; Kummerow, C.D.; Kojima, M.; Oki, R.; Nakamura, K.; Iguchi, T. The global precipitation measurement mission. *Bull. Am. Meteorol. Soc.* **2014**, *95*, 701–722. [CrossRef]
50. Prein, A.; Gobiet, A. Impacts of uncertainties in European gridded precipitation observations on regional climate analysis. *Int. J. Climatol.* **2017**, *37*, 305–327. [CrossRef] [PubMed]
51. Ma, L.; Zhang, T.; Frauenfeld, O.W.; Ye, B.; Yang, D.; Qin, D. Evaluation of precipitation from the ERA-40, NCEP-1, and NCEP-2 Reanalyses and CMAP-1, CMAP-2, and GPCP-2 with ground-based measurements in China. *J. Geophys. Res.* **2009**, *114*, D9. [CrossRef]
52. Reichle, R.H.; Liu, Q.; Koster, R.D.; Draper, C.S.; Mahanama, S.P.; Partyka, G.S. Land surface precipitation in MERRA-2. *J. Clim.* **2017**, *30*, 1642–1664. [CrossRef]
53. Bosilovich, M.G.; Lucchesi, R.; Suarez, M. MERRA-2: File Specification; GMAO Office Note No. 9 (Version 1.1). 2015; 73p. Available online: https://gmao.gsfc.nasa.gov/pubs/ (accessed on 30 November 2017).
54. Poli, P.; Hersbach, H.; Dee, D.P.; Berrisford, P.; Simmons, A.J.; Vitart, F.; Trémolet, Y. ERA-20C: An atmospheric reanalysis of the twentieth century. *J. Clim.* **2016**, *29*, 4083–4097. [CrossRef]
55. Hersbach, H.; Poli, P.; Dee, D. The observation feedback archive for the ICOADS and ISPD data sets. *ERA Rep. Ser.* **2015**, *18*, 74–85.
56. Adler, R.F.; Huffman, G.J.; Chang, A.; Ferraro, R.; Xie, P.P.; Janowiak, J.; Rudolf, B.; Schneider, U.; Curtis, S.; Bolvin, D.; et al. The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeorol.* **2003**, *4*, 1147–1167. [CrossRef]
57. Taylor, K.E. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.* **2001**, *106*, 7183–7192. [CrossRef]
58. Frich, P.; Alexander, L.V.; Della-Marta, P.; Gleason, B.; Haylock, M.; Klein, A.G.; Peterson, T. Observed coherent changes in climate extremes during the second half of the twentieth century. *Clim. Res.* **2002**, *19*, 193–212. [CrossRef]
59. Kioutsioukis, I.; Melas, D.; Zerefos, C. Statistical assessment of changes in climate extremes over Greece (1955–2002). *Int. J. Climatol.* **2010**, *30*, 1723–1737. [CrossRef]
60. Zolina, O.; Kapala, A.; Simmer, C.; Gulev, S.K. Analysis of extreme precipitation over Europe from different reanalyses: A comparative assessment. *Glob. Planet. Chang.* **2004**, *44*, 129–161. [CrossRef]
61. Trenberth, K.E. The impact of climate change and variability on heavy precipitation, floods, and droughts. *Encycl. Hydrol. Sci.* **2008**, 17. [CrossRef]
62. Nicótina, L.; Alessi Celegon, E.; Rinaldo, A.; Marani, M. On the impact of rainfall patterns on the hydrologic response. *Water Resour. Res.* **2008**, *44*. [CrossRef]
63. Diaz-Nieto, J.; Wilby, R.L. A comparison of statistical downscaling and climate change factor methods: Impacts on low flows in the River Thames, United Kingdom. *Clim. Chang.* **2005**, *69*, 245–268. [CrossRef]

© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

2.2 Paper II: Impact of reference data sets on climate model selection, bias correction and resulting climate change signals for precipitation. Journal of Hydrometeorology

Gampe, D. & Schmid, J. (2018). *Impact of precipitation reference data sets on climate model selection and resulting climate change signals.*

Paper II builds upon the analyses carried out in Paper I and assesses the uncertainty of model selection and bias correction of RCMs with a focus on observational uncertainty. The RCM ensemble presented in Paper III is extended as more simulations were available at this point. A simple evaluation scheme is constructed to evaluate the RCMs based on the capability to simulate the period 1989-2008. The entire RCM ensemble is then bias corrected using a well-established QM technique based on the observational data sets presented in Paper I. An uncertainty assessment is then carried out based on the variance decomposition method introduced in Paper III based on the four uncertainty sources included: GCM, RCM, RCP and observational uncertainty introduced by the bias correction. The study reveals large uncertainties introduced for absolute changes, especially for the projection of extremes where the uncertainties exceed the climate change signal. The paper adds to the field by bias correcting a RCM ensemble with an extensive set of reference data sets on the catchment scale that has not been carried out at this extent. It highlights the need to account for observational uncertainty in hydrological impact studies that is so far mostly neglected. This publication addresses the following research questions:

Q2: What is the relative contribution of observational uncertainty to the overall uncertainty in the hydro-climatic modeling chain?

Q3: What is the influence of observational uncertainty on post-processing, i.e. bias correction, climate model evaluation, model selection and the resulting future projections?

Author's contributions: The study was carried out by the D. Gampe who also prepared the first version of the manuscript. J. Schmid provided the MATLAB code for the bias correction, while all calculations were carried out by D. Gampe.
Status: submitted
Journal: Journal of Hydrometeorology (AMS)
Impact factor: 3.641

Impact of reference data sets on climate model selection, bias correction and resulting climate change signals for precipitation

David Gampe[1] and Josef Schmid[1]

(1) Department of Geography, Ludwig-Maximilians-Universität, Munich, Germany , david.gampe@lmu.de

Abstract: Gridded data sets for precipitation are of great importance to evaluate recent climate models and are frequently applied to select a subset of available models. As climate models are still prone to biases on the regional scale, gridded data sets are also essential to correct or adjust these biases. Various studies revealed considerable differences, i.e. observational uncertainty, in the available gridded data sets for precipitation, especially over complex terrain. This study focusses on the impacts of observational uncertainty on the evaluation, selection and bias correction of 15 Regional Climate Model (RCM) simulations provided through the EURO-CORDEX initiative over the alpine Adige catchment located in Northern Italy. Nine reference data sets originating from observations, reanalysis and remote sensing are applied to evaluate the performance of RCMs and select a subset based on validity. These reference data sets are then applied to bias correct the RCM ensemble using a standard quantile mapping method and the resulting changes in the projections are assessed. Results reveal only a minor impact on RCM selection, indicating that observational uncertainty is lower than model uncertainty. The influence of the choice of the reference grid on bias correction is negligible for the climate change signals, however model selection clearly influencing the projected change signals. As expected, the choice of the reference grid strongly influences future projections of precipitation even more pronounced for the extremes. The findings of this study highlight the need to account for observational uncertainty for bias correction of RCM simulations for impact modeling studies.

1. Introduction

Climate models serve as a primary tool to project future climate and are thus frequently used in impact models to assess future changes on the environment. Regional climate models (RCMs) that dynamically downscale global climate models (GCMs) provide high-resolution projections and are therefore of particular interest over complex terrain (Torma et al. 2015). An ensemble of RCMs is provided through the Coordinated Regional Downscaling Experiment (CORDEX) at a common grid with a horizontal resolution of 0.11° (~12 km, corresponding to the EUR-11 grid) over Europe (Jacob et al. 2014). Future scenarios are derived through Representative Concentration Pathways (RCPs) that result in a range of the increased radiative forcing of 2.6 to 8.5 Wm^{-2} for the year 2100 (Van Vuuren et al. 2011).

Despite the increased resolution, RCM simulations are still prone to systematic errors, i.e. model biases, when compared to regional observations (Christensen et al. 2008; Dosio, 2016; Rajczak et al. 2016; Smiatek et al. 2016). A correction of these biases is usually inevitable and applied when climate models are used for climate change impact assessment studies (Piani et al. 2009; Ehret et al. 2012; Muerth et al. 2013). To cope with this issue, several adjustment or correction techniques have been established. Out of these, distribution based approaches usually lead to more robust results (Chen et al. 2013a). Quantile mapping (QM) is a widely used technique to adjust climate model biases (Pierce et al. 2015; Addor et al. 2016) for precipitation, based on the distributions of historical data sets. After bias correction with QM the distribution of precipitation in the RCMs follows the distribution of the reference data set applied for bias correction.

In addition, it is often not feasible to consider the entire ensemble of available climate models in climate change impact assessment studies, mostly for computational reasons and the related difficulties to handle large data sets (Mendlik & Gobiet 2016). It is therefore necessary to define a subset of RCMs in a meaningful manner. Recent studies proposed a selection based on model independence (Mendlik & Gobiet 2016) or on a clustering approach based on climate change signals of various variables (Wilcke & Bärring 2016). Despite these statistical methods, model selection based on validity, i.e. the capability of the climate models to reproduce historical precipitation or temperature is still frequently applied (Pierce et al. 2009; Biemans et al. 2013; Kottlarski et al. 2017) possibly combined with analysis on model independence (Evans et al. 2013) or on climate change signals (Lutz et al. 2016).

Both, bias correction and model selection represent typical steps in climate change impact assessment studies. However, at least if a selection based on validity is performed, both require the use of reference data for the desired variable to be corrected or applied for selection. For precipitation a great variety of data sets is available. These can either be station data or gridded products derived from observations (Haylock et al. 2008; Isotta et al. 2014; Schamm et al. 2016), constructed through reanalysis (Compo et a. 2011; Dee et al. 2012; Bosilovich et al. 2015; Landelius et al. 2016; Poli et al. 2016), or originate from remote sensing (Kummerow et al. 1998; Ashouri et al. 2015). Available gridded products for precipitation differ not only on their source, but also in spatial and temporal resolution and domain size. It is widely known that these precipitation data sets are prone to errors, resulting from interpolation procedures or precipitation undercatch. Considerable differences exist for the general climatology (Palazzi et al. 2013; Isotta et al. 2015; Gampe & Ludwig 2017; Henn et al. 2018) and the representation of extremes (Herold et al. 2017).

Although the uncertainty related to gridded precipitation products is well known and widely acknowledged, the selection of the reference data set is often an arbitrary choice and the role of observational uncertainty often excluded from the analysis. For bias correction several studies were conducted to evaluate existing methods (Chen et al. 2013a; Lafon et al. 2013; Ruiz-Ramos et al. 2016) however rarely include the issue of reference grid uncertainty (Iizumi et al. 2017) and mostly focus on a single reference data set. For climate model

selection and evaluation great effort was conducted in comparing different methods over the past years (e.g. Schaller et al. 2011; Zubler et al. 2016) but only recently Kottlarski et al. (2017) added to the debate in the application of various reference grids for RCM selection based on a simple ranking scheme. The importance to consider observational uncertainty in precipitation for RCM evaluation over various areas over Europe was highlighted by Prein et al. 2017.

The presented study adds to the debate and complements previous research by assessing the role of observational uncertainty in the typical procedure of RCM selection and bias correction performed in many impact assessment studies. An alpine catchment serves as study site and a total of nine gridded precipitation products from various sources are included in the analysis. The three main research questions addressed can be defined as follows:

- What is the impact of the choice of the reference data set on bias detection and model selection?
- What is the impact of reference data set selection on bias correction and resulting future projections for precipitation?
- How large is the contribution of reference grid selection to the overall uncertainty?

A simple ranking scheme based on validity is introduced, where the performance of each RCM as evaluated by each of the included reference grids is assessed. The RCM ensemble is then bias corrected using each of the reference data sets and the resulting climate change signals for the selected models are presented. To quantify the uncertainty introduced by the selection of the reference grid, the relative contribution to the overall uncertainty is assessed by a variance decomposition approach.

2. Study Area & Data Sets

2.1. Study Area

This study is carried out over the catchment of the Adige River located in Northern Italy. The catchment, presented in Figure 1, covers an area of 12,100 km² ranging from the Southern Alps to the Adriatic Sea over a length of 409 km while passing three Italian provinces. About two thirds of the catchment is located over complex terrain in the Alps, with elevations up to 3800 m.a.s.l. The area is generally dominated by a humid climate with annual mean precipitation ranging from 500 mm over the inner-Alpine dry valleys, as the Venosta valley (Frei & Schär, 1998; Isotta et al. 2014), to 1600 mm with a strong peak in the summer months and comparatively dry winters (Chiogna et al. 2016). The hydrology of the Adige River is dominated by snow and glacier melting in the spring months and intense precipitation events in summer (Chiogna et al., 2016). The stream discharge serves for irrigation of intensive agriculture in summer and various touristic activities, especially in winter. Additionally, several reservoirs for power generation are located in the northern part of the catchment (Navarro-Ortega et al., 2015).

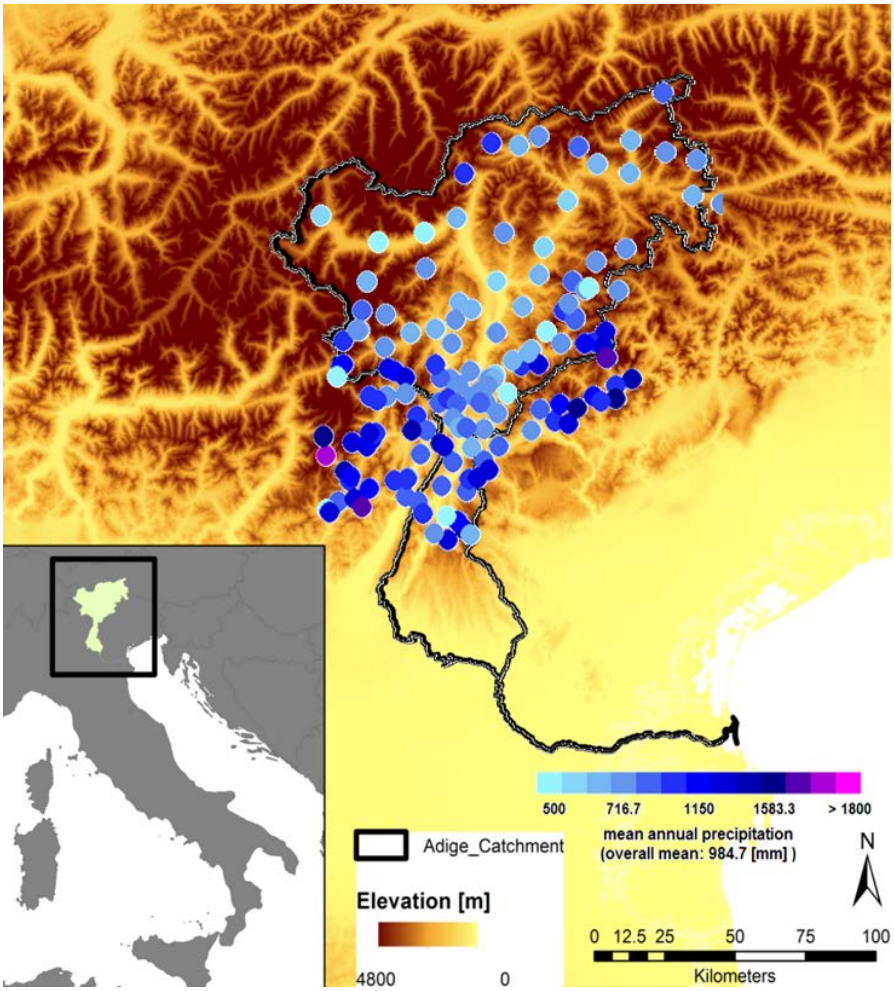


Figure 1: The Adige catchment located in Northern Italy. Circles show the precipitation gauges and their corresponding mean annual precipitation for the period 1989-2008.

2.2. Gridded Precipitation Data Sets

At this point, a multitude of gridded data sets for precipitation are available around the globe. The selection is usually determined by the requirements regarding spatial and temporal resolution as well as the location of the study area. In this study, extremely coarse data sets (>2.0° grid size) were not included due to the comparatively small size of the catchment. Additionally, as the data sets serve for bias correction of RCMs, the minimum temporal resolution was daily time steps. Monthly or seasonal data sets were hence also excluded from the analysis. Furthermore, to achieve a temporal coverage of at least 20 years, the period 1989-2008 was selected, corresponding to the maximum overlap of the high resolution data sets. Therefore, data sets with a later starting data than Jan. 1st 1989 or ending dates prior to Dec. 31st 2008 were rejected. This led to a selection of the nine data sets summarized in Table 1. These are briefly introduced in the following, while for detailed information the reader is referred to the corresponding references. An extensive comparison of these data sets was carried out over the study area in a previous study (Gampe & Ludwig, 2017).

Table 1: Gridded precipitation data sets included in this study with their corresponding resolution, temporal coverage, source and the reference publication. All data sets are available at daily time scale; the IDs correspond to the numbering in Figure 3 & Figure 4.

ID	Data set	Resolution	Temporal coverage	Type / Source	Reference
1	ADG-1KGPR	1 km	1989-2008	Observations	Gampe & Ludwig 2017
2	EURO4m-APGD	5 km	1971 - 2008	Observations	Isotta et al. 2014
3	MESAN	5 km	1989 - 2010	Downscaling / Reanalysis	Landelius et al. 2016
4	E-OBS v. 11	25 km	1950 - 2015	Observations	Haylock et al. 2008
5	PERSIANN-CDR	0.25° (~30 km)	1983 - present	Multisatellite (infrared), corrected	Ashouri et al. 2015
6	MERRA-2	0.5° latitude x 0.625° longitude (~50 km)	1980 - present	Reanalysis	Bosilovich et al. 2015
7	ERA-Interim	0.75° (~80 km)	1979 - present	Reanalysis	Dee et al. 2011
8	ERA-20C	125 km	1900-2010	Reanalysis	Poli et al. 2016
9	GPCC –FDD v1.0	1.0° (~100 km)	1988-2013	Observations	Schamm et al. 2016

Four observational data sets derived from interpolation of precipitation gauge measurements are included in this study at spatial resolutions ranging from 1 km to ~100 km. The highest resolution data set included is the local gridded precipitation data set over the Adige at 1 km grid size (ADG-1KGPR). This was derived from a relatively dense station network of over 150 stations in the catchment area, as presented in Figure 1. A combination of bilinear interpolation with elevation dependent regression was applied to construct a 1 km high resolution grid. Although the effective resolution is lower for some parts of the catchment, the data set showed robust results over the covered period 1989 – 2008 (Gampe & Ludwig, 2017). The second observational data set included is the Alpine Gridded Data Set constructed within the EURO4m project (EURO4m-APGD). It was derived from numerous stations over the Alpine area and provides daily precipitation for the period 1971-2008 at a 5 km grid (Isotta et al. 2014). Additionally, the widely used E-OBS data set (Haylock et al. 2008) available through the European Climate Assessment and Dataset (ECA&D) was included in version 11.0. E-OBS provides daily precipitation information for the European domain at a 25km grid, starting in the year 1951 and is updated frequently. Quality controlled stations with long data records are included in this data set, resulting in a considerably lower station density, compared to the high resolution grids mentioned, with approximately two stations per grid cell over the Alpine region (Isotta et al. 2015; Prein & Gobiet, 2017). The coarsest observational data set included is the Full Data Daily Product from the Global Precipitation Climatology Center (GPCC-FDD, Schamm et al. 2016) of the German Weather Service (DWD). This data set covers the period 1988-2013 and provides daily precipitation information on a 1° grid with a station density similar to E-OBS.

In addition to the observational data sets, three reanalysis products and one downscaling of a reanalysis data set are included ranging from a 5 to 125 km grid size. The first is the Modern Era Retrospective-analysis for Research and Applications in the second realization (MERRA-2) available through the National Aeronautics and

Space Administration (NASA). The data set provides precipitation information since 1980 at a ~50 km grid with ongoing updates and uses observed precipitation for the parameterization of the land surface scheme (Bosilovich et al. 2015). The two coarsest reanalysis products are available through the European Centre for Medium-Range Weather Forecasts (ECMWF). The ERA-Interim provides global precipitation information at ~80 km and assimilates various observational data sets in a 12hr cycle (Dee et al. 2011). The second ECMWF product is the Reanalysis for the 20th Century (ERA-20C) which only assimilates surface pressure and marine wind observations (Hersbach et al. 2015; Poli et al. 2016). ERA-20C is available at a ~125 km grid and offers precipitation information since 1900. The highest resolution product in this family is the Mesoscale Analysis (MESAN) that downscales the High Resolution Limited-Area Model (HIRLAM) to a 5 km grid (Landelius et al. 2016). The HIRLAM reanalysis uses ERA-Interim as boundary conditions and assimilates observational data on a 22 km grid (Dahlgren et al. 2016). MESAN provides precipitation information for the period 1989-2010.

The last data set included is the satellite-based Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN, Hsu et al. 1997) in the bias-corrected version PERSIANN-CDR. Geostationary satellite information of infrared brightness temperature is used in PERSIANN to derive precipitation patterns at ~30 km spatial resolution. PERSIANN-CDR corrects the precipitation amounts using monthly precipitation data from the Global Precipitation Climatology Project (GPCP) at 2.5° while keeping the high resolution and the spatial patterns of the original PERSIANN and with data record since 1983 (Ashouri et al. 2015).

2.3. Climate Model Data

The climate models included in this study comprise 15 simulations provided through the Coordinated Regional Climate Downscaling Experiment over the European domain (EURO-CORDEX; Jacob et al., 2013). Table 2 gives an overview on the included GCM-RCM combinations with the responsible institute. As the higher resolution realizations of those RCMs at 0.11° (12.5 km) show improvements over the 0.44° simulations (Prein et al. 2015) only the high resolution simulations were applied. With the exception of the variance decomposition presented in section 3.3 and Figure 9, only the simulations under RCP 8.5 are considered as the ambiguous climate change signals over the Adige catchment under RCP 4.5 (Gampe et al. 2016) would complicate the interpretation of the results in this study. Therefore, with the mentioned exceptions, all results presented and conclusions drawn refer to RCP 8.5 only.

All analyses and calculations presented in this study were carried out on a common grid. The EURO-CORDEX grid was selected as target grid. Following previous studies, e.g. Volosciuk et al. 2017, the bias correction was applied on the RCM grid. As it is desirable to keep the benefits of the higher resolution RCMs, a conservative remapping was performed on the reference data sets, which clearly penalizes the coarser grids. Nevertheless, this study is not meant as an evaluation of the precipitation reference data sets, but rather focusing on the consequences for bias correction and model selection.

Table 2: RCM ensemble of EURO-CORDEX simulations at 12 km resolution applied in this study under rcp 4.5 and 8.5.

Regional Climate Model (RCM)	General Circulation Model (GCM)
SMHI-RCA4	CNRM-CM5 HadGEM2-ES EC-EARTH-r12 MPI-ESM-LR IPSL-CM5A-MR
DMI-HIRHAM5	EC-EARTH-r3
KNMI-RACMO22E	EC-EARTH-r12 HadGEM2-ES
CLMcom-CCLM4-8-17	CNRM-CM5 HadGEM2-ES EC-EARTH-r12 MPI-ESM-LR
IPSL-INERIS-WRF331F	IPSL-CM5A-MR
MPI-CSC-REMO2009 (2 members)	MPI-ESM-LR

3. Methods

3.1. Model selection

Similar to previous studies (Pierce et al. 2009; Biemans et al. 2013; Kottlarski et al. 2017), model selection here is performed based on validity, i.e. the capability of the RCMs to reproduce the historical period. A relatively simple scheme is presented to derive the performance of each RCM and a ranking is performed. Although more sophisticated and complex approaches exist, the here presented method is sufficient for the purpose of this study as the focus is not to elaborate a sophisticated ranking scheme and the presented metrics and ranking led to satisfactory results. In addition it is crucial to mention, that the ranking is dependent on the selected error measures and the overall focus of the study and hence somewhat subjective (Kottlarski et al. 2017). The model selection is based on uncorrected RCM data under RCP 8.5 as derived from the ensemble of RCMs presented. The following selection procedure is calculated for each of the nine reference grids.

To assess the performance of the RCMs and to derive the ranking based on their capability of reproducing historical climate, several metrics were calculated. To compare data sets of the same length, the period 1989 – 2008 was chosen for the model selection. Model performance was evaluated on based on daily catchment precipitation derived from the common 0.11° grid. Each of the metrics was calculated for each reference grid *r*, each climate model *c* and each month *m*.

To assess the capability of the RCMs to reproduce the general climatology, the percent bias (PBIAS) was calculated defined as follows, where *RCM* refers to the evaluated RCM and *Ref* corresponds to the reference grid used for the evaluation:

$$\mathrm{PBias} = \frac{1}{\mathrm{m}} \sum_{\mathrm{m}=1}^{12} \frac{(\overline{\mathrm{RCM}}_{\mathrm{c,m}} - \overline{\mathrm{Ref}_{\mathrm{r,m}}})}{\overline{\mathrm{Ref}_{\mathrm{r,m}}}} \tag{1}$$

The overall bias based on annual precipitation instead of monthly time steps was calculated in a similar way. To further assess the performance, the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE) were computed based on the entire, ranked time series. In addition, to assess the performance of the upper end of the distribution, both metrics were computed for the 90th percentile as follows, 95th, 99th and 99.9th percentiles were calculated the same way.

$$\mathrm{MAE} = \frac{1}{\mathrm{m}} \sum_{\mathrm{m}=1}^{12} |\mathrm{P}^{90}(\mathrm{RCM}_{\mathrm{c,m}}) - \mathrm{P}^{90}(\mathrm{Ref}_{\mathrm{r,m}})| \tag{2}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{m}} \sum_{\mathrm{m}=1}^{12} |\mathrm{P}^{90}(\mathrm{RCM}_{\mathrm{c,m}}) - \mathrm{P}^{90}(\mathrm{Ref}_{\mathrm{r,m}})|} \tag{3}$$

In addition, to assess the variability in the reference and RCM data sets, the inter-annual variability of both, annual and monthly precipitation sums were calculated based on the standard deviations over the 20-year period.

For each of the metrics, and each of their variations (temporal or percentile based), a simple score system was applied to rank the models. In a first step, the resulting metrics were normalized individually so that the best performing model was awarded 100 % and the other models decrease accordingly. In a second step, the average score in percent was calculated as the simple mean over the calculated metrics. A score of 100% does hence not mean the model performs perfectly, but is rather the best-performing model relative to the rest of the ensemble. To reduce the ensemble to a feasible number of simulations, the best three and the best seven ranking RCMs based on each of the reference data sets were selected for the analyses.

3.2. Bias Correction Method

As RCMs are usually prone to biases on the regional or local scale, bias correction is, although controversy discussed, a necessary step (Muerth et al. 2013). Here a two-step correction was applied. First, to avoid artefacts in the RCM distributions a simple threshold was introduced to correct the drizzle effect, which produces too many days with very low precipitation (Teutschbein et al. 2013). In a second step a distribution based correction method based on the Daily Translation (DT) method (Mpelasoka & Chiew, 2009; Chen et al. 2013b) was applied to correct RCM precipitation. DT is similar to other QM bias correction methods and showed great for effective correction (Chen et al. 2013b). In contrary to Mpelasoka & Chiew (2009), the bias correction does not include an additional downscaling in this study, as all data sets were remapped to the RCM grid previously. In DT a relation between the distribution of precipitation in the reference grid and the historical RCM data set is established. The resulting correction factors for each percentile are then applied to correct the historical and future simulations of RCM precipitation. In this study the multiplicative correction factors are derived on a monthly basis and for integer percentiles. The period 1989 – 2008 served as calibration to derive the monthly correction factors which were then

applied to correct the RCM ensemble introduced for the historical period (1981 – 2010) and future period (2036 - 2065). The bias correction procedure was applied to the entire RCM ensemble under RCP 8.5 for each of the nine reference data sets.

3.3. Variance Decomposition

Four main sources of uncertainty can be identified in the design of the study: model uncertainty originating from GCMs and correspondingly uncertainty introduced through RCMs; scenario or radiative uncertainty, introduced through different radiative forcings, RCP 4.5 and RCP 8.5; and finally uncertainty introduced through post-processing. The latter here is the key interest and defined as the uncertainty introduced through the bias correction with different reference data sets. To quantity the relative contribution of each of the four sources to the overall uncertainty, the widely used approach of variance decomposition (Ferro 2004; Déqué et al. 2007; Déqué et al. 2012) is applied. The method will be briefly introduced here, while for a detailed presentation of the concept it is referred to Déqué et al. 2007.

Within this the variance in the climate response, i.e. the climate change signal defined as the difference in precipitation for the future period (2036 – 2065) and the reference period (1981 – 2010), is analyzed. The total variance of the ensemble is given by the sum of each of the four sources of uncertainty, where *G* is defined as the contribution from the GCMs, *R* the variance within the RCMs, *S* the scenario uncertainty and *B* the uncertainty attributed to bias correction:

$$V_{tot} = G + R + S + B \tag{4}$$

As neither of the sources of uncertainty appears solely the interaction terms of the four sources have to be considered making equation 4 more complex and more difficult to interpret:

$$V_{tot} = G + R + S + B + GR + GS + GB + RS + RB + SB + GRS + GRB + GSB + RSB + GRSB \tag{5}$$

The variance due to each source of uncertainty can hence be interpreted as the sum of all interaction terms with the corresponding variable. The uncertainty due to bias correction can thus be as:

$$V_B = B + GB + RB + SB + GRB + GSB + RSB + GRSB \tag{5}$$

Uncertainties due to GCM (G), RCM (R) and RCP or scenario (S) can be written accordingly.

The assumption of the total variance being the sum of each uncertainty component as in equations 4 & 5 is only valid for a complete matrix. The matrix is complete for scenarios, as all simulations are available under RCP 4.5 and 8.5 and for bias correction, as the entire ensemble was bias corrected with each reference data set. However, as presented in Table 2 not all GCM-RCM combinations are available and the matrix has to be artificially filled. We follow the simple, yet robust approach presented by Déqué et al. 2007 to solve this problem, where the calculation the climate responses was computed that the interaction term GRSB equals zero. Thus, the interaction terms are not correct by definition, however the goal of this study is to estimate the contribution of each source to the overall variance and not represent the interaction between them. The final variance hence represents the fraction of uncertainty of each source rather than the total variance. The interaction terms are therefore included in the final variance presented here and not explicitly mentioned. The sum of the overall variance exceeds 100% due to the interaction terms; therefore the final variance is normalized to 1 or 100% respectively to represent the fraction but allow for easier and more accessible comparison of the results. The variance decomposition was carried out based on monthly climate change signals in this study.

4. Results

Figure 2 shows the mean monthly precipitation sums over the Adige catchment for the EURO-CORDEX ensemble (grey bars, single models shown as black squares) and the nine selected reference grids (colored squares) for the period 1989-2008. In general, the RCMs show a wet bias for most months and the spread of the climate model ensemble exceeds the spread of the reference grids. Additionally, the spread of both, RCMs and reference grids, is larger during the summer period (JJA) compared to the winter months (DJF) due to convective and heavy precipitation events in JJA. The detection of a potential model bias is easier to interpret in DJF, where each member of the RCM ensemble shows higher precipitation compared to each reference grid, with the exception of E-OBS in December. In JJA the situation is different and the detection of model bias is more complex. Compared to ERA-Interim (blue square), most of the RCMs show a dry bias, while compared to the high resolution ADG-1KGPR (grey square), GPCC-FDD (green) and MERRA2 (dark brown) the majority of the included RCMs show a wet bias.

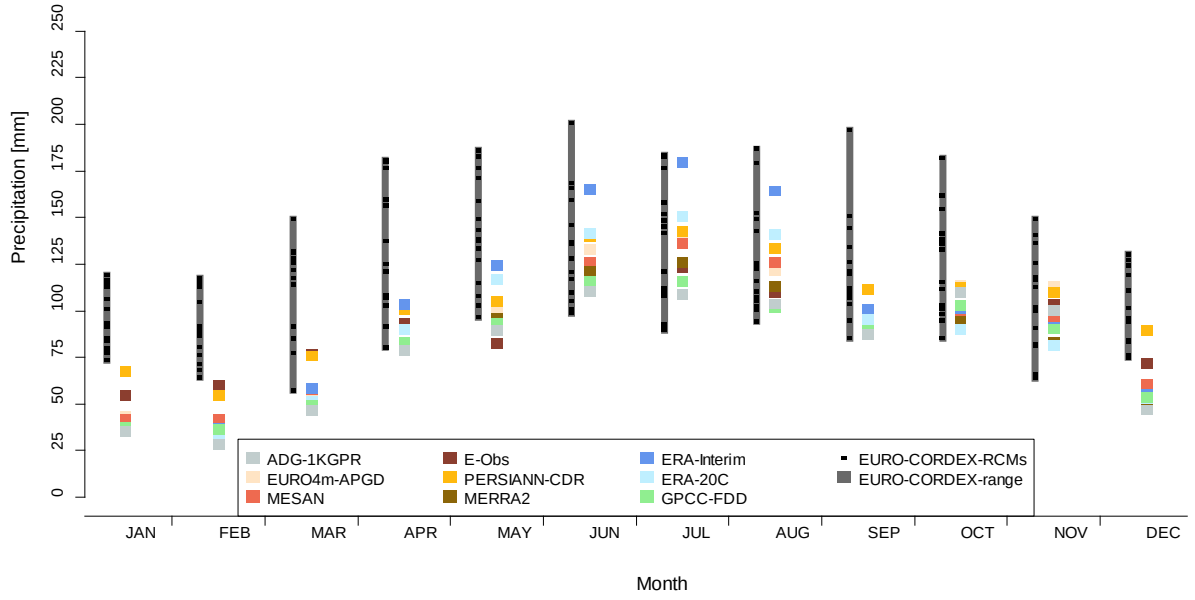


Figure 2: Uncorrected climatological monthly precipitation from EURO-CORDEX RCMs (black squares), the corresponding spread (dark gray bars) and the selected reference grids (colored squares) for the period 1989-2010.

To assess the RCM bias, Figure 3 shows the percent model bias based on mean annual precipitation for the years 1989-2008 for the RCM ensemble and the nine reference grids, numbering refers to Table 1. A positive bias, shown in red coloring, indicates an overestimation while a negative bias (blue coloring) refers to precipitation underestimation. Circles in the panels indicate a bias lower than the observational uncertainty. The latter was derived from the coefficient of variation (CV) which is calculated as the standard deviation normalized by the overall mean of the reference grids and given in percent. In this case the CV of the reference grids based on annual precipitation is around 10%. The average bias shown refers to the mean bias of each RCM over the reference data sets. Depending on the reference grid up to five RCMs (for PERSIANN-CDR, column 5) show biases within this range and can therefore not be considered as biased. As already indicated for monthly precipitation in Figure 2, most RCMs show a positive bias also for annual precipitation compared to the majority of reference grids. However, PERSIANN-CDR (5) and ERA-Interim (7) indicate underestimations of up to 20% for the RACMO22E simulations, while other reference data sets indicate a small positive bias for this RCM. Consequently, the average bias is close to zero as the positive biases are balanced with the negative biases. In contrary, other RCMs, such as CCLM4-8-17

and RCA4 show positive biases throughout the reference grids, independent from the driving GCM and show high average biases of up to 77%.

This is also reflected in the average score for model selection as presented in Figure 4 as RCMs with high annual bias show considerable lower scores. However, models with lower annual bias, e.g. KNMI-RACMO22E driven by the HadGEM2-ES, do not necessarily score better than other models. The results additionally are more homogenous than the annual bias presented in Figure 3. The least scoring RCMs show low performance throughout the reference data sets, while the best scoring RCMs are also similar for all reference data sets. As a consequence, for many reference data sets the same models are amongst the selected three or seven respectively. The only reference data set with considerable different scoring patterns is ERA-Interim, however the tendencies are similar. For each reference grid, the best three and seven RCMs are selected for the subsequent analyses of future changes.

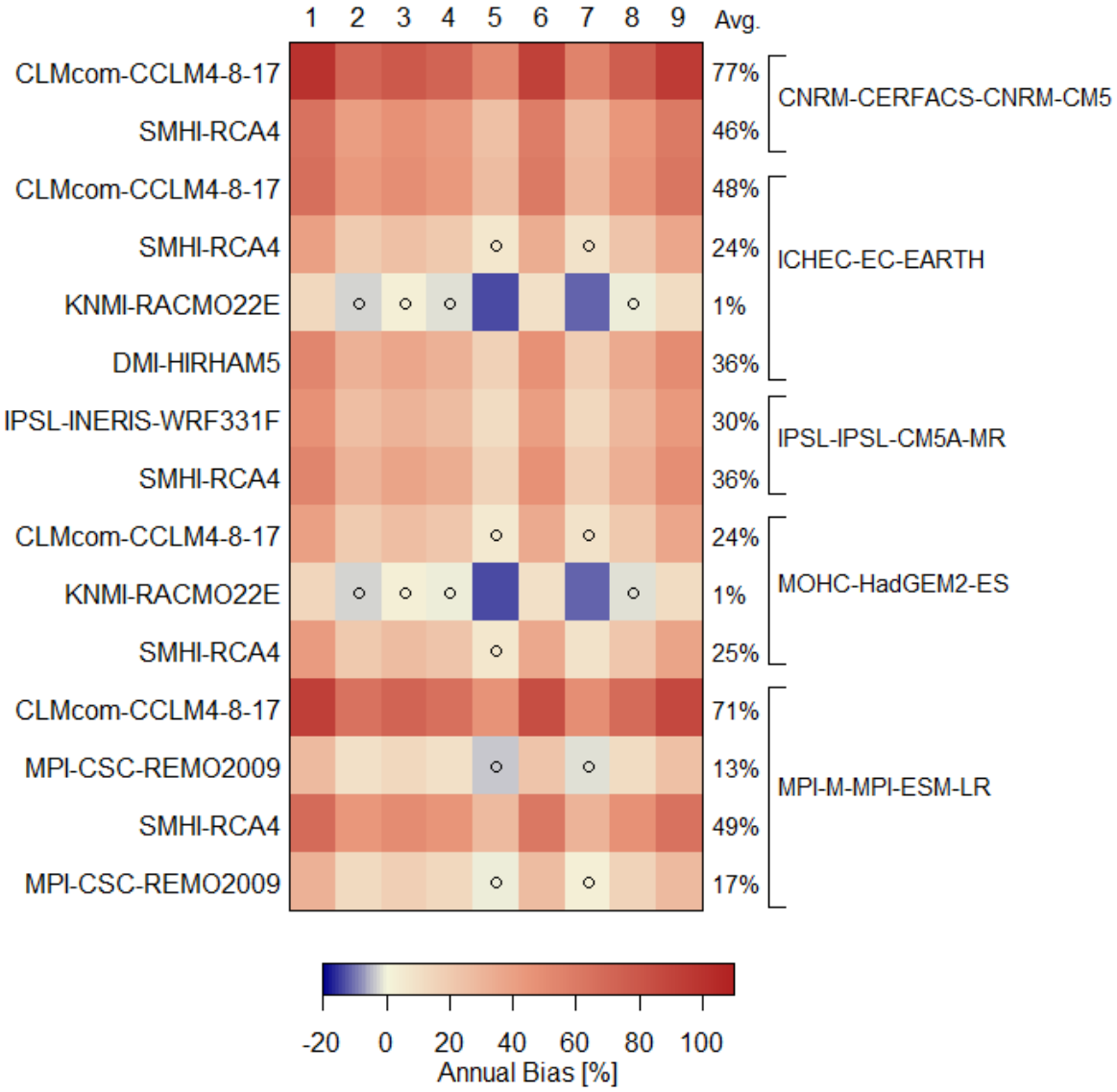


Figure 3: Model bias for each RCM with respect to each reference grid. Circles denote that the bias falls within the range of the observational uncertainty, which here is defined as the coefficient of variation of the climatological annual precipitation in the reference grids (+/- 10%).

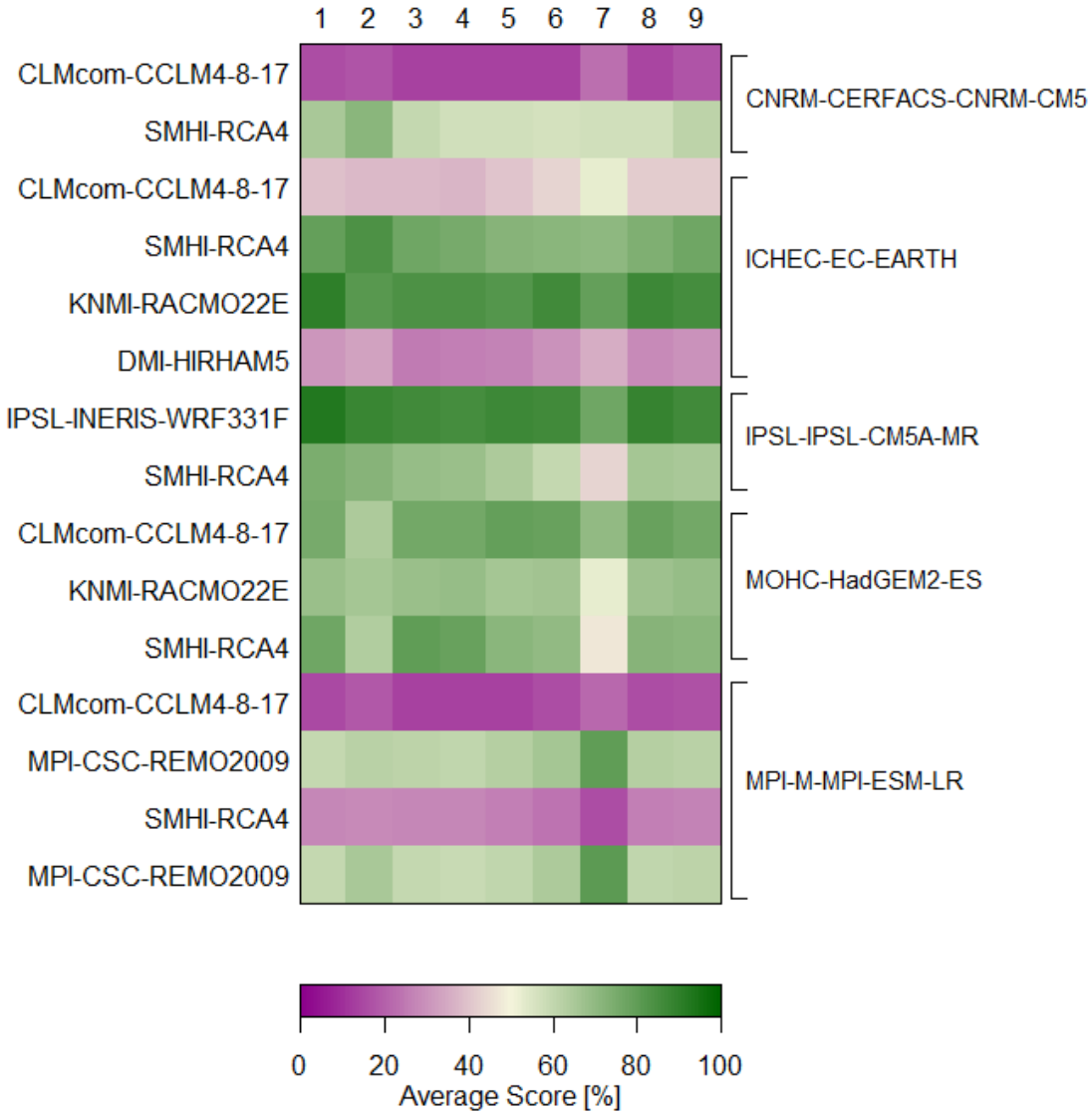


Figure 4: Average score (%) for each RCM and reference grid. The average score is based on several metrics and normalized from 0% (worst RCM) to 100% (best RCM).

Figure 5 shows the percentiles of daily precipitation for the original (grey) and bias corrected (red) RCM simulations with the corresponding reference products (panels a – i) used for bias correction. The positive bias detected in the climatological annual precipitation of the uncorrected RCMs (grey) can be clearly identified throughout the entire distribution for almost all reference products compared to the corresponding reference grid (black line). The corrected RCMs (red lines) agree better with each of the corresponding reference grids. The correction of the drizzle effect led to a better representation of the lower end of the distribution, hence very little or no precipitation. As expected, bias correcting the RCM ensemble to observed percentiles results in a considerable reduction of the RCM spread over the historical period. The highest differences between the bias-corrected RCM ensembles can be identified for low precipitation and the threshold of no-precipitation days, as well as the tail of the distribution characterizing extreme precipitation.

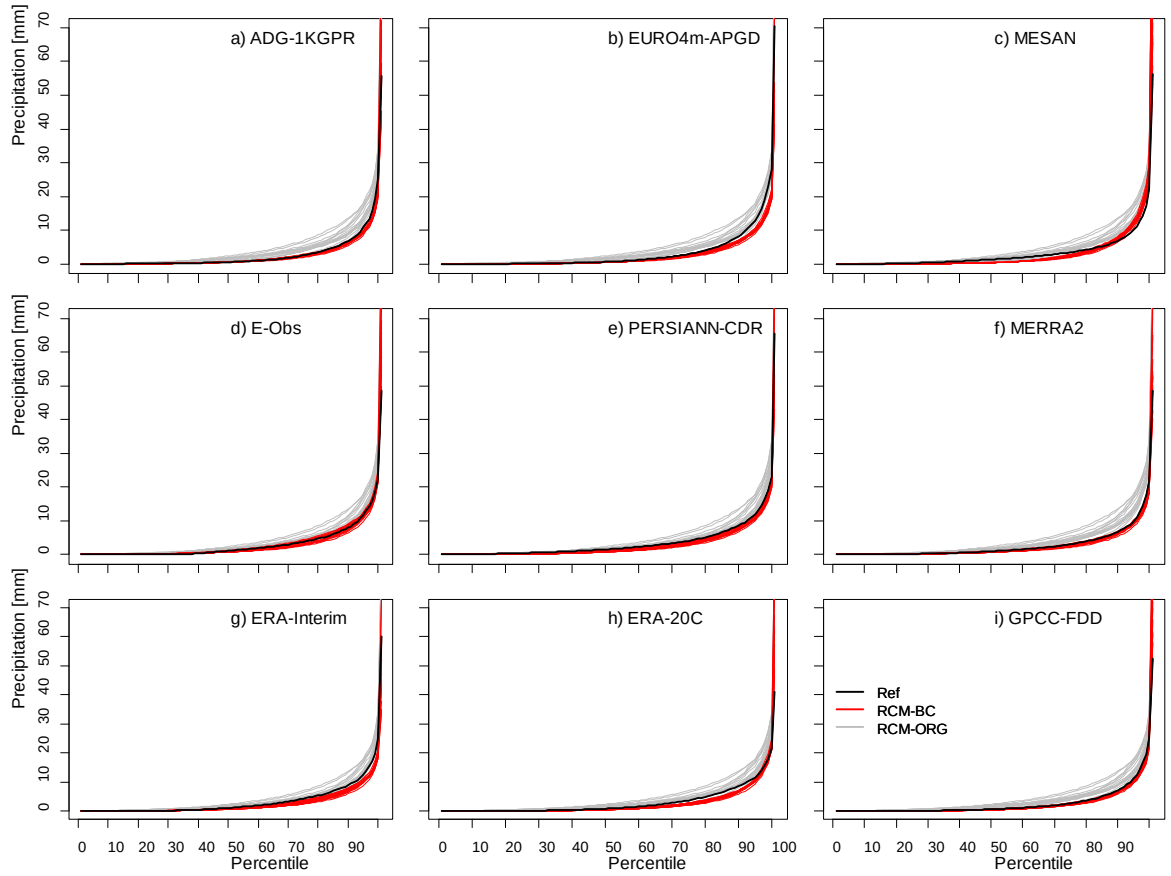


Figure 5: Percentiles of daily catchment precipitation for the selected reference grids (a-i), the original EURO-CORDEX RCMs (grey lines) and the bias corrected RCMs (red).

The best three and seven models evaluated against each of the reference grids identified through the model selection process were selected from the corresponding bias-corrected RCM ensemble. Figure 6 shows the resulting monthly climate change signals for the selected three (seven) models as black spread and squares (colored bars), sorted by the reference grid used for model selection and bias correction. The climate change signal here is defined as the percent change between the future period (2036-2065) compared to the historical period (1981-2010). As similar models were selected for each of the reference grids, the spread of the projected change in precipitation is also alike, especially if seven models are selected. For the winter months, the selection of three models captures the spread of seven models while for the summer months the results are ambiguous. Due to the smaller number of selected models, the spread is smaller for only three selected models. Additionally, depending on the reference grid a selection of only three models leads to different signs in projected changes of precipitation. This results in positive changes for some reference data sets (e.g. in August) while other data sets select a climate model that projects a decrease, such as for PERSIANN-CDR, MERRA2 and GPCC-FDD. Nevertheless, these differences are due to the model selection and not to bias correction, as the climate change signals are conserved throughout.

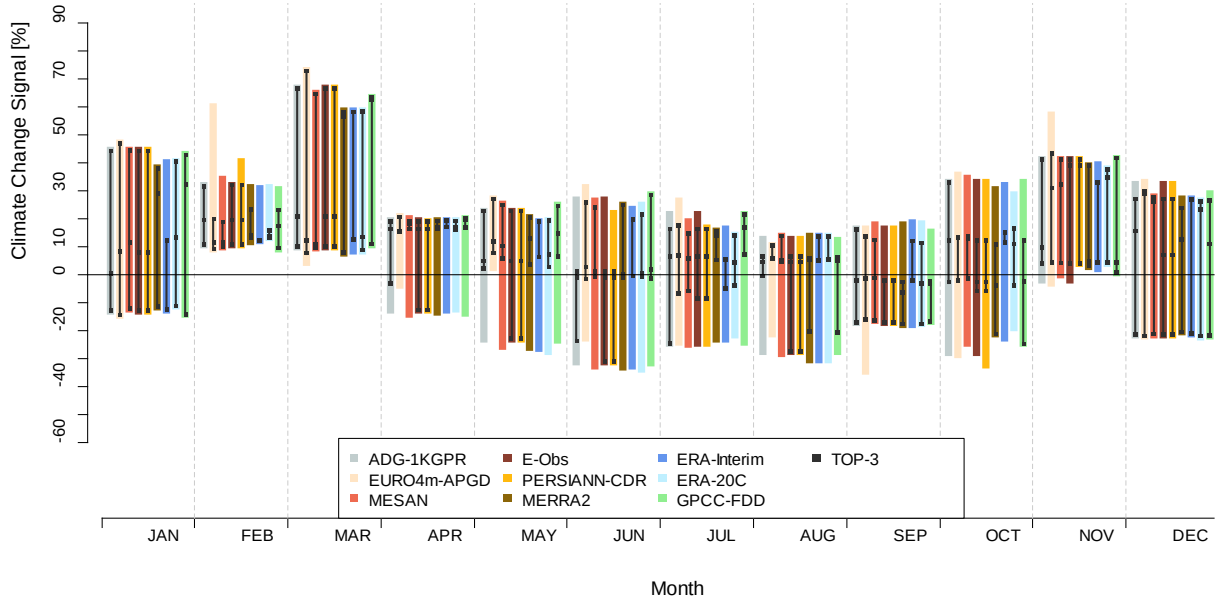


Figure 6: Monthly projected spread of climate change signals for the selected best-fit seven (colored bars) or three (black lines and squares) bias corrected RCMs. Coloring according to the corresponding reference grid applied for bias correction.

However, bias correction clearly changes the absolute values of projected precipitation. Figure 7 shows projected monthly precipitation for the future period (2036-2065) for the selected three (seven) models as black spread and squares (colored bars), sorted by the reference grid used for model selection and bias correction. Similar to Figure 6 the model spread in the summer months is higher than for the winter period. As the reference grids agree also more during the winter months (Figure 2), projected changes in this period are more robust with respect to the reference data set applied for bias correction. In summer in contrary, the impact of the reference data set is stronger. The maximum precipitation projected for August thus varies between 120mm (E-OBS) and 190mm (ERA-Interim). As the spread is similar for all reference grids, a higher maximum projection consequently means also higher projections for the minimum monthly precipitation. Bias correcting the percentiles leads to projections following the course of the reference data sets in the historical period: bias correction with products that show higher precipitation in the reference period (ERA-Interim, PERSIANN-CDR) results in higher projections. Both, the sequence of the data sets and similar differences to the historical period are reproduced in the bias corrected projections. The difference in August between E-OBS and ERA-Interim in projected precipitation is thus similar to the reference data sets. As expected, the selection of only three models results in considerable differences in the projections. On the one hand this might lead to an extremely small spread (e.g. ERA-20C in February) or on the other hand to even more amplified differences in the projected monthly temperatures (e.g. ERA-Interim and E-OBS in summer). As the model selection however is based on validity only, there is no systematic behavior and connection between model spread, absolute projected changes and the reference product.

To assess the impact of bias correction with various reference data sets on projected heavy (extreme) precipitation, Figure 8 shows the .95 (.99) quantile of daily catchment precipitation. As for the previous figures, the selected three (seven) models are shown as black spread (colored bars), sorted by the reference grid used for model selection and bias correction. For each reference data set the left bar corresponds to the reference period and the right bar shows projected changes for the future horizon. Independent from the data set used for bias correction, the RCMs agree in increased heavy and extreme precipitation. While the differences are not very distinct for the 95th percentile, the impact of the reference data set is higher for extreme precipitation. As

expected, bias correction with the high resolution observational data sets (ADG-1KGPR and EURO4m-APGD) results in higher extreme precipitation. However, the impact of spatial resolution is smoothened as the assessment is based on catchment mean rather than grid cell based and all data sets were remapped in the pre-processing. The differences originating from the reference data sets are as high as 30% and exceed the projected changes indicating that observational uncertainty exceeds the climate change signal in case of extreme precipitation.

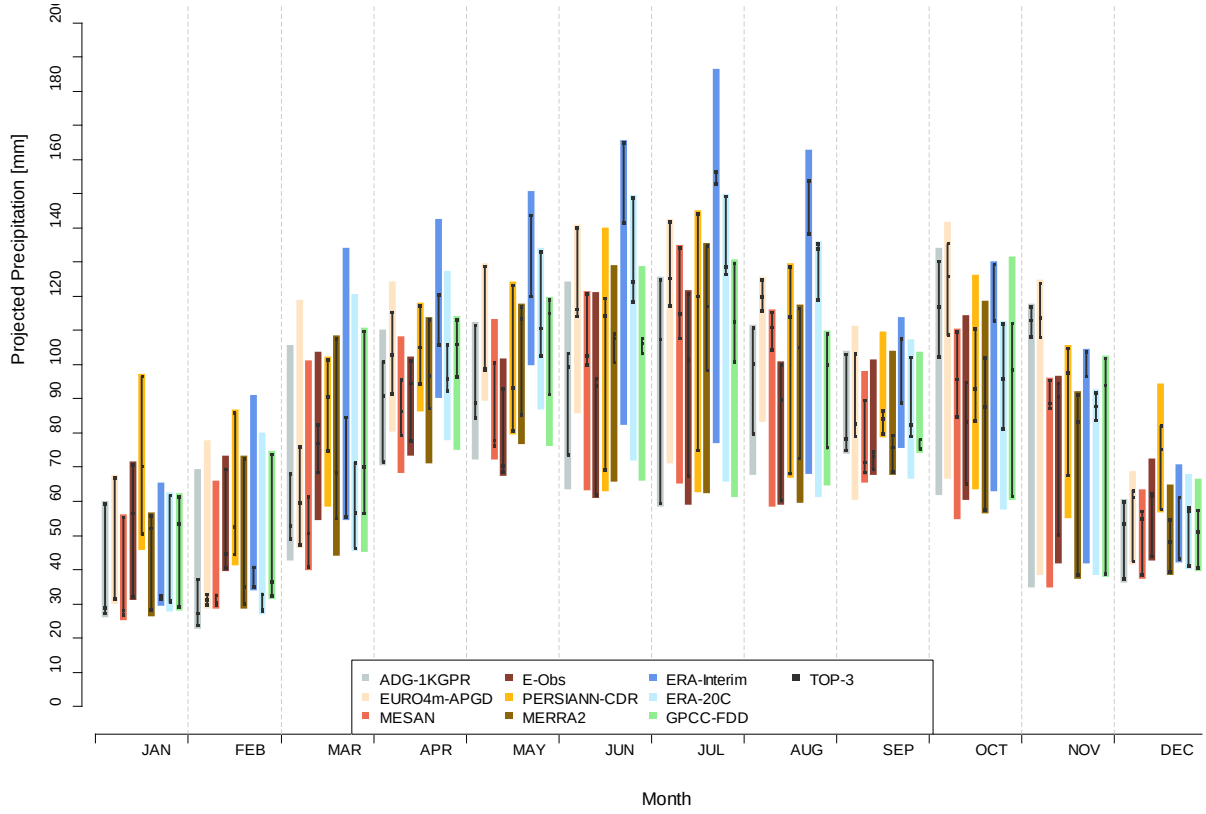


Figure 7: Monthly projected spread of monthly precipitation for the selected best-fit seven (colored bars) or three (black lines and squares) bias corrected RCMs for the period 2036 - 2065. Coloring according to the corresponding reference grid applied for the bias correction.

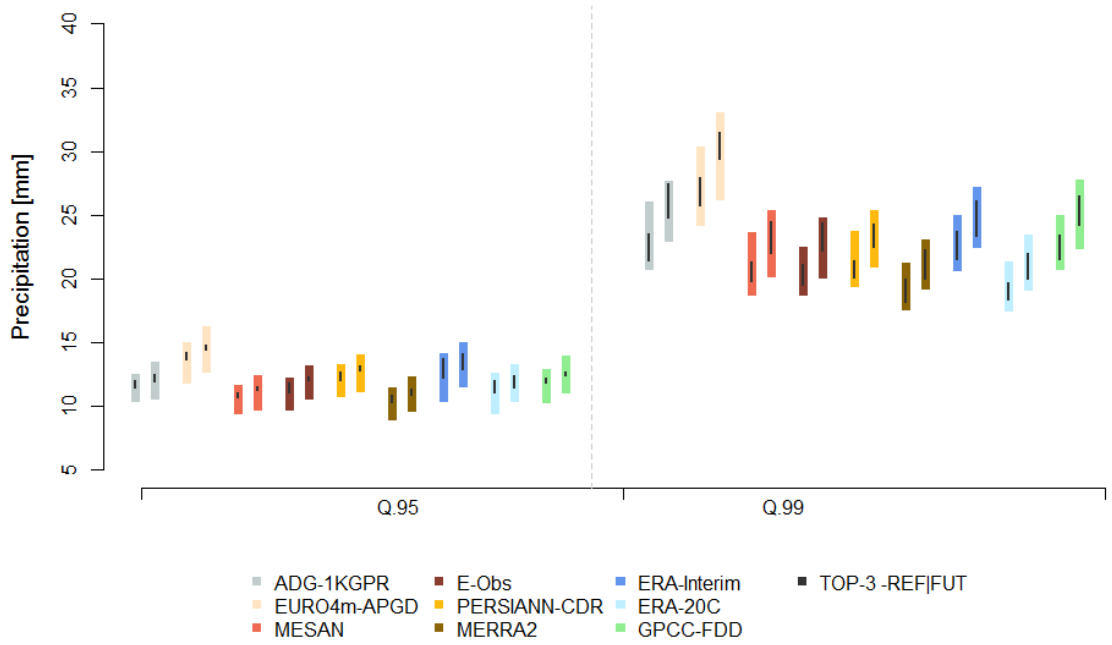


Figure 8: Catchment daily heavy (95th quantile, left) and extreme precipitation (99th quantile, right) for the seven (colored bars) and three (black bars) selected RCMs for the reference period (left bar) and future projections (right bar). Coloring refers to the corresponding reference grid.

To assess the overall uncertainty of reference data set selection on the climate change signal compared to model and scenario uncertainty, Figure 9 shows the percentage of the total variance attributed to each source of uncertainty as result of the variance decomposition per month. As mentioned, these results reveal only the relative contribution to the total variance. The relative contribution of the RCM is show in red, GCM in beige, RCP in light blue and bias correction, here originating from different reference data sets only, in dark blue. About 60% of the variance can be attributed to combined model uncertainty from RCM and GCM. The uncertainty introduced by the GCMs is usually larger than the contribution of the RCMs for all months. Scenario uncertainty introduced through the RCPs contributes comparably with GCM uncertainty in the range of 30%. Model uncertainty introduced by the RCMs is around or below 20% for most months with the exception of the winter months. The variance that can be attributed to reference data set selection for bias correction contributes only around 15% to the overall variance and thus less than the other three uncertainty sources assessed. The share is larger for the summer months, where the reference data sets show the most distinct differences.

5. Discussion

In this study the impacts of reference grid selection on climate model evaluation, model selection, bias correction and resulting climate change signals were presented for an alpine catchment. A selection of nine reference products from different sources were applied for this task on an ensemble of 15 RCMs provided through the EURO-CORDEX initiative. The results presented highlight the importance of uncertainty related to reference grid. Observational uncertainty overall is smaller than RCM related uncertainty, however it can be of similar magnitude, especially when the precipitation systems get more complex, e.g. through convective systems in summer. The findings demonstrate the importance of reference grid selection for climate model evaluation as the direction of bias (over- or underestimation of precipitation) is largely dependent on the reference data set used for evaluation. Strongest differences are identified for ERA-Interim and PERSIANN-CDR which already show strong

differences to the rest of the ensemble in a previous study (Gampe & Ludwig, 2017). This confirms the conclusions drawn in previous studies (Prein et al. 2017; Kottlarski et al. 2017) also on the catchment scale. However it is important to consider existing limitations in observational data sets such as undercatch (Rasmussen et al. 2012), interpolation errors (Addor & Fischer, 2015) and station density (Isotta et al. 2015). There is still a need for robust high-resolution data sets over larger domains for a better evaluation.

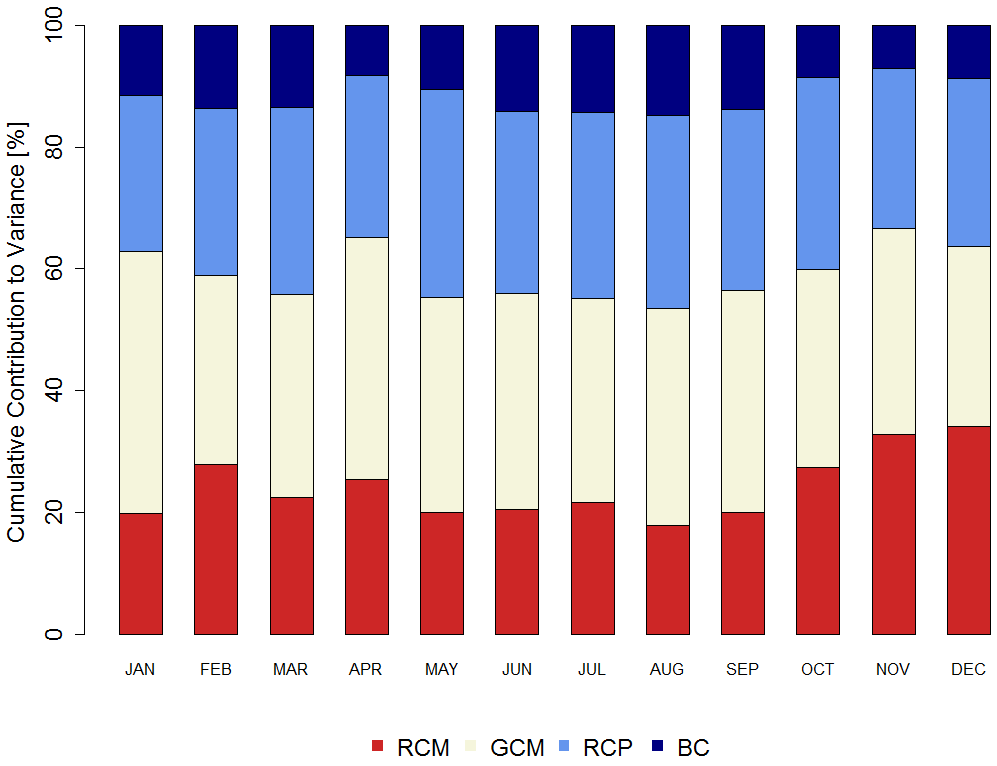


Figure 9: Monthly cumulative relative contribution of RCM (red), GCM (beige), RCP (light blue) and Bias Correction (BC).

It is important to highlight here, that both the RCMs as also the reference grids represent ensembles of opportunity and most of the data sets used in this study are not fully independent from each other. In case of the reference data sets they either partly use the same precipitation gauges for interpolation (all observational products + PERSIANN-CDR), or assimilate these information in some (reanalysis products) or serve as boundary conditions for a higher resolution data product (ERA-Interim and MESAN). However, this is a common problem when comparing different gridded precipitation products which cannot be avoided, but is nevertheless important to highlight. For the climate models the RCMs are often forced by the same GCM, however, the matrix is incomplete and similar schemes are implemented throughout the models. Additionally, the 20-year time period for this analysis is relatively short and additional uncertainty due to natural variability are not included in this study. Nevertheless, the general findings of this study will likely remain and the restricting lengths in available data set are a common problem. However, a bootstrapping method could be applied to simulate and account for natural variability in future studies similar to Addor & Fischer, 2015.

Despite the differences in the reference products, the RCMs score similarly for all reference grids, again with the exception of ERA-Interim. However, this is not negligible as depending on the number of models selected this will result in a different ensemble of RCMs chosen. The findings confirm the results presented by Kottlarski et

al. 2017 who also applied a fairly simple yet different scoring system. The similar scoring behavior of individual RCMs indicates again that model uncertainty exceeds observational uncertainty. Nevertheless, this also confirms the limitations of RCMs on the regional scale where fundamental difference arise and the climatology and variability is not correctly reflected over a given area for a given model. This emphasizes the need to account for plausible RCM inputs for bias correction as proposed by Maraun (2016). Different validity based selection can lead to partly severe changes in projections, especially if only a small subset of RCMs is selected as presented in Figure 6. The impact of model selection although strongly dependent on the number of models selected is thus considerable. Although QM also has an impact on the climate change signal (Maraun 2016), the differences in the presented signals largely originate from the selection process. This confirms the findings of Dosio et al. 2012 that the influence of bias correction is only small on projected change signals for mean precipitation and also expected as the same correction factors are applied for both periods.

However, bias correction with different reference products has severe impacts on absolute projected precipitation. The bias corrected RCM ensemble represents the features found for the reference grid, as wetter reference grids produce wetter projections. This highlights the importance of the reference grid on the quality of the bias correction procedure and the need for accurate observations and reference grids for plausible corrections. This effect becomes even more severe for extreme precipitation. The findings in this study as in Iizumi et al. 2017, the uncertainty originating from the reference data set exceeds the uncertainty from the climate model projections for bias corrected extremes. This is highly relevant for future studies focusing on extreme events and projections of changes in these.

The contribution of reference grid uncertainty in the presented framework to the overall uncertainty of the climate change signal is assessed by a variance decomposition approach. Four main sources of uncertainty were included in the analysis: model uncertainty originating from GCM and RCM, scenario uncertainty from the RCP and bias correction uncertainty through different reference data sets. As in previous studies (Déqué et al. 2007; Gampe et al. 2016) the contribution of the GCM dominates the uncertainty, followed by RCP and RCM. Bias correction uncertainty only contributes around 10 -15% to the overall variance with larger shares in summer. However, this is slightly misleading as this is only based on the climate change signals and observational uncertainty for absolute projection, especially for extremes is considerably larger.

This study confirms the outlooks by Prein et al. (2017) and Kottlarski et al. (2017) and highlights the need to account for observational uncertainty when bias correction is applied to climate model data. This is of special interest for studies focusing on projections of future extreme. As both, model selection and bias correction, are strongly dependent on the selected reference grid, the selection of these should be of special interest. These results indicate the necessity for a more thorough investigation of available data sets instead of arbitrary selection.

Acknowledgments

This work has been supported by the European Communities 7th Framework Programme Funding under Grant agreement no. 603629-ENV-2013-6.2.1-Globaqua.

References

Addor, N., Rohrer, M., Furrer, R., & Seibert, J. (2016). Propagation of biases in climate models from the synoptic to the regional scale: Implications for bias adjustment. *Journal of Geophysical Research: Atmospheres*, *121*(5), 2075-2089.

Addor, N., & Fischer, E. M. (2015). The influence of natural variability and interpolation errors on bias characterization in RCM simulations. *Journal of Geophysical Research: Atmospheres*, 120(19).

Ashouri, H., Hsu, K. L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B.R. & Prat, O. P. (2015). PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. Bulletin of the American Meteorological Society, 96(1), 69-83.

Biemans, H., Speelman, L.H., Ludwig, F., Moors, E.J., Wiltshire, A.J., Kumar, P., Gerten, D. & Kabat, P. (2013): Future water resources for food production in five South Asian river basins and potential for adaptation – a modeling study. *Sci. Total Environ*. 468–469: S117–S131

Bosilovich,M.G.; Lucchesi, R.; Suarez,M.(2015). MERRA-2: File Specification; GMAO Office Note No. 9 (Version 1.1). 73p. Available online: https://gmao.gsfc.nasa.gov/pubs/ (accessed on 30 December 2017).

Christensen, J. H., Boberg, F., Christensen, O. B., & Lucas-Picher, P. (2008). On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophysical Research Letters*, *35*(20).

Chen, J., Brissette, F. P., Chaumont, D., & Braun, M. (2013, a). Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America. *Water Resources Research*, *49*(7), 4187-4205.

Chen, J., Brissette, F. P., Chaumont, D., & Braun, M. (2013, b). Performance and uncertainty evaluation of empirical downscaling methods in quantifying the climate change impacts on hydrology over two North American river basins. *Journal of Hydrology*, *479*, 200-214.

Chiogna, G., Majone, B., Paoli, K. C., Diamantini, E., Stella, E., Mallucci, S., Lencioni, V., Zandonai, F. & Bellin, A. (2016). A review of hydrological and chemical stressors in the Adige catchment and its ecological status. *Science of The Total Environment*, *540*, 429-443.

Dahlgren, P., Landelius, T., Kållberg, P., & Gollvik, S. (2016). A high-resolution regional reanalysis for Europe. Part 1: Three-dimensional reanalysis with the regional HIgh-Resolution Limited-Area Model (HIRLAM). *Quarterly Journal of the Royal Meteorological Society*, *142*(698), 2119-2131.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. & Vitart, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, *137*(656), 553-597.

Dosio, A., Paruolo, P., & Rojas, R. (2012). Bias correction of the ENSEMBLES high resolution climate change projections for use by impact models: Analysis of the climate change signal. *Journal of Geophysical Research: Atmospheres*, *117*(D17).

Dosio, A. (2016). Projections of climate change indices of temperature and precipitation from an ensemble of bias-adjusted high-resolution EURO-CORDEX regional climate models. *Journal of Geophysical Research: Atmospheres*, *121*(10), 5488-5511.

Ehret, U., Wulfmeyer, V., Warrach-Sagi, K. and Liebert, J. “Should we apply bias correction to global and regional climate model data?”, Hydrol. Earth Syst. Sci., (2012), 16, 3391-3404.

Evans, J. P., Ji, F., Lee, C., Smith, P., Argüeso, D., & Fita, L. (2013). A regional climate modelling projection ensemble experiment-NARClim. *Geoscientific Model Development Discussions*, *6*(3).

Frei, C., & Schär, C. (1998). A precipitation climatology of the Alps from high-resolution rain-gauge observations. *International Journal of climatology*, *18*(8), 873-900.

Gampe, D., Nikulin, G., & Ludwig, R. (2016). Using an ensemble of regional climate models to assess climate change impacts on water scarcity in European river basins. *Science of the Total Environment*, *573*, 1503-1518.

Gampe, D., & Ludwig, R. (2017). Evaluation of Gridded Precipitation Data Products for Hydrological Applications in Complex Topography. *Hydrology*, *4*(4), 53.

Giorgi, F., Jones, C., & Asrar, G. R. (2009). Addressing climate information needs at the regional level: the CORDEX framework. *World Meteorological Organization (WMO) Bulletin*, *58*(3), 175.

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., & New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. Journal of *Geophysical Research: Atmospheres*, 113(D20).

Henn, B., Newman, A. J., Livneh, B., Daly, C., & Lundquist, J. D. (2018). An assessment of differences in gridded precipitation datasets in complex terrain. *Journal of Hydrology.*

Herold, N., Behrangi, A., & Alexander, L. V. (2017). Large uncertainties in observed daily precipitation extremes over land. *Journal of Geophysical Research: Atmospheres*, 122(2), 668-681.

Hsu, K. L., Gao, X., Sorooshian, S., & Gupta, H. V. (1997). Precipitation estimation from remotely sensed information using artificial neural networks. *Journal of Applied Meteorology*, *36*(9), 1176-1190.

Iizumi, T., Takikawa, H., Hirabayashi, Y., Hanasaki, N., & Nishimori, M. (2017). Contributions of different bias-correction methods and reference meteorological forcing data sets to uncertainty in projected temperature and precipitation extremes. *Journal of Geophysical Research: Atmospheres*.

Isotta, F. A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassegues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G.; Ratto, S.M & Munari, M. (2014). The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. *International Journal of Climatology*, *34*(5), 1657-1675.

Isotta FA, Vogel R, Frei C. (2015). Evaluation of European regional reanalyses and downscalings for precipitation in the Alpine region. *Meteorol. Z.* 24: 15–37.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M, Braun, A., Colette, A., Déqué, M., Georgievsk, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haenseler, A., Hempelmann, N., Jones, C, Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C-. Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Valentini, R., Vautar, R., Weber, B. and Yiou, P. (2014): EURO-CORDEX: new high-resolution climate change projections for European impact research. *Reg. Environ. Change*, vol. 14, 563-578.

Kotlarski, S., Keuler, K., Christensen, O. B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., van Meijgaard, E. ,Nikulin, G. ,Schär, C., Teichmann, C., Vautard, R., Warrach-Sagi, K. and Wulfmeyer, V. (2014): Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geoscientific Model Development*, 7(4), 1297-1333.

Kotlarski, S., Szabó, P., Herrera, S., Räty, O., Keuler, K., Soares, P. M., Cardoso, R.M., Bosshard, T., Pagé, C., Boberg, F., Gutiérrez, J. M., Isotta, F.A., Jaczewski, A., Kreienkamp, F., Liniger, M.A., Lussana, C. & Pianko- Kluczynska, K. (2017). Observational uncertainty and regional climate model evaluation: a pan-European perspective. *International Journal of Climatology*.

Kummerow, C., Barnes, W., Kozu, T., Shiue, J., & J. Simpson (1998). The tropical rainfall measuring mission (TRMM) sensor package. Journal of atmospheric and oceanic technology, 15(3), 809-817.

Lafon, T., Dadson, S., Buys, G., & Prudhomme, C. (2013). Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods. *International Journal of Climatology*, *33*(6), 1367-1381.

Landelius, T., Dahlgren, P., Gollvik, S., Jansson, A., & Olsson, E. (2016). A high-resolution regional reanalysis for Europe. Part 2: 2D analysis of surface temperature, precipitation and wind. Quarterly Journal of the Royal Meteorological Society, 142(698), 2132-2142.

Lutz, A. F., ter Maat, H. W., Biemans, H., Shrestha, A. B., Wester, P., & Immerzeel, W. W. (2016). Selecting representative climate models for climate change impact studies: an advanced envelope-based selection approach. *International Journal of Climatology*, *36*(12), 3988-4005.

Mendlik, T., & Gobiet, A. (2016). Selecting climate simulations for impact studies based on multivariate patterns of climate change. *Climatic change*, *135*(3-4), 381-393.

Mpelasoka, F. S., & F. H. S. Chiew (2009), Influence of rainfall scenario construction methods on runoff projections, *J. Hydrometeorol*., 10, 1168–1183.

Muerth, M., Gauvin St-Denis, B., Ricard, S., Velázquez, J. A., Schmid, J., Minville, M., Caya, D., Chaumont, D., Ludwig, R. & Turcotte, R. (2013). On the need for bias correction in regional climate scenarios to assess climate change impacts on river runoff. *Hydrology and Earth System Sciences Discussions*, 10205-10243.

Navarro-Ortega, A., Acuña, V., Bellin, A., Burek, P., Cassiani, G., Choukr-Allah, R., Dolédec, S., Elosegi, A., Ferrari, F., Ginebreda, A., Grathwohl, P., Jones, C., Ker Rault, P., Kok, K., Koundouri, P., Ludwig, R.P., Merz, R., Milacic, R., Muñoz, I., Nikulin, G., Paniconi, C., Paunovi, M., Petrovic, M., Sabater, L., Sabater, S., Skoulikidis, N.T., Slob, A., Teutsch, G., Voulvoulis, N. & Barceló, D., (2015). Managing the effects of multiple stressors on aquatic ecosystems under water scarcity. The GLOBAQUA project. *Science of the Total Environment*, *503*, 3-9.

Palazzi, E., Hardenberg, J., & A. Provenzale (2013). Precipitation in the Hindu-Kush Karakoram Himalaya: Observations and future scenarios. Journal of Geophysical Research: Atmospheres, 118(1), 85-100.

Piani, C., Haerter, J. O., & Coppola, E. (2010). Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoretical and Applied Climatology*, *99*(1-2), 187-192.

Pierce, D.W., Barnett, T.P., Santer, B.D. & Gleckler, P.J. (2009): Selecting global climate models for regional climate change studies. *Proc. Natl. Acad. Sci. U. S. A.* 106(21): 8441–8446.

Pierce, D. W., Cayan, D. R., Maurer, E. P., Abatzoglou, J. T., & Hegewisch, K. C. (2015). Improved bias correction techniques for hydrological simulations of climate change. *Journal of Hydrometeorology*, *16*(6), 2421-2442.

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D.G.H., Peubey, C., Thépaut, J.-N., Trémolet, Y., Holm, E.V., Bonavita, M., Isaksen, L. & Fisher, M. (2016). ERA-20C: An atmospheric reanalysis of the twentieth century. *Journal of Climate*, *29*(11), 4083-4097.

Prein, A. F., Gobiet, A., Truhetz, H., Keuler, K., Goergen, K., Teichmann, C. Fox Maule, C., van Mejgaard,E., Déqué,M., Nikulin,G., Vautard, R., Colette,A., Kjellström, E. & D. Jacob (2016): Precipitation in the EURO-CORDEX 0.11° and 0.44° simulations: high resolution, high benefits? *Climate dynamics*, *46*(1-2), 383-412.

Prein, A. F., & Gobiet, A. (2017). Impacts of uncertainties in European gridded precipitation observations on regional climate analysis. *International Journal of Climatology*, *37*(1), 305-327.

Rajczak, J., Kotlarski, S., Salzmann, N., & Schär, C. (2016). Robust climate scenarios for sites with sparse observations: a two-step bias correction approach. *International Journal of Climatology*, *36*(3), 1226-1243.

Rasmussen, R.; Baker, B.; Kochendorfer, J.; Meyers, T.; Landolt, S.; Fischer, A.P.; Black, J.; Thériault, J.M.;Kucera, P.; Gochis, D., Smith, C., Nitu, R., Hall, M., Ikeda, K. & Gutmann, E. (2012). How well are we measuring snow: The NOAA/FAA/NCAR winter precipitation test bed. *Bull. Am. Meteorol. Soc.*, 93, 811–829.

Ruiz-Ramos, M., Rodríguez, A., Dosio, A., Goodess, C. M., Harpham, C., Mínguez, M. I., & Sánchez, E. (2016). Comparing correction methods of RCM outputs for improving crop impact projections in the Iberian Peninsula for 21st century. *Climatic change*, 134(1-2), 283-297.

Schaller, N., Mahlstein, I., Cermak, J., & Knutti, R. (2011). Analyzing precipitation projections: A comparison of different approaches to climate model evaluation. *Journal of Geophysical Research: Atmospheres*, 116(D10).

Schamm, K., Ziese, M., Raykova, K., Becker, A., Finger, P., Meyer-Christoffer, A. & Schneider, U. (2016) GPCC Full Data Daily Version 1.0: Daily Land-Surface Precipitation from Rain Gauges Built on GTS Based and Historic Data. Available online: https://rda.ucar.edu/datasets/ds497.0/ (accessed on Dec 29 2017).

Smiatek, G., Kunstmann, H., & Senatore, A. (2016). EURO-CORDEX regional climate model analysis for the Greater Alpine Region: Performance and expected future change. *Journal of Geophysical Research: Atmospheres*, *121*(13), 7710-7728.

Teutschbein, C., & Seibert, J. (2013). Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions? *Hydrology and Earth System Sciences*, *17*(12), 5061.

Torma, C., Giorgi, F., & Coppola, E. (2015). Added value of regional climate modeling over areas characterized by complex terrain—Precipitation over the Alps. *Journal of Geophysical Research: Atmospheres*, *120*(9), 3957-3972.

Van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G.C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S.J. & Rose, S.K. (2011). The representative concentration pathways: an overview. Climatic change, 109(1-2), 5.

Volosciuk, C., Maraun, D., Vrac, M., & Widmann, M. (2017). A combined statistical bias correction and stochastic downscaling method for precipitation. *Hydrology and Earth System Sciences*, *21*(3), 1693-1719.

White, R. H., & Toumi, R. (2013). The limitations of bias correcting regional climate model inputs. *Geophysical Research Letters*, 40(12), 2907-2912.

Wilcke, R. and Bärring, L. (2016). Selecting regional climate scenarios for impact modelling studies, Environmental Modelling & Software, 78, 191-201.

Zubler, E. M., Fischer, A. M., Fröb, F., & Liniger, M. A. (2016). Climate change signals of CMIP5 general circulation models over the Alps–impact of model selection. International Journal of Climatology, 36(8), 3088-3104.

2.3 Paper III: Using an ensemble of regional climate models to assess climate change impacts on water scarcity in European river basins. Science of the Total Environment

Gampe, D., Nikulin, G., & Ludwig, R. (2016). *Using an ensemble of regional climate models to assess climate change impacts on water scarcity in European river basins.* Science of the Total Environment, 573, 1503-1518.

Paper III was elaborated within the project GLOBAQUA and funded under the European Union's Seventh Programme. This publication identifies the role of climate change as a potential pressure on aquatic ecosystems, i.e. four river basins in the Mediterranean. The paper focuses on the uncertainty related to climate projections and examines the relative contribution of model uncertainty from GCM and RCM as well as scenario uncertainty introduce by the two radiative forcing scenarios (RCP 4.5 & 8.5). It builds up on the previous study by including the entire high-resolution RCM ensemble available through the EURO-CORDEX initiative at this time. The uncertainty assessment is based on the contribution of each uncertainty source to the overall variance by means of a variance decomposition (Déqué et al., 2007, 2012). The hydrological consequences of the climate projections are derived from the RCM variables precipitation, evapotranspiration and surface runoff directly. This allows for a first step estimation of changes on the water balance even in data scarce areas. To further address the role of climate change as a potential stressor on the aquatic ecosystems, the Falkenmark indicator (Falkenmark, 1989; Jaeger et al., 2013; Schyns et al., 2015), defined as available water per capita, is presented and reveals increased water scarcity in most areas of the river basins. The following research questions are addressed in this paper:

Q4: What is the relative contribution of each of the uncertainty sources involved in recent regional climate projections on the catchment scale and for all the variables of the water balance in the downscaled CMIP5 climate projections?

Q6: Is the Mediterranean region likely to experience increased pressure on the water balance and thus likely to be exposed to increased water scarcity under future climate?

Author's contributions: All authors contributed to the concept of the study, while all calculations, data handling and analysis were performed by D. Gampe under the supervision of R. Ludwig. The first version of the manuscript was written by D. Gampe, while all authors contributed with valuable input and expertise to the final document.
Status: published
Journal: Science of the Total Environment (Elsevier)
Impact factor: 4.900

Science of the Total Environment 573 (2016) 1503–1518

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Using an ensemble of regional climate models to assess climate change impacts on water scarcity in European river basins

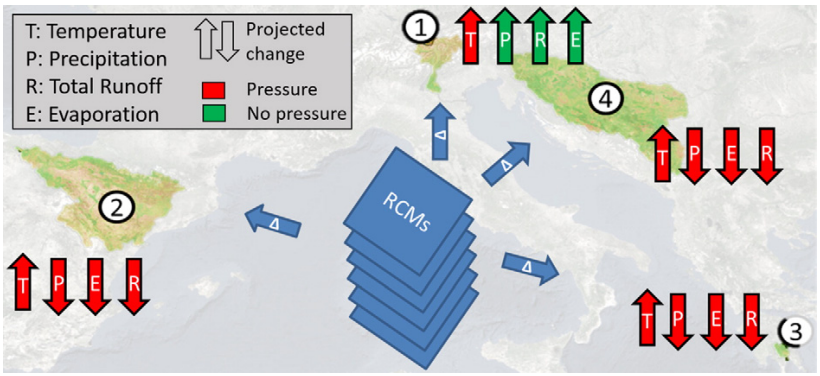
David Gampe [a,*], Grigory Nikulin [b], Ralf Ludwig [a]

[a] Department of Geography, Ludwig-Maximilians-Universität, Munich, Germany
[b] Rossby Centre, Swedish Meteorological and Hydrological Institute (SMHI), Norrkopping, Sweden

HIGHLIGHTS

- Ensemble of Regional Climate Models applied to assess changes in hydrological relevant variables to overcome data scarcity.
- Importance of high resolution climate model information to assess regional changes within the catchments and account for local topography evident.
- Strong decline in precipitation, evaporation and runoff increase water scarcity in the Ebro and Evrotas basin as well as parts of the Adige and Sava.
- Falkenmark indicator calculated using global data sets revealed moderate increase of water scarcity for the Ebro and parts of the Sava.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 14 June 2016
Received in revised form 6 August 2016
Accepted 7 August 2016
Available online 15 August 2016

Editor: D. Barcelo

Keywords:
EURO-CORDEX
Climate change
Hydrology
Water scarcity
Mediterranean river basins
Data scarcity

ABSTRACT

Climate change will likely increase pressure on the water balances of Mediterranean basins due to decreasing precipitation and rising temperatures. To overcome the issue of data scarcity the hydrological relevant variables total runoff, surface evaporation, precipitation and air temperature are taken from climate model simulations. The ensemble applied in this study consists of 22 simulations, derived from different combinations of four General Circulation Models (GCMs) forcing different Regional Climate Models (RCMs) and two Representative Concentration Pathways (RCPs) at ~12 km horizontal resolution provided through the EURO-CORDEX initiative. Four river basins (Adige, Ebro, Evrotas and Sava) are selected and climate change signals for the future period 2035–2065 as compared to the reference period 1981–2010 are investigated.

Decreased runoff and evaporation indicate increased water scarcity over the Ebro and the Evrotas, as well as the southern parts of the Adige and the Sava, resulting from a temperature increase of 1–3° and precipitation decrease of up to 30%. Most severe changes are projected for the summer months indicating further pressure on the river basins already at least partly characterized by flow intermittency. The widely used Falkenmark indicator is presented and confirms this tendency and shows the necessity for spatially distributed analysis and high resolution projections. Related uncertainties are addressed by the means of a variance decomposition and model agreement to determine the robustness of the projections. The study highlights the importance of high resolution climate projections and represents a feasible approach to assess climate impacts on water scarcity also in regions that suffer from data scarcity.

© 2016 Published by Elsevier B.V.

* Corresponding author.
E-mail address: david.gampe@lmu.de (D. Gampe).

http://dx.doi.org/10.1016/j.scitotenv.2016.08.053
0048-9697/© 2016 Published by Elsevier B.V.

1. Introduction

The Mediterranean region has been identified as one of the hot spots for climate change with a decline of precipitation up to over 50% in summer, severe increase in temperature, and an increased drought risk (Giorgi, 2006; Christensen et al., 2007; Field et al., 2012; Kjellström et al., 2013). Gosling and Arnell (2016) showed the importance of climate change and population increase on water scarcity on the global scale and concluded a larger fraction of the world is likely exposed to increased water scarcity. Hanasaki et al. (2012) pointed out the importance of population growth, economic development and climate change on global water scarcity. Additionally, global modeling revealed a strong decrease in runoff for many basins in the Mediterranean that will increase the pressure on already water scarce river basins (Haddeland et al., 2014). On a regional scale using the recent CMIP5 simulations, a similar reduction of freshwater was found recently by Koutroulis et al. (2016) for the eastern Mediterranean. However, assessment of water scarcity in Mediterranean basins, e.g. through hydrological modeling, is often hampered by available data. To overcome this problem, extensive field campaigns need to be conducted or remote sensing techniques are required to bridge data gaps (Serra et al., 2016, Meyer et al., 2016; Gampe et al., 2016). Most of these approaches are costly in terms of labor, computational demand and often still require additional data and expertise.

The EU-project '*GLOBAQUA - Managing the effects of multiple stressors on aquatic ecosystems under water scarcity*' started in February 2014 and aims to study the effects of water scarcity to foster the understanding of current management practices and identifying possible improvements in the management strategies. Therefore six river basins which are affected by water scarcity either due to climatological pressures or to high variability in rainfall or multiple conflicting water uses, were selected. Four of those six river basins have been chosen for extensive field work and will be subject for various impact modeling activities (Navarro-Ortega et al., 2015). These focus on the ecological status of the river ecosystems, to assess the role of emerging pollutants, other chemicals and geomorphological changes act as stressors for the aquatic ecosystems. In this study the role of climate change as a possible stressor will be examined. Therefore, these river basins, i.e. Adige (northern Italy), Ebro (Spain), Sava (Slovenia, Croatia, Serbia and Bosnia and Herzegovina) and Evrotas (Greece) are also chosen as case study areas also in this study.

Complex topography and/or small size of the basins demand for climate change information at regional to local scale, at high spatial resolution. The primary tools for providing future climate projections are coupled General Circulation Models (GCMs), which simulate climate changes under a range of possible future scenarios of greenhouse gas emissions. These stem from the scenarios provided through various Representative Concentration Pathways (RCPs) that span a range of the radiative forcing of 2.6 to 8.5 W/m^2 for the year 2100 (Van Vuuren et al., 2011). Over the last few years, about 20 modeling groups using >50 models have participated in the Coupled Model Intercomparison Project Phase 5 (CMIP5), thus generating a large multi-model ensemble of climate change simulations (Taylor et al., 2012). Present-day GCMs have spatial resolution of 100–250 km. Due to the additional constraint of providing an ensemble of projections over long time periods, GCMs cannot fulfil the requirements of high spatial detail required and are, therefore, generally supplemented with statistical or dynamical downscaling to produce future climate projections at regional scales. To meet the need for detailed climate change information, the World Climate Research Programme (WCRP) sponsors the Coordinated Regional Downscaling Experiment (CORDEX), which aims at developing high-quality, regionally specific, climate change projections for most land regions of the world (Giorgi et al., 2009; Jones et al., 2011). In this context simulations performed by numerous Regional Climate Models (RCMs) are provided at a common grid with a horizontal resolution of 0.11° (~12 km, corresponding to the EUR-11 grid) over Europe through the EURO-CORDEX initiative.

One of the main challenges when dealing with climate projections is the quantification of uncertainties, which can have different origins, such as emission scenario, model formulation and natural variability. Several different emission scenarios and climate models should be used to assess the uncertainties related to external forcing sampling a range of future possible climate outcomes (Jones and Nikulin, 2009).

This study, based on the Euro-CORDEX results, aims at providing an assessment of possible future climate change impacts on water scarcity in the four river basins mentioned, which are introduced in Section 2.1. A large RCM ensemble available under two RCPs is applied to determine the role of climate change as a possible stressor on the aquatic ecosystems mentioned. In contrary to most of the studies mentioned, water scarcity will be assessed at regional scale, using the Euro-CORDEX simulations at a horizontal resolution of 0.11°.

In a first step, the question whether there is a clear climate change signal detectable over four catchments in the Mediterranean for temperature and precipitation will be addressed in Section 3.1 and secondly how this translates into hydrological important variables as total runoff and evaporation in Section 3.2. These four variables are analyzed by using direct RCM output allowing for cost-effective, spatially distributed projections for future changes in water scarcity in these areas. Runoff and evaporation are crucial variables in the water balance and of upmost importance for water availability. Negative changes in runoff and actual evaporation can thus be interpreted as indications of increased water scarcity, as the limiting factor for evaporation in the Mediterranean is water availability. To further address water scarcity, the widely used Falkenmark indicator (Falkenmark, 1989) is calculated in Section 3.3; to solely address the impact of climate change, population changes are neglected in the scenario period. A discussion and the resulting consequences on water scarcity is included in Section 4 together with a variance decomposition to assess the share of GCM, RCM and radiative forcing on the explained variance of each variable and determine the related uncertainty.

2. Data sets & methods

2.1. Selected river basins

Fig. 1 shows the four river basins in the Mediterranean over which this study was carried out, which are also part of the GLOBAQUA project (Navarro-Ortega et al., 2015). The Adige River (1) in northern Italy, with a drainage basin of 12,100 km^2 and a length of 409 km rises in the Southern Alps and ends in the Adriatic Sea, passing three Italian provinces. The areas climate is characterized by dry winters and precipitation maximum in summer and fall, with a total annual precipitation ranging from 500 to 1600 mm yr^{-1}. The hydrology is dominated by snow and glacier melt in spring and the strong precipitation events in summer (Chiogna et al., 2016). Although the climate is rather humid throughout the catchment, periodical water scarcity in this basin is resulting from altered hydrology due to the operation of 30 major reservoirs used for power generation (Navarro-Ortega et al., 2015).

The second catchment is the Ebro River basin (2) located in Spain, ranging from the Pyrenees in the north to the Iberian mountains and the Mediterranean Sea with a total drainage area of 83,000 km^2. Precipitation ranges from Mediterranean conditions with 300 mm year^{-1} to mountainous, humid conditions with annual rainfall up to 2500 mm year^{-1} (Telesca et al., 2012). The Ebro is the largest Spanish river draining in the Mediterranean Sea. It is impacted by distinct seasonal hydrological characteristics with long low flow periods in summer causing severe water scarcity. The area is highly managed with many dams and channels used mainly for irrigation of the agricultural areas in the basin (Navarro-Ortega et al., 2015).

The Evrotas River in Greece (3) is the smallest basin within this study with an area of 2418 km^2 and a length of 90 km bordering the ranges of Taygetos and Parnon. Climate in the area is characterized by mild and humid winters and dry and hot summers and mean annual

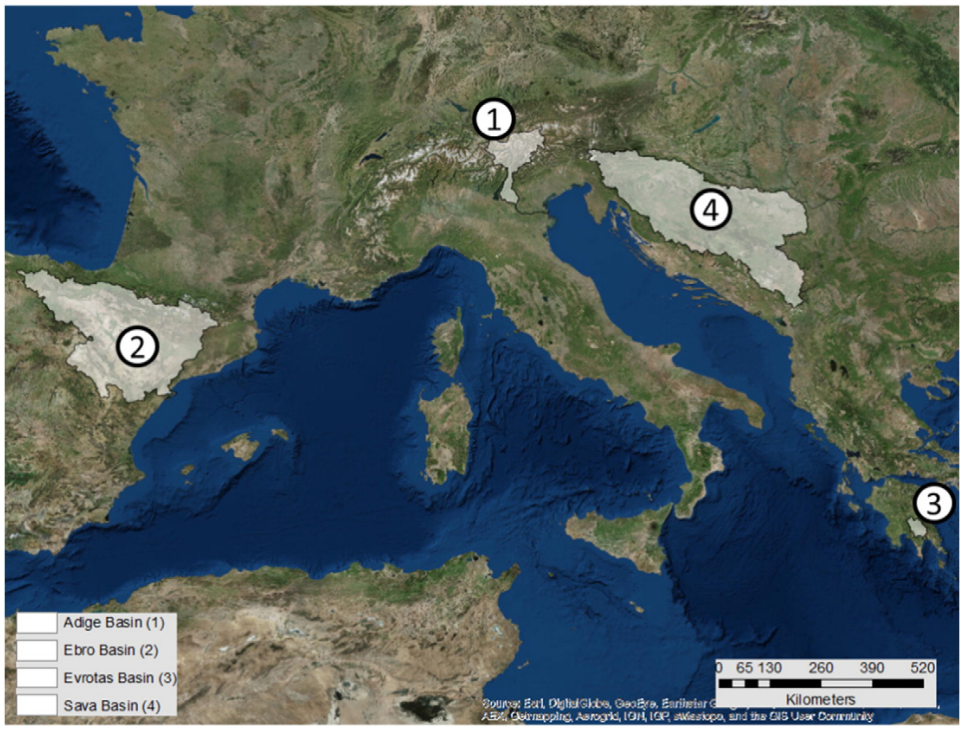


Fig. 1. Location of the selected four river basins: Adige (1), Ebro (2), Evrotas (3) and Sava (4).

precipitation of 803 mm year^{-1} originating mostly from rainfall events from October to March. The hydrology of the basin is dominated by intense precipitation, but also snow melt in spring and numerous karstic springs. Water scarcity in the river basin originates from dry periods in summer with occurring flow intermittency as well as ground water overexploitation to meet the agricultural irrigation needs (Skoulikidis et al., 2011, Navarro-Ortega et al., 2015).

The fourth catchment is the Sava River basin, representing the largest tributary to the Danube with a drainage area of 95,551 km^2. Before reaching the Danube, the Sava crosses the countries of Slovenia, Croatia, Bosnia and Herzegovina and Serbia over a length of 945 km (Milačič et al., 2010). The hydrology of the Sava is less variable compared to e.g. the Ebro, however similar pressures on the aquatic ecosystem are present (Navarro-Ortega et al., 2015). Hydrological behavior is dominated by snow melt in the northern part of the catchment. The annual mean precipitation is 1108 mm year^{-1}, yet with considerable spatial and seasonal variations (Levi et al., 2015).

2.2. RCM ensemble

An ensemble of high-resolution regional climate simulations, generated within the Euro-CORDEX activities, has been applied. The ensemble consists of eleven recent RCM simulations as shown in Table 1 as available by December 2015; abbreviations show here will be used in following figures. Four CMIP5 GCMs have been downscaled over Europe, in different combinations, by four RCMs at approx. 12 km resolution and under two RCPs – 4.5 and 8.5 (Table 1 and Jacob et al., 2013). The main focus here is on assessing future climate changes for the 2050 horizon (2035–2065), using 1981–2010 as the reference period for the variables summarized in Table 2.

In a first step, the two most common variables - 2 m temperature (tas) and precipitation (pr) are being analyzed. To estimate the hydrological consequences resulting from these climate changes in the case study areas, total runoff (mrro) and actual evaporation (evspsbl) were furthermore taken directly from the RCM output. This allows for a consistent analysis of future changes in a spatially explicit manner for all case study basins at the spatial resolution of the RCMs simulations (12 km). Therefore runoff in this study is based on the generated total runoff (mrro) by RCMs which is the excess of water per grid point without routing channels or translation to river discharge. In the following the terms runoff and the CMIP5 acronym mrro refer to these outputs whereas evaporation or evspsbl will refer to actual evaporation from RCMs and not potential evaporation.

Changes in precipitation are easy in their interpretation, as decreased precipitation likely increases water scarcity. Changes in runoff are in this study interpreted similar, without consideration of ground water storage and reservoirs. For evaporation this is more complex and dependent on the region and corresponding factors that limit evaporation. As in Mediterranean climate, hence for the Evrotas and most of the Ebro and Sava basin, as well as the southern part of the Adige, evaporation is usually limited by water availability; a decrease in this variable can be interpreted as increased water stress and indicate

Table 1
The ensemble of Euro-CORDEX simulations at 12 km resolution applied in this study. All simulations are available for rcp 4.5 and 8.5.

Regional Climate Model (RCM)	General Circulation Model (GCM)
SMHI-RCA4	CNRM-CM5
	HadGEM2-ES
	EC-EARTH-r12
	MPI-ESM-LR
DMI-HIRHAM5	EC-EARTH-r3
KNMI-RACMO22E	EC-EARTH-r12
	HadGEM2-ES
CLMcom-CCLM4-8-17	CNRM-CM5
	HadGEM2-ES
	EC-EARTH-r12
	MPI-ESM-LR

Table 2
Variables used in this study with their corresponding CMIP5 acronyms.

Variable	CMIP5 acronym	Unit original	Unit converted
2-meter air temperature	tas	K	°C
Total precipitation	pr	kg m^{-2} s^{-1}	mm
Total surface evaporation	evspsbl	kg m^{-2} s^{-1}	mm
Total runoff	mrro	kg m^{-2} s^{-1}	mm

increased water scarcity. In case of the mountainous areas of the Adige, Ebro and Sava, evaporation is not necessary limited by water availability, but rather other meteorological conditions such as air temperature and solar radiation. In these areas, increased evaporation can be interpreted as water loss to the atmosphere, hence increased water scarcity. The RCMs applied differ in their included land use, land-surface-,convection-, radiation- and planetary boundary layer scheme, as well as different soil initialization and spin-up (Kotlarski et al., 2014), causing different model results and climate sensitivities and different translations to evaporation and runoff.

3. Results

In this section the projected changes on the variables precipitation, 2 m air temperature, total runoff and evaporation are analyzed. Results are based on the climate change signals for the period 2036–2065 as compared to the reference period 1981–2010 are presented for the RCM ensemble shown in Table 1. Resulting changes in water scarcity are addressed by changes in runoff and evaporation and the Falkenmark indicator (Falkenmark, 1989) a widely used water scarcity indicator.

3.1. Changes in precipitation and air temperature

Fig. 2 shows the annual mean projected changes in precipitation [%] and 2 m air temperature for the four selected river basins, where numbers correspond to the different simulations, black coloring to rcp 4.5 and red coloring to rcp 8.5. As expected, a clear signal in temperature can be identified for all four basins ranging from 1 to 3 °C depending on simulation and catchment. The higher radiative forcing in rcp 8.5 results in generally higher temperature increase for all basins. As presented in Fig. 3, this applies to all catchments also on a monthly basis with a slight tendency to increased changes in summer and fall. The ensemble mean projects a mean monthly temperature increase between 1.4 and 2.7 °C depending on month of the year and catchment.

For precipitation, the signals are not so clear for the Adige (Fig. 2a), where in rcp 4.5 no clear direction of change can be identified, as four simulations project a decrease in precipitation, four show no change, whereas three simulations project an increase. For rcp 8.5 however, with the exception of the HadGEM2 driven CCLM run, all simulations project a slight increase in precipitation up to 10%. Similar results can be found for the Sava (Fig. 2d), however here the projections for rcp 8.5 are also somewhat ambiguous and most of the models are projecting no significant change in any direction.

Precipitation projections for the Evrotas (Fig. 2c) and Ebro (Fig. 2d) show a moderate to strong decrease up to 20% for most of the

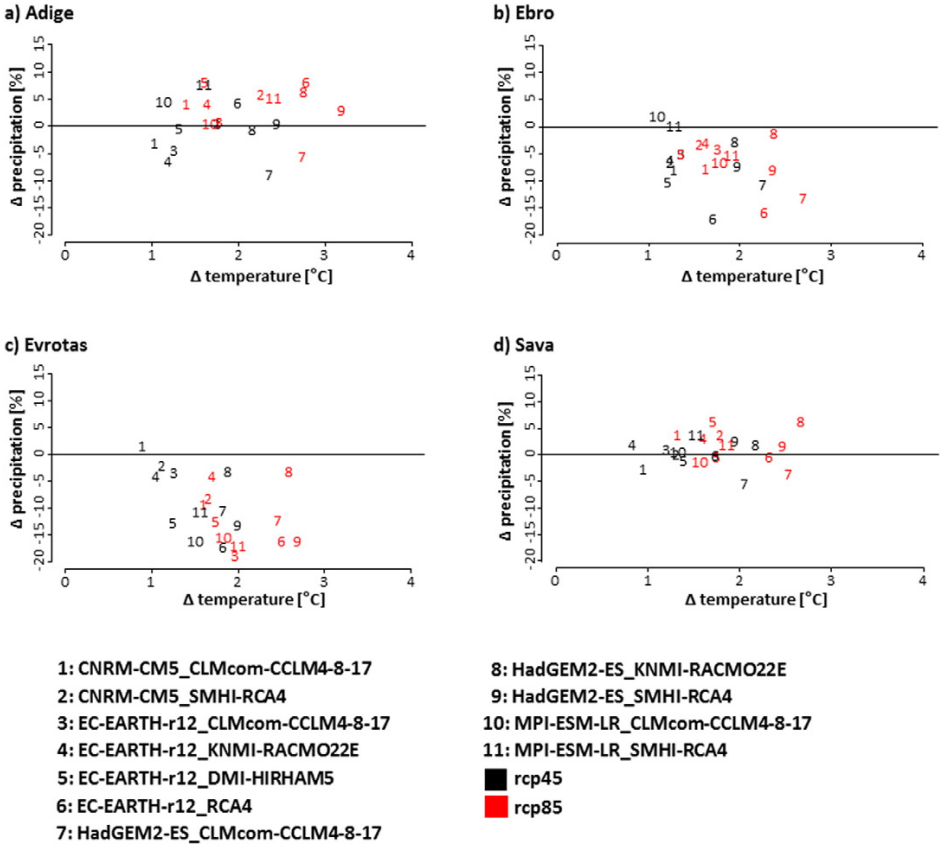


Fig. 2. Changes in annual mean precipitation [%] and air temperature [°C] over a) Adige, b) Ebro, c) Evrotas and d) Sava as projected by the EURO-CORDEX ensemble applied for the period 2035–2065 as compared to 1981–2010. Numbers represent various GCM-RCM combinations; red color refers to rcp 4.5, black to rcp 8.5.

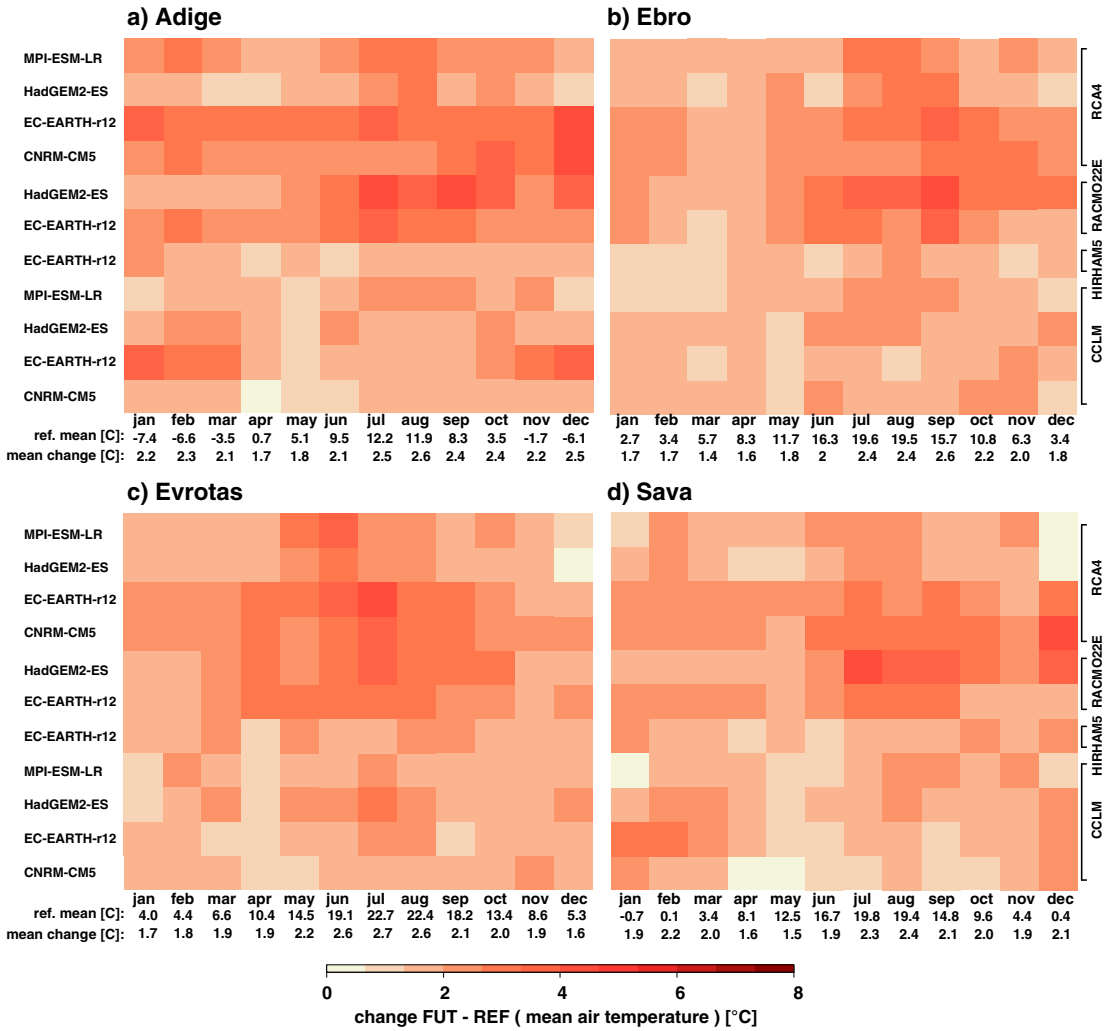


Fig. 3. Changes in mean 2 m air temperature (in °C) over a) Adige, b) Ebro, c) Evrotas and d) Ebro as projected by the EURO-CORDEX ensemble (RCP8.5) applied for each month in the period 2035–2065 as compared to 1981–2010.

simulations, only one resp. two projections reveal a slight increase in rcp 4.5 for the Evrotas resp. Ebro.

A general increase in precipitation over the Alps and a general decrease for the Mediterranean is well known and also found in other recent studies (Collins et al., 2013; Torma et al., 2015). The Sava catchment is located in the transition zone with increased precipitation in the northern, alpine part of the catchment and a decrease for the Southern Mediterranean.

Furthermore, the dependence of the driving GCM can be identified in the temperature changes, as the EC-EARTH driven simulations show the most moderate increase, whereas the HadGEM2 driven RCMs project the highest increase in response to a high sensitivity to radiative forcing. Fig. 4 shows monthly changes in precipitation for the four river basins and allows for a more detailed view on precipitation changes, grouped by RCM to address the impact of the regional model. Red colors represent decreased precipitation, while blue colors show increase. The mean changes for the ensemble are included below each figure for each month, with the corresponding absolute value for the reference period. There is a decrease in precipitation visible in summer and fall for all four catchments, however more pronounced for Ebro and Evrotas, which is partly due to low precipitation in the reference period. A general increase in winter precipitation can be identified for the Adige, Sava and with exceptions also for the Ebro. Only a few simulations show increased winter precipitation over the Evrotas. Changes in monthly precipitation over the Adige (Fig. 4a) range from a slight decrease in August and September (around 5%) to an increase in precipitation in winter and spring (up to 13%). For the Ebro (Fig. 4b), the ensemble mean only shows an increase in February and November (up to 5%) and strong decrease in summer (−16%).

Even more pronounced are the negative changes over the Evrotas, where the ensemble mean projects a decrease throughout the year, ranging from −3.7% (January) to −28.6% (August). With exception of the summer period (−1.7 to −12.7%) slight to moderate increase in precipitation is projected by the ensemble mean for the Sava (Fig. 4d) with a maximum in the winter months (13.6% in January).

The simulations downscaled with CCLM tend to show less sensitivity and more moderate changes – be it positive or negative - as compared to RCA4 and especially RACMO22E. Opposed changes for some months show the strong impact of the RCM on precipitation, which is especially visible in the winter months for the MPI-ESM-LR and HadGEM2 driven simulations downscaled with RCA4 and CCLM, respectively. These GCMs show high climate sensitivity (Andrews et al., 2012; Vial et al.,

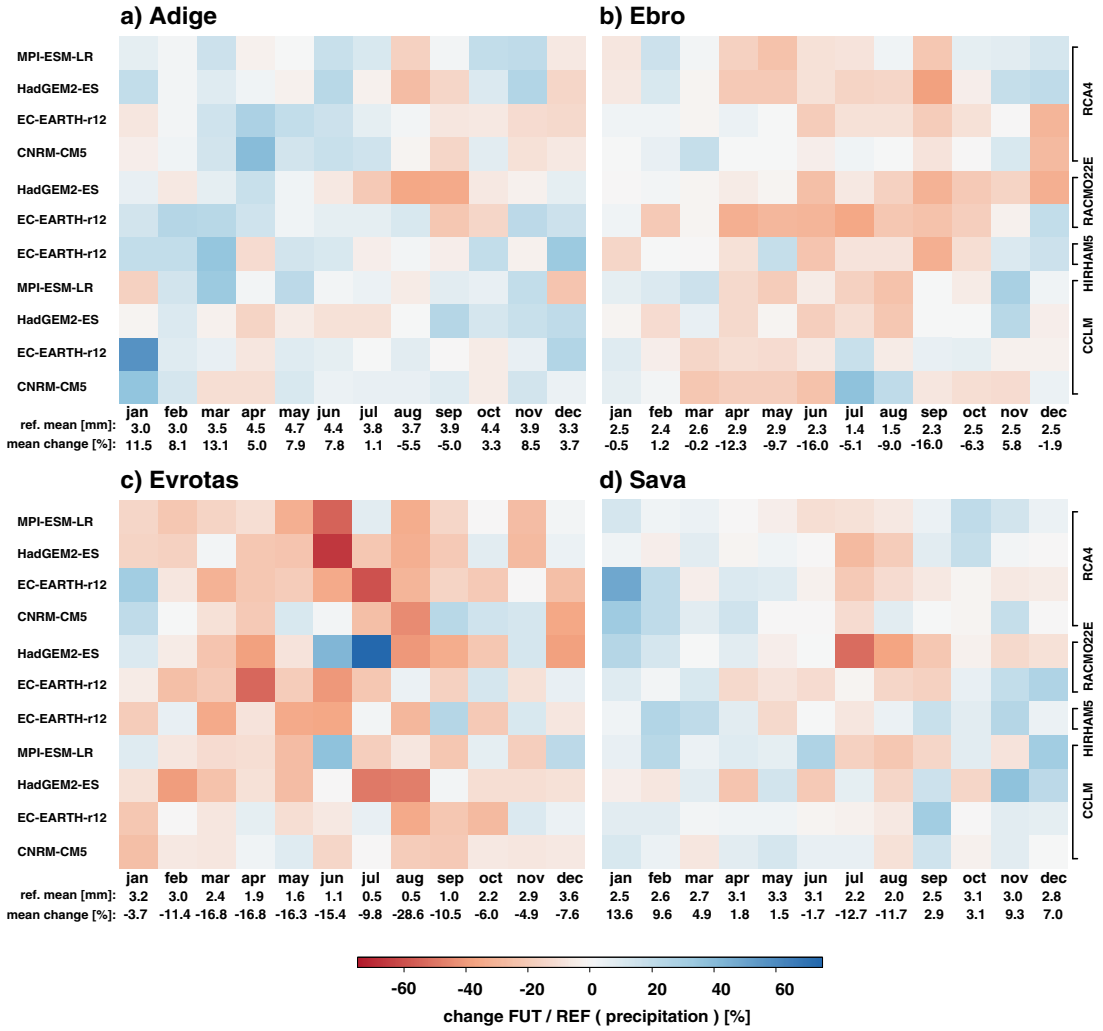


Fig. 4. Percentage changes in monthly mean precipitation over a) Adige, b) Ebro, c) Evrotas and d) Ebro as projected by the EURO-CORDEX ensemble (RCP8.5) applied for each month in the period 2035–2065 as compared to 1981–2010.

2013) which explains the stronger response of the simulations using those as boundary conditions.

3.2. Changes in evaporation and total runoff

Fig. 5 shows the change in evaporation the same way as already described for precipitation and temperature for the applied RCM ensemble. Most RCMs show increased evaporation over the Adige (a) and with exceptions in summer also for the Sava (d). Evaporation in these rather humid catchments is not limited by water availability, particularly in their northern parts. The ensemble mean projects up to 15% increase in evaporation throughout the year for the Adige and also for the Sava (with the exception of August). Strongest increase is projected for the winter months in both catchments, which follows the projections of precipitation (see Fig. 4) and stems from higher snow and glacier melting due to increased temperatures. In contrary to the Adige and most parts of the Sava, evaporation in the Ebro (b) and the Evrotas (c) is limited by water availability, hence indicating situations of water stress. Evaporation follows the changes in precipitation, which is likely to decrease (see Fig. 4) in the two catchments in spring, summer and fall. Temperature and slight precipitation increase cause evaporation to increase slightly in the winter months for most of the RCMs in the ensemble. The ensemble mean shows moderate decrease over the Ebro with maximum changes in August with a decrease of 8.1%. Changes over the Evrotas are more pronounced with a decrease of almost 15% in the summer months.

Simulations with the RCA4 tend to project increased evaporation in all catchments, with the exception of the summer months, whereas CCLM driven simulations rather project a moderate to strong decrease. This is likely linked to a general overestimation in cloud cover in the CCLM simulations (Jaeger et al., 2013; Pfeifroth et al., 2012) and a general positive cloud feedback for future climate conditions (Soden and Held, 2006). A strong cold bias of the RCA4 simulations over the Alps especially in the winter months likely leads to an underestimation of evaporation in the reference period resulting in a stronger positive change under future temperatures. Over three of the four catchments (Fig. 5a, c, d) and partially also over the Ebro (Fig. 5b) HIRHAM5 shows high sensitivity in evaporation changes with the strongest changes in spring over the Adige and the Sava and the strongest decrease in fall over these catchments. Additionally it is the only simulation showing increased evaporation over the Evrotas (Fig. 5c). A higher sensitivity of HIRHAM simulations due to overestimation of

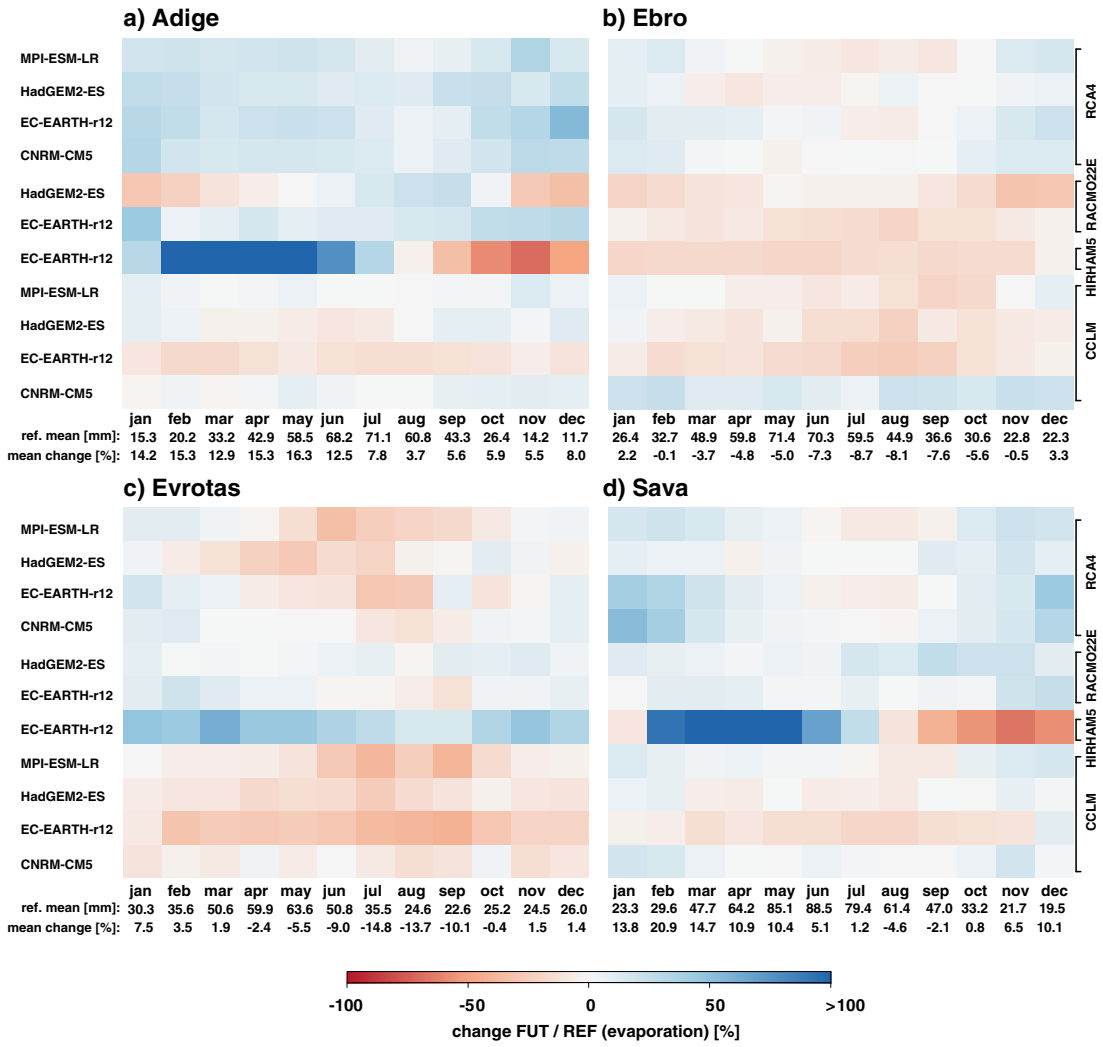


Fig. 5. Percentage changes in monthly evaporation over a) Adige, b) Ebro, c) Evrotas and d) Ebro as projected by the EURO-CORDEX ensemble (RCP8.5) applied for each month in the period 2035–2065 as compared to 1981–2010.

solar radiation resulting in less cloud cover was already observed by Hagemann et al. (2002) for version 4. Furthermore, soil moisture is not considered as multiple layers as in other RCMs, but rather as a bucket which strongly influences evaporation (Hagemann et al., 2002).

The RCM ensemble projects a mean increase in evaporation over the mountainous parts of the Adige (Fig. 6 a, e) and the Ebro (Fig. 6 c, g) catchment for both, rcp 4.5 and rcp 8.5, however more pronounced in the latter. Strongest decrease is projected for the southern part of the Ebro with up to 20% whereas for the Sava the mean decrease is homogenous throughout the catchment with slightly lower intensity of 10%. For the Evrotas, rcp 4.5 results in ambiguous results with slight changes in both directions (Fig. 6), however rcp 8.5 results in decrease throughout the basin (p).

Additionally, model agreement is presented to assess uncertainty in the projections. For the Adige, both under rcp 4.5 and 8.5 (b & f) high model agreement with nine or more models projecting increased evaporation can be identified. The same applies for the mountainous areas of the Ebro, whereas most parts of the basin are characterized by high model agreement in decrease, as zero or only a few models project an increase. Similar results can be found for the Evrotas, however only for rcp 8.5 (q); under rcp 4.5 (m) the models show ambiguous results. Model agreement over the Sava is less heterogeneous, as for rcp 4.5 some models also show decreased evaporation in parts of Bosnia and Herzegovina. Although the majority of simulations agree in increased evaporation, the ensemble mean indicates a slight decline. This can be attributed to the variations within the ensemble, as some RCMs project a strong decrease, causing misleading information in the ensemble mean. Stronger model agreement applies for the northern part of the basin where 8 or models project an increase in rcp 8.5, whereas rcp 4.5 shows, again, ambiguity in the model results.

Fig. 7 shows the monthly changes in total runoff for rcp 8.5 for the future period and all catchments. Increased runoff in the winter months can be identified over the Adige (a) and the Sava (d) which can be attributed to snow and glacier melt due to increased temperatures. Strong decrease in runoff is projected for the summer months over the Ebro (b) and the Evrotas (c) of around 30% from May to September. Lower decrease with 5–15% is projected for the summer months also in the Adige catchment. Over the Sava (d) the RCA4 simulations show the most distinct results in fall and the strongest decrease in summer, which is caused by slightly more moderate changes in evaporation (see Fig. 5d).

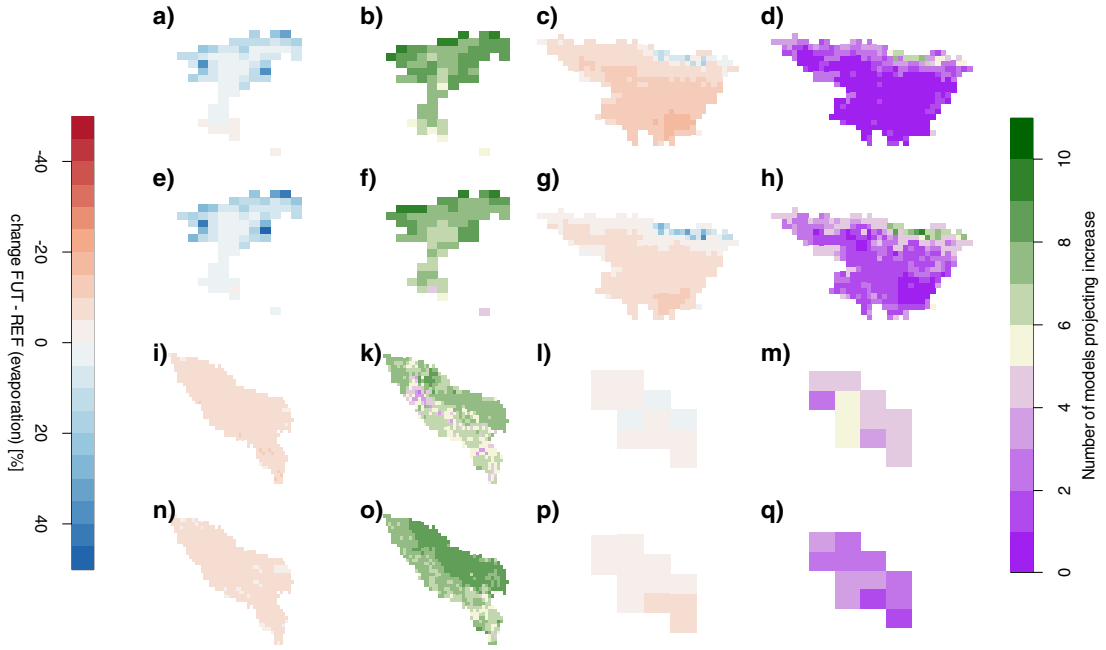


Fig. 6. Change in evaporation (evspsbl) for the four test sites as ensemble mean for rcp 4.5 (a, c, i, l) and rcp 8.5 (e, g, n, p) and corresponding number of models projecting an increase for rcp 4.5 (b, d, k, m) and rcp 8.5 (f, h, o, q) for the period 2035–2065 compared to 1981–2010.

The mean ensemble changes in runoff are shown in Fig. 8 together with the model agreement maps for all four catchments and both, rcp 4.5 and 8.5. For the Adige the projections show ambiguous results, as rcp 4.5 indicates a moderate increase of up to 10% (Fig. 8 a) with a high model agreement in decrease (b) of at least 8 models. Under rcp 8.5 (e) some areas in the catchment, especially the mountainous north show an increase in total runoff, however model agreement (f) shows heterogeneous and undetermined patterns as in some areas the model agreement in increase is high, others show a strong decrease, however in most of the catchment the overall agreement is low. In case of the Ebro, projected changes in runoff reveal a high heterogeneity in the change signal: in general a moderate to strong decline of up to 30% in the southern part of the basin is projected for both rcps (c & g). However, runoff projections show an increase in the central low lands in both rcps, more pronounced in rcp 8.5 (g). This area however shows, with some exceptions, the lowest model agreement, as four to six RCMs project an increase. Strong agreement in projected runoff decrease is visible for the mountainous north as well as the western and southern parts of the catchment.

Projections for the Sava indicate a strong decline throughout the catchment of 30% for both rcps (i & n). These changes seem more severe compared to the results presented in Fig. 7d, which is due to some RCMs projecting strong positive changes in some months, e.g. CNRM-CM5 driving RCA4 in summer. Model agreement for these changes is higher in rcp 4.5 (k) then in rcp 8.5 (o). Especially for the northern part of the basin, in the Slovenian Alps, the RCM ensemble shows ambiguous projections with five model projecting an increase in rcp 8.5 (o). For the Evrotas however a strong agreement on decreased runoff is given for both rcps (m & q) as none or at most three RCMs project an increase. Mean changes show a decline of 15–20% throughout the catchment, more pronounced under rcp 8.5 (p).

3.3. Falkenmark indicator

The Falkenmark indicator is widely used as a simple index to assess water scarcity on a larger scale (Brown and Matlock, 2011; Jaeger et al., 2013; Schyns et al., 2015). The indicator is defined as total annual runoff fraction available per capita. The classification of available water per capita to assess water scarcity as proposed by Falkenmark (1989), presented in Table 3, and was also applied here.

Although the indicator is widely used, obvious caveats arise from this measure, as the index does not take into account water reuse, technological improvement and, as only the annual runoff is applied, neither accounts for additional water from e.g. ground water aquifers. Additionally, important temporal variability and variation in water scarcity and system interactions are not represented by this indicator (Jaeger et al., 2013). Still, the index can be applied as a first step assessment of water scarcity and serve as a benchmark and serves still as indicator for decision making (Brown and Matlock, 2011) and assessment due to easy computation and availability of data (Jaeger et al., 2013). In the scope of this study, it is meant to serve as a spatial indicator for the need of adaptive water resources management in response to potential water scarcity due to climate change. Global population density information made available through the United Nations Environment Programme DEWA/GRID-Geneva was applied to calculate the total population within each catchment on the EUR-11 grid. Although grids are also available for 2005, the chosen density grid is based on population data for the year 1995 to represent the chosen reference period. As elaborated by Brown and Matlock, 2011, population change dominates changes in climate and other potential stressors when it comes to the assessment of available water per capita. The intention of applying the same population grid for future and reference period was to solely address the role of climate change as a possible stressor for aquatic ecosystems. Observed changes from the two population grids range from −2 to 10% already between the years 1995 and 2005. Additionally, the ensemble mean over all RCMs for total runoff (see Fig. 7) was used to calculate the indicator.

Fig. 9 shows the results for the indicator for each catchment and the reference period (a, e, i), rcp 4.5 (b, g, l) and rcp 8.5 (c, h, m) additionally, the corresponding population grid is shown in d, f and k. The Evrotas is excluded here, as the indicator highly follows the underlying population information and due to extremely low population (density), no water

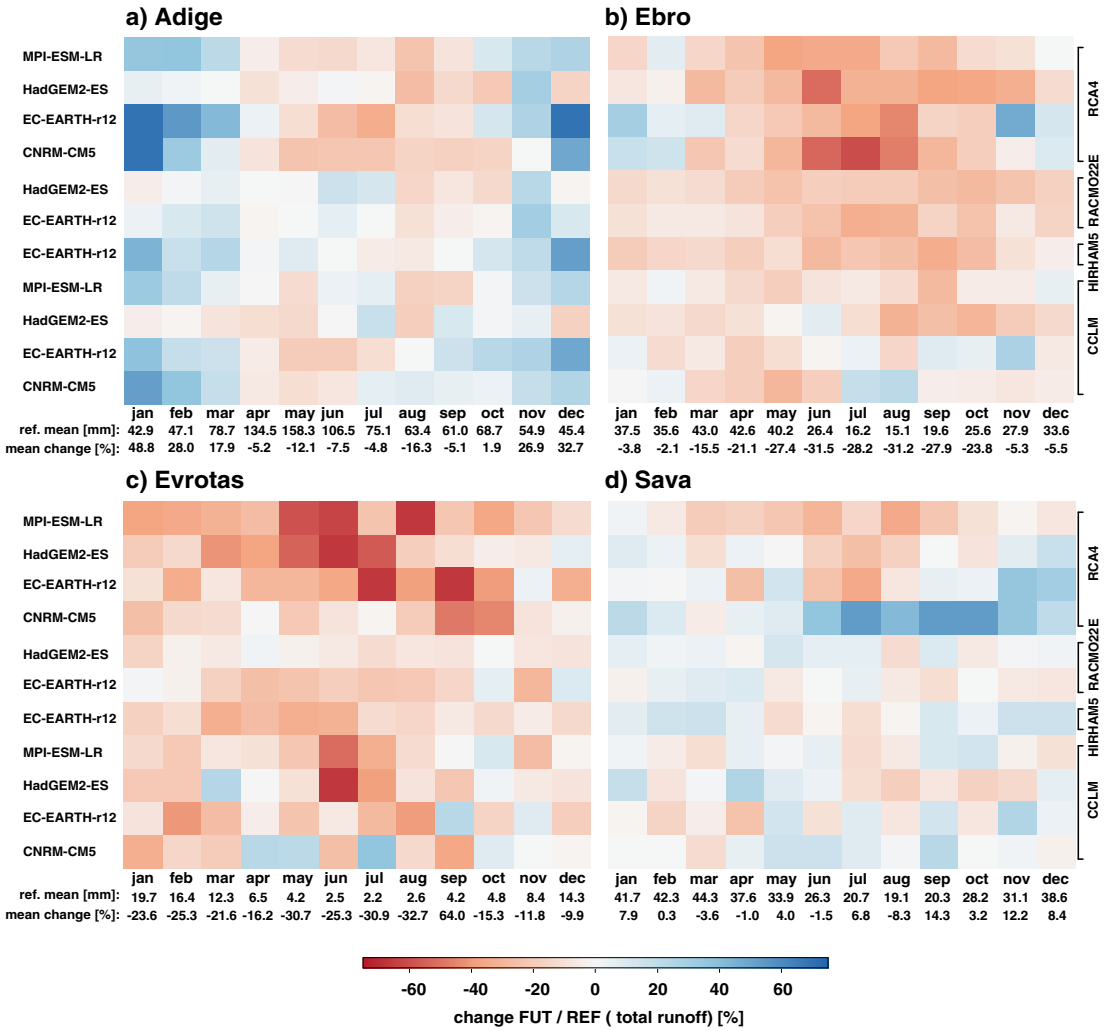


Fig. 7. Percentage changes in monthly total runoff over a) Adige, b) Ebro, c) Evrotas and d) Ebro as projected by the EURO-CORDEX ensemble (RCP8.5) applied for each month in the period 2035–2065 as compared to 1981–2010.

scarcity could be identified according to the classification in Table 3. Coloring in Fig. 9 follows the colors presented in Table 3. Grey areas for example represent grid cells with >1700 m^3 $year^{-1}$ water being available per person, hence corresponds to the class "no stress". Circles represent grid cells where the ensemble mean projects a decrease of >10% in rcp 4.5 or 8.5, respectively, as compared to the reference period. Grid cells marked with '+' represent model agreement in the ensemble, in that at least six out of the eleven RCM projections agree in the water scarcity category. This is only applied if stress is detected, hence if the available water per capita is <1700 m^3 $year^{-1}$.

The top row in Fig. 9 shows the results for the Adige for the reference period (a), rcp 4.5 (b), rcp 8.5 (c) and for comparison reasons the population map is included in d. Only a few grid cells show scarcity and follow the main cities in the area: Trento and Verona. Future projections show only one grid cell in rcp 4.5 and none for rcp 8.5, which follows the precipitation projections. However, the ensemble mean shows increased water scarcity with decreases in available water of >10% (circles in Fig. 9b & c). As already presented for precipitation, there is no significant agreement in the models on the water scarcity class.

In contrary, there is a stronger agreement on the categories for the Ebro (Fig. 9e–h). Several areas fall within the category 'absolute scarcity', mainly around the city of Zaragoza and in the east of the catchment. Projections show even stronger model agreement and high decrease for most of the areas already classified as 'stress' or less. As for the Adige, rcp 4.5 results in slightly higher water scarcity.

As shown in Fig. 9k, no spatially distributed population density is included in the data set for the countries of Bosnia and Herzegovina and Serbia, resulting in a homogenous distribution of population over large parts of the Sava catchment. This results for the reference period (i) in large areas of scarcity in Serbia, however, also some areas of Slovenia are classified as any of the three scarcity classes, as e.g. for the city of Ljubljana. Projections show a strong decrease in available water for most of the classified grid cells with higher agreement for rcp 4.5 (l) compared to rcp 8.5 (m) with more grid cells being classified as 'absolute water scarcity'.

4. Discussion and resulting water scarcity

Based on the results presented in this study, a clear climate change signal can be identified for all presented variables and both representative concentration pathways over the basins with the exception of precipitation over the Adige and the Sava in rcp 4.5. The general changes in

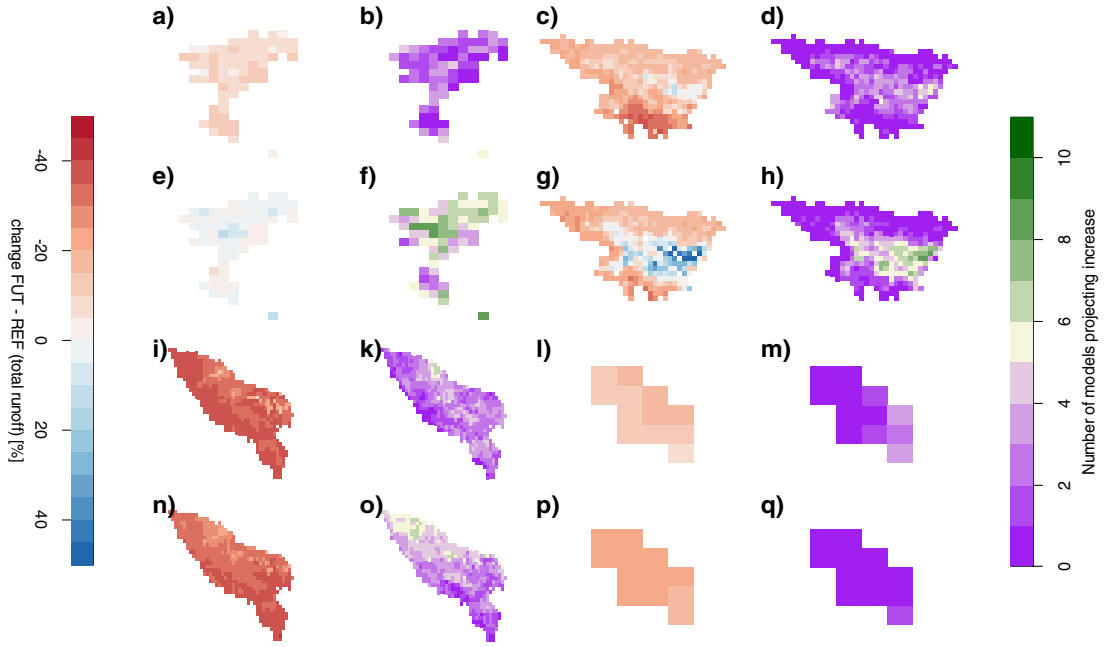


Fig. 8. Change in total runoff (mrro) for the four test sites as ensemble mean for rcp 4.5 (a, c, i, l) and rcp 8.5 (e, g, n, p) and corresponding number of models projecting an increase for rcp 4.5 (b, d, k, m) and rcp 8.5 (f, h, o, q) for the period 2035–2065 compared to 1981–2010.

precipitation with an increase over the Alps and a strong decline over the Mediterranean agree with results found in global or pan-European studies (Polade et al., 2014). However, local increase of precipitation over the mountainous regions of the Adige, Ebro and Sava confirm the importance of high resolution climate simulations that are capable of representing local topography (Gao et al., 2006). Changes in evaporation and total runoff revealed strong differences within the RCM ensemble that can be attributed to the individual RCM schemes for land surface and runoff generation. The differences are more structural for evaporation (Fig. 5) and more scattered for total runoff (Fig. 7). A decrease in runoff can be identified over three of the four basins with up to 30% which agrees with the results presented by Haddeland et al., 2014 for 2 °C warming level in the Mediterranean. The most important findings in each of the case studies are summarized in the following.

4.1. Findings in each of the catchments

4.1.1. Adige

Increased temperature and precipitation result in increased evaporation over the Adige, whereas no clear signal could be identified for annual mean runoff. However, monthly changes reveal a decrease in runoff for the summer months and an increase in winter. The latter can be explained by snow and glacier melt due to increased temperatures. These changes indicate a tendency toward increased water scarcity for the summer months especially for the southern part of the catchment. Although the Falkenmark indicator did not reveal water scarcity for most parts of the catchment, however, a decrease in water availability could be identified for the south. In this area a strong model agreement is present for both, runoff and evaporation, indicating an increased likelihood of water scarcity.

Table 3
Categories of water scarcity as defined by Falkenmark (1989). Coloring refers to color scheme applied in Fig. 9.

Available water per capita (m^3 yr^{-1})	Scarcity category
>1,700	No stress
1,000–1,700	Stress
500–1,000	Scarcity
<500	Absolute scarcity

4.1.2. Ebro

There is a clear signal for precipitation decrease over the Ebro in spring, summer and fall, resulting in lower evaporation and runoff decrease in these periods. Both changes indicate a strong increase in water scarcity over the basin. The strong signal in runoff for this period with up to 30% confirms this tendency. However, the increase in evaporation in the mountainous area in the north, resulting from increased precipitation and snow melt due to increased temperatures, causes increased runoff for the central low lands, which confirms the requirement and benefits of high resolution climate modeling. The Falkenmark indicator displays water scarcity for parts of the Ebro basin already for the reference period with a projected tendency to further increase in the future.

4.1.3. Evrotas

Strongest negative changes in precipitation are projected over the Evrotas basin with up to 30% in the summer months while some simulations show even more pronounced changes. This translates to strong changes in runoff throughout the year and decreased evaporation at least in spring, summer and fall. Differences in the RCMs are again most visible for this variable, as the CCLM simulations clearly show different projections for the winter months as compared to the rest of the ensemble. The ensemble mean shows a decrease in annual runoff of around 20%. These changes indicate increased water scarcity over the entire basin, especially in spring, summer and fall. High model agreement confirms this tendency, even though due to the small size of the basin only a few grid points are included in the analysis. The Falkenmark indicator did not deliver robust results, as it is dominated by the low population density in the catchment.

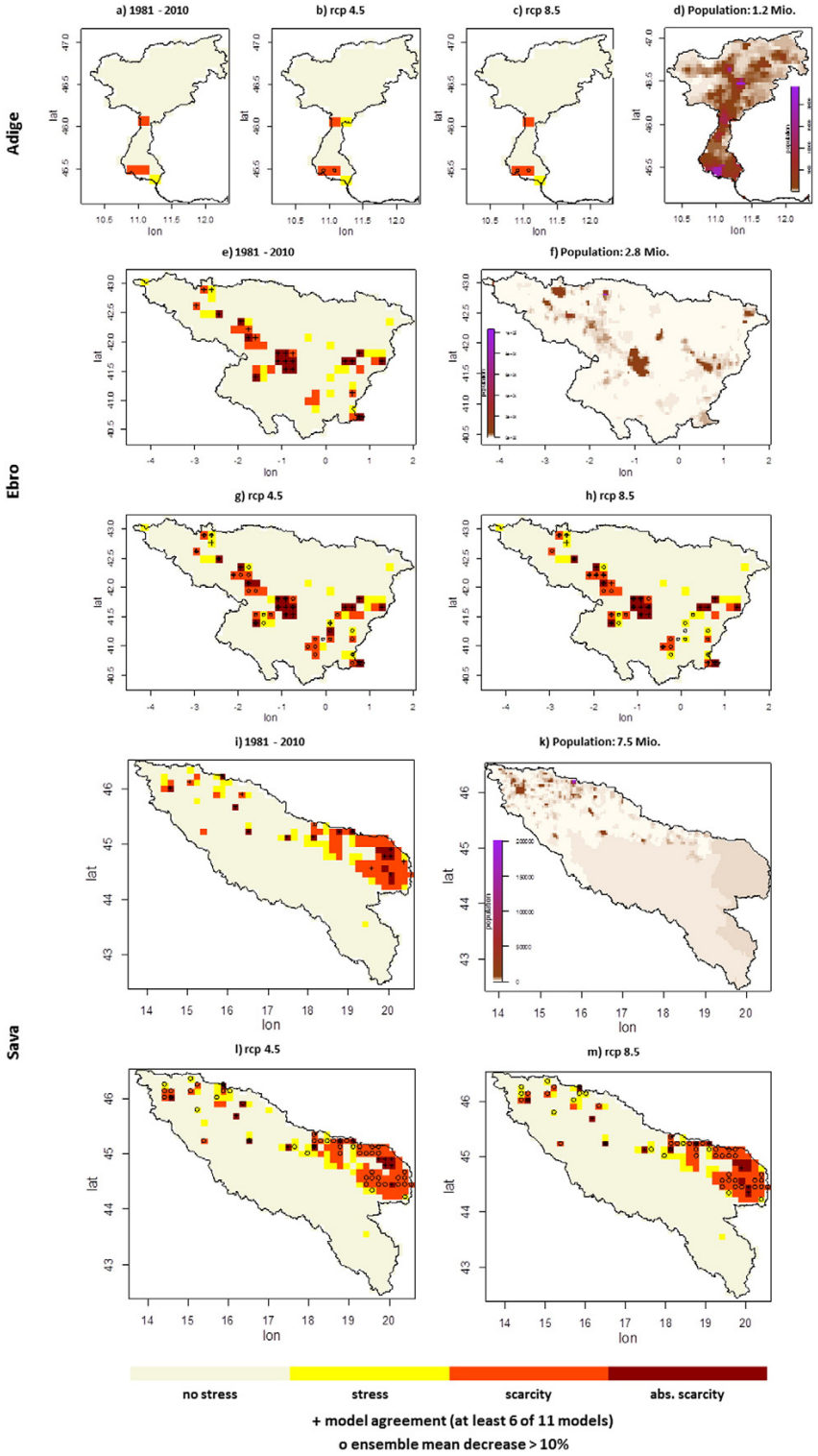


Fig. 9. Falkenmark indicator to assess water scarcity of the test sites for the reference period (a, e, i), rcp 4.5 (b, g, l) and rcp 8.5 (c, h, m). As the indicator is dependent on population, population maps are included (d, f, k). Evrotas is not shown here.

4.1.4. Sava

Results for the Sava are most ambiguous and difficult to interpret. Precipitation projections point toward slight increases on average for the entire catchment, resulting from increases over the alpine part in Slovenia and decreases over the Mediterranean south. As for the Adige, increased precipitation and temperature in winter translates to increased runoff and evaporation in these months. A decrease in precipitation causes decreased runoff in summer, whereas evaporation projections show indistinct results. Annual ensemble means however indicate a strong decrease of up to 30%, which is due to some RCMs projecting strong changes. High model agreement can only be identified over the southern part, indicating a tendency toward increased water scarcity over this area.

4.2. Variance decomposition

It is generally known and acknowledged that climate projections and resulting impact modeling activities are affected by many sources of uncertainty (Holzkaemper et al., 2015). One of the challenges in the interpretation of these information is the quantification of these uncertainties. In this study the uncertainties related to the climate change signal over each basin, defined here as the spread of the ensemble is quantified by applying a simple variance decomposition. This approach is widely used to assess the fraction of RCM, GCM and RCP to the total variance within an ensemble of climate models over a given area (Ferro, 2004; Déqué et al., 2007; Déqué et al., 2012). Following Déqué et al. (2007) and Déqué et al., (2012), three main sources of uncertainty can be identified in the ensemble applied in this study. Model uncertainty, as four different RCMs where applied here, boundary uncertainty due to different driving GCMs and scenario, or radiative uncertainty, as two different RCPs were applied. Here, the uncertainty in the climate response, defined as presented in the last section, averaged for all grid points over each river basin was assessed. The approach is described briefly in the following, for details on the calculation procedure it is referred to the studies mentioned.

The total variance can be written as the sum of the contributions of each contribution term as presented in Eq. (1). Here R is the part of the variance of the RCM, G the contribution of the GCM and S the

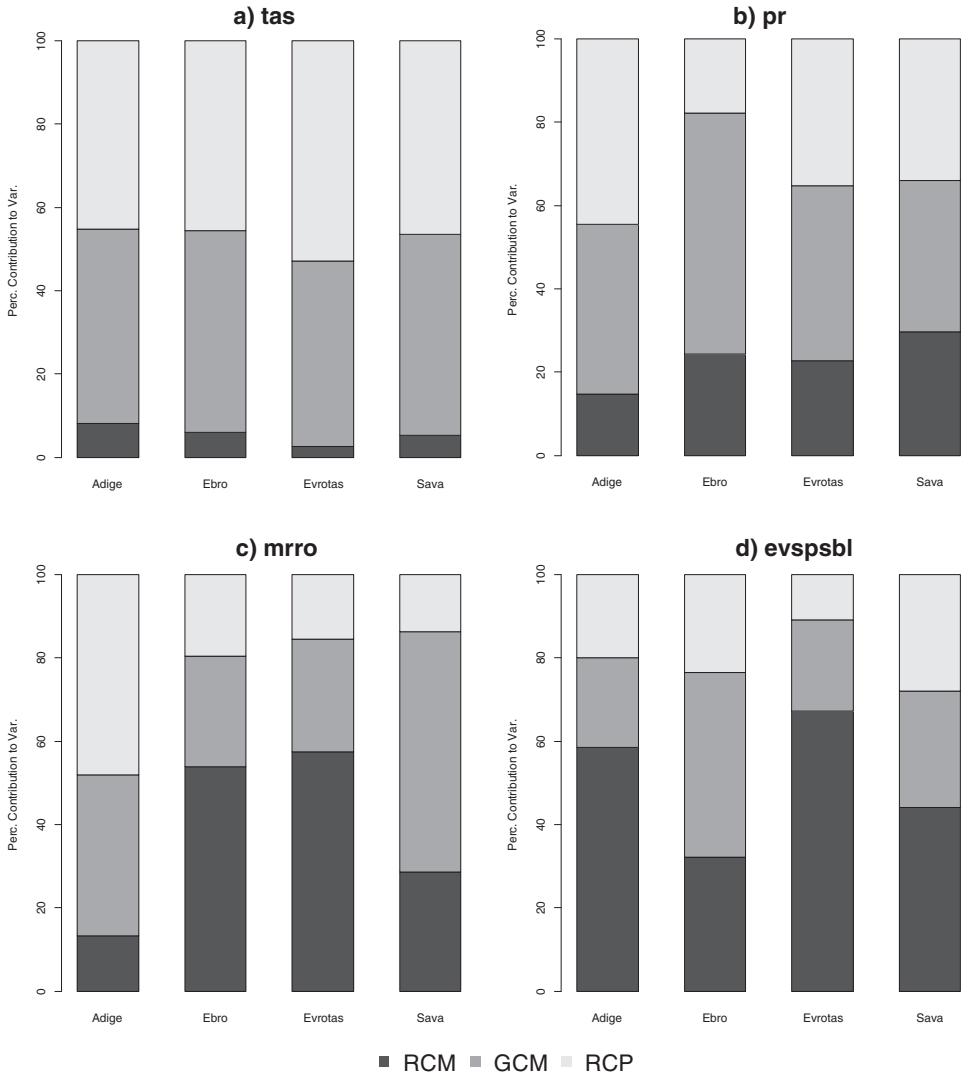


Fig. 10. Contribution to variance for RCM, GCM and RCP for the four variables analyzed: a) temperature (tas), b) precipitation (p), c) total runoff (mrro) and d) evaporation (evspsbl) in the climate change signals for the period 2035–2065 compared to 1981–2010 over the case study areas.

variance due to the scenario or RCP, respectively. Other terms are defined as interaction terms between the main contributions, where e.g. RG represents the interaction between RCM and GCM, others are defined similarly.

$$V = R + G + S + RG + RS + GS + RSG \quad (1)$$

The contribution of the RCM to the total variance can hence be written as sum of the main contribution of the RCMs and all corresponding interaction terms, G and R can be written in a similar form:

$$V(R) = R + RG + RS + RSG \quad (2)$$

Assuming the total variance as a simple sum as in Eq. (1) requires a complete matrix, meaning all simulation being available for all possible GCM-RCM combinations. As presented in Table 1 this is not the case in the present study hence the matrix needs to be filled. The simple, yet robust approach presented by Déqué et al. (2007) was applied to overcome this issue. The calculation of these artificial climate responses was performed in a way that the interaction term *RSG* equals zero. Therefore, the interaction terms are not represented correctly by definition, however the focus here is to get an estimate of the contribution to the variance of each source rather than the interaction between them. Additional, the final variance presented here is the fraction of each source of uncertainty rather than the total variance. Therefore in the following only the three main sources are presented calculated as presented in Eq. (2) and including all interaction terms. As the sum of these will exceed 100% due to the contribution of the interaction terms, the results are normalized to 1, or 100%, respectively to preserve the fraction but be easier comparable and accessible. Fig. 10 shows the contribution of RCM (black), GCM (grey) and RCP (light grey) to the annual variance as bar plots in percent for air temperature (a), precipitation (b), total runoff (c) and evaporation (d) over the four river basins. For temperature and precipitation the uncertainty introduced by the GCM and the RCP are larger than for the RCM. As expected from Fig. 2 and in agreement with previous studies (e.g. Hawkins and Sutton, 2009), the uncertainty introduced by the RCP contributes most to the total variance over the four river basins for temperature for this time period. In case of the Adige, scenario uncertainty represents the largest source of uncertainty in all variables except evaporation. The contribution of uncertainty

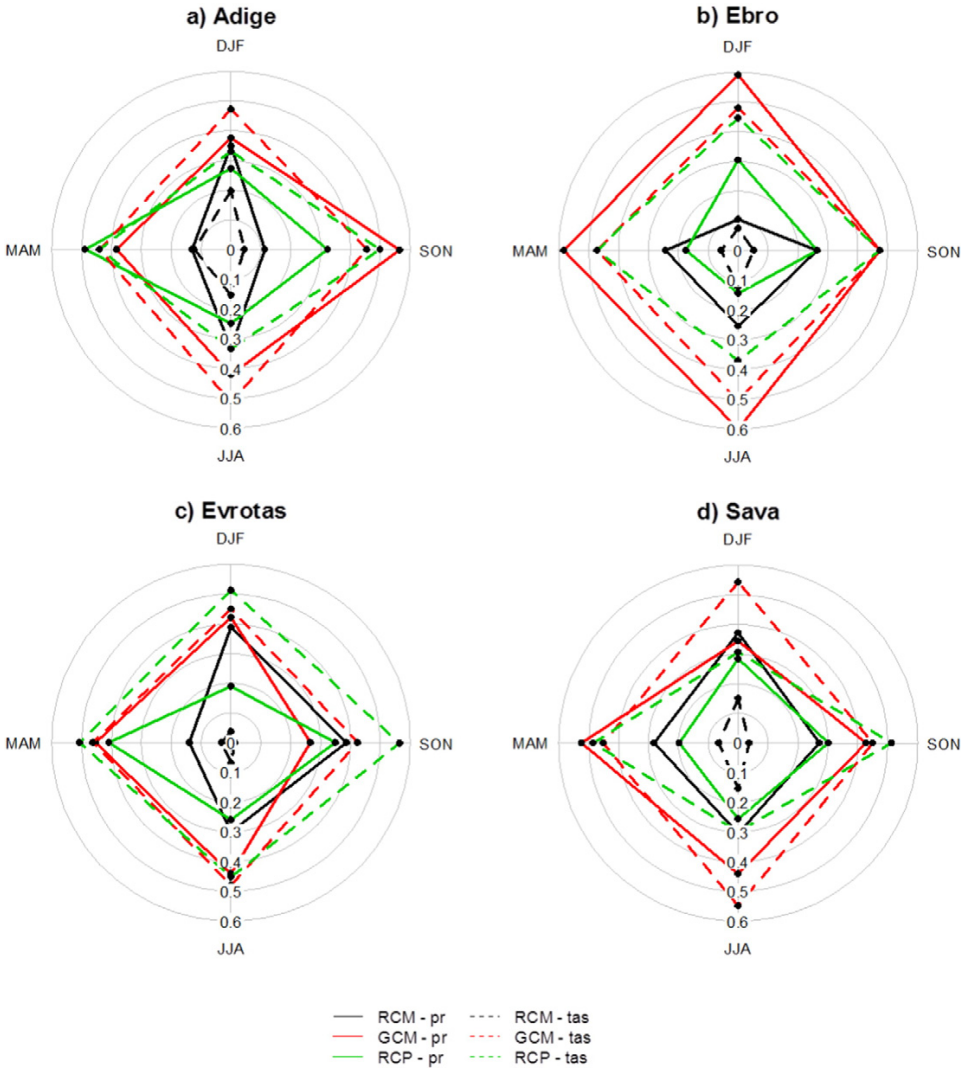


Fig. 11. Seasonal explained variance fraction for precipitation (solid) and air temperature (dashed) over the a) Adige, b) Ebro, c) Evrotas and d) Sava. Contribution of RCM is shown in black, GCM in red and the scenario (RCP) in green.

related to GCM dominates for precipitation. Uncertainty introduced by the RCM generally represents the largest fraction for evaporation and runoff which is due to the different structural differences of the RCMs for calculation of these processes.

The importance of not only regional but also seasonal assessment of the climate change signal was presented in previous sections. Therefore, variance decomposition was also carried out on seasonal climate change signals. Fig. 11 shows the seasonal uncertainty assessment for precipitation (solid line) and temperature (dashed), while evaporation (solid) and runoff (dashed) are presented in Fig. 12. As shown in Fig. 11, the contribution of the RCM is largest for summer precipitation in all four river basins, which is due to the higher resolution resulting in better representation of orography and convective events in RCMs compared to GCMs. However, in most cases, the absolute contribution to the variance is lower compared to GCM and RCP. Choice of GCM dominates the uncertainty in precipitation and temperature over the Adige in winter and summer, over the Sava with the exception of fall and over the Ebro in all seasons. The uncertainty introduced by the radiative forcing are dominate for temperature in fall and contribute largely to spring precipitation over the Adige and Evrotas.

As presented in Fig. 12 the fraction for RCM is usually the largest for evaporation (black, solid line) throughout the year in most of the river basins. The choice of RCP is more important for runoff than evaporation, where contribution to the variance is usually around 0.1 or lower. Following precipitation and temperature, the uncertainty related to the GCM mostly represent the largest fraction in the larger river basins Ebro and Sava also for evaporation and runoff, while the fraction of uncertainty of the RCM is more pronounced over the Adige and the Evrotas, hence the smaller basins.

5. Summary and conclusion

An ensemble of eleven simulations of four Regional Climate Models (RCMs) downscaling various General Circulation Models (GCMs) available for two Representative Concentration Pathways (RCPs) each, was applied over four river basins ranging from 2500 to 95,000 km^2 to assess climate change signals of four variables and resulting consequences for water scarcity. These variables, consisting of precipitation (pr), 2 m mean air temperature (tas), evaporation (evspsbl) and total runoff (mrro) were used from the RCM simulations in the original grid at

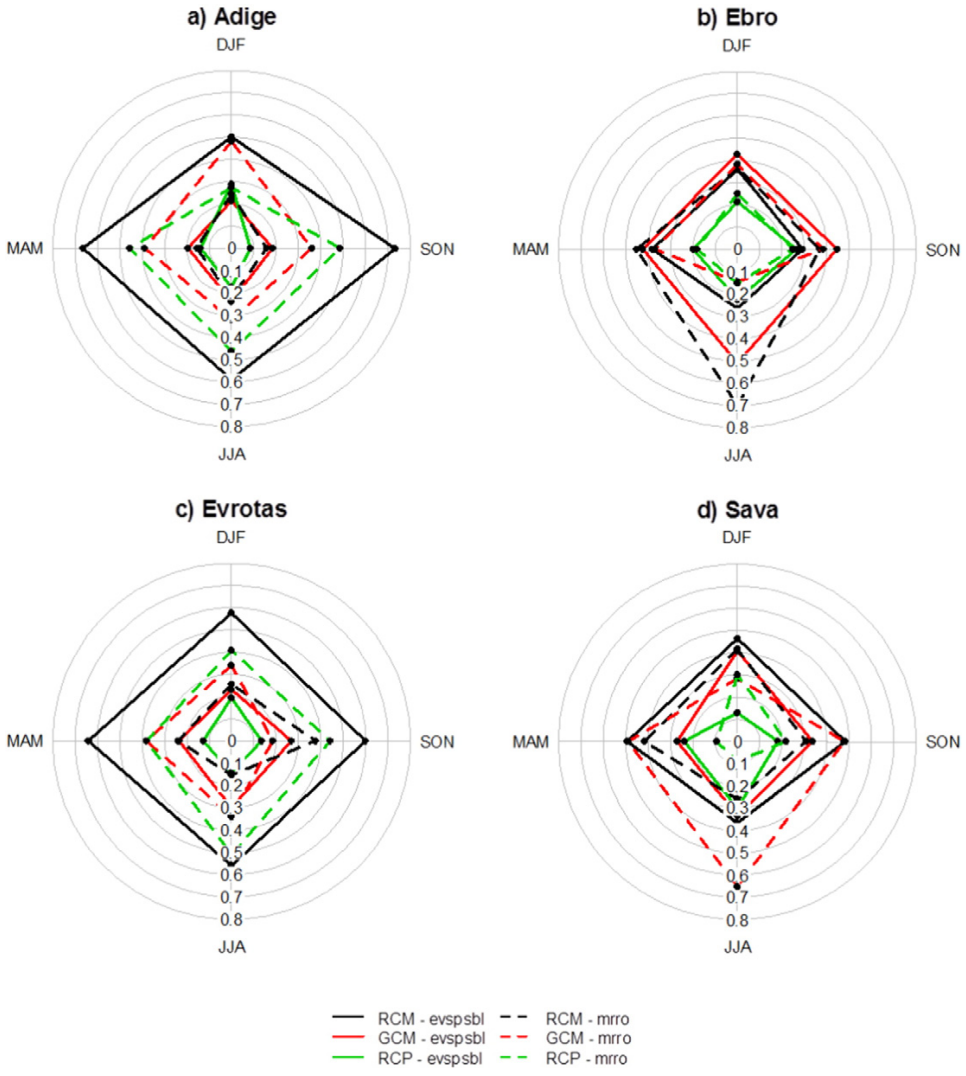


Fig. 12. Seasonal explained variance fraction for evaporation (solid) and runoff (dashed) over the a) Adige, b) Ebro, c) Evrotas and d) Sava. Contribution of RCM is shown in black, GCM in red and the scenario (RCP) in green.

0.11° (~12 km). The Falkenmark indicator was applied to determine future changes on water availability and to assess the impact of climate change on water scarcity dynamics. For this purpose the available water per capita was assessed based on the same population grid for reference and future calculations. The main findings of this study can be concluded as follows:

1) A clear climate change signal could be identified for all variables over the four basins, with the exception of northern parts of the Adige and the Sava. A consistent warming signal in mean air temperature is evident across all four case study basins. The ensemble shows an increase of 1.5 to 3° with all simulations agreeing in increase for all basins. Precipitation changes show decrease over the Ebro and Evrotas, slight increase over the Adige and ambiguous results over the Sava. Seasonal changes in evaporation and runoff were detected, showing the importance of a monthly assessment. Decreasing evaporation and runoff over all four basins was the main finding.
2) Small scale changes in parts of the basins, e.g. mountainous areas, confirm the importance of high resolution RCM projections that are capable of resolving relevant topographic features, to assess future changes and possible climate change impacts in mountainous areas: Precipitation increase and a resulting increase in evaporation are identifiable in the Adige, Ebro and Sava despite possible contrary projections for the basin mean.
3) Increased water scarcity can be concluded from the presented changes over the Ebro and the Evrotas, as well as the southern parts of the Adige and the Sava. For these areas, high model agreement indicates a higher probability for water scarcity. Decreased runoff and evaporation especially in summer indicate further pressure on the river basins that are already partly characterized by flow intermittency.
4) The Falkenmark indicator confirmed the importance of climate change as a stressor for the Ebro and the Sava and, to some extent, also for the Adige. The indicator was not presented for the Evrotas due to low population in the catchment.
5) Variance decomposition showed higher uncertainties in the GCMs and RCPs for temperature and precipitation compared to the RCM contribution. Convective events in summer however result in higher fraction of variance from RCMs in the summer months over the four basins. Different calculation schemes for evaporation and runoff as well as different land-surface- and radiation schemes in the RCMs result in larger fractions of variance for RCMs for these two variables.

The high horizontal resolution of the current generation of RCMs is capable of better representing topography acceptably and allows also for analyses on sub-catchment level. The use of a large ensemble of RCM projections is inevitable to detect climate change signals and separate them from natural variability. The study represents a computational effective method to assess future water scarcity over various river basins that can – in contrary to simulations with regional hydrological models – also be applied in data scarce basins. The abundance of RCM simulations available throughout the globe and the use of global data sets for population distribution permits water scarcity assessment in data scarce areas, not only limited to the Mediterranean.

Acknowledgments

This work has been supported by the European Communities 7th Framework Programme Funding under Grant agreement no. 603629-ENV-2013-6.2.1-Globaqua.

References

Andrews, T., Gregory, J.M., Webb, M.J., Taylor, K.E., 2012. Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models. Geophys. Res. Lett. 39, 9.

Brown, A., Matlock, M.D., 2011. A review of water scarcity indices and methodologies. The Sustainability Consortium. Univ. of Arkansas.

Chiogna, G., Majone, B., Cano Paoli, K., Diamantini, E., Stella, E., Malluci, S., Lencioni, V., Zandonai, F., Bellin, A., 2016. A review of hydrological and chemical stressors in the Adige catchment and its ecological status. Sci. Total Environ. 540, 429–443.

Collins, M., Knutt, R., Arblaster, J.M., Dufresne, J.L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W.J., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A.J., Wehner, M., 2013. Long-term climate change: projections, commitments and irreversibility. Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 1029–1136.

Christensen, J., Hewitson, B., Busuioc, A., Chen, A., Gao, X., Helt, I., Jones, R., Kolli, R., Kwon, W.-T., Laprise, R., Magaña Rueda, V., Mearns, L., Menéndez, C., Räisänen, J., Rinke, A., Sarr, A., Whetton, P., 2007. Regional climate projections. In: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L. (Eds.), Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.

Déqué, M., Rowell, D.P., Lüthi, D., Giorgi, F., Christensen, J.H., Rockel, B., Jacob, D., Kjellström, E., de Castro, M., van den Hurk, B., 2007. An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections. Clim. Chang. 81, 53–70.

Déqué, M., Somot, S., Sanchez-Gomez, E., Goodess, C.M., Jacob, D., Lenderink, G., Christensen, O.B., 2012. The spread amongst ENSEMBLES regional scenarios: regional climate models, driving general circulation models and interannual variability. Clim. Dyn. 38, 951–964.

Falkenmark, M., 1989. The massive scarcity threatening Africa-why isn't it being addressed. Ambio 18 (2), 112–118.

Ferro, C.A., 2004. Attributing Variation in a Regional Climate Change Modelling Experiment. EU Project PRUDENCE Tech. Rep.

Field, C.B., Barros, V., Stocker, T.F., Qin, D., Dokken, D.J., Ebi, K.L., Mastrandrea, M.D., Mach, K.J., Plattner, G.-K., Allen, S.K., Tignor, M., Midgley, P.M. (Eds.), 2012. Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, UK, and New York, NY, USA.

Gampe, D., Ludwig, R., Qahman, K., Afifi, S., 2016. Applying the triangle method for the parameterization of irrigated areas as input for spatially distributed hydrological modeling—assessing future drought risk in the Gaza strip (Palestine). Sci. Total Environ. 543, 877–888.

Gao, X., Pal, J.S., Giorgi, F., 2006. Projected changes in mean and extreme precipitation over the Mediterranean region from a high resolution double nested RCM simulation. Geophys. Res. Lett. 33, L03706.

Giorgi, F., 2006. Climate change hot-spots. Geophys. Res. Lett. 33, L08707.

Giorgi, F., Jones, C., Asrar, G.R., 2009. Addressing climate information needs at the regional level: the CORDEX framework. WMO Bull. 58, 175–183.

Gosling, S.N., Arnell, N.W., 2016. A global assessment of the impact of climate change on water scarcity. Clim. Chang. 134 (3), 371–385.

Haddeland, I., Heinke, J., Biemans, H., Eisner, S., Flörke, M., Hanasaki, N., Konzmann, M., Ludwig, F., Masaki, Y., Schewe, J., Stacke, T., Tessler, Z.D., Wada, Y., Wisser, D., 2014. Global water resources affected by human interventions and climate change. Proc. Natl. Acad. Sci. 111–119, 3251–3256.

Hagemann, S., Machenhauer, B., Christensen, O.B., Déqué, M., Jacob, D., Jones, R., Vidale, P.L., 2002. Intercomparison of water and energy budgets simulated by regional climate models applied over Europe. Report/MPI für Meteorologie, p. 338.

Hanasaki, N., Fujimori, S., Yamamoto, T., Yoshikawa, S., Masaki, Y., Hijioka, Y., Kainuma, M., Kanamori, Y., Masui, T., Takahashi, K., Kanae, S., 2012. A global water scarcity assessment under shared socio-economic pathways – part 2: water availability and scarcity. Hydrol. Earth Syst. Sci. Discuss. 9, 13933–13994.

Hawkins, E., Sutton, R., 2009. The potential to narrow uncertainty in regional climate predictions. Bull. Am. Meteorol. Soc. 90 (8), 1095–1107.

Holzkaemper, A., Klein, T., Seppelt, R., Fuhrer, J., 2015. Assessing the propagation of uncertainties in multi-objective optimization for agro-ecosystem adaptation to climate change. Environ. Model. Softw. 66, 27–35.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O.B., Bouwer, L.M., Braun, A., Colette, A., Déqué, M., Georgievsk, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haenseler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Valentini, R., Vautar, R., Weber, B., Yiou, P., 2013. EURO-CORDEX: new high-resolution climate change projections for European impact research. Reg. Environ. Chang. 14, 563–578.

Jaeger, W.K., Plantinga, A.J., Chang, H., Dello, K., Grant, G., Hulse, D., McDonnell, J.J., Lancaster, S., Moradkhani, H., Morzilla, A.T., Mote, P., Nolin, A., Santelmann, M., Wu, J., 2013. Toward a formal definition of water scarcity in natural-human systems. Water Resour. Res. 49, 4506–4517.

Jones, C., Nikulin, G., 2009. Understanding the time-axis in coupled climate models: uncertainty, natural variability and the need for an ensemble. Rossby Centre Newsletter 4–8 (May 2009).

Jones, C.G., Giorgi, F., Asrar, G., 2011. The coordinated regional downscaling experiment: CORDEX; an international downscaling link to CMIP5. CLIVAR Exchanges 56, 34–40.

Kjellström, E., Thejll, P., Rummukainen, M., Christensen, J.H., Boberg, F., Christensen, O.B.C., Fox, M.C., 2013. Emerging regional climate change signals for Europe under varying large-scale circulation conditions. Clim. Res. 56, 103–119.

Levi, L., Jaramillo, F., Andričević, R., Destouni, G., 2015. Hydroclimatic changes and drivers in the Sava River catchment and comparison with Swedish catchments. Ambio 44, 624–634.

Koutroulis, A.G., Grillakis, M.G., Daliakopoulos, I.N., Tsanis, I.K., Jacob, D., 2016. Cross sectoral impacts on water availability at +2 °C and +3 °C for east Mediterranean island states: the case of Crete. J. Hydrol. 532, 16–28.

Kotlarski, S., Keuler, K., Christensen, O.B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., van Meijgaard, E., Nikulin, G., Schär, C., Teichmann, C., Vautard, R., Warrach-Sagi, K., Wulfmeyer, V., 2014. Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. Geosci. Model Dev. 7 (4), 1297–1333.

Meyer, S., Blaschek, M., Duttmann, R., Ludwig, R., 2016. Improved hydrological model parametrization for climate change impact assessment under data scarcity—the potential of field monitoring techniques and geostatistics. Sci. Total Environ. 543, 906–923.

Milačič, R., Ščančar, J., Murko, S., Kocman, D., Horvat, M., 2010. A complex investigation of the extent of pollution in sediments of the Sava River. Part 1: selected elements. Environ. Monit. Assess. 163, 263–275.

Navarro-Ortega, A., Acuña, V., Bellin, A., Burek, P., Cassiani, G., Choukr-Allah, R., Dolédec, S., Elosegi, A., Ferrari, F., Ginebreda, A., Grathwohl, P., Jones, C., Ker Rault, P., Kok, K., Koundouri, P., Ludwig, R.P., Merz, R., Milacic, R., Muñoz, I., Nikulin, G., Paniconi, C., Paunovi, M., Petrovic, M., Sabater, L., Sabater, S., Skoulikidis, N.T., Slob, A., Teutsch, G., Voulvoulis, N., Barceló, D., 2015. Managing the effects of multiple stressors on aquatic ecosystems under water scarcity. The GLOBAQUA project. Sci. Total Environ. 503–504, 3–9.

Pfeifroth, U., Hollmann, R., Ahrens, B., 2012. Cloud cover diurnal cycles in satellite data and regional climate model simulations. Meteorol. Z. 21 (6), 551–560.

Polade, S.D., Pierce, D.W., Cayan, D.R., Gershunov, A., Dettinger, M.D., 2014. The key role of dry days in changing regional climate and precipitation regimes. Sci. Rep. 4.

Serra, P., Salvati, L., Queralt, E., Pin, C., Gonzalez, O., Pons, X., 2016. Estimating water consumption and irrigation requirements in a long-established Mediterranean rural community by remote sensing and field data. Irrig. Drain. http://dx.doi.org/10.1002/ird.1978.

Schyns, J.F., Hoekstra, A.Y., Booij, M.J., 2015. Review and classification of indicators of green water availability and scarcity. Hydrol. Earth Syst. Sci. 19, 4581–4608.

Skoulikidis, N.T., Vardakas, L., Karaouzas, I., Economou, A.N., Dimitriou, E., Zogaris, S., 2011. Assessing water stress in Mediterranean lotic systems: insights from an artificially intermittent river in Greece. Recent perspectives on temporary river ecology. Aquat. Sci. 73, 581–597.

Soden, B.J., Held, I.M., 2006. An assessment of climate feedbacks in coupled ocean-atmosphere models. J. Clim. 19 (14), 3354–3360 (Formularbeginn Formularende).

Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2012. An Overview of CMIP5 and the Experiment Design. Bull. Am. Meteorol. Soc. 93, 485–498.

Telesca, L., Vicente-Serrano, S.M., López-Moreno, J.I., 2012. Power spectral characteristics of drought indices in the Ebro river basin at different temporal scales. Stoch. Env. Res. Risk A. 27 (5), 1155–1170.

Torma, C., Giorgi, F., Coppola, E., 2015. Added value of regional climate modeling over areas characterized by complex terrain—precipitation over the Alps. J. Geophys. Res. 120 (9), 3957–3972.

Van Vuuren, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G.C., Kram, T., Krey, V., Lamarque, J.-F., 2011. The representative concentration pathways: an overview. Clim. Chang. 109, 5–31.

Vial, J., Dufresne, J.L., Bony, S., 2013. On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates. Clim. Dyn. 41 (11—12), 3339–3362.

2.4 Paper IV: Applying the Triangle Method for the parameterization of irrigated areas as input for spatially distributed hydrological modeling - Assessing future drought risk in the Gaza Strip (Palestine). Science of the Total Environment

Gampe, D., Ludwig, R., Qahman, K., & Afifi, S. (2016). *Applying the Triangle Method for the parameterization of irrigated areas as input for spatially distributed hydrological modeling—Assessing future drought risk in the Gaza Strip (Palestine).* Science of the Total Environment, 543, 877-888.

Paper IV was the outcome of a hydrological modeling study in the Gaza Strip carried out within the project CLIMB funded under the European Union's Seventh Programme. The publication highlights the challenges for hydrological climate change impact modeling studies in data scarce areas and serves as an example for the typical hydro-climatic modeling chain described in section 1.2. In the context of this study an R-package was compiled to derive evapotranspiration patterns from satellite imagery based on the triangle method approach elaborated by Jiang and Islam (1999). These patterns served to improve the parameterization of the irrigated areas in the hydrological model. The main motivation of this paper is to provide an as robust as possible assessment of future changes on the hydrology of a region under data scarcity. The study area, the Gaza Strip, was selected as it is used intensively for agriculture irrigated with ground water from the regional aquifer. A robust assessment of future drought risk and an estimation of climatic pressure on the ground water reserves is thus of high relevance. This study addresses the following research questions:

Q5: How can remote sensing data be applied to improve the robustness of a hydrological model while being computational efficient?

Q6: Is the Mediterranean region likely to experience increased pressure on the water balance and thus likely to be exposed to increased water scarcity under future climate?

Author's contributions: The concept of this study was developed by D. Gampe & R. Ludwig, while D. Gampe was responsible for the implementation and detailed work flow under the supervision of R. Ludwig. The hydrological modeling was carried out by D. Gampe. The first version of the manuscript was written by D. Gampe, while all authors contributed with valuable input and expertise to the final document.
Status: published
Journal: Science of the Total Environment (Elsevier)
Impact factor: 4.900

Science of the Total Environment 543 (2016) 877–888

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Applying the Triangle Method for the parameterization of irrigated areas as input for spatially distributed hydrological modeling — Assessing future drought risk in the Gaza Strip (Palestine)

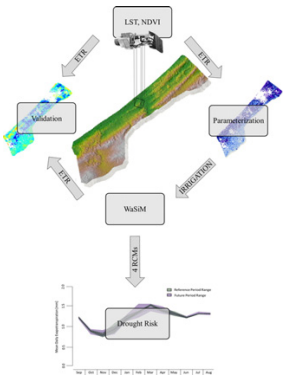
David Gampe [a,*], Ralf Ludwig [a], Khalid Qahman [b], Samir Afifi [b]

[a] Ludwig-Maximilians-Universität Munich, Department of Geography, Luisenstr. 37, 80333 Munich, Germany
[b] Islamic University of Gaza, Department of Civil Engineering, Al Talatiny St., P.O. Box 108, Gaza, Palestine

HIGHLIGHTS

- RCMs reveal increase in temperature & decrease in precipitation in the Mediterranean
- Climate change increases drought risk in the Gaza Strip
- Remotely sensed evapotranspiration used for parameterization and validation of WaSiM
- Future evapotranspiration index can be used as a robust indicator for drought risk
- Agricultural productivity cannot be maintained without adapting irrigation management

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 30 January 2015
Received in revised form 20 July 2015
Accepted 21 July 2015
Available online 14 August 2015

Editor: D. Barcelo

Keywords:
Climate change
Drought risk
Gaza
Remote sensing
Hydrological modeling
WaSiM
Triangle Method

ABSTRACT

In the Mediterranean region, particularly in the Gaza strip, an increased risk of drought is among the major concerns related to climate change. The impacts of climate change on water availability, drought risk and food security can be assessed by means of hydro-climatological modeling. However, the region is prone to severe observation data scarcity, which limits the potential for robust model parameterization, calibration and validation. In this study, the physically based, spatially distributed hydrological model WaSiM is parameterized and evaluated using satellite imagery to assess hydrological quantities. The Triangle Method estimates actual evapotranspiration (ETR) through the Normalized Difference Vegetation Index (NDVI) and land surface temperature (LST) provided by Landsat TM imagery. So-derived spatially distributed evapotranspiration is then used in two ways: first a subset of the imagery is used to parameterize the irrigation module of WaSiM and second, withheld scenes are applied to evaluate the performance of the hydrological model in the data scarce study area. The results show acceptable overall correlation with the validation scenes ($r = 0.53$) and an improvement over the usual irrigation parameterization scheme using land use information exclusively. This model setup is then applied for future drought risk assessment in the Gaza Strip using a small ensemble of four regional climate projections for the period 2041–2070. Hydrological modeling reveals an increased risk of drought, assessed with an evapotranspiration index, compared to the reference period 1971–2000. Current irrigation procedures cannot maintain the agricultural productivity under future conditions without adaptation.

© 2015 Elsevier B.V. All rights reserved.

* Corresponding author.
E-mail address: david.gampe@lmu.de (D. Gampe).

http://dx.doi.org/10.1016/j.scitotenv.2015.07.098
0048-9697/© 2015 Elsevier B.V. All rights reserved.

1. Introduction

According to current climate projections, the Mediterranean region will be affected by severe changes, both an increase in temperature as well as a decline in precipitation. The projections of climate models show a decline up to over 50% in the summer months, resulting in a rising frequency of drought events up to over 40% (Christensen et al., 2007).

Hydrological modeling can be applied to quantify these effects, provide decision support and develop effective water management strategies for the future. Usual procedures in hydrological model calibration range from the classical trial and error approach to sophisticated automized learning algorithms to fit a modeled output variable, typically discharge, to observations (Gupta et al., 1998). Statistical criteria, like the Root Mean Square Error or the Nash–Sutcliffe criterion (NSE) (Nash and Sutcliffe, 1970) are applied to validate the goodness of the model results. However, such approaches can be somewhat difficult for complex models or in heterogeneous catchments (Immerzeel and Droogers, 2008). Furthermore, if data for calibration and validation are scarce, this approach will add hardly quantifiable uncertainty to the results (Winsemius et al., 2009). The Predictions in Ungauged Basins (PUB) initiative (Sivapalan et al., 2003) increased the discussion on the work in ungauged basins (Winsemius et al., 2009). Possible solutions to calibrate a hydrological model in a data scarce area might be to apply the information of neighboring catchments as presented by Blöschl (2005), or to perform regional calibration by calibrating several catchments simultaneously (Parajka et al., 2007). However, both approaches imply knowledge of surrounding catchments, hence might not be applicable in many ungauged, data scarce areas.

Remotely sensed data can help find a solution to this problem as those data sets provide spatial information in an acceptable temporal resolution, which can be translated to hydrological variables. Schmugge et al. (2002) show general application of remote sensing to extract information on snow distribution, soil moisture and also water quality, Kite and Pietroniro (1996) provide an overview on this topic, Schultz (1993) used multispectral Landsat data to assimilate information for distributed hydrological modeling already two decades ago. Data assimilation for the parameterization and calibration of hydrological models is a more recent research field (Immerzeel and Droogers, 2008). Various studies assimilate remotely sensed soil moisture information in several hydrological applications (Boegh et al., 2004; Al-Shrafany et al., 2014).

Studies by Chen et al. (2005) and Immerzeel and Droogers (2008) focus on evapotranspiration, which tackles the traditional calculation of actual or real evapotranspiration (ETR) as an estimated fraction of potential evapotranspiration (ETpot) (Kite and Droogers, 2000). The latter compare evapotranspiration estimated from satellite information, hydrological models and field methods. Their results confirm the use of a remote sensing solution to estimate evapotranspiration, as the traditional field methods, FAO-27 and FAO-56 (Allen et al., 1998) showed ambiguous results. The two most widely used approaches to derive ETR from remote sensing are the Surface Energy Balance Algorithm for Land (SEBAL) and the Triangle Method.

SEBAL converts visible, near-infrared and thermal information to an estimate of evapotranspiration Bastiaansen et al. (1998a,b). The algorithm was applied for Landsat TM images by Kite and Droogers (2000), while Immerzeel and Droogers (2008) used MODIS information to derive evapotranspiration for the calibration of the Soil and Water Assessment Tool (SWAT) (Arnold et al., 1998).

In this study, another method to estimate evapotranspiration by remote sensing data is applied to parameterize the irrigation module hydrological model WaSiM (Schulla and Jasper, 2007) and to evaluate model performance. The Triangle Method, as presented by Price (1990), then established by Carlson et al. (1995) and Jiang and Islam (1999), was carried out for Landsat TM scenes to estimate the ETR for the Gaza strip. This method was chosen over SEBAL for reasons of required input data and will be briefly elaborated in Section 3. As water availability is the limiting factor for ETR, areas with high ETR during the dry summer months therefore have to be irrigated. Several of these Landsat TM scenes were then used to identify the irrigation areas and parameterize the model to distribute the amount of irrigated water over the catchment.

The objectives of this study are threefold: a) the Triangle Method was used to provide estimates of ETR to parameterize the irrigation module of WaSiM, as only irrigated areas show high evapotranspiration during the summer months. The procedure is presented in Section 3.3 and evaluated in Section 4.1; b) for lack of discharge data the performance of WaSiM will be assessed on the withheld scenes of the satellite derived evapotranspiration in Section 4.2; c) the WaSiM model setup is then driven with a small ensemble of regional climate model (RCM) simulations to assess future drought risk in the area as presented in Section 4.3. For a detailed analysis on the RCM-ensemble selection it is hereby referred to Deidda et al. (2013).

2. Study area

The Gaza Strip is located in the Eastern Mediterranean and forms, together with the West Bank, the Palestinian Autonomous Area, according to the Oslo agreement of 1993. The area covers 365 km^2 with a length of 35 km and width of 6 to 12 km (Baalousha, 2006). Sufficient supply of freshwater is a major concern for the rapidly growing population, which stands at approximately 1.5 million inhabitants and grew by 4.5% per year in the period 1997–2007. One third of the Gaza Strip is covered by urban or built up area, the largest city in the area being Gaza City with about 0.5 million inhabitants (Ajluni, 2010). Due to population growth, the total water demand in the Gaza Strip is strongly increasing. The current available resources do not satisfy the need of water, causing a huge deficit between water demand and supply (Qahman and Larabi, 2006).

Agriculture is the most important sector in terms of land cover, land use and water consumption in the Gaza Strip (Rusteberg et al., 2010). The main crops grown on the irrigated fields include tomatoes, potatoes, cucumbers, strawberries and melons. Furthermore, citrus plantations and olive orchards are widespread in the area. The climate conditions allow for more than one crop cultivation cycle per year.

The mean annual temperature for Gaza City is 20.1 °C and the mean annual precipitation 353 mm, resulting from events in the winter months, during the wet period October to April. The months May to September form the dry period, practically without any precipitation. A remarkable north-south gradient in precipitation is evident with 435 mm in the North, and 235 mm in the South as shown in Fig. 1. Differences in elevation are almost negligible as they range from sea level to 104 m. In the Gaza Strip no permanent surface water exists in form of streams and natural lakes. Only the Wadi Gaza could provide surface water during the winter months. The bed of the Wadi is characterized by a 1–2 m layer of unsorted Pleistocene gravel, originating from the mountainous areas in the Northern Negev and Hebron (MedWetCoast Project, 2001).

The major source of freshwater in the area is the coastal aquifer. This aquifer, hereafter referred to as Gaza Aquifer, covers a large area of about 2000 km^2, from the Carmel Mountains in the North to the Sinai Desert in the South with a width of 15–30 km (Baalousha, 2008). The aquifer provides freshwater for the entire Gaza Strip, and parts of Israel, including the metropolis of Tel Aviv–Jaffa in the north. The depth of the aquifer varies from 170 m at the Mediterranean coastline to just a few meters in the eastern parts. Main components of the aquifer are alluvial sandstone, with local limestone and chalk areas, underlain by a massive, impermeable clay layer, the Saqiya formation, developed in the Pliocene era, with depths of 400 to 1000 m (Baalousha, 2006). Older groups, like the Judea, Kurnub and Arad Group contain dolomite and sandstone and developed between the Jurassic and Cretaceous era (Assaf et al., 1998).

As mentioned before, data scarcity is hampering the calibration and validation of the hydrological model. Fig. 1 shows the location of the twelve precipitation gauges available within the Gaza Strip. However,

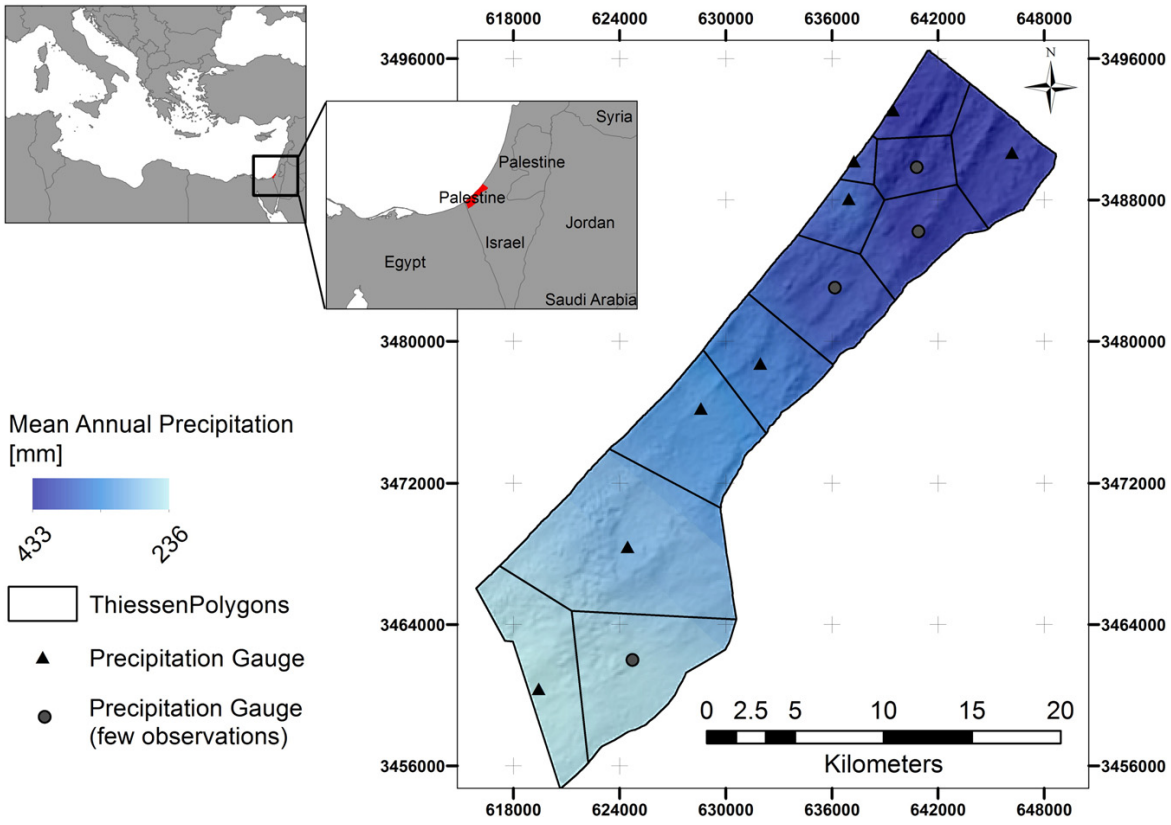


Fig. 1. Location of the Gaza Strip and the mean annual precipitation derived from inverse distance interpolation from given precipitation gauges. The Thiessen Polygons shown represent the subcatchments as implemented in the hydrological model. Projection: WGS-84, UTM zone 36 N.

four of those, marked as dark gray circles in the map, provide only a few years of observations. For the eight stations marked as triangles in Fig. 1 observations are available at daily scale for the period 1979 to 2003; the station in Gaza City provides data for the period 1973 to 2006. This is also the only station with temperature observations available (minimum and maximum at daily scale) for the period 1973 to 2006. As no permanent streamflow exists, no discharge gauge is existent, neither are observations for other meteorological variables. WaSiM – if run with Penman–Monteith evapotranspiration scheme – requires further variables as meteorological input such as shortwave downward radiation, relative humidity and mean wind speed. Therefore, a RCM run was chosen for the modeling procedure. The model was run with the bias-corrected and downscaled version of the RCA3 driven by ECHAM5 while the precipitation and temperature observations served as plausibility check for the reference period. The Thiessen Polygons are presented in Fig.1 as they are implemented in WaSiM as subcatchments to compare the RCM outputs with the observed precipitation.

A land use/land cover map was derived from classified multi-temporal SPOT-5 images, while a detailed soil map, stemming from intensive field campaigns, was provided by the Islamic University of Gaza. An overview of the Landsat TM scenes assessed for this study can be found in Table 1.

3. Method and hydrological model

3.1. The Triangle Method

The Triangle Method, originally published by Price (1990) then elaborated by Carlson et al. (1995) and Jiang and Islam (1999), is a widespread approach to estimate the latent heat flux λE or ETR (Batra et al., 2006; Stisen et al., 2008; Tang et al., 2010; Wang et al., 2006; Yiang and Wang, 2011). Various studies showed acceptable deviations from point-measurements of evaporation rates of +/−10–30% (Kalma et al., 2008).

This approach was chosen over the already mentioned SEBAL approach for reasons of data availability. SEBAL requires additional meteorological variables, such as relative humidity and wind speed, which were not available for this study site. Additionally, other surface parameters such as the Leaf Area Index (LAI) and albedo, are needed. These are assessed from remote sensing data for the SEBAL approach; however they add another source of uncertainty.

The aim of the Triangle Method is to estimate the evaporative fraction (EF) and λE of the energy balance. The basic idea is that high EF results in a pronounced cooling of the area and sunlit vegetation will therefore be cooler than bare soil under equal conditions (Carlson, 2007). ETR then was derived from remotely sensed thermal information

Table 1
Applied Landsat TM scenes (175/38 path/row) for parameterization and validation in this study.

Parameterization	Validation
1984/05/24	1998/06/16
1992/08/10	1998/07/02
1999/08/22	2000/06/05
2000/05/04	2002/06/11
2000/09/09	
2002/05/26	
2003/08/01	
2003/09/02	

and the Normalized Difference Vegetation Index (NDVI). As all components of the energy balance are dependent on the land surface temperature (LST) (Kalma et al., 2008), it is obvious to estimate LST from remote sensing data. Landsat TM data was used, as proposed by Roerink et al. (2000), due to relatively high spatial resolution and data availability. Other authors use NOAA AVHRR data (Jiang and Islam, 1999), MODIS data (Tang et al., 2010) or MSG-SEVIRI images (Stisen et al., 2008) for their studies.

In a first step, the radiances of the Landsat TM thermal band are converted to LST, so that each pixel has an assigned temperature in °C or K. This data was used to construct the LST–NDVI space, shown in Fig. 2, plotting LST, derived from the remote sensing images, versus the NDVI.

Priestley and Taylor (1972) proposed to substitute the resistance terms of the Penman equation by an empirical coefficient $_{\Phi}$, often also cited as α (Batra et al., 2006; Stisen et al., 2008):

$$\lambda E = \Phi\left[(Rn-G)\frac{\Delta}{\Delta+\gamma}\right] \quad (1)$$

where λE is the evapotranspiration (W m^{-2}), Φ the dimensionless Priestley–Taylor coefficient, Rn the net radiation (W m^{-2}), G the soil heat flux (W m^{-2}), Δ the slope of the saturated water vapor (kPa K^{-1}) and γ the psychrometric constant (kPa K^{-1}). According to Davies (1967), on the global scale, net radiation is about 55% of the global solar radiation. The soil heat flux G was estimated to be approximately seven per cent of the net radiation, following Crago and Brutsaert (1996). Fig. 2 shows the NDVI–LST plot for the image of May 4th 2000, where the blue line, the wet edge, represents ETpot, where ϕ_{max} is equal to ϕ. The maximum rate of ETR therefore can be found near ϕ_{max}, and the lowest evapotranspiration near ϕ_{min}. The gray line, the true dry edge, represents zero ET, while the red line, observed dry edge, represents limited evapotranspiration, hence ϕ_{min} (Stisen et al., 2008).

For detailed information on the concept and construction of the parameter ϕ, it is hereby referred to Priestley and Taylor (Priestley and Taylor, 1972). In the next step, the NDVI was used to calculate and scale ϕ between $^{\Phi}_{min}$ and Φ_{max}; while Batra et al. (2006) propose a linear interpolation between $NDVI_{min}$ and $NDVI_{max}$, Stisen et al. (2008) and Tang et al. (2010) propose a non-linear scaling, to increase ETR over higher vegetated areas compared to less vegetation-covered areas:

$$\Phi(i,\ min) = \Phi_{max}\left[\frac{NDVI_i - NDVI_{min}}{NDVI_{max} - NDVI_{min}}\right]^2 \quad (2)$$

where $\Phi_{(i,\ min)}$ is the minimum Φ for the grid point I, Φ_{max} is 1.26, $NDVI_i$ the NDVI for the vegetation class (NDVI value) i, $NDVI_{min}$ and $NDVI_{max}$ represent the minimum and maximum NDVI for the current scene.

The so scaled $\Phi_{(i,min)}$ was then used for calibration of the actual Priestley–Taylor coefficient for each pixel with the LST following the approach of Stisen et al. (2008) and Tang et al. (2010) to spread the NDVI–LST space according to the minimum (at $LST_{(i,max)}$ and $\Phi_{(i,min)}$) maximum evaporative cooling (at $LST_{(i,min)}$ and $\Phi_{(i,max)}$):

$$\Phi_i = \frac{LST(i,\ max) - T_i}{LST(i,\ max) - LST_{(i,\ min)}}(\Phi_{max} - \Phi(i,\ min)) + \Phi(i,\ min) \quad (3)$$

Here Φ_i represents the Priestley–Taylor coefficient for the NDVI value i, $LST_{(i,\ max)}$ and $LST_{(i,min)}$ the corresponding minimum and maximum LST and T_i the observed air temperature at I for the overpassing time. Batra et al. (2006) and Wang et al. (2006) propose only the use of $^{\Phi}_{max}$ instead of the term $(^{\Phi}_{max} - \Phi_{(i,min)}) + \Phi_{(i,min)}$. However, a sensitivity analysis (not presented here), of the various approaches

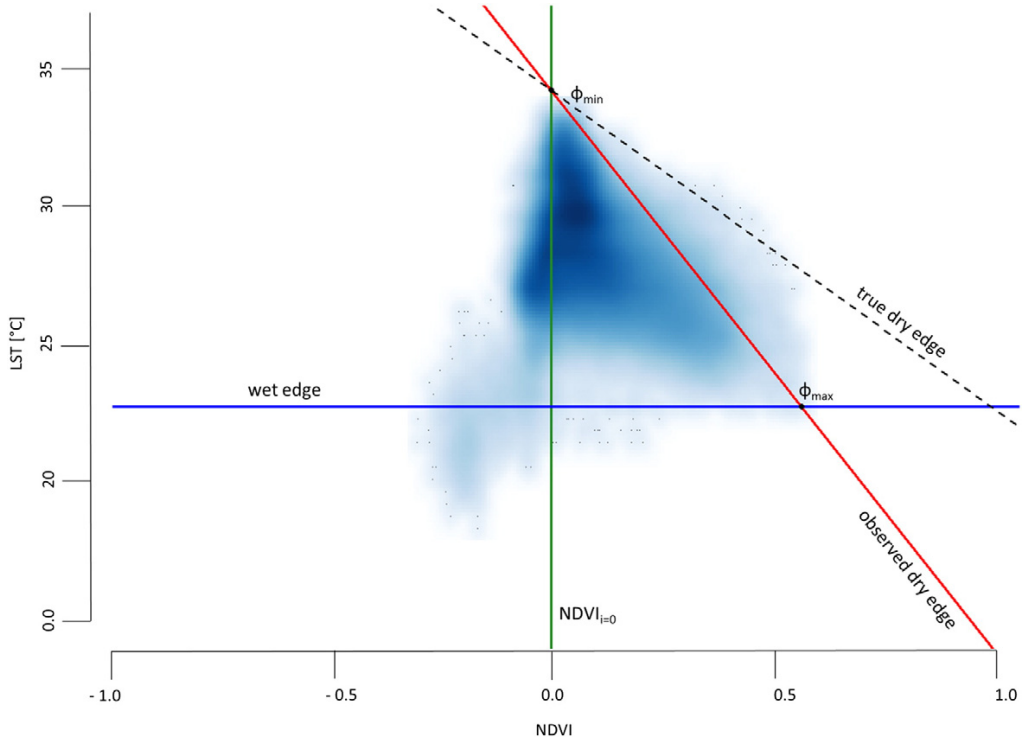


Fig. 2. Concept of the Triangle Method and the obtainment of the NDVI–LST plot as applied to the Landsat TM image of 2000/05/04 (for the Gaza Strip). Green line represents $NDIV_i$ for the vegetation class i = 0, the red line the observed dry edge, standing for limited ETR, the blue line the observed wet edge, representing potential evapotranspiration. Conception of the plot according to Batra et al. (2006); Stisen et al. (2008).

presented by different authors revealed only slight differences in resulting ETR.

In a final step, EF can be determined using the psychrometric constant and the atmospheric pressure:

$$EF = \Phi \frac{\Delta}{\Delta + \gamma} \quad (4)$$

λE can now be calculated in Wm^{-2}. To make the resulting ETR comparable to model outputs, a conversion to mm d^{-1} has to be applied as presented by Allen et al. (1998), where 1 W m^{-2} represents 0.0864 MJ m^{-3} d^{-1} which equals 0.408 mm d^{-1}.

3.2. The hydrological model WaSiM

The runoff and WAter balance SImulation Model (WaSiM), was developed by Schulla in 1997 and is continually updated. At the moment, two versions of the model exist: the early one using the TOPMODEL-approach for calculation of soil water flow, while the second and here applied version uses the Richards equation. WaSiM is distributed, deterministic and physically based (Schulla, 2012) and can be run in any spatial and temporal resolution, limited only by the input grids. In this application, the temporal resolution was set to one day, according to the temporal resolution of the meteorological forcing. Due to the small size of the Gaza Strip a 100 m spatial resolution was chosen, as it roughly matches the resolution of the used SRTM DEM (90 m). Simulations were carried out for two periods: a reference period from 1971 to 2000 and a future period 2041–2070 to assess future changes in the water balance and drought risk.

In WaSiM various schemes for the calculation of ETpot are implemented, such as Hamon, Haude and Wendling, however, Schulla (2012) recommends to apply the Penman–Monteith approach. To comply with this recommendation, the simulation was driven by the bias-corrected and downscaled version of the RCA3 forced by ECHAM5 as stated in Section 2. The ensemble for the forcing to assess future changes consists of ECHAM5/MPI OM forcing REMO (ECH-REM), RACMO2 (ECH-RMO) and RCA3 (ECH-RCA) additionally, also HadCM3 forcing the RCA3 (HCH-RCA). For details on the selection and bias-correction of these simulations it is referred to Deidda et al. (2013).

The structure of WaSiM allows the user to switch on or off several modules, according to the modeling needs in the specific case. The provided irrigation module is of utmost importance in this study and will therefore be described briefly.

Two different irrigation types can be implemented in WaSiM: either a daily quota for each pixel taken by groundwater or surface water based on an irrigation table, or irrigation in dependence on the actual soil water content of the given time step corresponding to a user defined threshold. The daily quota approach was chosen in this study and the parameterization of the irrigation table is defined in the next section. The resulting amount of irrigated water then is simply added to the precipitation for each day within the model. Additionally, soil and land use tables are required, containing the respective parameters. Based on these tables and parameters, ETR is calculated in dependence of the provided ETpot. In the following, the parameterization of the main modules is presented briefly, while for detailed model description it is hereby referred to Schulla (2012).

3.3. The parameterization strategy

Soil parameters were parameterized according to literature and assigned to the soil texture classes of a detailed digital soil map. Relevant vegetation parameters as the leaf area index (LAI), root depth or stomatal resistance were also derived from literature (Delalieux et al., 2008; Gómez-del-Campo, 2007; Raveh et al., 2003; Reichenstein et al., 2003; Saei et al., 2006; Sakcali and Ozturk, 2004; Villalobos et al., 2006) and assigned to the corresponding classes in the land use map (most importantly horticulture, olives, citrus and mixed agriculture).

Irrigation is a key factor for freshwater consumption in a highly agriculturally used semi-arid region such as the Gaza Strip. The annual irrigation quota for the year 2006 of 85.5 Mm^3, as reported by the Palestinian Ministry of Agriculture and the Palestinian Water Authority (PWA) Palestinian Water Authority (PWA) (2007), was implemented. It must be considered, however, that the real irrigation might exceed this number due to illegal wells and unregistered exceedance of the allowed quota. Therefore, the daily irrigation method (see Section 3.2) seemed more feasible than irrigation on demand, as the quota is usually applied in any case.

Two sets of irrigation scenarios were constructed: a) the 'classical' approach using the land use/land cover map to determine irrigated agriculture and b) using satellite imagery to derive evapotranspiration patterns and reconstruct the irrigated areas.

For a), scenarios were constructed to meet the land cover information derived from SPOT images and the different growing seasons as multiple crops can be grown at one field in the same year, and these cycles might well change over the years. A possible irrigation scenario contains e.g. 30% melon, 50% vegetables and 20% strawberries and other fruits. Comparison of ETR patterns – theoretically, each pixel of the classes 'horticulture', 'mixed agriculture', 'citrus', 'greenhouses' and 'olives' can be assumed as irrigated areas, however ETR shows different patterns – then leads to the parameterization of b).

The irrigation input for b) is based on the evapotranspiration patterns derived from the Triangle Method to distribute the irrigated water over the catchment. To determine the irrigated fields, satellite scenes were selected from May to September, when rain fed agriculture has come to a close as these months characterize the dry period. One third of the imagery, the months June and July, were used for validation purpose. These validation months were chosen to permit a validation period centered in dry months used for the parameterization before and after the validation period. A total of twelve cloud-free Landsat TM scenes (175/38 path/row, Table 1) met the criteria in and around the reference period 1971–2000.

The mean evapotranspiration in mm derived from these scenes was normalized for the parameterization procedure. The resulting irrigation input map, shown in Fig. 3, will be further analyzed in Section 4.2 and the remaining validation scenes compared to WaSiM results in Section 4.3.

4. Results

4.1. Parameterization of the irrigation input

The derived mean evapotranspiration grids for parameterization (see Table 1) were then compared with the WaSiM results for the corresponding months. Fig. 4 presents the absolute differences for the parameterization months (May, August and September) expressed as absolute difference [mm] for Triangle derived ETR – WaSiM-ETR. In this figure, the Triangle Method derived ETR represents the mean ETR derived by all Landsat TM scenes for parameterization while the WaSiM modeled ETR is calculated as long term daily mean evapotranspiration for the parameterization months. White areas in the plot represent urban areas, which were masked out, as they were also excluded in the parameterization file for irrigation. The deviations for the central part of the catchment are within an acceptable range of 1–2 mm and correspond to the patterns of the satellite derived image, which was the basic aim of this approach. This range is not constant over the entire catchment, causing high deviation in the southeast leading to an acceptable overall correlation ($r = 0.53$).

The bias of WaSiM modeled and satellite derived ETR stems from the fact that the satellite imagery used for this application represents clear-sky conditions and high radiation days only, due to the optical sensor. The resulting ETR of WaSiM is thus preferable lower than the Triangle

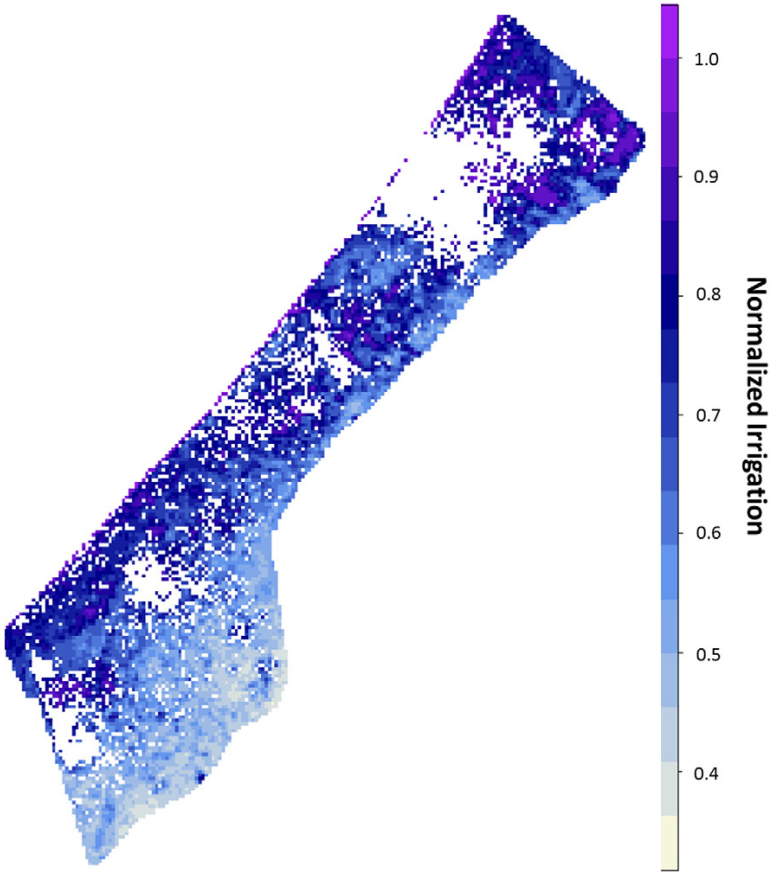


Fig. 3. Resulting irrigation parameterization input file according to the Triangle Method parameterization. White areas represent urban areas, which were masked out, as they do not represent irrigated areas.

Method derived. In this context, a strong statistical assessment of the absolute ETR in mm is not meaningful.

When compared to the study of Immerzeel and Droogers (2008), the correlation here is slightly lower. The reasons for this are twofold: a) in contrary to their study, WaSiM was not calibrated with the evapotranspiration patterns, but they were only used for parameterization. The reason for this is the lack of data to perform a standard calibration on a physically based model as WaSiM in the case of the Gaza Strip. And b) in the study presented here, several Landsat TM scenes were acquired to serve as a mean reference to be comparable with a long-term mean of climate model driven results. Considering this, the correlation ($r = 0.53$) deemed to be acceptable and the parameterization of WaSiM constitutes a satisfying setup. Significance of the correlation was tested successfully for the 95% confidence interval using Pearson, Spearman's rank and Kendall's rank.

4.2. Evaluation of model performance using ETR patterns

The months June and July were left out for validation purposes to compare WaSiM modeled ETR with the patterns derived from the Triangle Method. For these validation months, the deviations between WaSiM and Triangle derived ETR increased slightly throughout the catchment. The patterns of the WaSiM and triangle modeled ETR agree in the same way as for the calibration period for the validation months (Fig. 6 left). When aggregated to subcatchment scale, as presented in Fig. 5, the correlation increases ($r = 0.85$) and, although absolute values differ considerably, the patterns match satisfactorily. This correlation agrees with the results presented by Immerzeel and Droogers (2008). The bias in the absolute ETR in mm is explained in the previous section.

For a general validation of the approach, WaSiM was set up with three different irrigation schemes: a) the here presented approach using the Triangle Method as parameterization of the irrigation module, b) the classical approach of distributing the total amount of irrigated water over the (potentially) irrigated land use classes according to a land use classification and c) by applying no irrigation at all. The results are presented in Fig. 6 for the validation months June and July as long-term daily difference in [mm] from the triangle derived ETR patterns for these months. To better assess the patterns here and for visualization reasons, each pixel is based on a 5×5 pixel mean. Not surprisingly, the deviations for the non-irrigation setup are extremely high (Fig. 6, right), with a mean deviation of 3.24 mm compared to the triangle results. No specific patterns are identifiable for this setup and ETR is spread evenly over the entire catchment.

This applies partly also for the classical parameterization (center plot in Figs. 6 and 7), although the mean difference compared to the triangle validation ETR is considerably lower in this case as presented in Fig. 7. The triangle approach decreases the mean deviation further, to 1.95 mm (left plot in Figs. 6 and 7). This is not surprising, as the same method was used for the parameterization in this particular setup. Maximum differences occur at the boundaries of the catchment in the northwest, which are classified as “sand” in WaSiM, leading to extremely low ETR as the land use table was maintained for either irrigation setup. The overall mean for ETR is comparable for the classical parameterization

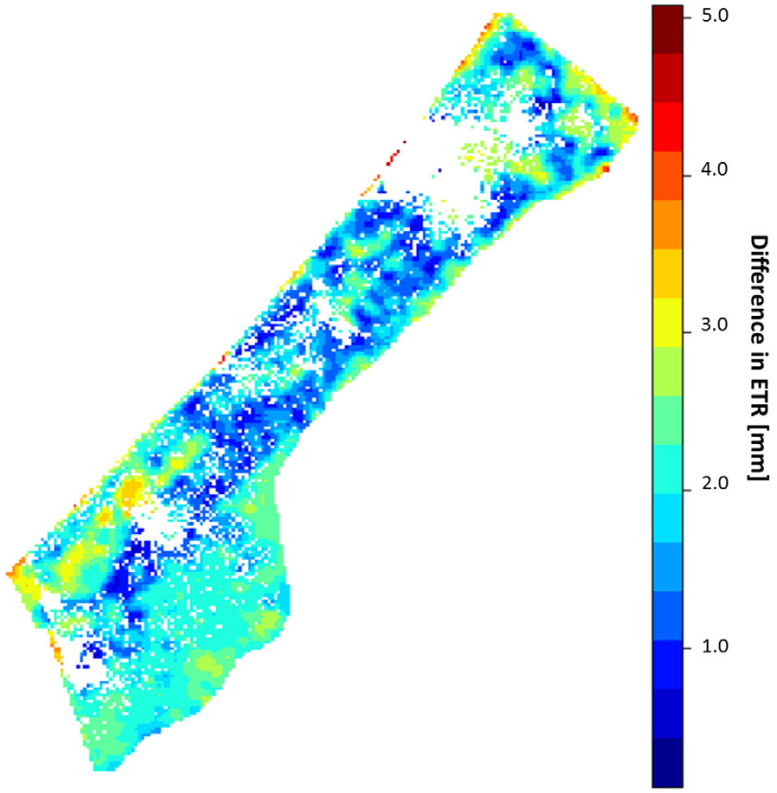


Fig. 4. Deviation for the calibration scenes for Triangle derived ETR – WaSiM modeled ETR.

and the triangle approach with around 1.41 mm per day for the months of June and July. However, the patterns for ETR follow the irrigation input of the Triangle Method, leading to low differences and high ETR in the central part of the Gaza Strip and in general a better representation, which is furthermore reflected by a lower standard deviation in Fig. 6. Reasons for this shift of 1.5–2 mm (Fig. 6) stem from the method itself, as an optical sensor was used to estimate ETR. The resulting evapotranspiration patterns are clearly biased, as only

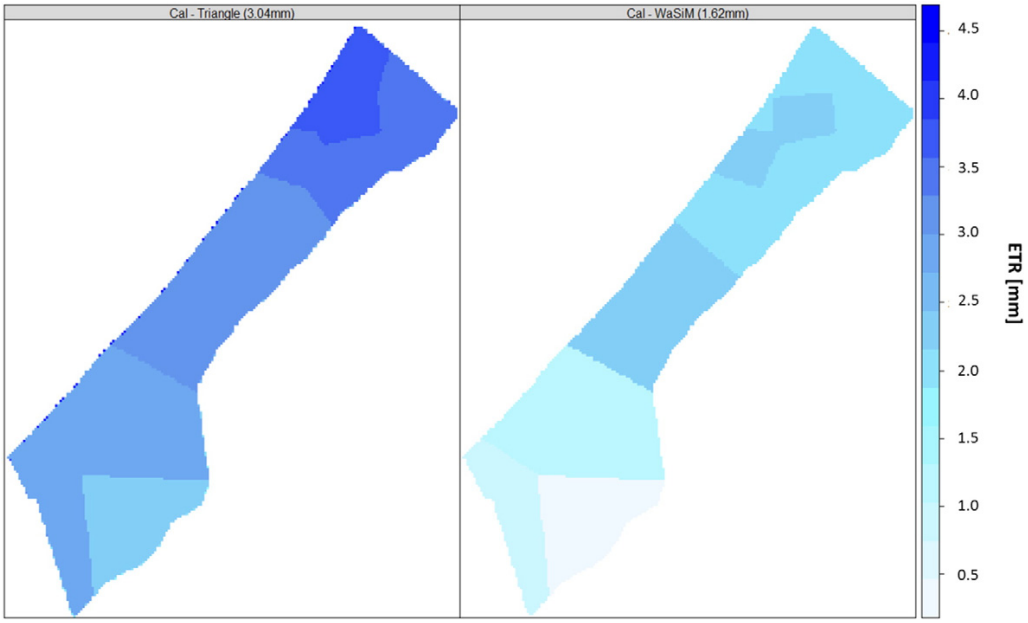


Fig. 5. ETR results per subcatchment for the calibration months for the Triangle Method (left) and WaSiM (right).

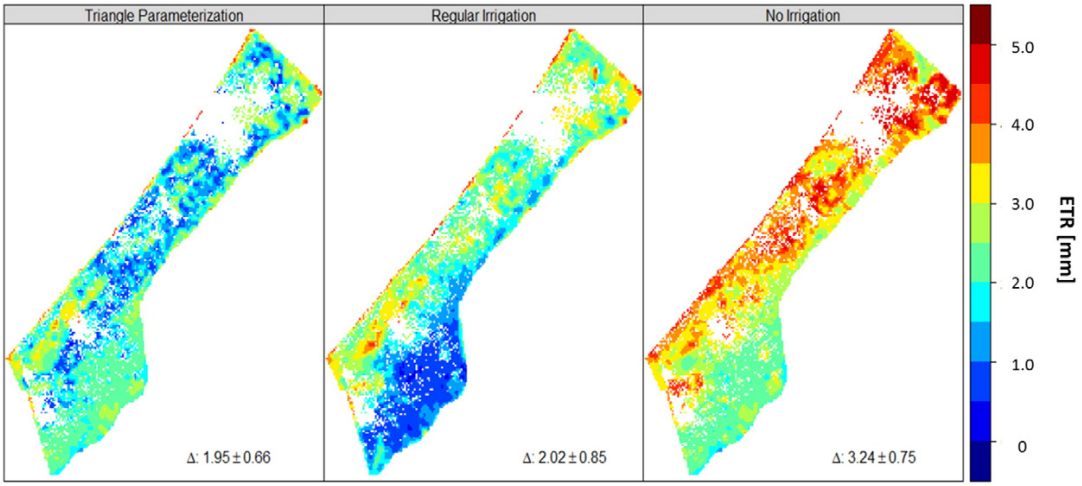


Fig. 6. Differences in ETR to triangle validation according to three WaSiM setups: The one with the irrigation input scaled according to the Triangle Method, with the irrigation patterns as in Fig. 3 (left), with regular irrigation, or the classical approach, meaning distributing the total irrigated water over all (potentially) irrigated land use classes (center) and finally the base line run without irrigation at all (right). Here, the long-term daily mean deviations for the validation months are shown as difference of triangle ETR – WaSiM ETR. Numbers represent mean difference from the triangle ETR and corresponding standard deviations. White areas represent urban areas, which were masked out here.

cloud-free high radiation days can be taken into account. A total match of ETR values can thus not be expected and would not even be desired. The correlation for the Triangle Method validation ETR is assessed in the Taylor diagram (Taylor, 2001) in Fig. 8.

The green circles represent the RMSE as compared to the validation ETR (dark green square), blue circles indicate the standard deviation for each grid, and the black lines correspond to the correlation with the validation ETR. At around 1 mm, RMSE is equally high for the classical irrigation (green dot) and the triangle derived (red dot), while slightly lower for the non-irrigated WaSiM simulation (blue dot). However, over- and underestimations throughout the catchment are balancing the RMSE. The correlation for the triangle irrigation is within an acceptable overall range ($r = 0.53$). Slightly negative correlations show the mismatch of the other two irrigation schemes, as already shown in Figs. 6 and 7.

4.3. Future drought risk Gaza — WaSiM results

As the correlation and deviation for both calibration and validation months showed acceptable results, the triangle derived irrigation scheme was chosen as input to assess future changes in the Gaza Strip. For this purpose, four different RCM–GCM combinations were chosen to provide climate scenarios. A detailed analysis of the performance of each RCM is given in Deidda et al. (2013).

Table 2 gives an overview on the changes projected by WaSiM for the small ensemble of four RCM combinations applied. The ECHAM5 driven RCMs agree in a decline of precipitation of 9–17%, resulting in a slight decrease of ETR of 1–3%. ETpot is likely to increase, as all four simulations agree in an increase of 5–10%. Runoff (Q) is projected to decrease by 8–26% depending on the simulation.

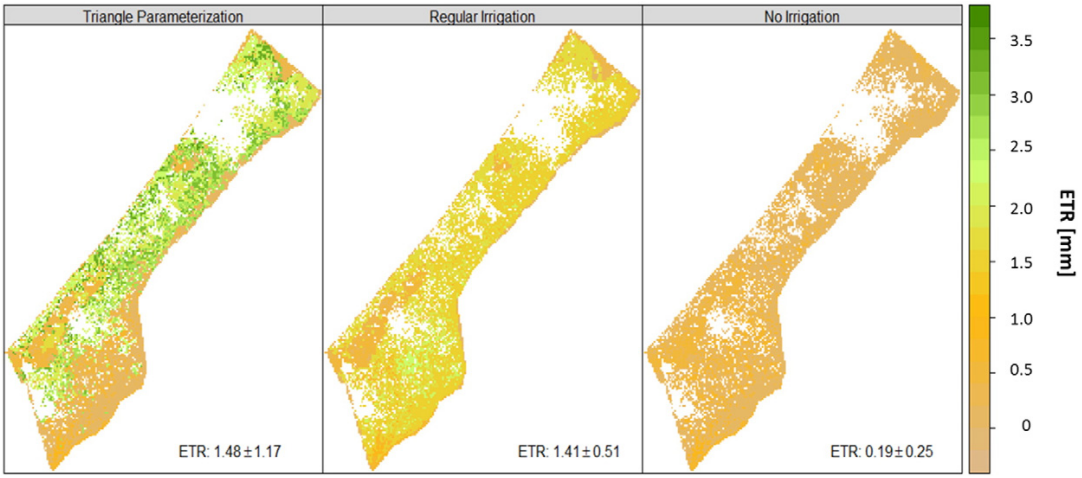


Fig. 7. Absolute ETR according to three WaSiM setups: The one with the irrigation input scaled according to the Triangle Method, with the irrigation patterns as in Fig. 3 (left), with regular irrigation, or the classical approach, meaning distributing the total irrigated water over all (potentially) irrigated land use classes (center) and finally the base line run without irrigation at all (right). Numbers represent long-term daily mean ETR for the validation months and corresponding standard deviations. White areas represent urban areas.

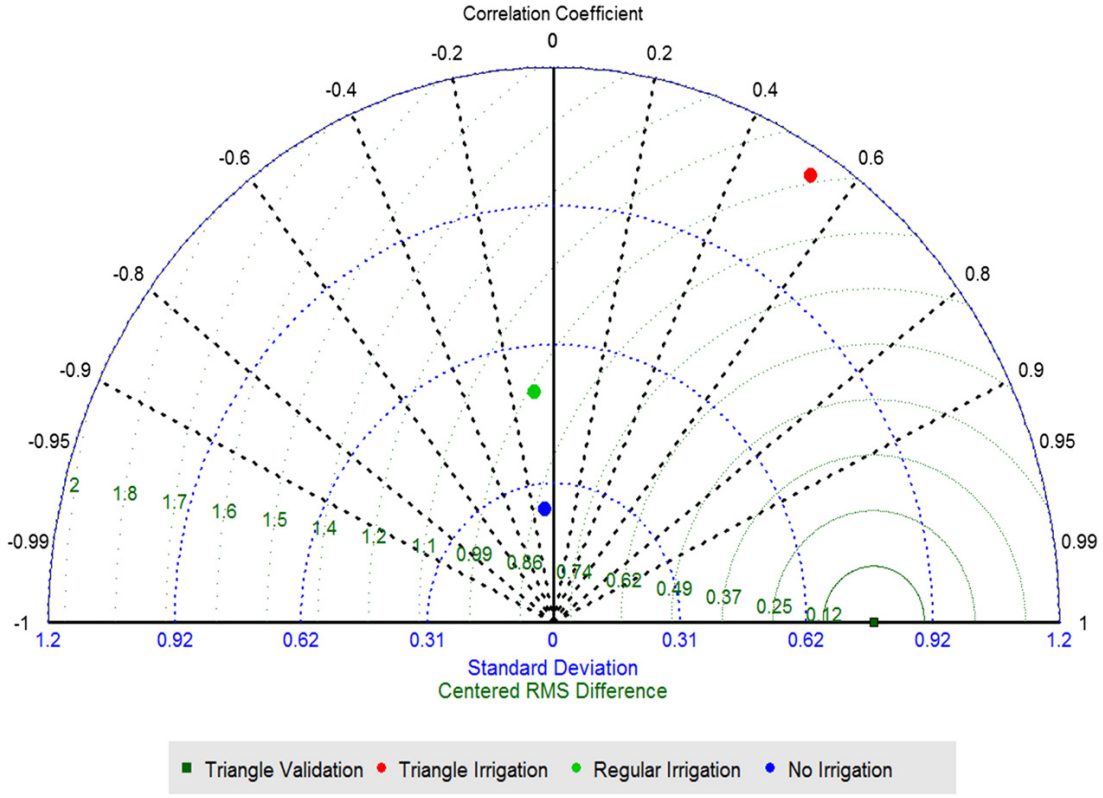


Fig. 8. Taylor diagram for the three different irrigation schemes, compared to the Triangle Method validation.

To address meteorological drought risks, Fig. 9 shows the long-term mean daily precipitation per month, which served as input for the hydrological model, as provided by the four RCM scenarios for both the reference (green band) and future (purple band) period. The model ensemble reveals that precipitation is likely to increase in fall, with a slight surplus in September and October, while throughout the rest of the year, a decrease is indicated. All in all, the projected decline of precipitation reported by Christensen et al. (2007) can also be found here. One of four RCMs, RCA3 forced by the HadCM3 GCM shows a slight increase in annual precipitation (Table 2), mainly resulting from an increase in January. However, the three RCMs driven by ECHAM5 agree in a decline in annual precipitation with a pronounced decrease in January. The mean derived from the four scenarios, shown as solid (dotted) black line for reference (future) period, results in a decline of 10–20% as also summarized in Table 2.

As a consequence, ETR, shown in Fig. 10, is likely to decrease in winter months. However, the HadCM3 scenario driven run indicates an increase in ETR in February, originating from precipitation increase in January. The general distribution, however, is preserved also for the future simulations, with maximum evapotranspiration in February and March, due to water availability, increased temperatures and higher radiation as compared to December and January. To estimate the impact on drought risk, an evapotranspiration index (ETI) using ETR and ETpot is calculated from the ratio of actual to potential evapotranspiration (see Eq. (5)).

$$\mathrm{ETI} = \frac{\mathrm{ETR}}{\mathrm{ET_{pot}}} \tag{5}$$

As the main limitation for ETR is water availability a decreased ETI indicates increased water stress. Fig. 11 presents the change in ETI from future to reference in percent. Therefore, values >100% represent an increased share of ETR for the future period. The ETI is likely to decrease throughout the year, with the exception of November, for most parts of the Gaza Strip. Only for the southeast an increase in ETI is projected, stemming from a slight increase in ETR in the respective summer months. As presented in Fig. 10, the ETR in this area was already relatively low; therefore small absolute changes can result in extreme relative changes. However, ETI for the central part of the Gaza Strip, which is the main area for irrigated agriculture, is projected to decrease throughout the year. In the summer months, July to September, these effects are not too pronounced, as irrigation management remains unchanged in the model setup for both periods. The months with only limited irrigation or without irrigation such as February and March, show considerably lower ETI for the future period due to the precipitation projections. The southern part of the Gaza Strip benefits from the

Table 2
Overview over long-term annual results for WaSiM model results for the four applied RCMs. The results are presented as difference for the future period 2041–2070 compared to the reference period 1971–2000 in percent. Presented are changes in precipitation (P), actual evapotranspiration (ETR), potential evapotranspiration (ETpot) and runoff (Q). The same irrigation input was chosen for reference and future periods.

RCM	P [%]	ETR [%]	ETpot [%]	Q [%]
ECH–REM	−12.4	−3.1	+5.8	−22.9
ECH–RMO	−16.9	−5.1	+5.5	−26.1
ECH–RCA	−9.1	−1.3	+6.9	−22.8
HCH–RCA	+5.6	+4.4	+10.5	−8.6

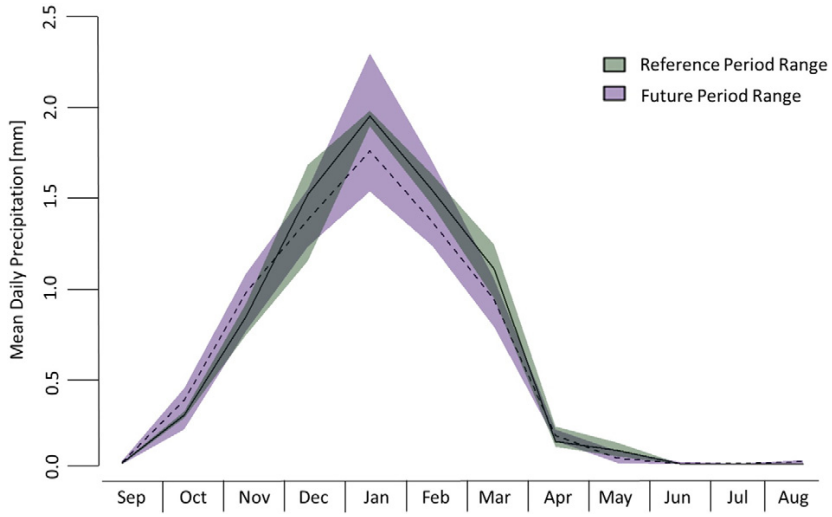


Fig. 9. Long-term daily mean precipitation results for the future (2041–2070) and reference (1971–2000) periods according to the chosen 4-RCM-ensemble and WaSiM simulations. Solid line represents the mean of the reference simulations, dashed line the future period.

slight increase in precipitation during the fall months and evapotranspiration is consequently projected to increase here as well.

The presented results for ETI predict an increased drought risk for the Gaza Strip in the future period, as the share of actual to potential evapotranspiration is likely to decline. This increase will be most severe in the transition months February and March. To compensate this general decline and to balance the water need to maintain agricultural productivity, the demand for future irrigation is increasing.

5. Discussion and conclusion

Irrigation demand and supply is a highly relevant component for the management of increasingly scarce freshwater resources in the Gaza Strip. Acquiring a better understanding of the potential impacts of climate change on water availability and drought risk is essential for the adaptation of water resources management in the region. This, however, requires the application of a meteo-hydrological modeling chain, which remains a challenge under the given conditions of substantial data scarcity.

To overcome the problem of data scarcity and to set up a hydrological model in an ungauged basin, a remote sensing approach is introduced and applied to support the parameterization process. The Triangle Method by Price (1990), refined and elaborated by Carlson et al. (1995) and Jiang and Islam (1999), was adapted for an application in the Gaza Strip to parameterize the irrigation input of the hydrological model WaSiM. This approach was chosen over comparable methods, such as the mentioned SEBAL, for reasons of data availability, as SEBAL requires information on additional meteorological parameters, e.g. wind speed and relative humidity. As shown in this study, the parameterization using the irrigation patterns derived from remotely sensed ETR is possible and improves the model performance compared to a classical parameterization approach. However, data availability is still a limiting factor. It is not clear how many Landsat TM – or comparable – scenes are necessary to derive robust results. Furthermore, the

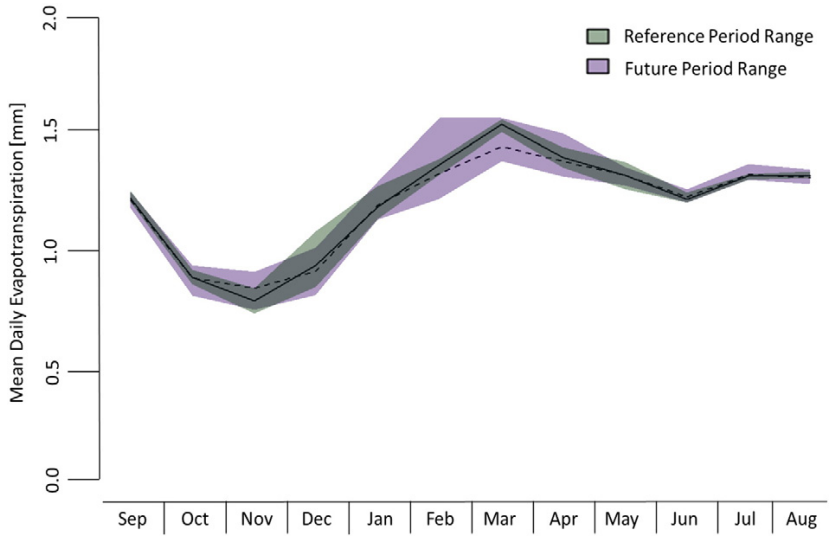


Fig. 10. Long-term daily mean ETR results for the future (2041–2070) and reference (1971–2000) periods according to the chosen 4-RCM-ensemble and WaSiM simulations. Solid line represents the mean of the reference simulations, dashed line the future period.

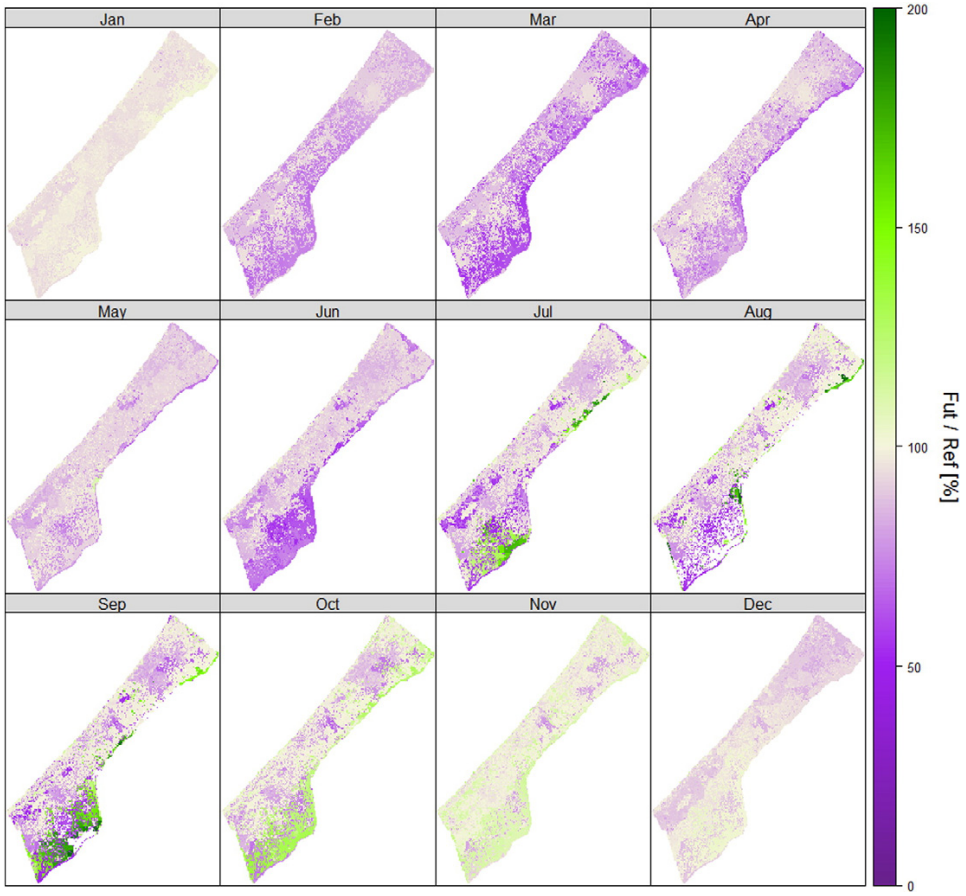


Fig. 11. Change in the evapotranspiration index (ETI), defined as in Eq. (5) for the difference of future (2041–2070) and reference (1971–2000) period in percent. Changes > 100 represents a higher ETI for the future period, thus a larger share of ETR on ETP. Simulations shown here are for the ECHAM5 driven RCA3 simulation.

performance of the Triangle Method is limited to several applications (Batra et al., 2006; Stisen et al., 2008; Tang et al., 2010; Wang et al., 2006; Yiang and Wang, 2011), however, their studies showed acceptable results and the method seems robust. Nevertheless, ground truth information of ETR would be necessary in each specific case to evaluate the performance.

In the context of this study, the R-package 'TriangleMethod' was written, containing all necessary calculation steps to efficiently perform the calculations for Landsat TM and ETM + which allow easy processing of a large data base of remote sensing images. In future research, this could allow for a detailed analysis to answer the question on how many scenes are actually required to provide a robust prediction.

An optical sensor was used for the estimation of ETR; thus, only high radiation, cloud free days are analyzed to estimate ETR which leads to a general overestimation of ETR. This main limitation of the method is unavoidable but nevertheless important when comparing the Triangle Method ETR with model data. A match of WaSiM-modeled and Triangle Method ETR would thus not be desirable. Comparison with the WaSiM-modeled ETR showed acceptable results for most parts of the catchment. Compared to the classical approach, assigning the irrigated water according to the land use classes, could not reproduce the patterns in the validation process, as shown in the Taylor diagram in Fig. 8 as well as in Figs. 6 & 7. On a subcatchment level, the results of the Triangle derived parameterization revealed satisfactory results.

Finally, the setup of the hydrological model was used to assess future changes in evapotranspiration and to assess future drought risk. Projected decreases in precipitation cause a general decline of actual/real evapotranspiration, as water is the only limiting factor in this area. Contrary to the ECHAM5 driven RCMs, the HadCM3 GCM driven RCA 3 model projects slight increases in precipitation and evapotranspiration in consequence.

Furthermore, the calculated ETI, the ratio of actual to potential evapotranspiration, was presented as a possible and robust indicator to assess drought risk. The results presented for the ECHAM5 driven RCA3 model run show a decline in ETI, consequently indicating an increased drought risk in the area. The other RCMs applied agree in this decline as the results, presented in Table 2, indicate. Current irrigation inputs compensate this effect to a certain degree in the summer months, however not to the full extent.

To maintain current agricultural productivity, irrigation demand is likely to increase in the future also for the already highly irrigated months in the dry season. Furthermore, the projected decline in precipitation also during the wet season in winter and an increase in ETpot are likely to raise the demand for irrigation also for these months. Current irrigation management is already unsustainably, using vast amounts of non-renewable ground water from the Gaza Aquifer. It is very likely that continued population growth and climate change will contribute conjointly to an ever increased exploitation of the remaining ground water resources. An increased demand for irrigation in the future will also increase the pressure on the ground water aquifer in this area. To avoid a further exploitation of the resources while preserving or even enhancing current agricultural productivity, irrigation quotas must either be reduced, more efficient irrigation techniques must be employed or alternative water resources must be made available.

Acknowledgment

The funding for this research through the FP7-project CLIMB by the European Commission (grant number 244151) is gratefully acknowledged.

References

Ajluni, S., 2010. West Bank & Gaza Strip Population Census of 2007. UNRWA.
Allen, R., Pereira, L., Raes, D., Smith, M., 1998. Crop evapotranspiration — guidelines for computing crop water requirements. FAO Irrigation and Drainage Paperp. 56.
Al-Shrafany, D., Rico-Ramirez, A., Han, D., Bray, M., 2014. Comparative assessment of soil moisture estimation from land surface model and satellite remote sensing based on catchment water balance. Meteorol. Appl. 21, 521–534.
Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modeling and assessment part I: model development. J. Am. Water Resour. Assoc. 34, 73–89.
Assaf, K., Ben-Zvi, M., Clarke, J., El-Naser, H., Kesslerm, S., Landers, M., Nuseibeh, M., Wipperfurth, C., 1998. Overview of Middle East water resources of Palestinian, Jordanian and Israeli interest. EXACT Middle East Water Data Banks Project.
Baalousha, H., 2006. Desalination status in the Gaza Strip and its environmental impacts. Desalination 196, 1–12.
Baalousha, H., 2008. Analysis of nitrate occurrence and distribution in groundwater in the Gaza Strip using major ion chemistry. Glob. NEST 10 (3), 337–349.
Bastiaansen, W.G.M., Menenti, M., Feddes, R.A., Holtslag, A.A.M., 1998a. The Surface Energy Balance Algorithm for Land (SEBAL): part 1 formulation. J. Hydrol. 212–213, 198–212.
Bastiaansen, W.G.M., Pelgrum, H., Wang, J., Ma, Y., Moreno, J., 1998b. The Surface Energy Balance Algorithm for Land (SEBAL): 2. Validation. J. Hydrol. 212–213, 213–229.
Batra, N., Islam, S., Venturini, V., Bisht, G., Liang, L., 2006. Estimation and comparison of evapotranspiration from MODIS and AVHRR sensors for clear sky days over the southern Great Plains. Remote Sens. Environ. 103, 1–15.
Blöschl, G., 2005. Rainfall-runoff modelling of ungauged catchments. Encycl. Hydrol. Sci. 2061–2080 (Chichester, U.K.).
Boegh, E., Thorsen, M., Butts, M.B., Hansen, S., Christiansen, J.S., Abrahamsen, P., Hasager, C.B., Jensen, N., Van der Keur, P., Refsgaard, J.C., Schelde, K., Soegaard, H., Thomsen, A., 2004. Incorporating remote sensing data in physically based distributed agro-hydrological modeling. J. Hydrol. 287, 279–299.
Carlson, T., 2007. An overview of the "Triangle Method" for estimating surface evapotranspiration and soil moisture from satellite imagery. Sensors 7, 1612–1629.
Carlson, T.N., Capehart, W.J., Gillies, R.R., 1995. A new look at the simplified method for remote-sensing of daily evapotranspiration. Remote Sens. Environ. 54, 161–167.
Chen, J.M., Chen, X., Weimin, J., Geng, X., 2005. Distributed hydrological model for mapping evapotranspiration using remote sensing inputs. J. Hydrol. 305, 15–39.
Christensen, J., Hewitson, B., Busuioc, A., Chen, A., Gao, X., Helt, I., Jones, R., Kolli, R., Kwon, W.-T., Laprise, R., Magaña Rueda, V., Mearns, L., Menéndez, C., Räisänen, J., Rinke, A., Sarr, Whetton, P., 2007. Regional climate projections. In: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L. (Eds.), Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.
Crago, R., Brutsaert, W., 1996. Daytime evaporation and the self-preservation of the evaporative fraction and the Bowen ratio. J. Hydrol. 178, 241–255.
Davies, J., 1967. A note on the relationship between net radiation and solar radiation. Q. J. R. Meteorol. Soc. 93, 109–115.
Deidda, R., Marrocu, M., Caroletti, G., Pusceddu, G., Langousis, A., Lucarini, V., Puliga, M., Speranza, A., 2013. Regional climate models' performance in representing precipitation and temperature over selected Mediterranean areas. Hydrol. Earth Syst. Sci. 17, 5041–5059.
Delalieux, S., Somers, B., Hereijgers, S., Verstraeten, W.W., Keulemans, W., Coppin, P., 2008. A near-infrared narrow-waveband ratio to determine Leaf Area Index in orchards. Remote Sens. Environ. 112, 3762–3772.
Gómez-del-Campo, M., 2007. Effect of water supply on leaf area development, stomatal activity, transpiration, and dry matter production and distribution in young olive trees. Aust. J. Agric. Res. 58, 1–7.
Gupta, V.K., Sorooshian, S., Yapo, P.P., 1998. Towards improved calibration of hydrologic model: multiple noncommensurable measures of information. Water Resour. Res. 34, 751–763.
Immerzeel, W.W., Droogers, P., 2008. Calibration of a distributed hydrological model based on satellite evapotranspiration. J. Hydrol. 349, 411–424.
Jiang, L., Islam, S., 1999. A methodology for estimation of surface evapotranspiration over large areas using remote sensing observations. Geophys. Res. Lett. 26, 2773–2776.
Kalma, J., McVicar, T., McCabe, M., 2008. Estimating land surface evaporation: a review of methods using remotely sensed surface temperature data. Surv. Geophys. 29, 421–469.
Kite, G.W., Droogers, P., 2000. Comparing evapotranspiration estimates from satellites, hydrological models and field data. J. Hydrol. 229, 3–18.
Kite, G.W., Pietroniro, A., 1996. Remote sensing applications in hydrology. Hydrol. Sci. 41, 563–592.
MedWetCoast Project, 2001. Wadi Gaza wetland area — the hydrological study. The Draft Report.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models, part I—a discussion of principles. J. Hydrol. 10, 282–290.
Palestinian Water Authority (PWA), 2007. Agricultural and municipal water demand in Gaza governorates for 2006. Strategic planning directorate, water resources directorate.
Parajka, J., Blöschl, G., Merz, R., 2007. Regional calibration of catchment modes: potential for ungauged catchments. Water Resour. Res. 43, W06406.
Price, J.C., 1990. Using spatial context in satellite data to infer regional scale evapotranspiration. IEEE Trans. Geosci. Remote Sens. 28, 940–948.
Priestley, C., Taylor, R., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Mon. Weather Rev. 100 (2), 81–92.
Qahman, K., Larabi, A., 2006. Evaluation and numerical modeling of seawater intrusion in the Gaza aquifer (Palestine). Hydrogeol. J. 14, 713–728.
Raveh, E., Cohen, S., Raz, T., Yakir, D., Grava, A., Goldschmidt, E., 2003. Increased growth of young citrus trees under reduced radiation load in a semi-arid climate. J. Exp. Bot. 54 (381), 365–373.
Reichenstein, M., Rey, A., Freibauer, A., Tenhunen, J., Valentini, R., Banza, J., Casals, P., Cheng, Y., Grünzweig, J.M., Irvine, J., Joffre, R., Law, B.E., Loustau, D., Miglietta, F., Oechel, W., Ourcival, J.-M., Pereira, J.S., Ponti, F., Qi, Y., Rambal, S., Rayment, M., Romanya, J., Rossi, F., Tedeschi, V., Tirone, G., Xu, M., Yakir, D., 2003. Modeling temporal and large-scale spatial variability of soil respiration from soil water availability, temperature and vegetation productivity indices. Glob. Biochem. Cycles 17 (4) (15–1 – 15–14).
Roerink, G., Su, Z., Menenti, M., 2000. S-SEBI: a simple remote sensing algorithm to estimate the surface energy balance. Phys. Chem. Earth Part B 25 (2), 147–157.
Rusteberg, B., Azizur Rahman, M., Abu Saada, M., Rabi, A., Sauter, M., 2010. Water management strategies analysis using multiple criteria decision techniques towards sustainable development of northern Gaza Strip. BALWOIS Conference — Ohrid, Macedonia.
Saei, A., Zamani, Z., Talaie, A.-R., Fatahi, R., 2006. Influence of drought stress periods on olive (*Olea europaea* L.cv. Zard) leaves stomata. Int. J. Agric. Biol. 8 (4), 430–443.
Sakcali, M.S., Ozturk, M., 2004. Eco-physiological behaviour of some Mediterranean plants as suitable candidates for reclamation of degraded areas. J. Arid Environ. 57, 1–13.
Schmugge, T.J., Kustas, W.P., Ritchie, J.C., Jackson, T.J., Rango, A., 2002. Remote sensing in hydrology. Adv. Water Resour. 25, 1367–1385.
Schulla, J., 2012. Model Description WaSiM.
Schulla, J., Jasper, K., 2007. Model Description WaSiM-ETH.
Schultz, G.A., 1993. Hydrological modeling based on remote sensing information. Adv. Space Res. 13 (5) (149 – (5)166).
Sivapalan, M., Takeuchi, K., Franks, S.W., Gupta, V.K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J.J., Mendiondo, E.M., O'Connell, P.E., Oki, T., Pomeroy, J.W., Schertzer, D., Uhlenbrook, S., Zehe, E., 2003. IAHS decade on Predictions in Ungauged Basins (PUB), 2003–2012: shaping an exciting future for the hydro-logical sciences. Hydrol. Sci. J. 48, 857–880.
Stisen, S., Sandholt, I., Nørgaard, A., Fensholt, R., Høgh Jensen, K., 2008. Combining the Triangle Method with thermal inertia to estimate regional evapotranspiration — applied to MSG-SEVIRI data in the Senegal River basin. Remote Sens. Environ. 112, 1242–1255.
Tang, R., Li, Z.-L., Tang, B., 2010. An application of the Ts - VI Triangle Method with enhanced edges determination for evapotranspiration estimation from MODIS data in arid and semi-arid regions: implementation and validation. Remote Sens. Environ. 114, 540–551.
Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. 106, 7183–7192.
Villalobos, F.J., Testi, L., Hidalgo, J., Pastor, M., Orgaz, F., 2006. Modelling potential growth and yield of olive (*Olea europaea* L.) canopies. Eur. J. Agron. 24, 296–303.
Wang, K., L., Z., Cribb, M., 2006. Estimation of evaporative fraction from a combination of day and night land surface temperatures and NDVI: a new method to determine the Priestley–Taylor parameter. Remote Sens. Environ. 102, 293–305.
Winsemius, H.C., Schaefli, B., Montanari, A., Savenije, H.H.G., 2009. On the calibration of hydrological models in ungauged basins: a framework for integrating hard and soft hydrological information. Water Resour. Res. 45, W12422.
Yiang, J., Wang, Y., 2011. Estimating evapotranspiration fraction by modeling two-dimensional space of NDVI/albedo and day-night land surface temperature difference: a comparative study. Adv. Water Resour. 34, 512–518.

3 Conclusion & Outlook

This thesis focused on the assessment of uncertainties in the hydro-climatic modeling chain and a robust assessment of future changes on the hydrology of selected river basins over heterogeneous landscapes in the Mediterranean and the Alpine region. The presented results confirm the findings of previous studies (e.g Christensen et al., 2007) and highlight that the Mediterranean region is likely to experience severe changes in the water balance under climate change. The presented publications however also highlight the various sources of uncertainty included in climate projections and along the hydro-climatic uncertainty cascade. The main findings of this thesis are presented in the following conclusion together with the research questions presented in section 1.4.

Q1: How well do gridded precipitation products perform on the catchment scale especially over complex topography?

The results presented in this thesis confirm the results presented by Prein and Gobiet (2017) also on the catchment scale and indicate a considerable uncertainty introduced by the selection of the reference grid. The results presented by Gampe and Ludwig (2017) highlight the need to account for observational uncertainty and the importance for a thorough evaluation and selection of the reference grid instead of an arbitrary selection as in most impact assessment studies. The analyzed data sets reveal a large spread for annual precipitation with deviations of up to 30% and show even more pronounced differences for monthly precipitation. The uncertainty exceeds the spread introduced purely through different spatial resolutions throughout the year. Uncertainties increase for heavy precipitation and consecutive dry and wet days, where also the impact of spatial resolution is more important. As presented in Gampe and Ludwig (2017), the spread of the climate projections generally exceeds observational uncertainty, with the exception of the summer months, where both are of similar magnitude. This indicates the importance to account for observational uncertainty. Based on the presented results, it is recommended to consider multiple data sets originating from multiple sources, to account for source specific weaknesses as undercatch, and construct an ensemble similar to climate projections also for the reference data sets. It is highly relevant to evaluate the data sets first to exclude unrealistic data sets.

Q2: What is the relative contribution of observational uncertainty to the overall uncertainty in the hydro-climatic modeling chain?

Variance decomposition revealed only a minor contribution of bias correction on the overall variance of 10-15%. However, this is only based on climate change signals and therefore somewhat misleading as it likely underestimates the role of observational uncertainty. The contribution of the GCMs to the overall uncertainty dominates with 30-40%, followed by the RCP with 20-30%. Similar to the uncertainty introduced by the RCMs, observational uncertainty is larger for the summer months, where also the applied reference data sets showed the highest spread.

Q3: What is the influence of observational uncertainty on post-processing, i.e. bias correction, and climate model evaluation and model selection?

Despite large differences in the reference grids, the influence of observational uncertainty on RCM evaluation is less pronounced than previously expected. This confirms the higher uncertainty arising from climate projections compared to observational uncertainty and indicates that RCM bias exceeds observational uncertainty. Nevertheless, the choice of the reference does influence model selection based on validity and may cause different climate change signals and also the direction of change due the difference in selected models. Bias correction with different reference data sets did not have a significant impact on the climate change signals but severely influences the magnitude of future precipitation. This becomes even more pronounced for heavy precipitation and highlights the need to consider observational uncertainty in future studies, especially if an assessment of future extremes is of key interest.

Q4: What is the relative contribution of each of the uncertainty sources involved in recent regional climate projections on the catchment scale and for all the variables of the water balance in the downscaled CMIP5 climate projections?

The results presented in Gampe et al. (2016b) confirm the results of previous studies that uncertainties introduced by the GCM also dominate at the catchment scale for the most recent climate projections and the increased uncertainty share of the RCM for precipitation compared to temperature. However, scenario uncertainty is of similar magnitude as model uncertainty from the GCMs for both, precipitation and temperature. This indicates that scenario uncertainty might have been underestimated in previous studies that are based on the SRES emissions. These findings have strong implications for model selection as the choice of GCM and RCP are crucial to capture the uncertainties of the EURO-CORDEX ensemble. For the variables surface runoff and evaporation RCM uncertainty dominates in most cases, which was expected as these variables are driven by the different parameterizations of the RCMs and thus not directly impacted by the GCM. However, the findings indicate a considerably lower spread introduced by the RCMs over the mountainous basins (Adige and Sava).

Q5: How can remote sensing data be applied to improve the robustness of a hydrological model while being computational efficient?

Hydrological modeling using high-resolution climate projections can be a valuable tool to assess the impacts of future climate on the hydrology of given areas. However, data scarcity hampers the robustness of hydrological models and can be considered an omnipresent problem in many Mediterranean catchments. To overcome this issue, the thesis explored two potential strategies: increase the robustness of hydrological models using remote sensing and the direct application of RCM outputs to estimate changes on the water balance without the need of additional data. A relatively simple approach to estimate actual evapotranspiration from remote sensing imagery, originally elaborated by Jiang and Islam (1999), was applied to parameterize irrigated areas in the Gaza Strip. Results reveal

a more realistic representation and increased robustness in the model setup. In the course of this study an R-package was compiled to allow for an efficient processing of remote sensing images to derive evapotranspiration also over other regions and river basins.

Q6: Is the Mediterranean region likely to experience increased pressure on the water balance and thus likely to be exposed to increased water scarcity under future climate?

The direct use of RCM outputs for evaporation and surface runoff provides promising estimates of future changes on the water balance and can thus be applied as a first assessment of future water scarcity. While this does not represent the correct magnitudes - due to model biases in the RCMs - the model ensemble shows high agreement over most of the areas which indicates robustness in the projected changes. The approach is therefore recommended as a first estimation of future changes on the water balance but does not allow for detailed recommendations. The results presented by Gampe et al. (2016a) indicate increased pressure on the water resources to maintain the current agricultural productivity in the Gaza Strip also under future climate. As presented in Gampe et al. (2016b) the climate projections indicate that areas already affected by water scarcity, such as the Ebro river in Spain and the Evrotas in Greece, are likely to experience increased scarcity in the future.

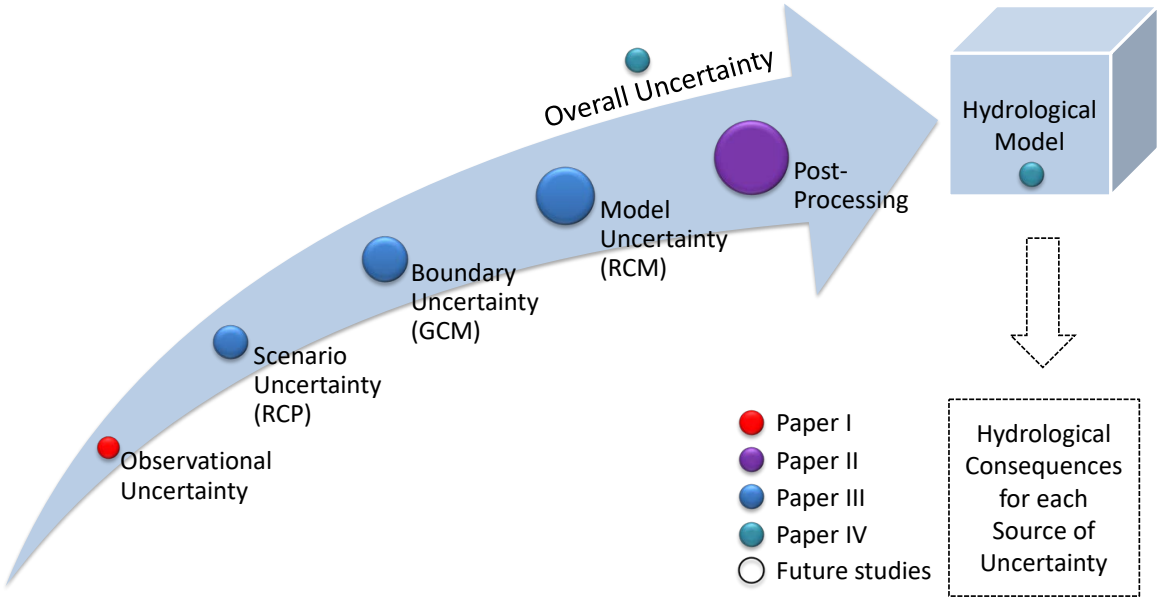


Figure 5: Updated structure for future activities to assess the hydrological consequences of each source of uncertainty. Additionally, the application of a hydrological model will allow for a comparison with the changes projected by the RCMs as presented in Paper III.

Based on the results presented, several research questions for future studies emerge. As presented in Figure 5, a detailed assessment of the hydrological consequences is necessary to fully answer the role of several uncertainty sources for the hydro-climatic modeling chain. As bias correction with multiple reference data sets revealed large differences for future extreme precipitation it is essential to know how this translates to peak flows and

future runoff projections. Additionally, different bias correction methods shall be applied to evaluate the role of the reference grid against the uncertainties arising from the bias correction method. Although there is a large variety of reference data sets available there is still a strong need for a high resolution data set on the continental or even global scale for a more detailed evaluation of climate simulations.

Additionally, detailed hydrological models are required to evaluate the estimates of surface runoff and evapotranspiration provided by current RCMs. This is essential in order to further evaluate to what degree RCM outputs can be used directly and how the results can be interpreted. As precipitation outputs of RCMs are likely biased, this would also allow to estimate model biases for surface runoff and evapotranspiration and help to improve RCMs and increase their credibility.

References

Addor, N. and Fischer, E. M. (2015). The influence of natural variability and interpolation errors on bias characterization in RCM simulations. *Journal of Geophysical Research: Atmospheres*, 120(19).

Addor, N., Rohrer, M., Furrer, R., and Seibert, J. (2016). Propagation of biases in climate models from the synoptic to the regional scale: Implications for bias adjustment. *Journal of Geophysical Research: Atmospheres*, 121(5):2075–2089.

Alexander, L. V. and Arblaster, J. M. (2017). Historical and projected trends in temperature and precipitation extremes in Australia in observations and CMIP5. *Weather and Climate Extremes*, 15:34–56.

Arnell, N. W. and Gosling, S. N. (2016). The impacts of climate change on river flood risk at the global scale. *Climatic Change*, 134(3):387–401.

Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., and Prat, O. P. (2015). PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bulletin of the American Meteorological Society*, 96(1):69–83.

Barkhordarian, A., von Storch, H., and Bhend, J. (2013). The expectation of future precipitation change over the Mediterranean region is different from what we observe. *Climate Dynamics*, 40(1-2):225–244.

Bellprat, O., Kotlarski, S., Lüthi, D., and Schär, C. (2012). Exploring perturbed physics ensembles in a regional climate model. *Journal of Climate*, 25(13):4582–4599.

Bellprat, O., Kotlarski, S., Lüthi, D., and Schär, C. (2013). Physical constraints for temperature biases in climate models. *Geophysical Research Letters*, 40(15):4042–4047.

Bengtsson, L., Hagemann, S., and Hodges, K. I. (2004). Can climate trends be calculated from reanalysis data? *Journal of Geophysical Research: Atmospheres*, 109(D11).

Beniston, M., Stephenson, D. B., Christensen, O. B., Ferro, C. A., Frei, C., Goyette, S., Halsnaes, K., Holt, T., Jylhä, K., Koffi, B., Palutikof, J., Schöll, R., Semmler, T., and Woth, K. (2007). Future extreme events in European climate: an exploration of regional climate model projections. *Climatic Change*, 81(1):71–95.

Biemans, H., Speelman, L., Ludwig, F., Moors, E., Wiltshire, A., Kumar, P., Gerten, D., and Kabat, P. (2013). Future water resources for food production in five South Asian river basins and potential for adaptation - A modeling study. *Science of the Total Environment*, 468:S117–S131.

Bird, D. N., Benabdallah, S., Gouda, N., Hummel, F., Koeberl, J., La Jeunesse, I., Meyer, S., Prettenthaler, F., Soddu, A., and Woess-Gallasch, S. (2016). Modelling climate change impacts on and adaptation strategies for agriculture in Sardinia and Tunisia using AquaCrop and value-at-risk. *Science of The Total Environment*, 543:1019–1027.

Bowden, J. H., Otte, T. L., Nolte, C. G., and Otte, M. J. (2012). Examining interior grid nudging techniques using two-way nesting in the WRF model for regional climate modeling. *Journal of Climate*, 25(8):2805–2823.

Brigode, P., Oudin, L., and Perrin, C. (2013). Hydrological model parameter instability: A source of additional uncertainty in estimating the hydrological impacts of climate change? *Journal of Hydrology*, 476:410–425.

Brown, J. F. and Pervez, M. S. (2014). Merging remote sensing data and national agricultural statistics to model change in irrigated agriculture. *Agricultural Systems*, 127:28–40.

Casanueva, A., Herrera, S., Fernández, J., and Gutiérrez, J. (2016). Towards a fair comparison of statistical and dynamical downscaling in the framework of the EURO-CORDEX initiative. *Climatic Change*, 137(3-4):411–426.

Castro, C. L., Pielke, R. A., and Leoncini, G. (2005). Dynamical downscaling: Assessment of value retained and added using the Regional Atmospheric Modeling System (RAMS). *Journal of Geophysical Research: Atmospheres*, 110(D5).

Chen, J., Brissette, F. P., Chaumont, D., and Braun, M. (2013). Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America. *Water Resources Research*, 49(7):4187–4205.

Chen, J., Brissette, F. P., and Leconte, R. (2012). Coupling statistical and dynamical methods for spatial downscaling of precipitation. *Climatic Change*, 114(3-4):509–526.

Chen, T., Ren, L., Yuan, F., Yang, X., Jiang, S., Tang, T., Liu, Y., Zhao, C., and Zhang, L. (2017). Comparison of spatial interpolation schemes for rainfall data and application in hydrological modeling. *Water*, 9(5):342.

Christensen, J. H., Boberg, F., Christensen, O. B., and Lucas-Picher, P. (2008). On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophysical Research Letters*, 35(20).

Christensen, J. H., Hewitson, B., Busuioc, A., Chen, A., Gao, X., Held, R., Jones, R., Kolli, R. K., Kwon, W., Laprise, R., Magana Rueda, V., Mearns, L., Menendez, C., Raisanen, J., Rinke, A., Sarr, A., and Whetton, P. (2007). Regional climate projections. In Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., and Miller, H., editors, *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, chapter 11, pages 847–940. University Press.

Collins, M., Booth, B. B., Bhaskaran, B., Harris, G. R., Murphy, J. M., Sexton, D. M., and Webb, M. J. (2011a). Climate model errors, feedbacks and forcings: a comparison of perturbed physics and multi-model ensembles. *Climate Dynamics*, 36(9-10):1737–1766.

Collins, W., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C., Joshi, M., Liddicoat, S., Martin, G., O'Connor, F., Rae, J., Senior, C., Sitch, S., Totterdell, I., Wiltshire, A., and Woodward, S. (2011b). Development and evaluation of an Earth-System model-HadGEM2. *Geoscientific Model Development*, 4(4):1051–1075.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P., Kruk, M. C., Kruger, A. C., Marshall, J., Maugeri, M., Mok, H. Y., Nordli, O., Ross, T., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J. (2011). The twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society*, 137(654):1–28.

Cudennec, C., Leduc, C., and Koutsoyiannis, D. (2007). Dryland hydrology in Mediterranean regions - a review. *Hydrological Sciences Journal/Journal des Sciences Hydrologiques*, 52(6):1077–1087.

Daly, C., Slater, M. E., Roberti, J. A., Laseter, S. H., and Swift, L. W. (2017). High-resolution precipitation mapping in a mountainous watershed: ground truth for evaluating uncertainty in a national precipitation dataset. *International Journal of Climatology*, 37(S1):124–137.

Dee, D. P., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E., Isaksen, L., Kåhlenberg, P., Köler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitard, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656):553–597.

Deidda, R., Marrocu, M., Caroletti, G., Pusceddu, G., Langousis, A., Lucarini, V., Puliga, M., and Speranza, A. (2013). Regional climate models' performance in representing precipitation and temperature over selected Mediterranean areas. *Hydrology and Earth System Sciences*, 17(12):5041.

Déqué, M., Rowell, D., Lüthi, D., Giorgi, F., Christensen, J., Rockel, B., Jacob, D., Kjellström, E., De Castro, M., and van den Hurk, B. (2007). An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections. *Climatic Change*, 81(1):53–70.

Déqué, M., Somot, S., Sanchez-Gomez, E., Goodess, C., Jacob, D., Lenderink, G., and Christensen, O. (2012). The spread amongst ENSEMBLES regional scenarios: regional climate models, driving general circulation models and interannual variability. *Climate Dynamics*, 38(5-6):951–964.

Deser, C., Phillips, A., Bourdette, V., and Teng, H. (2012). Uncertainty in climate change projections: the role of internal variability. *Climate Dynamics*, 38(3-4):527–546.

Devia, G. K., Ganasri, B., and Dwarakish, G. (2015). A review on hydrological models. *Aquatic Procedia*, 4:1001–1007.

Di Luca, A., de Elía, R., and Laprise, R. (2012). Potential for added value in precipitation simulated by high-resolution nested regional climate models and observations. *Climate Dynamics*, 38(5-6):1229–1247.

Dosio, A. (2016). Projections of climate change indices of temperature and precipitation from an ensemble of bias-adjusted high-resolution EURO-CORDEX regional climate models. *Journal of Geophysical Research: Atmospheres*, 121(10):5488–5511.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J. (2012). HESS Opinions "Should we apply bias correction to global and regional climate model data?". *Hydrology and Earth System Sciences*, 16(9):3391–3404.

Falkenmark, M. (1989). The massive water scarcity now threatening Africa: why isn't it being addressed? *Ambio*, 18(2):112–118.

Ferro, C. A. (2004). Attributing variation in a regional climate change modelling experiment. *EU Project PRUDENCE Tech. Rep.*

Feser, F., Rockel, B., von Storch, H., Winterfeldt, J., and Zahn, M. (2011). Regional climate models add value to global model data: a review and selected examples. *Bulletin of the American Meteorological Society*, 92(9):1181–1192.

Fischer, E. M., Beyerle, U., and Knutti, R. (2013). Robust spatially aggregated projections of climate extremes. *Nature Climate Change*, 3(12):1033.

Flaounas, E., Drobinski, P., Vrac, M., Bastin, S., Lebeaupin-Brossier, C., Stéfanon, M., Borga, M., and Calvet, J.-C. (2013). Precipitation and temperature space–time variability and extremes in the Mediterranean region: evaluation of dynamical and statistical downscaling methods. *Climate Dynamics*, 40(11-12):2687–2705.

Flint, L. E. and Flint, A. L. (2012). Downscaling future climate scenarios to fine scales for hydrologic and ecological modeling and analysis. *Ecological Processes*, 1(1):2.

Foley, A. (2010). Uncertainty in regional climate modelling: A review. *Progress in Physical Geography*, 34(5):647–670.

Fowler, H. J., Blenkinsop, S., and Tebaldi, C. (2007). Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology*, 27(12):1547–1578.

Frei, C., Schöll, R., Fukutome, S., Schmidli, J., and Vidale, P. L. (2006). Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models. *Journal of Geophysical Research: Atmospheres*, 111(D6).

Gampe, D. and Ludwig, R. (2017). Evaluation of gridded precipitation data products for hydrological applications in complex topography. *Hydrology*, 4(4):53.

Gampe, D., Ludwig, R., Qahman, K., and Afifi, S. (2016a). Applying the Triangle Method for the parameterization of irrigated areas as input for spatially distributed hydrological modeling - Assessing future drought risk in the Gaza Strip (Palestine). *Science of the Total Environment*, 543:877–888.

Gampe, D., Nikulin, G., and Ludwig, R. (2016b). Using an ensemble of regional climate models to assess climate change impacts on water scarcity in European river basins. *Science of the Total Environment*, 573:1503–1518.

Gampe, D., Schmid, J., and Ludwig, R. (2018). Impact of precipitation reference data sets on climate model selection and resulting climate change signals. *Journal of Hydrometeorology*. submitted.

García-Díez, M., Fernández, J., and Vautard, R. (2015). An RCM multi-physics ensemble over Europe: multi-variable evaluation to avoid error compensation. *Climate Dynamics*, 45(11-12):3141–3156.

Gerstengarbe, F.-W., Hoffmann, P., Österle, H., and Werner, P. C. (2015). Ensemble simulations for the RCP8.5-Scenario. *Meteorologische Zeitschrift*, 1(2015):147–156.

Gettelman, A. and Morrison, H. (2015). Advanced two-moment bulk microphysics for global models. part i: Off-line tests and comparison with other schemes. *Journal of Climate*, 28(3):1268–1287.

Giorgi, F. and Gutowski, W. J. (2015). Regional dynamical downscaling and the CORDEX initiative. *Annual Review of Environment and Resources*, 40.

Giorgi, F. and Gutowski, W. J. (2016). Coordinated experiments for projections of regional climate change. *Current Climate Change Reports*, 2(4):202–210.

Giorgi, F., Hewitson, B., Christensen, J., Hulme, M., Von Storch, H., Whetton, P., Jones, R., Mearns, L., Fu, C., Arritt, R., Bates, B., Boer, G., Buishand, A., Castro, M., Chen, D., Cramer, W., Crane, R., Crossley, J. F., Dehn, M., Dethloff, K., Dippner, J., Emori, S., Francisco, R., Fyfe, J., Gersetengarbe, F. W., Gutowski, W., Gyalistras, D., Hanssen-Bauer, I., Hante, M., Hassell, D. C., Heimann, D., Jack, C., Jacobeit, J., Kato, H., Katz,

R., Kauker, F., Knutson, T., Lal, M., Landsea, C., Laprise, R., Leung, L. R., Lynch, A. H., May, W., McGregor, J. L., Miller, N. l., Murphy, J., Ribalaygua, J., Rinkte, A., Rummukainen, M., Semazzi, F., Walsh, K., Werner, P., Widmann, M., Wilby, R., Wild, M., and Xue, Y. (2001). *Climate Change 2001: The Scientific bases*, volume 110, chapter Regional climate information - evaluation and projections. Geological and Atmospheric Sciences Publications.

Giorgi, F., Jones, C., and Asrar, G. R. (2009). Addressing climate information needs at the regional level: the CORDEX framework. *World Meteorological Organization (WMO) Bulletin*, 58(3):175–183.

Giuliani, M. and Castelletti, A. (2016). Is robustness really robust? How different definitions of robustness impact decision-making under climate change. *Climatic Change*, 135(3-4):409–424.

Gómez-Navarro, J., Montávez, J., Jerez, S., Jiménez-Guerrero, P., and Zorita, E. (2012). What is the role of the observational dataset in the evaluation and scoring of climate models? *Geophysical Research Letters*, 39(24).

Graham, L. P., Hagemann, S., Jaun, S., and Beniston, M. (2007). On interpreting hydrological change from regional climate models. *Climatic Change*, 81(1):97–122.

Gunkel, A., Shadeed, S., Hartmann, A., Wagener, T., and Lange, J. (2015). Model signatures and aridity indices enhance the accuracy of water balance estimations in a data-scarce Eastern Mediterranean catchment. *Journal of Hydrology: Regional Studies*, 4:487–501.

Gutowski Jr, W. J., Giorgi, F., Timbal, B., Frigon, A., Jacob, D., Kang, H.-S., Raghavan, K., Lee, B., Lennard, C., Nikulin, G., O'Rourke, E., Rixen, M., Solman, S., Stephenson, T., and Tangang, F. (2016). WCRP coordinated regional downscaling experiment (CORDEX): a diagnostic MIP for CMIP6. *Geoscientific Model Development*, 9(11):4087.

Hagemann, S., Chen, C., Clark, D., Folwell, S., Gosling, S. N., Haddeland, I., Hannasaki, N., Heinke, J., Ludwig, F., Voss, F., and Wiltshire, A. (2013). Climate change impact on available water resources obtained using multiple global climate and hydrology models. *Earth System Dynamics*, 4:129–144.

Harder, P. and Pomeroy, J. W. (2014). Hydrological model uncertainty due to precipitation-phase partitioning methods. *Hydrological Processes*, 28(14):4311–4327.

Hattermann, F., Vetter, T., Breuer, L., Su, B., Daggupati, P., Donnelly, C., Fekete, B., Flörke, F., Gosling, S. N., Hoffmann, P., Liersch, S., Masaki, Y., Motovilov, Y., Müller, C., Samaniego, L., Stacke, T., Wada, Y., Yang, T., and Krysnaova, V. (2018). Sources of uncertainty in hydrological climate impact assessment: a cross-scale study. *Environmental Research Letters*, 13(1):015006.

Hawkins, E. and Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, 90(8):1095–1107.

Hawkins, E. and Sutton, R. (2010). The potential to narrow uncertainty in projections of regional precipitation change. *Climate Dynamics*, 37(1-2):407–418.

Hawkins, E. and Sutton, R. (2012). Time of emergence of climate signals. *Geophysical Research Letters*, 39(1).

Haylock, M., Hofstra, N., Klein Tank, A., Klok, E., Jones, P., and New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research: Atmospheres*, 113(D20).

Henn, B., Newman, A. J., Livneh, B., Daly, C., and Lundquist, J. D. (2018). An assessment of differences in gridded precipitation datasets in complex terrain. *Journal of Hydrology*, 556:1205–1219.

Herger, N., Abramowitz, G., Knutti, R., Angélil, O., Lehmann, K., and Sanderson, B. M. (2018). Selecting a climate model subset to optimise key ensemble properties. *Earth System Dynamics*, 9(1):135.

Herrera, S., Gutiérrez, J., Ancell, R., Pons, M., Frías, M., and Fernández, J. (2012). Development and analysis of a 50-year high-resolution daily gridded precipitation dataset over Spain (Spain02). *International Journal of Climatology*, 32(1):74–85.

Herrero, A., Gutiérrez-Cánovas, C., Vigiak, O., Lutz, S., Kumar, R., Gampe, D., Huber-García, V., Ludwig, R., Batalla, R., and Sabater, S. (2018). Multiple stressor effects on biological quality elements in the ebro river: Present diagnosis and predicted responses. *Science of the Total Environment*, 630:1608–1618.

Isotta, F. A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassegues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S. M., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G., and Vertacnik, G. (2014). The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. *International Journal of Climatology*, 34(5):1657–1675.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Vautard, R., Weber, B., and Yiou, P. (2014). EURO-CORDEX: new high-resolution climate change projections for European impact research. *Regional Environmental Change*, 14(2):563–578.

Jaeger, W. K., Plantinga, A. J., Chang, H., Dello, K., Grant, G., Hulse, D., McDonnell, J., Lancaster, S., Moradkhani, H., Morzillo, A., Mote, P., Nolin, A., Santelmann, M., and Wu, J. (2013). Toward a formal definition of water scarcity in natural-human systems. *Water Resources Research*, 49(7):4506–4517.

Jeon, S., Paciorek, C. J., and Wehner, M. F. (2016). Quantile-based bias correction and uncertainty quantification of extreme event attribution statements. *Weather and Climate Extremes*, 12:24–32.

Jiang, L. and Islam, S. (1999). A methodology for estimation of surface evapotranspiration over large areas using remote sensing observations. *Geophysical Research Letters*, 26(17):2773–2776.

Joseph, R., Smith, T. M., Sapiano, M. R., and Ferraro, R. R. (2009). A new high-resolution satellite-derived precipitation dataset for climate studies. *Journal of Hydrometeorology*, 10(4):935–952.

Karimi, P. and Bastiaanssen, W. G. (2015). Spatial evapotranspiration, rainfall and land use data in water accounting–Part 1: Review of the accuracy of the remote sensing data. *Hydrology and Earth System Sciences*, 19(1):507–532.

Karl, T. R. and Trenberth, K. E. (2003). Modern global climate change. *Science*, 302(5651):1719–1723.

Kay, J., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J., Bates, S., Danabasoglu, G., Edwards, J., Holland, M., Kushner, P., Lamarque, J.-F., Lawrence, D., Lindsay, K., Middleton, A., Munoz, E., Neale, R., Oleson, K., Polvani, L., and Vertenstein, M. (2015). The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bulletin of the American Meteorological Society*, 96(8):1333–1349.

Kienzle, S. W., Nemeth, M. W., Byrne, J. M., and MacDonald, R. J. (2012). Simulating the hydrological impacts of climate change in the upper North Saskatchewan River basin, Alberta, Canada. *Journal of Hydrology*, 412:76–89.

Knutti, R., Masson, D., and Gettelman, A. (2013). Climate model genealogy: Generation CMIP5 and how we got there. *Geophysical Research Letters*, 40(6):1194–1199.

Knutti, R. and Sedláček, J. (2013). Robustness and uncertainties in the new CMIP5 climate model projections. *Nature Climate Change*, 3(4):369.

Kotlarski, S., Keuler, K., Christensen, O. B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., Van Meijgaard, E., , Nikulin, G., Schär, C., Teichmann, C., Vautard, R., Warrach-Sagi, K., and Wulfmeyer, V. (2014). Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geoscientific Model Development*, 7(4):1297–1333.

Kotlarski, S., Szabó, P., Herrera, S., Räty, O., Keuler, K., Soares, P. M., Cardoso, R. M., Bosshard, T., Pagé, C., Boberg, F., Guitiérrez, J., Isotta, F., Jaczewski, A., Kreienkamp, F., Liniger, M. A., Lussana, C., and Pianko-Kliczýnska, K. (2017). Observational uncertainty and regional climate model evaluation: a pan - European perspective. *International Journal of Climatology*.

Kunnath-Poovakka, A., Ryu, D., Renzullo, L., and George, B. (2016). The efficacy of calibrating hydrologic model using remotely sensed evapotranspiration and soil moisture for streamflow prediction. *Journal of Hydrology*, 535:509–524.

Lamarque, J.-F., Shindell, D. T., Josse, B., Young, P., Cionni, I., Eyring, V., Bergmann, D., Cameron-Smith, P., Collins, W. J., Doherty, R., Dalsoren, S., Faluvegi, G., Folberth, G., Ghan, S. J., Horowitz, L. W., Lee, Y. H., MacKenzie, I. A., Nagashima, T., Naik, V., Plummer, D., Righi, M., Rumbold, S. T., Schulz, M., Skeie, R. B., Stevenson, D., Strode, S., Sudo, K., Szopa, S., Voulgarakis, A., and Zeng, G. (2013). The Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): overview and description of models, simulations and climate diagnostics. *Geoscientific Model Development*, 6:179–206.

Liu, Y. and Gupta, H. V. (2007). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Research*, 43(7).

Ludwig, R., Soddu, A., Duttmann, R., Baghdadi, N., Benabdallah, S., 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. (2010). Climate-induced changes on the hydrology of Mediterranean basins - A research concept to reduce uncertainty and quantify risk. *Fresenius Environmental Bulletin*, 19(10A):2379–2384.

Maraun, D. and Widmann, M. (2015). The representation of location by a regional climate model in complex terrain. *Hydrology and Earth System Sciences*, 19(8):3449–3456.

Maslin, M. and Austin, P. (2012). Uncertainty: Climate models at their limit? *Nature*, 486(7402):183.

Masson, D. and Knutti, R. (2011). Climate model genealogy. *Geophysical Research Letters*, 38(8).

Maure, G. A., Pinto, I., Ndebele-Murisa, M. R., Muthige, M., Lennard, C., Nikulin, G., Dosio, A., and Meque, A. O. (2018). The southern African climate under 1.5° and 2° C of global warming as simulated by CORDEX models. *Environmental Research Letters*.

Meehl, G. A., Covey, C., Taylor, K. E., Delworth, T., Stouffer, R. J., Latif, M., McAvaney, B., and Mitchell, J. F. (2007). The WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bulletin of the American Meteorological Society*, 88(9):1383–1394.

Mendlik, T. and Gobiet, A. (2016). Selecting climate simulations for impact studies based on multivariate patterns of climate change. *Climatic Change*, 135(3-4):381–393.

Meyer, S., Blaschek, M., Duttmann, R., and Ludwig, R. (2016). Improved hydrological model parametrization for climate change impact assessment under data scarcity - The potential of field monitoring techniques and geostatistics. *Science of The Total Environment*, 543:906–923.

Mora, C., Frazier, A. G., Longman, R. J., Dacks, R. S., Walton, M. M., Tong, E. J., Sanchez, J. J., Kaiser, L. R., Stender, Y. O., Anderson, J. M., Ambrosino, C., Fernandez-Silva, I., Giuseffi, L., and Giambelluca, T. (2013). The projected timing of climate departure from recent variability. *Nature*, 502(7470):183.

Moradkhani, H. and Sorooshian, S. (2009). General review of rainfall-runoff modeling: model calibration, data assimilation, and uncertainty analysis. In *Hydrological modelling and the water cycle*, pages 1–24. Springer.

Muerth, M., Gauvin St-Denis, B., Ricard, S., Velázquez, J., Schmid, J., Minville, M., Caya, D., Chaumont, D., Ludwig, R., and Turcotte, R. (2013). On the need for bias correction in regional climate scenarios to assess climate change impacts on river runoff. *Hydrology and Earth System Sciences Discussions*, 17:1189–1204.

Nakicenovic, N., Alcamo, J., Grubler, A., Riahi, K., Roehrl, R., Rogner, H.-H., and Victor, N. (2000). *Special report on emissions scenarios (SRES), a special report of Working Group III of the intergovernmental panel on climate change*. Cambridge University Press.

Navarro-Ortega, A., Acuña, V., Bellin, A., Burek, P., Cassiani, G., Choukr-Allah, R., Dolédec, S., Elosegi, A., Ferrari, F., Ginebreda, A., Grathwohl, P., Jones, C., Ker Rault, P., Kok, K., Koundouri, P., Ludwig, R. P., Merz, R., Milacic, R., Muñoz, I., Nikulin, G., Paniconi, C., Paunović, M., Petrovic, M., Sabater, L., Sabater, S., Skoulikidis, N., Slob, A., Teutsch, G., Vouvoulis, N., and Barceló, D. (2015). Managing the effects of multiple stressors on aquatic ecosystems under water scarcity. the globaqua project. *Science of the Total Environment*, 503:3–9.

Nikulin, G., Kjellström, E., Hansson, U., Strandberg, G., and Ullerstig, A. (2011). Evaluation and future projections of temperature, precipitation and wind extremes over Europe in an ensemble of regional climate simulations. *Tellus A*, 63(1):41–55.

Palazzi, E., Hardenberg, J. v., and Provenzale, A. (2013). Precipitation in the hindukush karakoram himalaya: Observations and future scenarios. *Journal of Geophysical Research: Atmospheres*, 118(1):85–100.

Piani, C., Haerter, J., and Coppola, E. (2010). Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoretical and Applied Climatology*, 99(1-2):187–192.

Pierce, D. W., Barnett, T. P., Santer, B. D., and Gleckler, P. J. (2009). Selecting global climate models for regional climate change studies. *Proceedings of the National Academy of Sciences*, 106(21):8441–8446.

Pierce, D. W., Cayan, D. R., Maurer, E. P., Abatzoglou, J. T., and Hegewisch, K. C. (2015). Improved bias correction techniques for hydrological simulations of climate change. *Journal of Hydrometeorology*, 16(6):2421–2442.

Prein, A. F. and Gobiet, A. (2017). Impacts of uncertainties in European gridded precipitation observations on regional climate analysis. *International Journal of Climatology*, 37(1):305–327.

Prein, A. F., Gobiet, A., and Truhetz, H. (2011). Analysis of uncertainty in large scale climate change projections over Europe. *Meteorologische Zeitschrift*, 20(4):383–395.

Prudhomme, C., Giuntoli, I., Robinson, E. L., Clark, D. B., Arnell, N. W., Dankers, R., Fekete, B. M., Franssen, W., Gerten, D., Gosling, S. N., , Hagemann, S., Hannah, D., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Yoshihide, W., and Wisser, D. (2014). Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. *Proceedings of the National Academy of Sciences*, 111(9):3262–3267.

Räisänen, J. and Räty, O. (2013). Projections of daily mean temperature variability in the future: cross-validation tests with ENSEMBLES regional climate simulations. *Climate Dynamics*, 41(5-6):1553–1568.

Rasmussen, R., Baker, B., Kochendorfer, J., Myers, T., Landolt, S., Fisher, A., Black, J., Theriault, J., Kucera, P., Gochis, D., Smith, C., Nitu, R., Hall, M., Ikeda, K., and Gutmann, E. (2012). The NOAA/FAA/NCAR Winter Precipitation Test Bed: how well are we measuring snow. *Bulletin of the American Meteorological Society*, 93:811–829.

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin III, F. S., Lambin, E. F., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H. J., Nykvist, B., de Wit, C. A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P. K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R. W., Fabry, V. J., Hansen, J., Walker, B., Livermann, D., Richardson, K., Crutzen, P., and Foley, J. A. (2009). A safe operating space for humanity. *Nature*, 461(7263):472.

Rummukainen, M. (2016). Added value in regional climate modeling. *Wiley Interdisciplinary Reviews: Climate Change*, 7(1):145–159.

Ruti, P. M., Somot, S., Giorgi, F., Dubois, C., Flaounas, E., Obermann, A., Dell'Aquila, A., Pisacane, G., Harzallah, A., Lombardi, E., Ahrens, B., Akhtar, N., Alias, A., Arsouze, T., Aznar, R., Bastin, S., Bartholy, J., Béranger, K., Beuvier, J., Bouffies-Cloché, S., Brauch, J., Cabos, W., Calmanti, S., Clavet, J.-C., Carillo, A., Conte, D., Coppola, E., Djurdjevic, V., Drobinski, P., Elizalde-Arellano, A., Gaetner, M., Galán, P., Gallardo, C.,

Gualdi, S., Goncalves, M., Jorba, O., Jorda, G., L'Heveder, B., Lebeaupin-Brossier, C., Li, L., Liguori, G., Lionello, P., Maciás, D., Nabat, P., Önol, B., Raikovic, B., Ramage, K., Sevault, F., Sannino, G., Struglia, M. V., Sanna, A., Torma, C., and Vervatis, V. (2016). MED-CORDEX initiative for Mediterranean climate studies. *Bulletin of the American Meteorological Society*, 97(7):1187–1208.

Sanford, T., Frumhoff, P. C., Luers, A., and Gulledge, J. (2014). The climate policy narrative for a dangerously warming world. *Nature Climate Change*, 4(3):164.

Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N. W., Clark, D. B., Dankers, R., Eisner, S., Fekete, B. M., Colón-González, F. J., Gosling, N., Kim, H., Liu, X., Masaki, Y., Portmann, F., Satoh, Y., Stacke, T., Tang, Q., Wada, Y., Wisser, D., Albrecht, T., Frieler, K., Piontek, F., Warszawski, L., and Kabat, P. (2014). Multimodel assessment of water scarcity under climate change. *Proceedings of the National Academy of Sciences*, 111(9):3245–3250.

Schmidhuber, J. and Tubiello, F. N. (2007). Global food security under climate change. *Proceedings of the National Academy of Sciences*, 104(50):19703–19708.

Schneider, U., Ziese, M., Meyer-Christoffer, A., Finger, P., Rustemeier, E., and Becker, A. (2016). The new portfolio of global precipitation data products of the Global Precipitation Climatology Centre suitable to assess and quantify the global water cycle and resources. *Proceedings of the International Association of Hydrological Sciences*, 374:29.

Schulla, J. and Jasper, K. (2007). *Model description WaSiM-ETH*. PhD thesis, Institute for Atmospheric and Climate Science, Swiss Federal Institute of Technology, Zürich.

Schyns, J. F., Hoekstra, A. Y., and Booij, M. J. (2015). Review and classification of indicators of green water availability and scarcity. *Hydrology and Earth System Sciences*, 19(11):4581.

Scott, D., Hall, C. M., and Stefan, G. (2012). *Tourism and climate change: Impacts, adaptation and mitigation*. Routledge.

Screen, J. A. and Simmonds, I. (2010). The central role of diminishing sea ice in recent Arctic temperature amplification. *Nature*, 464(7293):1334.

Seiller, G. and Anctil, F. (2014). Climate change impacts on the hydrologic regime of a Canadian river: comparing uncertainties arising from climate natural variability and lumped hydrological model structures. *Hydrology and Earth System Sciences*, 18(6):2033.

Seiller, G., Anctil, F., and Perrin, C. (2012). Multimodel evaluation of twenty lumped hydrological models under contrasted climate conditions. *Hydrology and Earth System Sciences*, 16(4):p–1171.

Shrestha, M., Wang, L., Koike, T., Tsutsui, H., Xue, Y., and Hirabayashi, Y. (2014). Correcting basin-scale snowfall in a mountainous basin using a distributed snowmelt model and remote-sensing data. *Hydrology and Earth System Sciences*, 18(2):747–761.

Sillmann, J., Kharin, V., Zhang, X., Zwiers, F., and Bronaugh, D. (2013a). Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate. *Journal of Geophysical Research: Atmospheres*, 118(4):1716–1733.

Sillmann, J., Kharin, V., Zwiers, F., Zhang, X., and Bronaugh, D. (2013b). Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections. *Journal of Geophysical Research: Atmospheres*, 118(6):2473–2493.

Silvestro, F., Gabellani, S., Rudari, R., Delogu, F., Laiolo, P., and Boni, G. (2015). Uncertainty reduction and parameter estimation of a distributed hydrological model with ground and remote-sensing data. *Hydrology and Earth System Sciences*, 19(4):1727.

Solomon, S., Plattner, G.-K., Knutti, R., and Friedlingstein, P. (2009). Irreversible climate change due to carbon dioxide emissions. *Proceedings of the National Academy of Sciences*, 106(6):1704–1709.

Sunyer, M., Sørup, H. J. D., Christensen, O. B., Madsen, H., Rosbjerg, D., Mikkelsen, P. S., and Arnbjerg-Nielsen, K. (2013). On the importance of observational data properties when assessing regional climate model performance of extreme precipitation. *Hydrology and Earth System Sciences*, 17(11):4323.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4):485–498.

Tebaldi, C. and Knutti, R. (2007). The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 365(1857):2053–2075.

Teutschbein, C. and Seibert, J. (2010). Regional climate models for hydrological impact studies at the catchment scale: a review of recent modeling strategies. *Geography Compass*, 4(7):834–860.

Teutschbein, C. and Seibert, J. (2012). Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology*, 456:12–29.

Teutschbein, C. and Seibert, J. (2013). Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions? *Hydrology and Earth System Sciences*, 17(12):5061–5077.

Themeßl, M., Gobiet, A., and Leuprecht, A. (2011). Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *International Journal of Climatology*, 31(10):1530–1544.

Van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., and Nakicenovic, N. (2011). The representative concentration pathways: an overview. *Climatic Change*, 109(1-2):5.

Velázquez, J., Schmid, J., Ricard, S., Muerth, M., St-Denis, B. G., Minville, M., Chaumont, D., Caya, D., Ludwig, R., and Turcotte, R. (2013). An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources. *Hydrology and Earth System Sciences*, 17(2):565–578.

Voldoire, A., Sanchez-Gomez, E., y Mélia, D. S., Decharme, B., Cassou, C., Sénési, S., Valcke, S., Beau, I., Alias, A., Chevallier, M., Déqué, M., Deshayes, J., Douville, H., Fernandez, E., Madec, G., Maisonnave, E., Moine, M.-P., Planton, S., Saint-Martin, D., Szopa, S., Tyteca, S., Alkama, R. Belamari, S., Braun, A., Coquart, L., and Chauvin, F. (2013). The CNRM-CM5. 1 global climate model: description and basic evaluation. *Climate Dynamics*, 40(9-10):2091–2121.

von Storch, H. and Zwiers, F. W. (1999). *Statistical analysis in climate research*. Cambridge University Press, Cambridge.

Vrac, M., Drobinski, P., Merlo, A., Herrmann, M., Lavaysse, C., Li, L., and Somot, S. (2012). Dynamical and statistical downscaling of the French Mediterranean climate: uncertainty assessment. *Natural Hazards and Earth System Sciences*, 12(9):2769.

Wada, Y., Wisser, D., Eisner, S., Flörke, M., Gerten, D., Haddeland, I., Hanasaki, N., Masaki, Y., Portmann, F. T., Stacke, T., Tessler, Z., and Schewe, J. (2013). Multi-model projections and uncertainties of irrigation water demand under climate change. *Geophysical Research Letters*, 40(17):4626–4632.

Whetton, P., Hennessy, K., Clarke, J., McInnes, K., and Kent, D. (2012). Use of Representative Climate Futures in impact and adaptation assessment. *Climatic Change*, 115(3-4):433–442.

Wilcke, R. A. and Bärring, L. (2016). Selecting regional climate scenarios for impact modelling studies. *Environmental Modelling & Software*, 78:191–201.

Wilcke, R. A. I., Mendlik, T., and Gobiet, A. (2013). Multi-variable error correction of regional climate models. *Climatic Change*, 120(4):871–887.

Wood, A. W., Leung, L. R., Sridhar, V., and Lettenmaier, D. (2004). Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic change*, 62(1-3):189–216.

Xu, C.-y., Widén, E., and Halldin, S. (2005). Modelling hydrological consequences of climate change - progress and challenges. *Advances in Atmospheric Sciences*, 22(6):789–797.

Yoon, J.-H., Ruby Leung, L., and Correia, J. (2012). Comparison of dynamically and statistically downscaled seasonal climate forecasts for the cold season over the United States. *Journal of Geophysical Research: Atmospheres*, 117(D21).