Improvements, Uncertainties and Scale Dependencies of High Resolution Climate Simulations

Von der Fakultät Nachhaltigkeit
der Leuphana Universität Lüneburg zur Erlangung des Grades

Doktor der Naturwissenschaften
- Dr. rer. nat. -

genehmigte Dissertation von
Bastian Eggert

geboren am 09.07.1984 in Winsen (Luhe)

Eingereicht am: 18.10.2017
Mündliche Verteidigung (Disputation) am: 23.08.2018

Erstbetreuerin und -gutachterin: Prof. Dr. Daniela Jacob
Zweitgutachter: Prof. Dr. Günther Groß
Drittgutachter: Prof. Dr-Ing. Peter Fröhle

Die einzelnen Beiträge des kumulativen Dissertationsvorhabens sind oder werden ggf. inkl. des Rahmenpapiers wie folgt veröffentlicht:

1 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., Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F. F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., Yiou, P., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F. F., Teichmann, C., Valentini, R., Vautard, R., Weber, B. and Yiou, P.: EURO-CORDEX: new high-resolution climate change projections for European impact research
Published Reg. Environ. Chang., 14(2), 563–578, doi:10.1007/s10113-013-0499-2, **2013**

2 **Eggert, B.**, Berg, P., Haerter, J. O., Jacob, D. and Moseley, C.: Temporal and spatial scaling impacts on extreme precipitation
Published Atmos. Chem. Phys., 15(10), 5957–5971, doi:10.5194/acp-15-5957-2015, **2015**

3 Haerter, J. O., **Eggert, B.**, Moseley, C., Piani, C. and Berg, P.: Statistical precipitation bias correction of gridded model data using point measurements
Published Geophys. Res. Lett., 42(6), 1919–1929, doi:10.1002/2015GL063188, **2015**

4 **Eggert, B.**, Jacob, D., Haensler, A.: Uncertainties in snow and precipitation projections in the northern Alps: the role of model biases
Submitted J. Geophys. Res. Atmos. **2017**

Veröffentlichungsjahr: 2018

Abstract

Implications of a changing climate have become increasingly perceptible. Stakeholders in economy and politics recognize these changes and demand for temporal and spatially highly resolved climate projections that describe future changes at their respective locations, or for their communities. In order to make informed decisions, assessments on the robustness and significance of the projected changes are also needed. Technical progress in computing and storage capacities, as well as global partnerships that coordinate the science and applications of climate modeling activities, allow to conduct highly resolved climate simulations, and to compile homogeneous ensembles to estimate uncertainties of these projections that can be categorized into scenario uncertainties, natural variability and model uncertainties. However, in order to increase the temporal and spatial resolution of the simulations, climate models have to be adapted to the desired scales, which is a grand challenge for the modeling community. Also, there is a need for high temporal and spatial resolution of long-term observational data sets, to enable the modeling groups to validate their climate models by comparing historic climate simulations to observations.

The overall aim of this PhD-thesis is to develop empirical probabilistic frameworks, that help to quantify the impacts of temporal and spatial scale dependencies and model uncertainties of climate projections regarding precipitation-dependent parameters. These studies can be the basis for the development of products that increase the value of climate projections for impact-research and decision-makers.

The thesis is structured in four journal articles. Article one is the first study that analyzed climate projections from the spatially highly resolved regional climate model (RCM) ensemble EURO-CORDEX. Additionally, the significance and the robustness of the projected changes are analyzed, and improvements related to the higher horizontal resolution of the new data set are discussed. A major finding is that RCM simulations provide higher daily precipitation intensities which are missing in the global climate model (GCM) simulations, and, that they show a significantly different climate change of daily precipitation intensities, with a smoother shift from low towards high intensities.

The second article elaborates on impacts of temporal and spatial aggregation on extreme precipitation intensities. By combining radar data with cloud observations, the different temporal and spatial scaling behavior of stratiform and convective type precipitation events can be analyzed for the first time. The separation between convective and stratiform type events also allows to quantify the contribution of convective events to the extremes. Further, it is shown that temporal averaging has similar effects on the precipitation distribution as spatial averaging. Associated pairs of temporal and spatial resolutions that show comparable intensity distributions are identified. This knowledge can be used to improve data storage, to setup measuring campaigns and to bring together data sets that are stored at different resolutions. The latter issue is the focus of the third paper.

In order to use climate model output for impact assessments, a bias-adjustment towards the observed climatology (also referred to as "bias correction") is often inevitable. In many regions of the world the gauge station density is not sufficient to represent the spatial variability of precipitation within the gridbox size of regional climate models. In some cases only single

station data is available. Using precipitation data from radar observations, a gauge station network, and a spatially highly resolved regional climate model, the third paper optimizes the process that finds associated temporal and spatial scales (see second article). This information is used to develop a method that adjusts point measurements to the temporal and spatial scale of a previously defined model grid. The study shows that this procedure can be used to improve bias-adjustment methods in areas with a low gauge station density.

It is known that the EURO-CORDEX ensemble overestimates precipitation and shows a common cold bias in the Alpine region. The fourth article evaluates how these biases are changing the temperature distribution, and the temperature dependency, of precipitation-frequencies. These biases are a source of uncertainty that is not captured by the robustness tests performed in the first article. A probabilistic-decomposition-framework is developed to quantify the impact of these biases on precipitation-frequency changes, and to investigate causes for the ensemble spread. Because of the high economic relevance of snow for that region, the article distinguishes between total precipitation- and snowfall-frequencies. Two different intensity thresholds are used in order to analyze if heavy precipitation events deviate from the average.

Article Summary

The first article: EURO-CORDEX: new high-resolution climate change projections for European impact research

The first article introduces the high-resolution climate change projection ensemble created within EURO-CORDEX. The paper was the first that compared large-scale climate change patterns of the EURO-CORDEX and the ENSEMBLES data set. Besides temperature and precipitation, also complex climate indices are analyzed regarding future projected changes. The paper analyses how the comparatively high spatial resolution of EURO-CORDEX is improving the climate change projections. A strong focus is set on assessing the robustness and the inherent uncertainties of the projections considering different climate change scenarios. For this paper new kinds of figures were developed, that show not just the ensemble mean statistics but also illustrate the significance and the robustness of the given information. It is found that the large-scale patterns of mean-temperature and precipitation changes are similar in both ensembles. For heat waves, the way the parameter is defined has a stronger influence on the investigated changes, than differences found between the scenarios or time periods. However, the EURO-CORDEX ensemble shows a reduced northward shift of Mediterranean drying evolution, and a slightly stronger mean precipitation increase over most of Europe. Further, more detailed spatial patterns are observed in the high-resolution data, that can be connected to better resolved physical processes and surface characteristics. A comparison between the EURO-CORDEX ensemble and the driving Global Climate Models (GCMs), revealed that the Regional Climate Models (RCMs) simulate higher daily precipitation intensities, and that they show a significantly different change signal of daily precipitation intensities, resulting in a smoother shift in the intensity distribution from weak to high intensities.

The second article: Temporal and spatial scaling impacts on extreme precipitation

The second article elaborates on differences in the precipitation distributions of differently resolved data sets, that are caused by aggregation effects. For this analysis a very highly resolved radar dataset with a 1 km horizontal and 5 min temporal resolution is used. The data set covers entire Germany over the period 2007-2008. It is investigated separately for different regions and seasons as well as for events of convective and stratiform type. The first part of the article concentrates on the intensity reduction due to spatial and temporal smoothing of the data set when aggregated to lower resolutions. Type-dependent reduction factors are calculated for different seasons and regions. With up to 30% higher reduction factors for the convective type compared to stratiform extremes, these differences exceeded all other observed seasonal and regional differences within one type. In a second step, the effect of these different reduction factors on the contribution of each type to extreme events as a whole is analyzed. It is found that the temporal and spatial resolution as well as the chosen threshold that is used to identify the extremes have a profound influence on the contribution of convective events to the extremes. A third study presented in this article compares the ratio of area to duration reduction factors for convective and stratiform events. This ratio can be used to associate temporal and spatial scales. The matching scales can be thought of as the average time an event will stay in a grid box of the matching size before it is advected out of the box. This ratio mainly depends on the type of event. Comparing this ratio over different scales, a self-affine behavior is identified. This indicates fractal properties of precipitation. In the last analysis of the paper, the entire precipitation intensity distribution is analyzed at different scales. Matching pairs of temporal and spatial resolutions are identified that show similar intensity distributions.

Also the information loss due to temporal and spatial aggregation is investigated. It is found that the information loss due to temporal aggregation is disproportionately high in many extreme precipitation studies compared to spatial aggregation effects.

The third article: Statistical precipitation bias correction of gridded model data using point measurements

With increasing model resolution the demands for temporal and spatial higher resolved observation data increases. Especially precipitation events that show strong spatial and temporal inhomogeneities need a dense station network in order to capture these events properly. For validation and bias-adjustment of climate models, it is important to capture the variability preferably at least on the scale of the model grid. Therefore, the demands for the station network increase with model resolution and can not always be fulfilled. On the basis of the Taylor hypothesis of frozen turbulence, this article takes account of the connection between temporal and spatial scales as described in the second article. The study introduces a concept in which this information is used to improve state-of-the-art bias correction methods in areas where model and observation data sets are not available at the same resolution. This is especially interesting if only point measurements are available. In this case the station will always be spatially higher resolved as a grid average from the model output. A nonparametric method is developed that is called scale-adaptation statistical bias correction (SABC). The study shows that the results of the bias correction can be improved if a lack of spatial information is partly compensated by using high temporal information or vice versa.

The fourth article: Uncertainties in snow and precipitation projections in the northern Alps: the role of model biases

Precipitation, especially when falling as snow, is a key element of the Alpine climate system. Simulations and observations are known to contain errors regarding snowfall- and precipitation-frequencies in this region. It is likely that observations underestimate snowfall, while climate models tend to overestimate precipitation especially in the winter season. Based on EURO-CORDEX data the fourth article analyzes how these errors impact the temperature distribution and the temperature dependent occurrence distribution of snowfall- and precipitation-frequencies. Differences are found between the shape of the observed and the simulated distributions. The impact of these biases on past and projected changes is analyzed. A framework is developed in which changes are decomposed into two parts: changes in the temperature distribution and changes in the temperature dependency of the event occurrence. Past changes are analyzed and compared to observations allowing to quantify the impact of model biases on these changes. The observations show an increase in (heavy) precipitation-frequency above the model spread at most altitudes. The main reason is a common cold bias in the surface-temperature of all ensemble members. A bias in the surface-temperature dependency of these events is a second reason. This bias also impacts snowfall-frequency changes. The study shows that mean snowfall-fractions at a specific temperature depend on the regional climate model, or snowfall discrimination method. These differences impact the temperature dependency of snowfall-frequencies and influence the simulated changes. After investigating the impact of model errors on past climate changes, future changes are analyzed using the RCP4.5 scenario. This analysis is performed with and without bias-adjustment. Major parts of the differences in the projected changes that are caused by the bias-adjustment, can be related to biases identified in the past period.

Acknowledgement

Contents

1. Introduction

1.1 Motivation

Climate change has become broadly accepted as a major threat to society (*Rosenzweig et al.*, 2007). In recent years, the focus of the scientific discourse has changed from the question whether the global climate is changing, to the questions what these changes mean for humans and the environment at a local level, and how to adapt to the imminent changes (e.g. *Pfeifer et al.* (2015); *Cortekar et al.* (2016); *Hackenbruch et al.* (2017); *Lemos et al.* (2012)). As climate change becomes more and more evident (*Cubasch et al.*, 2013), the information derived from climate projections is increasingly becoming part of decision-making processes. Especially for precipitation related parameters this leads to high demands regarding the temporal and spatial resolution of the projections. In addition, a better understanding of the uncertainties associated with these projections is becoming increasingly important, the more far-reaching the connected decisions are (*Hawkins and Sutton*, 2009).

Against this backdrop, the overall aim of this PhD-thesis is to develop empirical probabilistic frameworks that help to quantify the impacts of temporal and spatial scale dependencies and model uncertainties of climate projections regarding precipitation dependent parameters.

1.1.1 Temporal and spatial scales of climate simulations

To simulate future climate scenarios, global models are deployed, which are driven by different greenhouse gas concentration pathways. However, today's computing performance is still not sufficient to conduct large ensembles of global climate model simulations at spatial grid resolutions below ∼100 km (*Taylor et al.*, 2012), which are needed for uncertainty analyses. At best, the coarse resolution results in a lack of information on the sub-grid scale spatial variability of the model output parameters, whereby the precise spatial distribution is not captured, and the distribution functions of spatial inhomogeneous parameters are smoothed. The latter leads to an information loss particularly on the tails of the distribution, which prevents an analysis of local extreme values. In addition to the loss of information, decisive processes that are necessary to represent the regional or local climate may not be resolved by the coarse grid. On the one hand, this can lead to the fact that future changes in local processes are not taken into account (*Di Luca et al.*, 2012a,b; *Torma et al.*, 2015). On the other, these unresolved processes can also have an influence on the resolved scales (*Mesinger et al.*, 2012).

Regional climate models (RCMs) have been developed in order to better resolve atmospheric-processes at the regional to local scale. These models use data from GCM simulations as lateral constraints and dynamically refine the global data sets for selected regions (*EURO-CORDEX community et al.*, 2017). In order to be able to assess uncertainties of climate projections, homogeneous RCM ensembles were created. Widely used RCM ensembles for the European continent are ENSEMBLES (*Hewitt and Griggs*, 2004), and a comparatively new data set from the EURO-CORDEX (*EURO-CORDEX community et al.*, 2017) project. Technological advances in computation and storage capacities allowed to significantly increase the spatial resolution of RCM simulations in recent years. While the ENSEMBLES climate projections

have a spatial resolution between 25 and 50 km, the EURO-CORDEX ensemble has a spatial resolution of ∼12.5 km. However, these resolutions still do not match the needs of local impact studies. For many practical applications, climate data is needed at the local scale, meaning that the mean amplitude, the temporal variability and the long term trends of the analyzed variables should be comparable with data obtained from point measurements.

A further increase in the temporal and spatial resolutions of the climate models is a grand challenge for the community. Besides the high demands on computational infrastructures, modelers have to adapt their models to the desired scales. This adaptation process includes improvements in the resolution of the surface boundary information, including the topography, the soil type, land cover and land use data sets (e.g. *Suklitsch et al.* (2011)). Also improvements in the physical parameterizations, like the implementation of a 3D turbulence scheme, a 3D radiation or an advanced precipitation scheme may be relevant. In addition the hydrostatic balance, a simplification made in many regional climate models, has to be revoked for higher resolved simulations.

Today only a few long-term climate model simulations at convection-permitting-scales (∼1- 2 km) are available (*Kendon et al.*, 2014). Most of these simulations are conducted over small domains and span only a few decades. Therefore impact modelers often depend on statistical downscaling methods. However, as discussed above, climate models do not necessarily detect changes that are caused by insufficiently resolved processes and statistical methods are usually not capable of capturing these deficits either.

Because of the high inhomogeneity of precipitation in time and space, precipitation is strongly smoothed when aggregated over large areas and time scales. To describe the average decrease of precipitation intensities due to spatial and temporal aggregation (*Bacchi and Ranzi*, 1996), areal reduction factors (ARFs) and intensity-duration functions (IDFs), have previously been used. These aggregation effects result in a strong scale dependency of precipitation dependent parameters and their associated impacts. Therefore a specification of the analyzed spatial and temporal resolution is needed for the assessment of precipitation extremes, for example as defined by an intensity threshold. Whereas precipitation extremes relevant to large catchments may be analyzed at a daily time scale, using climate projections conducted at ∼100 km horizontal resolution (*Arnbjerg-Nielsen et al.*, 2013), the analysis of urban drainage systems requires input data of at least 10 km spatial resolution and temporal resolutions from minutes to hours (*Arnbjerg-Nielsen et al.*, 2013; *De Toffol et al.*, 2009). To analyze soil erosion, precipitation information is needed at even smaller scales (*Mueller and Pfister*, 2011).

Besides the statistical aggregation effects that change the precipitation distributions depending on the analyzed scale, also the major physical processes that cause precipitation events may change. A common separation of precipitation events is made between convective and stratiform types. Convection is associated with local radiative surface heating that results in a buoyantly unstable atmosphere (*Houze*, 1997). Stratiform type precipitation stems from large-scale frontal systems and comparatively weak and uniform up-lifting. Because of the reduced spatial and temporal extent of convective type precipitation events, these types of events will be strongly affected by aggregation effects. Therefore highly resolved spatial and temporal scales are needed to analyze convective type extremes. First studies that use high-resolution model simulations suggested, that heavy precipitation at high temporal resolutions will increase strongly in a future climate and suppose that convective events contribute the most to this increase (*Kendon et al.*, 2014; *Muller et al.*, 2011; *Attema et al.*, 2014).

In this thesis two different approaches are used to compare precipitation distributions at different

scales. In the first article the two spatially differently resolved ensemble data sets ENSEMBLES and EURO-CORDEX are compared. Both the differences in the spatial patterns of the climate change signal, and the differences in the precipitation distribution are discussed. However, the identified differences between these two data sets can have a variety of reasons, that can not be clearly distinguished, e.g. new model developments and different greenhouse gas emission scenarios. The second article elaborates on the statistical differences of precipitation extremes between differently resolved data sets. For this study, a temporal and spatially highly resolved radar data set is aggregated in time and space. The data set is separately analyzed for different regions and seasons, as well as for convective and stratiform type events.

Regarding temporal and spatial scale dependencies of precipitation events, the following two research questions are addressed:

- How strong do temporal and spatial scale dependencies influence the intensity distributions of convective and stratiform type precipitation events, and at which temporal and spatial scales are extreme events dominantly convective type events?

- Has temporal aggregation similar effects on the precipitation distribution as spatial aggregation, and can a lack of spatial information be partly compensated by using high temporal information, or vice versa.

1.1.2 Uncertainties in climate model simulations

Besides the need for highly resolved data sets, also uncertainty estimates are needed in order to make informed decisions. According to *Hawkins and Sutton* (2009) (cf. *Foley* (2010)) these uncertainties can be grouped into three primary categories: scenario uncertainty, internal variability of the climate system, and model uncertainty.

Scenario uncertainties are mainly associated with human action caused by unknown greenhouse gas emissions and land use changes. Other factors are uncertainties from external forcing like volcanic eruptions or solar variability (*Foley*, 2010). Internal (natural) variability of the climate system, refers to deterministic and random fluctuations that are not connected to any forcing (*Hawkins and Sutton*, 2009). This uncertainty is also known as sampling uncertainty, because it results from a too small sample size. Climate is estimated by averaging over a finite number of years (commonly 20 to 30 years). This time-average may not always be sufficient to capture the natural variability (*Déqué et al.*, 2007). Model uncertainties are associated with incomplete understanding and simplified formulations. This category includes the above mentioned scale dependencies that cause downscaling uncertainties. Differences in the climate projections that are caused by any of these uncertainties can be amplified or diminished by climate feedback mechanisms (*Foley*, 2010). *Hawkins and Sutton* (2009) point out that the relative importance of each source of uncertainty varies with prediction lead time, and with spatial and temporal scales. Whereas most uncertainties come from the fact that future greenhouse gas concentrations are not known for predictions of the end of the century, model uncertainty and internal variability are the dominant contributions for predictions of the next few decades.

Most regional climate modeling studies, that analyze uncertainties use large model ensembles and consider different greenhouse gas emission scenarios (e.g. *Gobiet et al.* (2014)). The different ensemble members implement different methods to discretize the dynamic equations and different parameterizations to represent sub-grid scale effects. These differences results in an ensemble spread of the projected changes which can be used to approximate the sensitivity of the climate signal to possible model deficiencies, and thus to describe model uncertainties. Also the uncertainties caused by natural variability of the climate system are captured if different

GCM simulations are used as lateral boundary condition. However, knowledge gaps or insufficient temporal-spatial resolutions are only two examples that could lead to common errors in most models. Therefore uncertainties related to model errors cannot always be captured by the ensemble spread. The Alpine region is an example for an area where most of the models show similar biases. In this area, most models show a cold bias in the surface-temperature in a range from -0.8 to -1.9 K, and a wet precipitation bias between 14.8 % in summer and 41.5 % in winter (*Smiatek et al.*, 2016). This raised the following research question that is addressed in the fourth article.

- How strong are total-precipitation and snowfall projections influenced by these model biases and can this information be used to complement ensemble robustness tests.

1.2 Aims and objectives

In order to answer the research questions from section 1.1 the following aims and objectives were defined in the individual articles.

Aim 1: To analyze climate change projections of the EURO-CORDEX ensemble including the robustness and significance of the signal with a focus on precipitation (article #1).
Objectives:

- To analyze the EURO-CORDEX data set using different greenhouse gas emission scenarios, focusing on temperature and precipitation dependent parameters.

- To identify improvements and benefits of the spatially highly resolved simulations compared to earlier RCM ensembles.

- To analyze the robustness and the significance of the projected changes, and to combine this information with information about the amplitude of the signal.

Aim 2: To analyze the temporal and spatial scaling behavior of extreme precipitation events, and to quantify how strong convective type events contribute to the extremes (article #2)
Objectives:

- To quantify the reduction of extreme precipitation events due to temporal and spatial aggregation depending on the type of the precipitation event, the regional and seasonal characteristics, and the threshold used to detect extreme events.

- To identify how the differences in reduction factors affect the contribution of the convective and stratiform type to extreme events as a whole, dependent on the scale, region, season and the threshold chosen.

- To analyze the temporal and spatial scaling behavior of convective and stratiform extreme precipitation events, and to establish a connection between temporal and spatial aggregation effects.

- Assessing aggregation effects in the entire precipitation distribution to quantify the information loss, and to optimize data storage.

Aim 3: To assess the potential use of temporal and spatial scaling relations of precipitation to improve bias-adjustment methods, when only point measurements are available (article #3)
Objectives:

- To identify matching pairs of temporal and spatial resolutions that show the best agreement in the precipitation distribution.

- To perform a scale-adjustment for two exemplary cases in order to associate the scales of the model simulations with point measurements.

- To bias-adjust the model data before and after a scale-adaptation by applying quantile mapping, and to compare the performance between the scale-adapted bias-adjustment and the non-scale-adapted bias adjustment.

Aim 4: To quantify uncertainties in future climate projections of total-precipitation- and snowfall-frequencies, that result from common biases in the model simulations (article #4)
Objectives:

- Identify and analyze model biases in the temperature distribution and the temperature dependency of snowfall- and precipitation-frequencies.

- Analyze the impact of different snowfall discrimination methods on the temperature dependency of snowfall and its impact on the observed changes.

- Develop a probabilistic-decomposition-framework to analyze impacts of model biases on the past and projected changes of precipitation- and snowfall-frequencies, at different altitudes.

1.3 Research approach

To pursue the aims and objectives from section 1.2, the following research approaches are used.

Aim: 1 To analyze climate change projections of the EURO-CORDEX ensemble including the robustness and significance of the signal, with a focus on precipitation (article #1)

An ensemble analysis is performed using daily data of all land areas of the EURO-CORDEX domain. The ensemble includes as many RCM and GCM model combinations from the EURO-CORDEX database as available at that time, in order to capture the entire ensemble spread. All models are equally weighted and only models that show clear errors are excluded. The analysis is based on seven different RCMs that were driven by five different GCMs. To include uncertainties caused by the different greenhouse gas emission scenarios different RCPs are compared. Using different RCM-GCM combinations, in total nine RCP4.5 simulations, and ten RCP8.5 simulations are analyzed. To study how the ensemble mean is influenced by the domain size, and by the neglection of certain simulations, sensitivity studies are performed. To get a broad overview of the ensemble quality, different indices are analyzed. These indices include temperature, precipitation, heavy precipitation, dry spells and heat waves. Further indices are calculated for a set of sub-regions that were classified based on *Metzger et al.* (2005). All of these indices have a high impact on infrastructure, agriculture and human health.

To identify a possible added value due to the high resolution, the results of the ∼12.5 km EURO-CORDEX simulations are compared to the ENSEMBLES data set (*Hewitt and Griggs*, 2004). The ENSEMBLES data set consists of 20 transient RCM simulations that include simulations with a 50 km and 25 km grid size. A comparison of large scale climate change patterns is conducted. For a more detailed analysis of the differences in the spatial patterns of EURO-CORDEX and ENSEMBLES, a spatial correlation of the mean annual temperature and annual total precipitation is performd between EURO-CORDEX

and ENSEMBLES. To analyzed if the high spatial resolution of the RCMs improve the precipitation distribution, and if this has an effect on the projections, a comparison of the precipitation intensity distribution between the RCM's and the driving models is performed.

A method later published in *Pfeifer et al.* (2015) is adopted, to identify regions where the ensemble shows robust and significant changes. The robustness test regards the agreement of the simulations in terms of the direction of the changes. Regions in which more than 66% (according to the *Cubasch et al.* (2013) associated with a likely outcome) of the projections agree in direction of change are called robust. Whether a climate change signal is significant or not, is assessed by applying the two-sided Mann-Whitney-Wilcoxon test. This test is robust against outliers and does not require the assumption of normal distributions that is needed e.g. for the t-test.

Aim: 2 To analyze the temporal and spatial scaling behavior of extreme precipitation events and to quantify how strong convective type events contribute to the extremes (article #2)

The strong inhomogeneity of precipitation in space and time leads to a dependency of precipitation intensities on the temporal and spatial resolution of the data set. The higher precipitation extremes found in the spatially higher-resolved simulations, are identified as an important added value in the first article. For a more detailed comparison of precipitation events at different grid resolutions, knowledge about the scaling behavior of precipitation as well as the processes behind the events is needed. The second paper analyzes the scaling behavior of extreme precipitation events using radar data. For this kind of study it is most important to capture the shape, the size and the velocity of these events. Station observations are often too dispersed to sufficiently capture these properties. It has been shown that radar data deliver good results in deriving area reduction factors (*Bacchi and Ranzi* (1996); *Arnbjerg-Nielsen et al.* (2013)). The study is based on the RADOLAN-RY radar composite from the German Weather Service. This data set is provided on an 1 km horizontal and 5 min temporal resolution. To study how scaling impacts affect the intensity distributions, the radar grid points are aggregated in time, i.e., $\Delta t \in \{5, 10, 15, 20, 30, 45, 60, 120, 180, 240, 360\}$ min, and in space over square grid box areas with linear dimensions $\Delta x \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 25, 50\}$ km, including all possible pairs $\{t, x\}$. Using only one data set that is aggregated to different scales makes it possible to analyze intensity changes that only occur due to the change in grid resolution. Like for the first article, different percentile thresholds are used to identify precipitation extremes. Compared to a fixed threshold, this has the advantage that the analysis can be applied to different scales and over large regions, that show different precipitation characteristics.

An objective of this study is the analysis of the different scaling behaviors of convective and stratiform type events and to identify at which resolution the transition from stratiform to convective type dominated extremes occurs. In order to analyze this transition the data is split in stratiform and convective type conditions. This separation is done using cloud observations obtained from the Met Office Integrated Data Archive System (MIDAS). After the data separation, the number of convective and stratiform events with intensities above a common percentile threshold are counted. This allows to assess the relative likelihood of a certain precipitation type to cause extreme precipitation.

A comparison of the area against the time reduction factors gives information about the relevance of space compared to time aggregation. By matching identical precipitation

extremes the study associates pairs of temporal and spatial resolution. Dividing the spatial resolution by the temporal resolution defines a velocity (v_{eff}), which is used to generalize the Taylor hypothesis (*Taylor et al.* (2012)). Presuming a precipitation event with a constant intensity in the analyzed time frame, v_{eff} would be scale independent. Deviations from this assumption can result in situations where temporal scales change disproportionately strong compared to spatial scales, and v_{eff} becomes scale dependent (self-affinity; *Deidda* (2000)).

Changing the resolution has an impact on the entire distribution function. If a dataset has to be transferred to a lower resolution, e.g. to save storage space, it is important to know how strong the entire distribution will be affected. To provide an estimate of the information loss due to the aggregation process, a measure similar to the Perkins skill score (*Perkins et al.* (2007)) is used. This measure allows a quantification of the overlap between two intensity probability density functions (PDF) at different horizontal and temporal resolutions. The overlap value is referred to as PDF overlap.

Aim: 3 To assess the potential use of temporal and spatial scaling relations of precipitation, to improve bias-adjustment methods when only point measurements are available (article #3)

With increasing model resolution, gridded observations at a sufficient spatial resolution are not always at hand. Often, only point measurements exist with a poor spatial coverage. The basic idea of this study is that information about the spatial variability in the surrounding area of a gauge station is contained in the temporal variability. Therefore, lowering the temporal resolution can be a good estimate to mimic an area average. This procedure is called scale adaptation.

That temporal and spatial scales can be connected to compensate missing spatial information with an increased temporal resolution or vise versa, is an observation already made by *Taylor* (1938), known as the Taylor hypothesis of frozen turbulence. The identification of these matching pairs is a main objective in order to pursue the third aim of this thesis. For this objective, the procedure used to identify matching pairs, that was developed in the second article, is optimized. Three different methods are tested to measure the PDF-agreement: The intensity-weighted and the non-intensity-weighted PDF-overlap, as well as the Kolmogoroff-Smirnov test. For further analysis, the intensity-weighted PDF overlap is applied. The intensity weighting gives more emphasis to high precipitation intensities but ensures that the entire distribution is considered. First, the detection is done using observational data. As detailed information about the temporal and spatial scales of precipitation events are needed, the aggregated radar data sets from the second article are taken. In a second step the process is repeated using a regional climate model simulation. For this step, COSMO-CLM (*Doms and Schättler* (2002)) simulations driven with ERA-Interim (*Dee et al.* (2011)) are used. The simulation is provided at a 7 km horizontal and 1 h temporal resolution.

To assess how scale adaptation can improve the results of bias-adjustment methods, the regional model simulation is bias-adjusted in three different ways, all based on empirical quantile mapping (*Gudmundsson et al.* (2012)). In article three, the term "bias correction" is used. However recent discussions within the EURO-CORDEX community suggest to speak of bias-adjustment instead, to emphasize that the corrected data is not free from errors. For the bias-adjustment procedure, a dense rain gauge network over southwestern Germany is applied. Besides the good station network the area is particularly well suited

because it covers topographically less variable sites in the Rhine Valley as well as more complex topographic regions in the Black Forest. An eight year time period with the most dense station network is selected (1997-2004). Stations with more than 10% missing data are discarded. Because the model is driven with ERA-Interim data, a close match between the model and the simulations is ensured.

The first adjustment is designed as a benchmark. An entire station network is taken to analyze the performance of the bias-adjustment procedure. For this benchmark gridboxes with less than three stations are not considered. This first adjustment is assumed to be the optimal solution, with enough stations to capture the spatial variability within the gridbox. The second adjustment takes only a single station per gridbox to identify the performance if only one station is available. The third adjustment also takes only one station, but the quantile mapping is applied after the scales of the station and the simulation are adjusted. After the quantile mapping, all three data sets are compared to the station network. Besides the PDF-agreement also the 99.9th percentile and the dry period fraction are compared. Two different test cases are calculated. In the first case the model data is aggregated to 28 km spatial and 1 h temporal resolution. In this case the data is bias-adjusted using 1 h gauge data. For the second test case the model data is aggregated to a 77 km spatial and 1 h temporal resolution and the gauge station is assumed to have only daily data.

Aim: 4 To quantify uncertainties in future climate projections of precipitation- and snowfall-frequencies that result from common biases in the model simulations (article #4)

Precipitation and especially snowfall are key elements of the Alpine climate system. Changes in this parameters will have a high environmental and economic impact. It is known that most regional climate models show a cold temperature bias in the region, and overestimate precipitation especially in the winter season. An objective to pursue aim four is to identify how these biases modify the temperature distribution, and to analyze if the temperature dependency of snowfall- and precipitation-frequencies is subject to model biases. Assuming a future temperature increase it can be expected that a bias in the temperature dependency will impact the projected changes. This objective is assessed by comparing the ensemble data to observations. For this comparison the ensemble median as well as the ensemble spread (5th to 95th percentile) is applied to analyze the ensemble agreement and to identify if the observations are within the range of the ensemble. For both parameters two different thresholds are used to analyze if heavy events deviate from the average. The temperature dependency of an event is identified by calculating the probability of the event occurrence at different temperatures.

Gridded snowfall observations at a daily temporal resolution are not available for the studied region, therefore snowfall has to be estimated. Different approaches exist to determine the precipitation phase (solid or liquid). Because atmospheric data is often not available at different heights, the most common methods estimate snowfall by applying an empirical relationship between snowfall-fractions and surface-temperature (*Ye et al.*, 2013). However, including information about the atmospheric conditions that act on the falling hydrometeors e.g. relative humidity and air-temperatures at different heights, has a large potential to improve these estimates *Feiccabrino et al.* (2015). Depending on the discrimination method a different percentage of total precipitation will be counted as snow. This differences are especially high at temperatures close to 0 °C where the most and the heaviest snowfall is to be expected. Therefore the second objective is, to asses how different discrimination methods modify the temperature dependency of snow-

fall and how this effects past and projected climate changes. To analyze these differences two substantial different discrimination methods are used. The first method is a simple surface-temperature threshold scheme derived from snow-flux observations (*Feiccabrino et al.*, 2013), that proved to show good results over large regions of northern Europe (*O'Gorman et al.*, 2014). This method is applied to the Alpine gridded precipitation data set EURO4M-APGD (*Isotta et al.*, 2013) using surface-temperature data from E-OBS (*Haylock et al.*, 2008) and HISTALP (*Chimani et al.*, 2011). The two different temperature observations allow to account for uncertainties resulting from scale mismatches between the different data sets. The second method uses snowfall statistics (snowfall-intensity-fraction and snowfall-frequency-fraction) from the model ensemble. It includes a dependency on precipitation intensity and altitude.

Because clear differences in the surface-temperature and precipitation distributions are found between the model ensemble and the observations, a third objective is to develop a probabilistic-decomposition-framework to analyze the impact of these biases on the climate change signal of precipitation- and snowfall-frequencies. The framework is based on the law of total probability. The occurrence probability of an event is decomposed into the probability that an event will occur at a specific surface-temperature (temperature dependency) and the probability that a specific surface-temperature will occur (temperature distribution). On the basis of this distinction, frequency changes of precipitation and snowfall can be directly related to changes in the temperature distribution. The spatially highly resolved precipitation data set covers the period 1971-2008. Splitting this period into two nineteen year time slices (1971-1989 and 1990-2008) allows to compare observed changes against simulated changes. Differences are identified and related to the observed model biases to quantify their impacts. In total, the study separates between eleven different terms that together explain the different changes in the simulations and the observations. After analyzing the model biases and their impacts on past climate changes, projected climate changes are analyzed by the end of the century. The analysis is conducted with and without bias-adjustment of the temperature distribution and the temperature dependency of the events, in order to identify how the bias in the distributions affects the projected changes.

1.4 Results and discussion

Aim: 1 To analyze climate change projections of the EURO-CORDEX ensemble including the robustness and significance of the signal with a focus on precipitation (article #1)

Article one is the first ensemble study that analyzes the EURO-CORDEX data set using different greenhouse gas emission scenarios (Representative Concentration Pathways *Moss et al.* (2010)). For RCP8.5 and RCP4.5 similar change patterns are found for surface-temperature and precipitation, with more pronounced changes in RCP8.5. Surface-temperature projections indicate a greater warming in Southern Europe and towards the north east in the range of 2.5-5.5 °C for RCP8.5, and between 1.0-4.6 °C for RCP4.5. An increase in mean annual precipitation is found for most parts of central and northern Europe and a decrease in the south. The climate projections indicate a reduction of precipitation intensities below 9 mm/day and an increase for all intensities above this threshold. This shift towards more intense precipitation events is most noticeable in RCP8.5. The analysis of impact indices shows that for almost all indices the changes are stronger in RCP8.5 than for RCP4.5. Also more pronounced differences are found for temperature-based indices

than for precipitation-based indices. Changes in dry spells show substantial differences between the two scenarios only by the end of the century. For heat waves stronger differences between RCP8.5 and RCP4.5 are identified. However, the definition of heat waves is found to dominate regional change patterns more than the differences in scenario or time period.

An objective of this study is to identify improvements and benefits of the high-resolution of EURO-CORDEX compared to earlier ensemble data sets. Overall comparisons against the ENSEMBLES data set (A1B scenario) show similar spatial patterns for temperature and precipitation changes as well as for all related indices. The spatial correlations of temperature and precipitation changes between RCP8.5 and A1B are very high. For the mid of the century correlation coefficients between 0.82-0.97 for temperature changes and between 0.59-0.92 for precipitation changes are found. The magnitude of temperature and precipitation changes for the A1B scenario are mostly in-between RCP4.5 and RCP8.5. However, the new EURO-CORDEX ensemble shows more detailed spatial patterns, related to better-resolved physical processes like convection and heavy precipitation and to better representation of surface characteristics and their spatial variability. Differences are especially found for heavy precipitation changes. The RCPs project strongest heavy precipitation changes of up to 35 % in central and eastern Europe, whereas A1B only projects changes of up to 25 % within this region. Comparisons between EURO-CORDEX and the driving GCM ensemble show that the RCMs provide higher daily precipitation intensities, which are not observed in the GCM simulations. The higher precipitation intensities add value to the projections, because of the better representation of the right tail of the precipitation distribution results in a different climate change of daily precipitation intensities, with a smoother shift from weak to moderate and high intensities.

Validation of the robustness and the significance of the projected climate change signal shows that a statistically significant and robust warming is found for all land parts of Europe. For mean annual precipitation the increase for most parts of central and northern Europe as well as the decrease over southern Europe is found to be mostly robust and significant. Between these two different patterns a transition zone is located, where changes are neither robust nor significant. Projected changes in the 95th percentile of the length of dry spells are mostly robust but significant only in parts of southern Europe. Changes in heat waves are mostly robust and significant, except of parts in northern Europe where the robustness and significance of the change considerably depend on the definition of the indices.

Aim: 2 To analyze the temporal and spatial scaling behavior of extreme precipitation events and to quantify how strong convective type events contribute to the extremes (article #2)

A main objective is to quantify the reduction due to temporal and spatial aggregation of extreme precipitation events. The analysis is conducted depending on the precipitation type, the regional and seasonal characteristics and the percentile used to detect extreme events, which has not been done before. Compared to stratiform events, convective extremes show clear regional and seasonal differences, with the strongest extremes in southern Germany during summer. Analyzing the duration (DRF) and the area (ARF) reduction factors, up to 30% higher values are found for convective compared to stratiform extremes, exceeding all other observed seasonal and regional differences within one type. After the separation of convective and stratiform events, regional and seasonal differences are only observed in the area reduction factors of convective type events. This proves the importance of distinguishing between these events, e.g. for statistical downscaling

exercises.

Analyzing the contribution of the convective type to extreme events as a whole, it is found that the contribution is strongest in summer and over southern Germany. Besides the expected seasonal and regional differences, also the resolution and the intensity threshold show an impact. Convective type events contribute more to the extremes when the resolution or the intensity threshold is increased. At a fixed horizontal resolution of about 10 km (∼EURO-CORDEX grid size) changing the temporal resolution or the intensity threshold can change the dominant precipitation type accountable for the extremes.

Another objective of article two is to analyze the temporal and spatial scaling behavior of convective and stratiform extreme precipitation events, and to establish a connection between space and time. By matching identical precipitation extremes, the study associates pairs of temporal and spatial resolution, which define an effective velocity defined as (v_{eff}). For constant v_{eff} the Taylor hypothesis would be obeyed. However, v_{eff} of convective and stratiform extreme precipitation decreases with increasing spatial scale, with similar exponents for both precipitation types. The main scaling difference between convective and stratiform events can be described by a constant scaling factor. This scaling factor leads to about 1.75 times higher advection velocities for stratiform than for convective events. Analyzing the associated temporal resolution at horizontal scales close to the EURO-CORDEX grid size, shows that temporal resolutions of approximately 20 to 25 min are needed in order to avoid an imbalance between duration and area reduction effects.

To assess aggregation effects in the entire precipitation distribution, the PDFs from different temporal and spatial resolutions are compared with each other by calculating the overlap of these PDFs hereafter referred to as PDF-overlap (*Perkins et al.*, 2007). It is found that the impact of temporal aggregation strongly depends on the spatial scale of the data and vice versa. For example, PDF changes that are observed when the temporal resolution is decreased from 5 min to 2 h at 50 km horizontal resolution are quantitatively comparable with PDF changes when going from 5 to 30 min at 10 km horizontal resolution, or from 5 to 10 min at 2 km horizontal resolution. Comparisons between a certain reference resolution to all other aggregated resolutions show a ridge with PDF-overlap values close to 1 (values go from 0 to 1 with 1 being a perfect match). Using a 60 min temporal and 10 km spatial reference resolution the ridge ranges from 5 min and 25 km to 120 min at 1km resolution for convective type events, and from 5 min and 25 km to 90 min at 1 km resolution for stratiform events. This proves that associated pairs of temporal and spatial resolution can also be found, when the entire distribution is considered.

Aim: 3 To assess the potential use of temporal and spatial scaling relations of precipitation, to improve bias-adjustment methods when only point measurements are available (article #3)

Article three analyzes if pairs of temporal and spatial resolution, that show similar aggregation effects (see article two), can be used to improve bias-adjustment methods. The first objective is to optimize the procedure to identify these pairs for the application of bias-adjustment and to apply the procedure in two case studies. Three different methods are tested to measure the PDF-agreement. The intensity-weighted and the non-intensity-weighted PDF-overlap as well as the Kolmogoroff-Smirnov test. It is found that all methods give similar results, albeit with varying degree of noise. The intensity-weighted PDF-overlap showed the best performance. Therefore, it is used for all further evaluations and

revered to as PDF agreement.

To test whether the scale-adjustment improves bias-adjustment methods, two different cases are analyzed. In the first example, model data at a 28 km spatial and 1 h temporal resolution is bias-adjusted using station data at a 1h resolution. In this case, the scale-adjustment indicates that the temporal resolution of the station data needs to be aggregated to 3 h resolution, in order to mimic an area average of the 28 km model grid (highest PDF-aggrement). Because the temporal resolution of the station data needs to be reduced, this case is referred to as model-limited correction. For the second case a coarser model resolution of 77 km spatial and 1 h temporal is assumed and bias-adjusted using daily station data. Because of the low temporal resolution of the station data, in this case the model data has to be aggregated to a 19 h temporal resolution (Gauge-limited).

Quantile mapping is applied to scale-adjusted and non-scale-adjusted data, to analyze possible improvements of the scale-adaptation. First, an adjustment using at least three stations in each gridbox is performed as a benchmark. This represents the "ideal correction" obtainable when sufficient stations were available to calculate a spatial average. A good agreement is found between the bias-adjusted data and the observations with a PDF agreement above 95% for all gridboxes. In the next step, bias-adjustment is performed assuming that only a single station is available for bias correction. In the Model-limited case, this correction significantly shifts the original distributions to more extreme intensities. The point information is spatially too highly resolved for the model grid, causing unrealistic extremes. Repeating the bias-adjustment using the aggregated 3 h station observations results in an increase $> 10\%$ in the PDF agreement between the observed gridbox average and the corresponding bias-adjusted model values. In addition, improvements are found in the dry period fraction and the extremes. Analogous improvements in all three parameters are also shown for the Gauge-limited case. In that case, the model data can be disaggregated back to hourly data after the bias-adjustment. It is found that especially when gauge density is low scale-adjustment may allow for substantially improved bias correction at essentially no cost in terms of model output, data storage, or mathematical complexity.

Aim: 4 To quantify uncertainties in future climate projections of total-precipitation- and snowfall-frequencies that result from common biases in the model simulations (article #4)

Within the Alpine region, regional climate models are known to contain common biases in the simulated surface-temperature and precipitation fields. The fourth article develops a new framework to analyzes how these errors impact the temperature distribution, and the temperature dependent occurrence distribution, of snowfall- and precipitation-frequencies.

Analyzing the surface-temperature distribution, it is found that the center of the observed distribution is, compared to the simulations, shifted towards higher temperatures. Also, the shape of the distributions differ. The simulated distributions are stronger skewed to the right with a more pronounced peak at 0 °C. The overestimation of days with a mean surface-temperature of 0 °C could be caused by an overestimation of snowfall days as well as the overall cold bias in the region, that may lead to an enhanced snow-cover.

For the surface-temperature dependency of precipitation-frequencies ($P(I|T)$) three important observations are made. First, the slope of the $P(I|T)$ distributions is not constant. Therefore, the sensitivity of the precipitation-frequency on temperature changes depends on the temperature distribution itself. Second, within the temperature range of $\sim$0 °C

to 12 °C, the slopes of the observed and simulated $P(I|T)$ distributions point to opposite directions. This modifies the response of precipitation changes to changes in the temperature distribution. Third, a kink between 0 to 1 °C is found in the model ensemble that is not present in the observed distribution. This kink could be caused by parameterizations that describe interactions between liquid- and ice-cloud-content and should be further investigated.

An objective of the study is to analyze the impact different snowfall discrimination methods have on the temperature dependency of snowfall and to quantify resulting uncertainties in the observed climate change signal. Depending on the discrimination method, the temperature range with the highest snowfall occurrence shifts from -2 °C using snowfall-fraction statistics from the model ensemble, to 0 °C when a surface-temperature threshold scheme developed by (*Feiccabrino et al.*, 2013) is applied. The shift towards higher temperatures clearly attenuates the projected snowfall decrease. Further, comparisons against observations indicate that the models likely underestimate snowfall-fractions in a temperature range close to 0 °C. An explanation for this underestimation may be missing information about the sub-grid orography. The model grid only provides an average altitude for each gridbox. Parts of the area within one gridbox will, however, be located higher as the mean which will likely result in colder temperatures and higher snowfall probabilities.

A further objective is to develop a framework to analyze the impact of the observed model biases on the climate change signal of precipitation- and snowfall-frequencies. This task is addressed in two steps. First, the simulated past climate change signal is compared to observations. Second, future climate projections are analyzed with and without bias-adjustment of the distributions.

Analyzing changes in the surface-temperature probability distribution between the periods 1971-1989 and 1990-2008, it is found that the simulations underestimate a reduction that is observed in the probability distribution between -5 to 3°C as well as an increase above 12 °C. Also no clear changes in the temperature dependent occurrence probability of total-precipitation events are simulated, whereas the observations show an increase at temperatures above 2 °C and decreasing below. For past changes in total-precipitation-frequency, the observations indicate an increase above the spread of the model ensemble at most altitudes. On average the observations show an increase of ∼3 % for precipitation- and ∼9 % for heavy precipitation frequencies. The precipitation increase in the simulations is close to 1 % for both intensity thresholds. The main cause for these differences is the cold temperature bias. The different climate change signal in the temperature distribution played only a minor role. Because of this cold bias, temperatures above ∼10 °C occur less often in the simulations. Precipitation-frequencies are however especially increasing at days with mean surface temperatures above ∼10 °C. Past snowfall-frequency changes were mainly driven by temperature changes. For heavy snowfall-frequencies also non-temperature dependent changes become important.

The study highlights that the different snowfall discrimination methods clearly influence the observed snowfall-frequency changes. The spread between the different observational data sets is in the same order of magnitude as the changes themself. Compared to observations, the simulations underestimate the decrease in snowfall-frequency at all altitudes. At low altitudes these differences are mainly caused by an underestimation of the decrease in the precipitation probability at cold temperatures. At high altitudes the different changes in the temperature distribution are the main reason. It is discussed that the enhanced

reduction of days close to 0 °C may be a result of feed-back processes with a decrease in snow cover. The high precipitation intensities and the cold model bias might have prevented this feedback loop in the simulations.

Future projections show no clear signal for the change in the precipitation-frequency. The ensemble mean increases less than 1 %. Heavy precipitation-frequencies are projected to increase by about 10 %. In both cases the bias in the temperature dependency of these events results in a decrease in the event frequency caused by the projected warming. After the bias-adjustment the temperature-precipitation relation is changing which results in a stronger total increase in precipitation- and heavy precipitation-frequencies.

Like past snowfall-frequency changes, future projections indicate that changes in these indices are mainly driven by temperature changes. Stronger changes are projected for snowfall-frequencies as for heavy snowfall-frequencies. The main reason for this is an increase in the temperature dependent occurrence probability for heavy events, that counteracts the warming induced decrease of snowfall events.

1.5   Conclusion

The first article highlights a strong agreement between the large scale climate change patterns of the new high-resolved EURO-CORDEX ensemble and the ENSEMBLES data set. The finding that neither model improvements, changes in grid size, nor the use of different greenhouse gas scenarios lead to drastic changes in these patterns increases the credibility of the results.

Tests for significance and robustness of the projected changes reveal that in most areas where significant changes are detected these changes are also robust. This indicates a strong agreement between the different climate models in the direction of the projected changes. However, it is important to clarify that the robustness test does not include all sources that could cause uncertainties in the projected changes. The fact that most models agree gives comfort in the results, but it does not necessarily mean that the models are right or that the results are also valid at different temporal and spatial scales.

In order to apply climate projections for decisions-making, knowledge about the inherent uncertainties is essential (*Foley*, 2010). Article one is a first step towards providing this additional information together with the climate signal. Results of this thesis suggest that additional information about uncertainties and scale dependencies can be gained from the EURO-CORDEX ensemble, which should be further researched and included in new developed products, in order to increase the value of climate projections for impact-research and decision-makers. The following conclusions about guidelines of data usability, as well as the uncertainties arising from common model biases suggest that the definition of the indices, as well as the temporal and spatial scales of the provided data sets, have to be tailored to the questions asked.

Guidance for data usability

Two important findings concerning the usability of the projections are highlighted. It is found, that care has to be taken when the data is transformed to different temporal and spatial scales. Article one and two both indicate that changes in the resolution of the data set impact the intensity distribution of precipitation. Significantly different changes are projected by the RCMs compared to the lower resolved GCMs. Also, it is found that the ratio between convective and stratiform type extremes is changing depending on the resolution of the data set. The impacts of extreme precipitation events vary enormously with scale (*Blöschl and Sivapalan*

(1995); *Arnbjerg-Nielsen et al.* (2013); *Mueller and Pfister* (2011)). As climate model output is mostly not available at the desired scales, the simulations are often statistically downscaled. It is questionable whether these procedures are able to reflect scale dependencies in the projected changes. Especially for precipitation that is strongly inhomogeneous in time and space, it should be considered to complement the information about future climate projections with a scale range of validity in order to prevent misuse.

A second finding is that the projected changes of several parameters indicates a dependency on their definitions. In the first article the definition of a heat wave dominated the regional change pattern. In the second article, the threshold used to define precipitation extremes influenced the ratio of convective vs. stratiform extreme events and in the fourth study the snowfall discrimination method clearly modified the change signal of snowfall-frequencies. The results highlight the importance of clearly defining the index that needs to be tailored to the sector of interest. For the tourism sector it might for example be more sufficient, if days with sleet are not counted as snow days, even though the amount of snow exceeds a certain threshold.

Complementing the robustness test with additional model uncertainties resulting from common model biases

An analysis of past changes within the Alpine region demonstrated that for certain parameters the observations are outside of the model spread (article 4). Common model errors are a major source for these differences. In earlier studies, little focus is set on the question how errors in the models and the observations effect regional climate change signals. However, especially for climate change analysis for the near future, or under a fixed global warming threshold, these errors may be of utmost importance to estimate the underlying uncertainties.

Detection of model biases

To detect model biases and their possible impacts on the climate change signal, articles one, two and four stress the importance of analyzing the entire distribution of the indices in question. In many cases, changes in the shape or at the tails of the distribution can have a stronger impact on society as changes in the mean. Further, it is found that a correlation analysis between different parameters is needed to detect model deficiencies, e.g. a correct representation of the temperature dependency is identified to be a key issue for changes in precipitation dependent parameters.

Bias-adjustment is mostly applied in order to make climate model results usable for impact analysis. Article four emphasizes the utility of bias-adjustment methods for model development by using the adjusted data sets to quantify the impacts of the identified biases. However, for this kind of application it is important that deviations in the projected changes can be directly related to a certain change made by the adjustment method. Many bias-adjustment methods are applied as a "black box", making it impossible to explain resulting impacts.

Nexus of time and space

A major part of the analysis in article two and three investigates the connections between temporal and spatial aggregation effects of (extreme) precipitation. From these studies, the following two conclusions are drawn.

First, aggregation in time has similar effects on the intensity distribution of precipitation than aggregation in space. Therefore it is possible to use highly resolved temporal information to estimate spatial variability and vice versa. This needs to be kept in mind when bringing data sets with different temporal and spatial scales together. In article two it is shown that even in

regions with moderate topographic variability, point measurements can be adjusted to the scale of gridded data sets by taking a lower time average to mimic the area mean. However, this method can only be used to adjust statistical changes in the intensity distribution that apply over the entire gridbox area. Spatial information about the sub-scale variability e.g. caused by strong orographic gradients as well as temporal patterns like the diurnal cycle cannot be restored or can even get lost by the averaging.

Second, the smoothing effect caused by temporal aggregation depends on the spatial scale of the data set and vice versa. The good agreement in the change signal between the EURO-CORDEX ensemble and the ENSEMBLES data set is partly a result of the low temporal resolution used. By applying daily data most of the high spatial variability will be smoothed. The analysis indicates that the temporal resolution of regional model output needs to be reconsidered. The temporal output of RCMs is often low compared to the horizontal grid size. For variables like precipitation, that show strong inhomogeneities in time, this results in a pronounced loss of information about the variability of these events.

1.6   Future perspectives

This thesis focused on the quantification of uncertainties and scale dependencies of observed and simulated precipitation dependent indices. The analysis are mostly exemplary and conducted for specific regions and variables. Extending the analysis using additional data can improve confidence in the results. Most of the methods and the results can likely be transferred to other regions and variables. The scaling analysis from the second article for example could also be applied to other variables with a strong inhomogeneity in space and time e.g. wind-speed or tracer concentrations. In addition to the expansion of the analysis mentioned above, the following suggestions for future research are proposed, based on the work presented in this study.

Postprocessing information to aid in decision-making

Adding information about the robustness and the significance to the results, is an important step towards postprocessing climate projections for impact-modelers and stakeholders (e.g. *Foley* (2010)). However, this thesis highlights, that further information about scale dependencies and model uncertainties can be gained from the RCM ensembles. In the first article, the climate change signal is detected to be robust if 66% of all models agree in the direction of change. This method does not give any information about whether the projections can also be trusted quantitatively. Including a quantitative robustness test, that is defined over the ensemble spread, could make the information more versatile. As the robustness test only checks for the direction of change, the meaning of the test becomes less clear in areas where only minor changes are projected. In these areas the robustness test is often rejected although most models agree that the changes are small. A solution for this problem might be to consider all areas as robust, where 66% of all models agree that the changes are not significant. However, a quantitative approach of the robustness-test is likely a more sophisticated solution.

The analysis of the first and the second article shows that the climate change signal of precipitation is likely to be scale dependent. If this is true, providing an information about the valid scales at which the information from a model projection can safely be used, should be thought about. The studies indicate that a better representation of the precipitation intensity distribution at spatially higher resolved data sets, has a major impact on the climate projections. Therefore, it could be possible to estimate the valid scales as scales with similar intensity distri-

butions. However these scales may differ for different variables, regions, and seasons. Further analysis is needed to test if this behavior can also be detected for other variables.

Article four demonstrates, that common model biases have the potential to strongly modify the climate signal of the entire ensemble. It should be considered to include this information in the robustness test e.g. by marking areas with strong biases. In extreme cases, these areas might need to be excluded from the analysis.

Analysing scale dependencies in the climate change signal of precipitation extremes

Article two analyzes the intensity distributions of convective and stratiform type precipitation events at different scales. It is found that the contribution of convective type events to the extremes is changing depending on the temporal and spatial scale, the region and the season. Based on these results the following next steps are proposed. First, it should be analyzed if the changing contribution of convective events is impacting the observed temperature scaling of precipitation events. This would be expected following the results from *Berg et al.* (2013). Secondly, the results of this analysis should be confirmed using climate model simulation that were conducted at different grid sizes. For this test the temporal output needs to be very high in all simulations in order to avoid stronger duration than area reduction effects. Most studies that compare highly resolved data sets still use hourly or even daily temporal resolutions that prove too be to coarse for this kind of analysis.

Nexus of time and space

The different temporal and spatial aggregation effects on convective and stratiform type events make it difficult to statistical downscale precipitation. Assumptions have to be made about the size and the duration of precipitation events as well as about the sub-grid temporal and spatial variability. For most impact-assessments precipitation data is needed at scales well below the gridsize of most climate model simulations. Downscaling precipitation data while the model is running would have two major benefits. First, the very high temporal information at run time can be used to estimate the spatial variability of the event. Secondly (not subject of the articles), most models calculate the area of precipitation within a grid-box at each model time step in order to parameterize evaporation of falling rain. Using this information about the size of the event would be consistent with the model physics.

The similar impact of temporal and spatial aggregation on the precipitation distribution is also used in the second article, to adjust the scales between point measurements and model data. This method could further be improved if the different aggregation effects of convective and stratiform type events are considered. Further it should be investigated if this information can also be used to improve very highly resolved gridded observation data sets.

Additional model output

This thesis highlights the benefits of an improved representation of the precipitation distribution for climate change analysis. Further it is shown that temporal aggregation has a strong impact on the distribution, as precipitation events are strongly inhomogeneous in time. Increasing the temporal output frequency of climate models leads to a considerable increase in data volume which makes it difficult to store and handle the files. Introducing a new model output format, that only stores information, e.g. in daily or monthly histograms, would be a solution to store information about the temporal variability with a comparatively low increase in the required storage space. Information on the chronological sequence including e.g. the diurnal cycle, are not captured. However, using histograms, sub-hourly information could effectively be stored and processed, which would be a large benefit for studies that analyze extreme events.

Bibliography

Arnbjerg-Nielsen, K., P. Willems, J. Olsson, S. Beecham, A. Pathirana, I. Bülow Gregersen, H. Madsen, and V.-T.-V. Nguyen, Impacts of climate change on rainfall extremes and urban drainage systems: a review., *Water Sci. Technol.*, *68*(1), 16–28, doi:10.2166/wst.2013.251, 2013.

Attema, J. J., J. M. Loriaux, and G. Lenderink, Extreme precipitation response to climate perturbations in an atmospheric mesoscale model, *Environ. Res. Lett.*, *9*(1), 14,003, doi:10.1088/1748-9326/9/1/014003, 2014.

Bacchi, B., and R. Ranzi, On the derivation of the areal reduction factor of storms, *Atmospheric Research*, *42*(1-4), 123–135, doi:10.1016/0169-8095(95)00058-5, 1996.

Berg, P., C. Moseley, and J. O. Haerter, Strong increase in convective precipitation in response to higher temperatures, *Nature Geoscience*, *6*(3), 181–185, doi:10.1038/ngeo1731, 2013.

Blöschl, G., and M. Sivapalan, Scale issues in hydrological modelling: A review, *Hydrological Processes*, *9*(3-4), 251–290, doi:10.1002/hyp.3360090305, 1995.

Chimani, B., R. Böhm, C. Matulla, and M. Ganekind, Development of a longterm dataset of solid/liquid precipitation, *Adv. Sci. Res.*, *6*, 39–43, doi:10.5194/asr-6-39-2011, 2011.

Cortekar, J., S. Bender, M. Brune, and M. Groth, Why climate change adaptation in cities needs customised and flexible climate services, *Clim. Serv.*, *4*, 42–51, doi:10.1016/j.cliser.2016.11.002, 2016.

Cubasch, U., D. Wuebbles, D. Chen, M. C. Facchini, D. Frame, N. Mahowald, and J.-G. Winther, Introduction in Climate Change 2013, *Intergov. Panel Clim. Chang. 2013 Phys. Sci. Basis. Contrib. Work. Gr. I to Fifth Assess. Rep. Intergov. Panel Clim. Chang.*, pp. 119–158, doi:10.1017/CBO9781107415324.007, 2013.

De Toffol, S., a. N. Laghari, and W. Rauch, Are extreme rainfall intensities more frequent? Analysis of trends in rainfall patterns relevant to urban drainage systems., *Water Sci. Technol.*, *59*(9), 1769–76, doi:10.2166/wst.2009.182, 2009.

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart, The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597, doi:10.1002/qj.828, 2011.

Deidda, R., Rainfall downscaling in a space-time multifractal framework, *Water Resources Research*, *36*(7), 1779–1794, doi:10.1029/2000wr900038, 2000.

Déqué, M., D. P. Rowell, D. Lüthi, F. Giorgi, J. H. Christensen, B. Rockel, D. Jacob, E. Kjellström, M. Castro, and B. Hurk, An intercomparison of regional climate simulations for

Europe: assessing uncertainties in model projections, *Clim. Change*, *81*(S1), 53–70, doi: 10.1007/s10584-006-9228-x, 2007.

Di Luca, A., R. Elía, and R. Laprise, Potential for small scale added value of RCM's downscaled climate change signal, *Clim. Dyn.*, doi:10.1007/s00382-012-1415-z, 2012a.

Di Luca, A., R. Elía, and R. Laprise, Potential for added value in temperature simulated by high-resolution nested RCMs in present climate and in the climate change signal, *Clim. Dyn.*, doi:10.1007/s00382-012-1384-2, 2012b.

Doms, G., and U. Schättler, A description of the nonhydrostatic regional model LM, Part I: Dynamics and numerics, *2*, 225–235, 2002.

EURO-CORDEX community, R. Benestad, A. Haensler, B. Hennemuth, T. Illy, D. Jacob, E. Keup-thiel, S. Kotlarski, G. Nikulin, J. Otto, D. Rechid, K. Sieck, S. Sobolowski, P. Szabó, G. Szepszo, C. Teichmann, R. Vautard, T. Weber, and G. Zsebeházi, *Guidance for EURO-CORDEX climate projections data use*, available at http://www.euro-cordex.net/imperia/md/content/csc/cordex/euro-cordex-guidelines-version1.0-2017.08.pdf, version: 1.0, 2017.

Feiccabrino, J., D. Gustafsson, and A. Lundberg, Surface-based precipitation phase determination methods in hydrological models, *Hydrology Research*, *44*(1), 44, doi:10.2166/nh.2012.158, 2013.

Feiccabrino, J., W. Graff, A. Lundberg, N. Sandström, and D. Gustafsson, Meteorological Knowledge Useful for the Improvement of Snow Rain Separation in Surface Based Models, *Hydrology*, *2*(4), 266–288, doi:10.3390/hydrology2040266, 2015.

Foley, A. M., Uncertainty in regional climate modelling: A review, *Prog. Phys. Geogr.*, *34*(5), 647–670, doi:10.1177/0309133310375654, 2010.

Gobiet, A., S. Kotlarski, M. Beniston, G. Heinrich, J. Rajczak, and M. Stoffel, 21st century climate change in the European Alps-A review, *Sci. Total Environ.*, *493*, 1138–1151, doi: 10.1016/j.scitotenv.2013.07.050, 2014.

Gudmundsson, L., J. B. Bremnes, J. E. Haugen, and T. Engen-Skaugen, Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations &ndash a comparison of methods, *Hydrology and Earth System Sciences*, *16*(9), 3383–3390, doi: 10.5194/hess-16-3383-2012, 2012.

Hackenbruch, J., T. Kunz-Plapp, S. Müller, and J. Schipper, Tailoring Climate Parameters to Information Needs for Local Adaptation to Climate Change, *Climate*, *5*(2), 25, doi:10.3390/cli5020025, 2017.

Hawkins, E., and R. Sutton, The potential to narrow uncertainty in regional climate predictions, *Bull. Am. Meteorol. Soc.*, *90*(8), 1095–1107, doi:10.1175/2009BAMS2607.1, 2009.

Haylock, M. R., N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones, and M. New, A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006, *J. Geophys. Res. Atmos.*, *113*(20), doi:10.1029/2008JD010201, 2008.

Hewitt, C. D., and D. J. Griggs, Ensembles-based predictions of climate changes and their impacts (ENSEMBLES), eos 85:566, *Tech. Rep. December 2004*, 2004.

Houze, R. A., Stratiform Precipitation in Regions of Convection: A Meteorological Para-

dox?, *Bull. Am. Meteorol. Soc.*, *78*(10), 2179–2196, doi:10.1175/1520-0477(1997)078⟨2179: SPIROC⟩2.0.CO;2, 1997.

Isotta, F. A., C. Frei, V. Weilguni, B. Rudolf, V. Pavan, C. Cacciamani, G. Antolini, S. M. Ratto, M. Munari, S. Micheletti, V. Bonati, C. Lussana, C. Ronchi, E. Panettieri, G. Marigo, M. Per˘, P. Lass, B. Rudolf, V. Pavan, C. Cacciamani, G. Antolini, S. M. Ratto, M. Munari, S. Micheletti, V. Bonati, C. Lussana, C. Ronchi, E. Panettieri, G. Marigo, and G. Verta˘, 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. Clim.*, doi:10.1002/joc.3794, 2013.

Kendon, E. J., N. M. Roberts, H. J. Fowler, M. J. Roberts, S. C. Chan, and C. A. Senior, Heavier summer downpours with climate change revealed by weather forecast resolution model, *Nat. Clim. Chang.*, *4*(7), 570–576, doi:10.1038/nclimate2258, 2014.

Lemos, M. C., C. J. Kirchhoff, and V. Ramprasad, Narrowing the climate information usability gap, *Nat. Clim. Chang.*, *2*(11), 789–794, doi:10.1038/nclimate1614, 2012.

Mesinger, F., K. Veljovic, M. J. Fennessy, and E. L. Altshuler, Value Added in Regional Climate Modeling: Should One Aim to Improve on the Large Scales as Well?, pp. 201–214, doi: 10.1007/978-3-7091-0973-1, 2012.

Metzger, M. J., R. G. H. Bunce, R. H. G. Jongman, C. A. Mücher, and J. W. Watkins, A climatic stratification of the environment of Europe, *Global Ecology and Biogeography*, *14*(6), 549–563, doi:10.1111/j.1466-822x.2005.00190.x, 2005.

Moss, R. H., J. A. Edmonds, K. A. Hibbard, M. R. Manning, S. K. Rose, D. P. van Vuuren, T. R. Carter, S. Emori, M. Kainuma, T. Kram, G. A. Meehl, J. F. B. Mitchell, N. Nakicenovic, K. Riahi, S. J. Smith, R. J. Stouffer, A. M. Thomson, J. P. Weyant, and T. J. Wilbanks, The next generation of scenarios for climate change research and assessment, *Nature*, *463*(7282), 747–756, doi:10.1038/nature08823, 2010.

Mueller, E. N., and A. Pfister, Increasing occurrence of high-intensity rainstorm events relevant for the generation of soil erosion in a temperate lowland region in Central Europe, *Journal of Hydrology*, *411*(3-4), 266–278, doi:10.1016/j.jhydrol.2011.10.005, 2011.

Muller, C. C. J., P. a. O'Gorman, L. E. L. Back, P. O'Gorman, and L. E. L. Back, Intensification of Precipitation Extremes with Warming in a Cloud-Resolving Model, *J. Clim.*, *24*(11), 2784–2800, doi:10.1175/2011JCLI3876.1, 2011.

O'Gorman, P. A., P. A. O'Gorman, and P. A. O'Gorman, Contrasting responses of mean and extreme snowfall to climate change, *Nature*, *512*(7515), 416–418, doi:10.1038/nature13625, 2014.

Perkins, S. E., A. J. Pitman, N. J. Holbrook, and J. McAneney, Evaluation of the AR4 Climate Models' Simulated Daily Maximum Temperature Minimum Temperature, and Precipitation over Australia Using Probability Density Functions, *Journal of Climate*, *20*(17), 4356–4376, doi:10.1175/jcli4253.1, 2007.

Pfeifer, S., K. Bülow, A. Gobiet, A. Hänsler, M. Mudelsee, J. Otto, D. Rechid, C. Teichmann, and D. Jacob, Robustness of Ensemble Climate Projections Analyzed with Climate Signal Maps: Seasonal and Extreme Precipitation for Germany, *Atmosphere (Basel).*, *6*(5), 677–698, doi:10.3390/atmos6050677, 2015.

Rosenzweig, C., G. Casassa, D. Karoly, A. Imeson, C. Liu, A. Menzel, S. Rawlins, T. Root, B. Seguin, and P. Tryjanowski, Assessment of observed changes and responses in natural

and managed systems, *Clim. Chang. 2007 Impacts, Adapt. Vulnerability*, pp. 79–131, doi: Milandreclimat, 2007.

Smiatek, G., H. Kunstmann, and A. Senatore, EURO-CORDEX regional climate model analysis for the Greater Alpine Region: Performance and expected future change, *J. Geophys. Res. Atmos.*, *121*(13), 7710–7728, doi:10.1002/2015JD024727, 2016.

Suklitsch, M., H. Truhetz, A. F. Prein, and A. Gobiet, NHCM-1: Non-hydrostatic climate modelling Part II Current state of selected cloud resolving regional climate models and their error characteristics, *Tech. Rep. 40*, 2011.

Taylor, G. I., The spectrum of turbulence, *Proceedings of the Royal Society of London*, *164*(919), 476–490, 1938.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, An overview of CMIP5 and the experiment design, *Bull. Am. Meteorol. Soc.*, *93*(4), 485–498, doi:10.1175/BAMS-D-11-00094.1, 2012.

Torma, C., F. Giorgi, and E. Coppola, Added value of regional climate modeling over areas characterized by complex terrain—Precipitation over the Alps, *J. Geophys. Res. Atmos.*, pp. 3957–3972, doi:10.1002/2014JD022781.Received, 2015.

Ye, H., J. Cohen, and M. Rawlins, Discrimination of Solid from Liquid Precipitation over Northern Eurasia Using Surface Atmospheric Conditions*, *J. Hydrometeorol.*, *14*(4), 1345–1355, doi:10.1175/JHM-D-12-0164.1, 2013.

2. Article 1: EURO-CORDEX: new high-resolution climate change projections for European impact research

EURO-CORDEX: new high-resolution climate change projections for European impact research

**Daniela Jacob · Juliane Petersen · Bastian Eggert · Antoinette Alias · Ole Bøssing Christensen ·
Laurens M. Bouwer · Alain Braun · Augustin Colette · Michel Déqué · Goran Georgievski ·
Elena Georgopoulou · Andreas Gobiet · Laurent Menut · Grigory Nikulin · Andreas Haensler ·
Nils Hempelmann · Colin Jones · Klaus Keuler · Sari Kovats · Nico Kröner · Sven Kotlarski ·
Arne Kriegsmann · Eric Martin · Erik van Meijgaard · Christopher Moseley · Susanne Pfeifer ·
Swantje Preuschmann · Christine Radermacher · Kai Radtke · Diana Rechid · Mark Rounsevell ·
Patrick Samuelsson · Samuel Somot · Jean-Francois Soussana · Claas Teichmann · Riccardo Valentini ·
Robert Vautard · Björn Weber · Pascal Yiou**

Abstract A new high-resolution regional climate change ensemble has been established for Europe within the World Climate Research Program Coordinated Regional Downscaling Experiment (EURO-CORDEX) initiative. The first set of simulations with a horizontal resolution of 12.5 km was completed for the new emission scenarios RCP4.5 and RCP8.5 with more simulations expected to follow. The aim of this paper is to present this data set to the different communities active in regional climate modelling, impact assessment and adaptation. The EURO-CORDEX ensemble results have been compared to the SRES A1B simulation results achieved within the ENSEMBLES project. The large-scale patterns of changes in mean temperature and precipitation are similar in all three scenarios, but they differ in regional details, which can partly be related to the higher resolution in EURO-CORDEX. The results strengthen those obtained in ENSEMBLES, but need further investigations. The analysis of impact indices shows that for RCP8.5, there is a substantially larger change projected for temperature-based indices than for RCP4.5. The difference is less pronounced for precipitation-based

D. Jacob (✉) · J. Petersen · B. Eggert · A. Haensler ·
N. Hempelmann · A. Kriegsmann · C. Moseley · S. Pfeifer ·
S. Preuschmann · D. Rechid · C. Teichmann · B. Weber
Climate Service Center (CSC), Helmholtz-Zentrum Geesthacht,
Fischertwiete 1, 20095 Hamburg, Germany
e-mail: daniela.jacob@hzg.de

D. Jacob · J. Petersen · C. Moseley · C. Radermacher ·
D. Rechid · C. Teichmann
Max Planck Institute for Meteorology (MPIM), Bundesstr. 53,
20146 Hamburg, Germany

A. Alias · A. Braun · M. Déqué · E. Martin · S. Somot
Météo-France/CNRS, Centre National de Recherches
Météorologiques (CNRM-GAME), 42, Avenue Gaspard
Coriolis, 31057 Toulouse Cedex 1, France

O. B. Christensen
Danish Climate Centre, Danish Meteorological Institute (DMI),
Lyngbyvej 100, 2100 Copenhagen, Denmark

L. M. Bouwer
Institute for Environmental Studies, Vrije Universiteit,
De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands

L. M. Bouwer
Deltares, PO Box 177, 2600 MH Delft, The Netherlands

A. Colette
Institut National de l'Environnement Industriel et des Risques
(INERIS), 60550 Verneuil-Halatte, France

G. Georgievski · K. Keuler · K. Radtke
Environmental Meteorology, Brandenburg University of
Technology, Burger Chaussee 2, 03013 Cottbus, Germany

E. Georgopoulou
Institute for Environmental Research and Sustainable
Development (IERSD), National Observatory of Athens,
118 10 Thissio, Greece

A. Gobiet
Wegener Center for Climate and Global Change and Institute for
Geophysics, Astrophysics, and Meteorology, University of Graz,
Brandhofgasse 5, 8010 Graz, Austria

L. Menut
Laboratoire de Météorologie Dynamique (IPSL), Ecole
Polytechnique, 91128 Palaiseau Cedex, France

indices. Two effects of the increased resolution can be regarded as an added value of regional climate simulations. Regional climate model simulations provide higher daily precipitation intensities, which are completely missing in the global climate model simulations, and they provide a significantly different climate change of daily precipitation intensities resulting in a smoother shift from weak to moderate and high intensities.

Introduction

Climate impact assessments and the development of regional to local-scale adaptation strategies require the availability of high-resolution climate change scenarios, including an assessment of their robustness and their inherent uncertainties. The WCRP Coordinated Regional Downscaling Experiment (CORDEX, http://wcrp-cordex.ipsl.jussieu.fr/; Giorgi et al. 2006) provides an internationally coordinated framework to improve regional climate scenarios. This includes harmonisation of model evaluation activities in the individual modelling centres and the generation of multi-model ensembles of regional climate projections for the land-regions worldwide.

As part of the global CORDEX framework, the EURO-CORDEX initiative (http://www.euro-cordex.net/) provides regional climate projections for Europe at 50 km (EUR-44) and 12.5 km (EUR-11) resolution, thereby complementing coarser resolution data sets of former activities like, e.g., PRUDENCE and ENSEMBLES. The regional simulations are downscaling the new CMIP5 global climate projections (Taylor et al. 2012) and the new representative concentration pathways (RCPs) (Moss et al. 2010; van Vuuren et al. 2011). Twenty-six modelling groups contributing 11 different regional climate models, partly in different model configurations, actively support EURO-CORDEX.

In its initial phase, EURO-CORDEX mainly focussed on model evaluation in present-day climate (e.g., Vautard et al. 2013; Kotlarski et al. 2013). So far more than 30 evaluation simulations have been conducted. Further activities include the coordinated analysis of future climate simulations, the joint analysis of dynamical and empirical–statistical methods and the design of suitable bias correction techniques to tailor EURO-CORDEX data for direct application in climate impact research. Particular emphasis is put on the construction of a simulation matrix that covers uncertainty in emission scenarios, the driving global climate model and the downscaling method in the best affordable manner. Here, we present the first results of the high-resolution (EUR-11) future climate simulations from EURO-CORDEX.

We used for comparison the results obtained from the regional climate projections in the FP6 ENSEMBLES project (Hewitt and Griggs 2004; van der Linden and Mitchell 2009). A number of regional limited-area models were used to downscale transient global climate projections over Europe at a 25 and 50 km resolution over the second half of the twentieth century and along the twenty-first century. The ENSEMBLES climate projections were carried out under the assumptions of the Special Report on Emission Scenario (SRES) A1B scenario (IPCC 2000). This scenario follows the storyline of the IPCC family of A1 scenarios. It assumes a rapid economic growth and development of technologies, with a worldwide population peaking in the middle of the twenty-first century, and a

G. Nikulin · C. Jones · P. Samuelsson
Swedish Meteorological and Hydrological Institute (SMHI), Folkborgsvägen 17, 601 76 Norrköping, Sweden

S. Kovats
London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9 SH, UK

N. Kröner · S. Kotlarski
Institute for Atmospheric and Climate Science, ETH Zurich, Universitätsstraße 16, 8092 Zurich, Switzerland

E. van Meijgaard
Royal Netherlands Meteorological Institute (KNMI), Wilhelminalaan 10, 3732 GK De Bilt, The Netherlands

M. Rounsevell
David Kinloch Michie Chair of Rural Economy & Environmental Sustainability, Institute of Geography & the Lived Environment, School of GeoSciences, The University of Edinburgh, Drummond Street, Edinburgh EH8 9XP, UK

J.-F. Soussana
INRA, UR874, Grassland Ecosystems Research (UREP), Clermont-Ferrand, France

R. Valentini
Department of Forest Science and Environment, University of Tuscia, Viterbo, Italy

R. Valentini
Euro–Mediterranean Center for Climate Change, Via De Lellis, 01100 Viterbo, Italy

R. Vautard · P. Yiou
Laboratoire des Sciences du Climat et de l'Environnement, IPSL, CEA/CNRS/UVSQ, Gif sur Yvette, France

balanced use of energy resources. This scenario leads to a rapid increase in fossil CO_2 emissions until 2050 and a decrease afterwards. As compared to other SRES scenarios, the CO_2 emissions lie in the middle of the scenario range.

EURO-CORDEX scenario simulations use the new Representative Concentration Pathways (RCPs) defined for the Fifth Assessment Report of the IPCC (Moss et al. 2010). In contrast to the SRES scenarios, RCP scenarios do not specify socioeconomic scenarios, but assume pathways to different target radiative forcing at the end of the twenty-first century. For instance, scenario RCP8.5 assumes an increase in radiative forcing of 8.5 W/m2 by the end of the century relative to pre-industrial conditions.

A comparison between the climate effects of SRES and RCP scenarios (Rogelj et al. 2012) indicates that the A1B scenario leads to a global mean temperature increase in the likely range of 2.8–4.2 °C, which approximates to RCP6 and lies clearly between RCP4.5 and RCP8.5.

The aim of the present study is to present a new high-resolution (12.5 km) data set from a multi-model multi-scenario ensemble of regional climate simulations for impact research. Such a high resolution has not been reached before in previous climate model projections. The ENSEMBLES project covered RCM simulations for Europe with a maximum resolution of 25 km, and in PRUDENCE, all simulations were done on a 50 km grid.

The analysis carried out here is directed towards regional climatic changes in Europe, addressing the differences of mean changes in annual mean temperature and total precipitation for the scenarios A1B, RCP4.5 and RCP8.5. In addition to the mean values, a range of climate indices important for climate impact studies in different sectors were calculated for sub-regions of Europe including: heavy precipitation events, dry spells and heat waves.

Experimental setup, statistical methods and definitions of indices

Experimental setup

In total, nine simulations have been done for RCP4.5 and ten for RCP8.5. They were all performed on the EURO-CORDEX domain, expect one simulation, which covered only the MED-CORDEX domain. The size and location of the CORDEX domains can be seen on the CORDEX web page (http://wcrp-cordex.ipsl.jussieu.fr/). The EURO-CORDEX domain covers all countries in the European Union, but it does not map perfectly to the Europe region defined for the IPCC Fifth Assessment Report. For the eastern part of Turkey, unfortunately, no regional model projections are available.

A summary of the grid configuration and differences in the parameterisation schemes for the participating regional models (ALADIN5.1: Colin et al. 2010, Herrmann et al. 2011, CCLM: Rockel et al. 2008, HIRHAM: Christensen et al. 1998, RACMO2: Meijgaard van et al. 2012, RCA4: Samuelsson et al. 2011, Kupiainen et al. 2011, REMO: Jacob et al. 2012, WRF Version 3.3.1: Skamarock et al. 2008) is given in the supplementary material (Table s1). The RCP scenarios, the driving GCMs and the driven RCMs as well as the simulation length are listed in the supplementary material (Table s2a). Seven different RCMs and five different GCMs have been used in this study. Two of the RCMs were driven by four/five different GCMs; five GCM-RCM chains did simulate both RCP scenarios. They all provide data at least until the mid of the century. Eight RCP4.5 simulations and nine RCP8.5 simulations had reached the end of the century.

Two additional sensitivity studies have been performed to study the influence of the smaller domain size used by one model (supplementary material, Figure s6) and the effect of using one RCM more often than others (supplementary material, Figure s7).

From the ENSEMBLES data set, 20 transient RCM simulations reaching the end of the century are used for most of the analysis (supplementary material, Table s2b). Maximum and minimum temperature values were only available from nine simulations. The ensemble includes simulations on 50 and 25 km grid scale, but most simulations were carried out on a 25 km grid (16 out of 20 and 7 out of 9). The coarser resolution simulations were included to enlarge the sample size for the statistical analyses. A sensitivity study using only the 25 km simulations for the analyses of changes in the mean fields of temperature and precipitation and in the indices showed that there are only minor differences in the horizontal pattern of the meteorological parameters as well as in the regions with a significant and/or robust change (not shown).

Statistical methods

All regional changes have been analysed for the land areas of the EURO-CORDEX domain using daily data for three time periods: 1971–2000, 2021–2050 and 2071–2100. For changes in all outcome measures (mean annual and seasonal temperature, total annual and seasonal precipitation, heavy precipitation, heat waves and dry spells), significance and robustness were tested using a method adapted from Pfeifer et al. (2013). This method identifies regions with relatively strong and robust climate changes from an ensemble of climate change simulations. It can be applied to simulation results on regular model grids or to data aggregated onto larger regions. The robustness of the information given by the ensemble of climate projections is

analysed using two statistical tests. The first test regards the agreement of the simulations in terms of the direction of the changes. In the second test, the significance of the projected changes in each simulation is assessed by applying the Mann–Whitney–Wilcoxon test. Only regions, which pass both tests, are identified as regions with robust projected changes.

The stringency of the tests can be adjusted by the choice of several parameters, for example the significance level, or by shifting the percentage of simulations, which have to pass the individual tests. We visualise the regions with relatively strong and robust climate changes. Regions with non-significant changes are treated in the same way as regions where the models strongly disagree on the direction of the changes. It is thus by purpose not possible to distinguish between regions where to a certain confidence no significant changes are projected to occur and regions where the projected changes cannot be given with sufficient confidence. Insignificant changes are not equivalent to the lack of agreement in the projected changes and should be interpreted differently, depending on the context.

First, the direction of change was assessed. Changes in regions in which more than 66 % of the models agree in the direction of change were called robust changes. For the analyses of the significance, two different data preparation methods were used. For total precipitation and maximum, mean and minimum temperature, one value each per year and grid box was calculated (30 values per simulation per grid box). For heat waves, dry spells and heavy precipitation, only one value exists for each simulation and grid box, since they are defined over a 30-year period. Again they were calculated for all three time periods. For significance testing, the Mann–Whitney–Wilcoxon test was applied to the entire ensemble data set for the respective time periods.

For the climate change signal of mean annual temperature and total annual precipitation, a spatial correlation is calculated for the respective parameters of the SRES A1B and RCP8.5 simulation results for the end of the century based on the five sub-regions. For this analysis, all parameters for both scenarios were remapped to the ENSEMBLES grid with 25 km resolution.

Definitions of impact-relevant indices

In addition to the calculation of mean temperature and precipitation changes, we focussed on high-impact phenomena. Heavy precipitation, dry spells and heat waves were calculated according to the following definitions:

- "Heavy precipitation" is defined as the intensity of the heavy precipitation events defined as the 95th percentile of daily precipitation (only days with precipitation >1 mm/day are considered).

- "Dry spells" are defined as periods of at least 5 consecutive days with daily precipitation below 1 mm. The 95th percentile of the length of all identified dry spells is considered in this analysis and referred to as "extended dry spells."

- For the analysis of the change in number of heat waves, we used two different definitions. In the first definition, heat waves were considered as periods of more than three consecutive days exceeding the 99th percentile of the daily maximum temperature of the May to September season of the control period (1971–2000). The second definition, representing more extreme heat wave events, is based on the WMO definition (Frich et al. 2002). Heat waves were defined as periods of more than 5 consecutive days with daily maximum temperature exceeding the mean maximum temperature of the May to September season for the control period (1971–2000) by at least 5 °C.

A range of indices was selected based on known climate impacts in Europe for infrastructure, agriculture and human health (Alcamo et al. 2007). The impact indices are calculated as means for 5 European sub-regions used in the European chapter of the forthcoming IPCC Fifth Assessment Report. The sub-regional classification is based on Metzger et al. (2005) (supplementary material, Fig. s1) and selected to represent 5 different climate and environmental characteristics: Alpine, Atlantic, Continental, Northern and Southern Europe. The list of impact index definitions is given in http://cccma.seos.uvic.ca/ETCCDI/list_27_indices.shtm.

For all ensembles, the "likely" range is defined as range within the 17th and 83rd percentile of projected changes.

Results

Mean changes of temperature and precipitation

To date, the regional climate change projections provided by the EU-FP6 ENSEMBLES multi-model ensemble for the SRES A1B scenario are considered as state-of-the-art for European climate impact research. Therefore, we compared this data set to the new regional EURO-CORDEX data set for RCP8.5 and RCP4.5. Figure 1 shows the ensemble mean of the mean annual temperature and total annual precipitation change until the end of the century. Figure 1b, d shows a robust and statistically significant warming, with regional differences, in the range of 1–4.5 °C for RCP4.5 and of 2.5–5.5 °C for RCP8.5. These ranges encompass the warming range projected for the A1B scenario, where temperature increases between 3 and
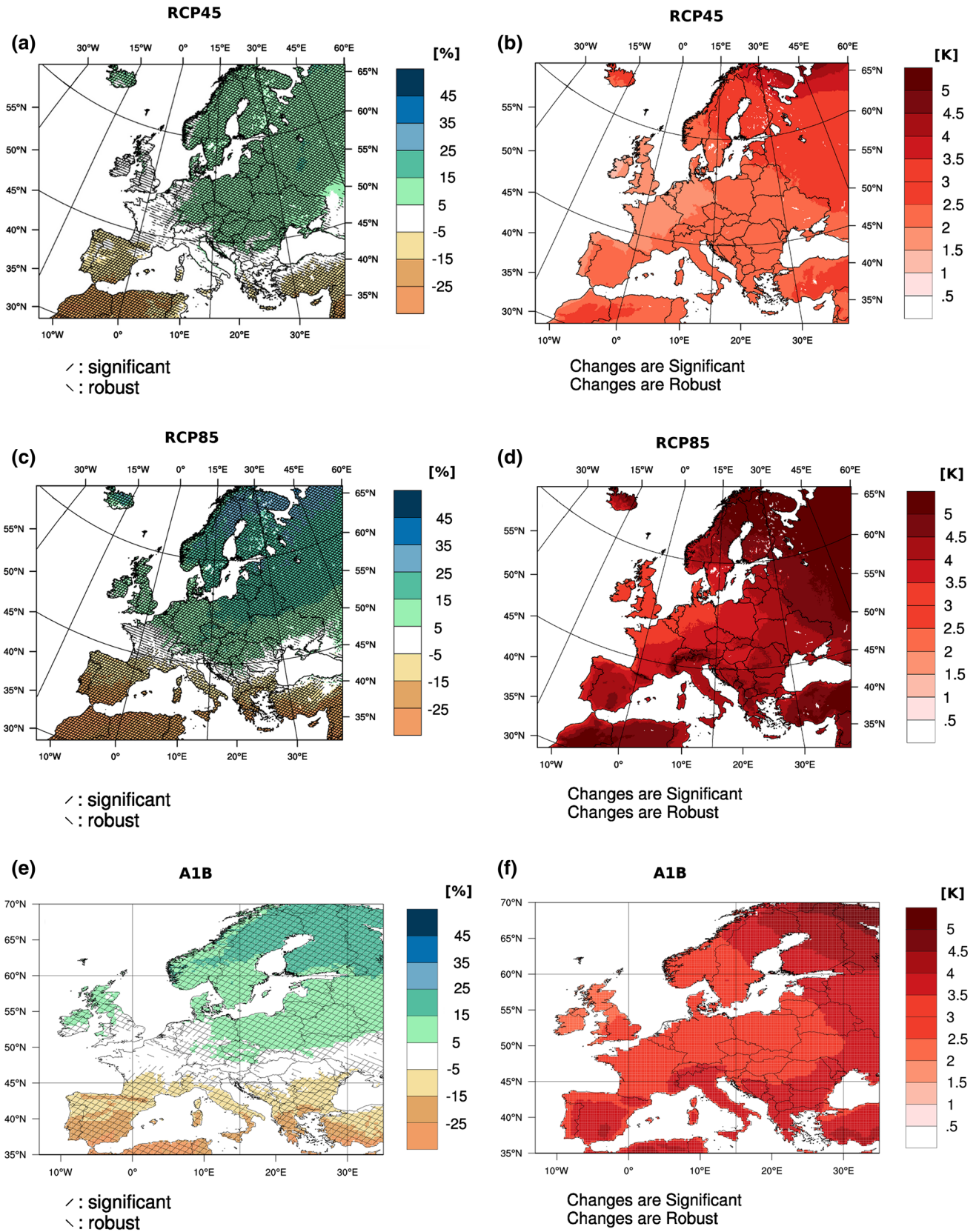
Fig. 1 Projected changes of total annual precipitation (%) (*left*) and annual mean temperature [K] (*right*) for 2071–2100 compared to 1971–2000, for A1B (**e**, **f**), RCP8.5 (**c**, **d**) and RCP4.5 (**a**, **b**) scenarios. *Hatched areas* indicate regions with robust and/or statistical significant change (**a**, **c**, **e**). Changes are robust and significant across the entire European continent (**b**, **d**, **f**)

Table 1 Spatial correlation of SRES A1B and RCP8.5 emission scenarios for changes in mean annual temperature and annual total precipitation of the sub-regions for the time periods 2021–2050 and 2071–2100

Spatial correlation of RCP8.5 and SRES A1B	Climate parameter			
	Mean annual temperature		Annual total precipitation	
	2021–2050	2071–2100	2021–2050	2071–2100
Alpine	0.88	0.95	0.92	0.94
Atlantic	0.82	0.98	0.87	0.94
Continental	0.94	0.96	0.72	0.92
Northern	0.97	0.97	0.59	0.81
Southern	0.90	0.89	0.71	0.96

4.5 °C (Fig. 1f). The projected spatial patterns are very similar in all scenarios with greater annual mean warming in Southern Europe and towards the northeast. Under RCP8.5, large parts of Northern Scandinavia, Eastern Europe and the Alpine ridge might be exposed to a warming of more than 4.5 °C compared to 1971–2000, which could be avoided by RCP4.5.

Associated with the large increase in temperature in RCP8.5 are robust changes in annual precipitation. The ensemble mean projects a statistically significant increase in large parts of Central Europe and Northern Europe of up to about 25 % and a decrease in Southern Europe. A zone with small changes, which are not significant (however, partially robust in the sign of change), indicates where the climate change signals change the sign (Fig. 1c, white areas). The pattern of the changes is very similar for RCP4.5, but less pronounced (Fig. 1a). The spatial pattern for A1B precipitation changes qualitatively agrees with the described changes for RCP4.5 and RCP8.5, and the magnitude of the changes mostly lies in-between the two RCPs. However, differences in the spatial patterns are seen over the British Isles, Benelux and Germany (Fig. 1e).

For mean temperature and precipitation change, a spatial correlation has been done between RCP8.5 and A1B results (Table 1). For all sub-regions, the spatial correlation between SRES A1B and RCP8.5 is very high, with 0.82–0.97 for temperature changes and 0.59–0.92 for precipitation changes depending on the region for mid-century. Towards the end of the century, the correlation is even stronger for both parameters.

Seasonal changes of mean temperature and precipitation are shown in the supplementary material (Fig. s2-s5). The seasonal temperature change signals show more regional heterogeneity than the annual mean. The zone between regions in which precipitation increases in the north and decreases in the south shifts southwards in summer and northwards in winter.

Seasonal mean changes for heavy precipitation

The projected seasonal mean changes in heavy precipitation for the three emission scenarios are relatively similar, but some regional differences are visible (Figs. 2 (RCP8.5), 3 (A1B), 4 (RCP4.5)). Most obvious differences are the increased regional detail in the RCP8.5 and RCP4.5, which is related to the higher horizontal resolution of about 12.5 km for the RCPs compared to 25 km for A1B, for which more homogeneous changes are calculated. The annual cycle of changes in heavy precipitation is similar in all three scenarios, but the amplitude of the change is stronger in RCP8.5 than in A1B in several regions. The results for RCP8.5 include a possible decrease in heavy summer precipitation by about 25 % over some parts of the Iberian Peninsula and Southern France, accompanied by regional increases in parts of Spain and Portugal. For winter, RCP8.5 projects strongest increases in heavy precipitation (up to 35 %) in Central and Eastern Europe, whereas A1B projects changes up to 25 % only in this region. Only for some parts in Scandinavia, A1B shows similar values as RCP8.5.

Important regional differences in heavy precipitation are projected for the RCP4.5 scenario. Compared to RCP8.5, the seasonal patterns of change are similar, but the amount of change is much smaller (up to 15 % in large areas with isolated spots up to 25 %) and—besides isolated regions in Southern Europe (mostly along coastlines)—no decrease in heavy precipitation is indicated.

GCM simulations tend to underestimate the high precipitation intensities (Sun et al. 2006). An improved distribution of high precipitation intensities is an important advantage of regional climate simulations. Figure 5 shows the relative frequencies of daily precipitation intensities of an ensemble of five GCM simulations and the corresponding regional downscaling experiments for the reference period 1971–2000 analysed over a central land region of the EURO-CORDEX domain (45°N–50°N and 2°E–17°E). For this analysis, data from all grid cells were taken into account. The distribution illustrates that the GCMs generally produce more precipitation intensities with up to 12 mm/day. The RCMs, in contrast, show higher intensities. Strong intensities above 30 mm/day do hardly occur in the GCM simulations. Figure 6 shows the temporal changes of the precipitation frequencies between the near future period and the reference period for both scenarios and the GCM and the RCM ensembles. Striking is that both model types reduce the number of weak precipitation intensities below 9 mm/day in both scenarios and increase the relative frequencies in all higher intensity classes. This shift in daily precipitation intensities, however, turns out much more moderate in the RCM than in the GCM simulations of both scenarios, the RCP4.5 and the RCP8.5.
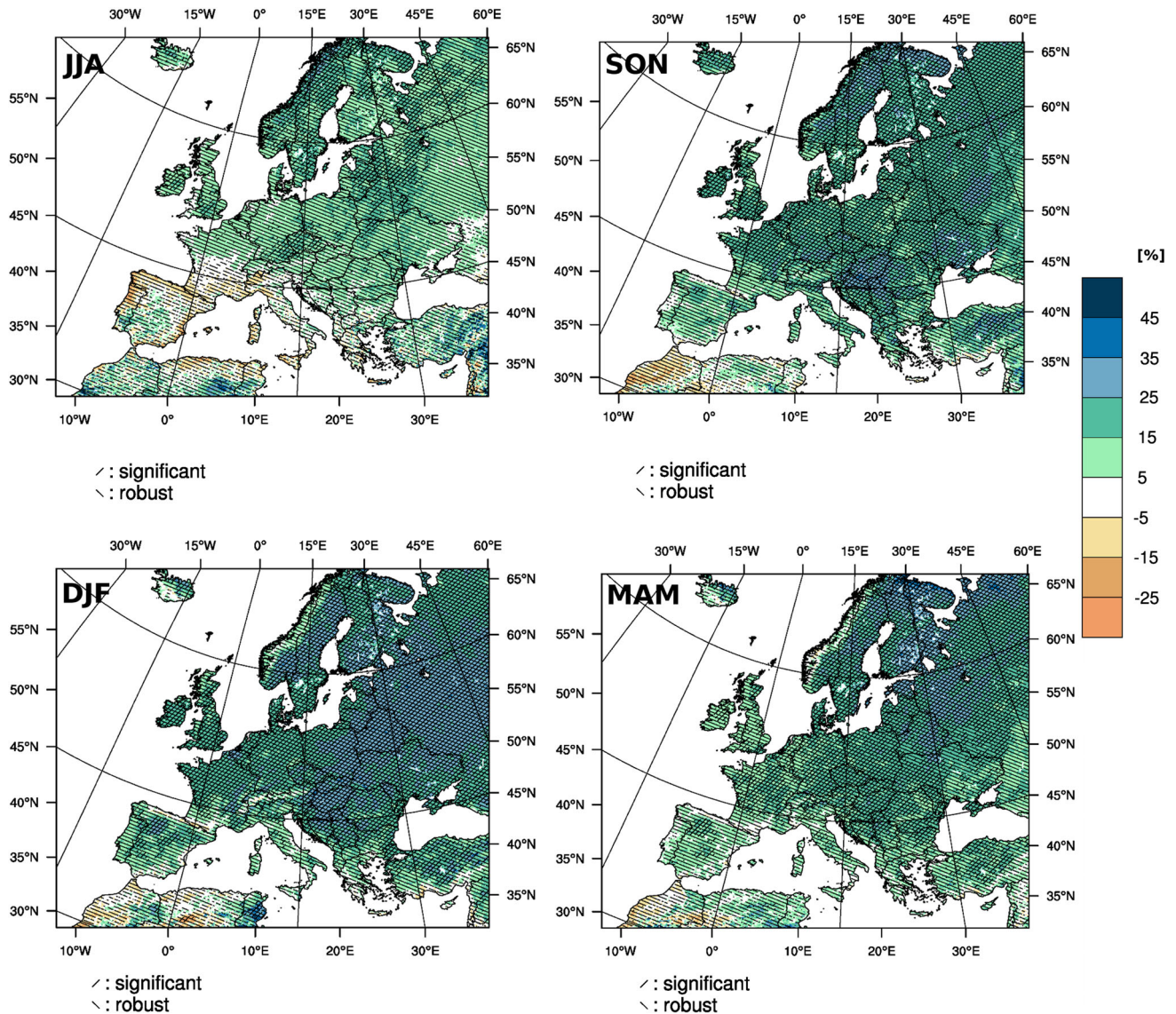
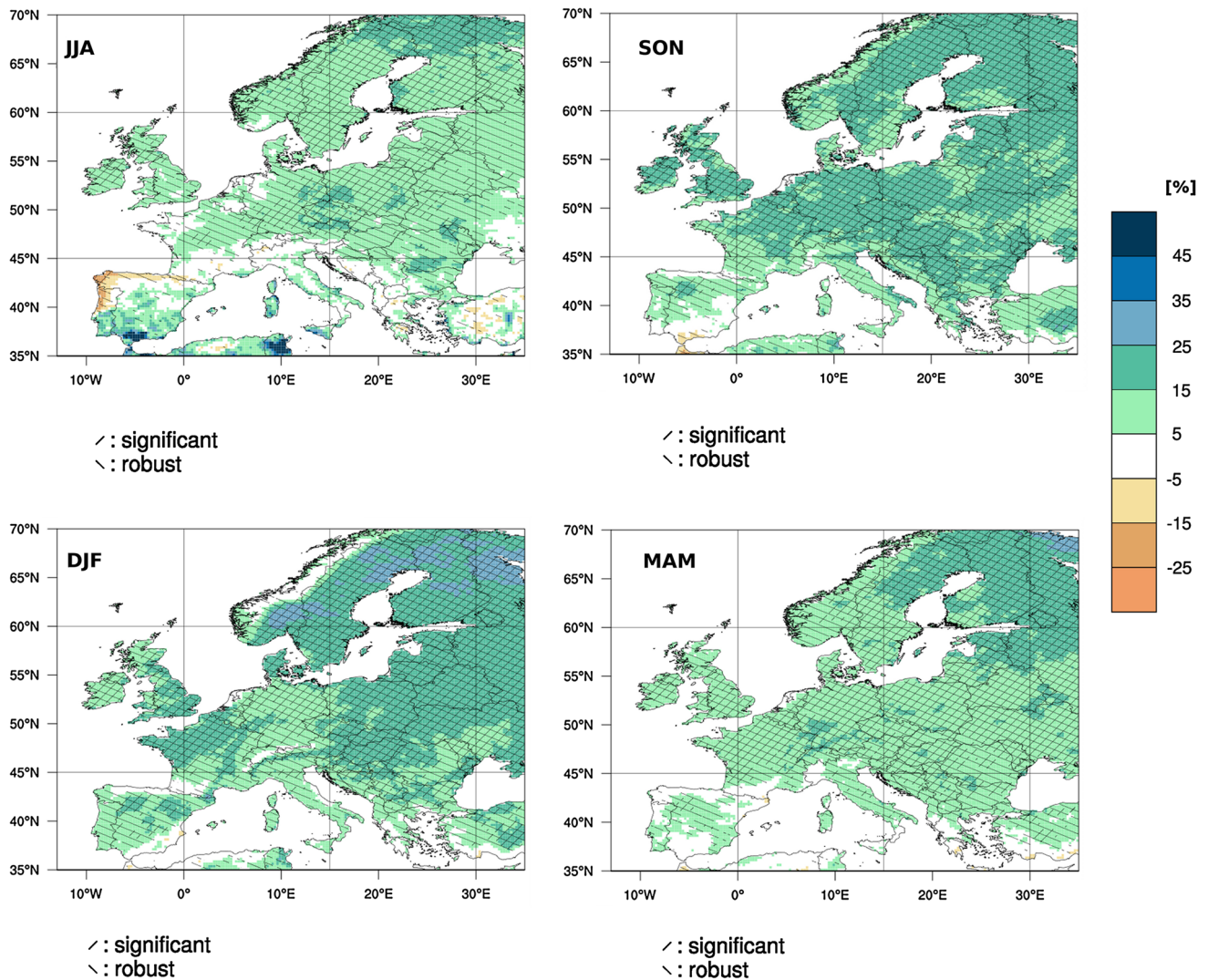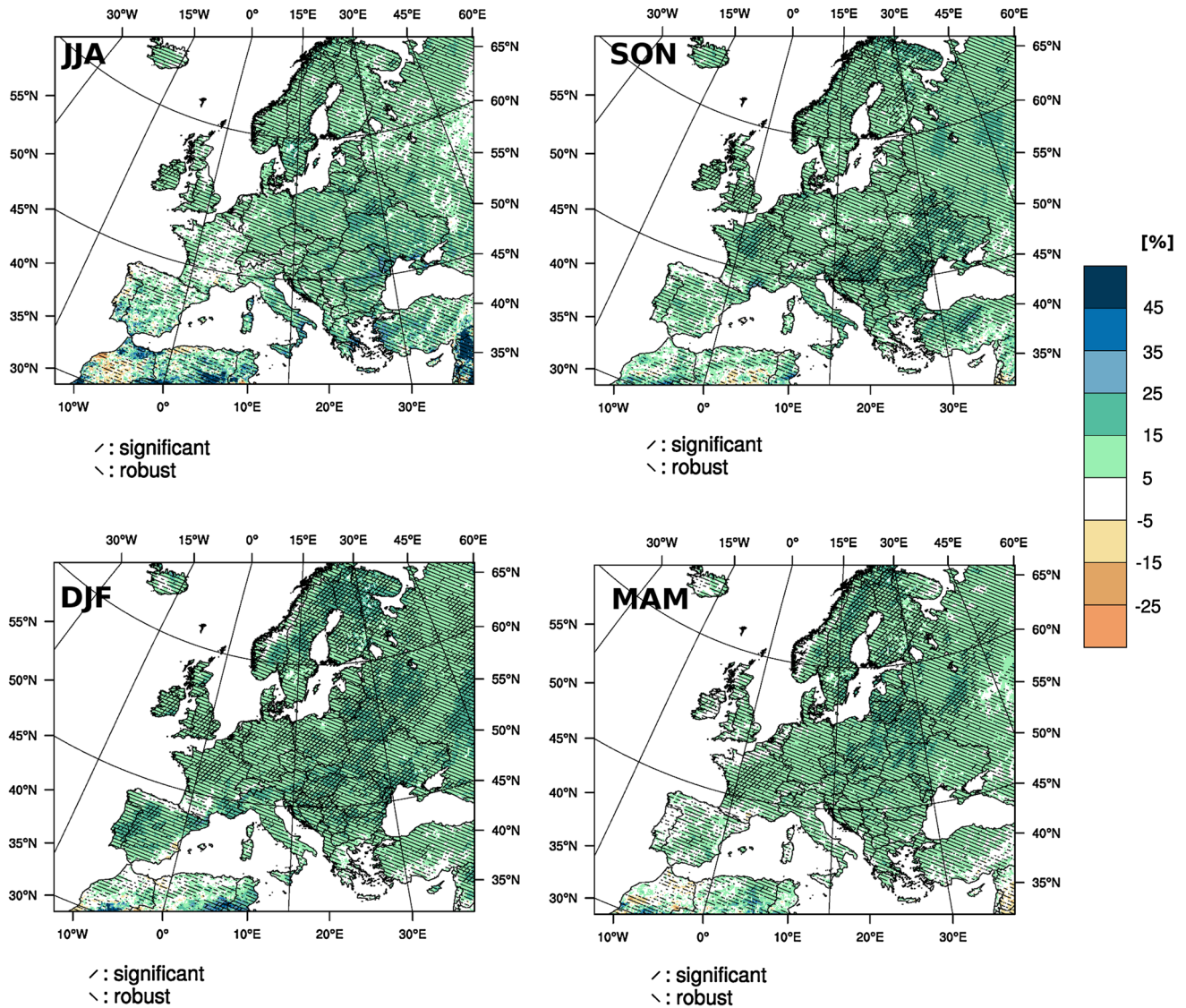**Fig. 2** Projected seasonal changes of heavy precipitation (%) based on the RCP8.5 scenario for 2071–2100 compared to 1971–2000. *Hatched areas* indicate regions with robust and/or statistical significant change

The frequency changes of the RCM simulations are especially in the range between 10 and 20 mm/day less than half of the GCM changes. Above 30 mm/day, however, the increase in the RCM ensembles exceeds the climate change signal of the GCMs. Figure 6 also demonstrates that both effects—the reduction in weak intensities and the increase in strong intensities—are more pronounced in the RCP8.5 scenario. The analyses of the processes leading to the different behaviour are beyond the scope of this paper and will be studied in a separate paper.

This analysis proves two effects of an increased resolution, which can be regarded as an added value of regional climate simulations. On the one hand, the RCMs provide higher daily precipitation intensities, which are completely missing in the GCM simulations, and on the other hand,

they provide a significantly different climate change of daily precipitation intensities resulting in a smoother shift from weak to moderate and high intensities.

Mean length of dry spells

Projected changes in the 95th percentile of the mean length of dry spells are shown in Fig. 7 for A1B (top), RCP8.5 (middle) and RCP4.5 (bottom) for 2021–2050 and 2071–2100 with respect to 1971–2000. For the early period, the change patterns are very similar in all scenarios (left row), although the number of simulations taken into account for each scenario ensemble is different. Beside some common features in South-West Europe, substantial differences in the projected changes for dry spells lengths

Fig. 3 Projected seasonal changes of heavy precipitation (%) based on the A1B scenario for the period 2071–2100 compared to 1971–2000. *Hatched areas* indicate regions with robust and/or statistical significant change

are visible until 2100. For RCP4.5 and A1B, a small increase in the length of extended dry spells is projected for Central Europe, which is more pronounced in A1B. A decrease in the length of extended dry spells is calculated in A1B for parts of Scandinavia. This feature is extended towards the Alps in the RCP8.5, in which the number of dry spells increases (not shown). This means that under RCP8.5 more but shorter dry spells are projected in the alpine region. For regions with a large increase in the length of extended dry spells, the number of dry spells is decreasing (not shown).

Mean number of heat waves

Projected changes in the mean number of heat waves during May–September are presented in Fig. 8, for RCP4.5

and RCP8.5, for the two future time periods and for two different definitions of heat waves. From the upper four panels, displaying the p99-heat wave definition, it is obvious that with less warming (see "Mean changes of temperature and precipitation" section) in RCP4.5, the increase in number of heat waves is smaller than in RCP8.5. This is more pronounced towards the end of the century (Fig. 8c, d) than for the earlier time period (Fig. 8a, b). For both scenarios, the increase is strongest in Southern Europe, but towards the end of the century the number of heat waves increases all over Europe. The number of heat waves for Southern Europe is projected to increase by more than 45. The increase is mostly robust and significant. The change in the number of heat waves considerably depends on the definition (thresholds and duration), which is used. Therefore, a second definition was

Fig. 4 Projected seasonal changes of heavy precipitation (%) based on the RCP4.5 scenario for the period 2071–2100 compared to 1971–2000. *Hatched areas* indicate regions with robust and/or statistical significant change

used based on that developed by the World Meteorological Organization (see "Definitions of impact-relevant indices" section). Under this definition, the increase in the mean number of heat waves is much less (Fig. 8e, f). For the WMO-heat wave definition not a single heat wave is detected in the ensemble mean for the reference period as well as for mid of the century, because the criteria are much stricter. Also the duration of the heat wave is two days longer than in the p99-heat wave definition. For RCP8.5, meaning under the strongest projected warming, towards the end of the century, an increase is only projected for some parts of Southern Europe with additional 5 to more than 9 heat waves. The increase is significant and robust south of 55° latitude.

Indices by sub-region

Projected changes of several impact indices, which could be of interest for impact studies in different sectors, are listed in Tables 2 (A1B) and 3 (RCP4.5 and RCP8.5) for the 5 sub-regions. For all sub-regions and indices, the median shifts into the same direction, independent of the scenario. For almost all indices, a substantially larger change in the median is projected in RCP8.5, compared to RCP4.5, however, the likely ranges frequently overlap. Exceptions to this are annual total precipitation in the Atlantic and Continental sub-regions, tropical nights in the Northern sub-region and the cold spell duration everywhere. Here the projected changes in RCP4.5 and RCP8.5
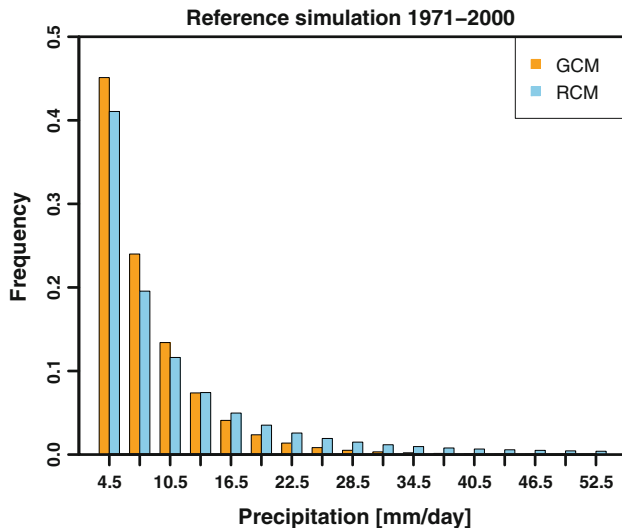
Fig. 5 Frequencies of daily precipitation intensities of an ensemble of five GCM simulations and the corresponding RCMs for the reference period 1971–2000 over a central land region of the EURO-CORDEX domain (45°N–50°N and 2°E–17°E)
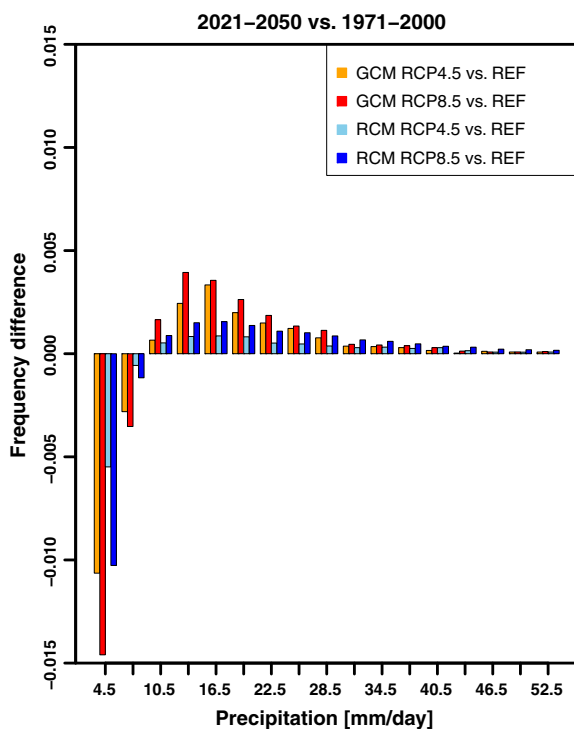
Fig. 6 Temporal changes of the precipitation frequencies between the near future period (2021–2050) and the reference period (1971–2000) for RCP4.5 and RCP8.5 and the GCM and the RCM ensembles

are rather similar. Differences between the RCP scenarios are most pronounced for growing season length and warm spell duration index, with no overlap between the likely ranges over all sub-regions (even the full range seldom overlaps).

The median change in A1B is generally centred within RCP4.5 and RCP8.5. For some cases like annual total precipitation in the Southern sub-region or tropical nights in the Atlantic, Northern and Southern sub-regions, however, the median change in A1B is even stronger than the median change in the RCP8.5 and only for change in annual total precipitation in the Continental sub-region, the median change in A1B is lower than the median change in RCP4.5. The spreads of the projected changes defined as the likely ranges are generally the same between RCP4.5 and RCP8.5 or slightly larger in RCP8.5. Exceptions to this are frost days in the Continental sub-region, tropical nights in the Northern sub-region and total rainfall amount above the 99th percentile of daily rain (wet days only) in the Southern sub-region. Here the likely ranges of projected changes are larger in the RCP4.5 scenario.

Conclusions and outlook

Regional climate change patterns for Europe projected by the high-resolution regional climate change ensemble within the EURO-CORDEX initiative for the new emission scenarios RCP4.5 and RCP8.5 have been compared to the state-of-the-art regional climate change data set from ENSEMBLES. The overall spatial patterns for temperature and precipitation changes and related indices are similar. There is a large degree of consistency between the ensembles of the three emission scenarios. The results clearly strengthen the previous findings obtained from the ENSEMBLES data set; however, there are some important new findings.

Climate projections from the new ensemble indicate a reduced northwards shift of Mediterranean drying evolution and slightly stronger mean precipitation increases over most of Europe. This is consistent with the two evaluation papers by Vautard et al. (2013) and Kotlarski et al. (2013), for reanalysis-driven hindcast simulations, where it was found that higher resolution leads to more precipitation in the climate models.

The high-resolution in the EURO-CORDEX simulations is clearly visible in the change pattern for heavy precipitation events. This spatial information is potentially very useful for climate impact studies. The more detailed spatial patterns in the high-resolution simulations can be related to better resolved physical processes like convection and heavy precipitation, and due to better representation of surface characteristics and their spatial variability. This needs to be elaborated further to understand the physical processes involved and the robustness of the pattern.

For some regions, especially the Alps, a connection between the changes in length of extended dry spells and the number of all dry spells (not only the longest ones) is
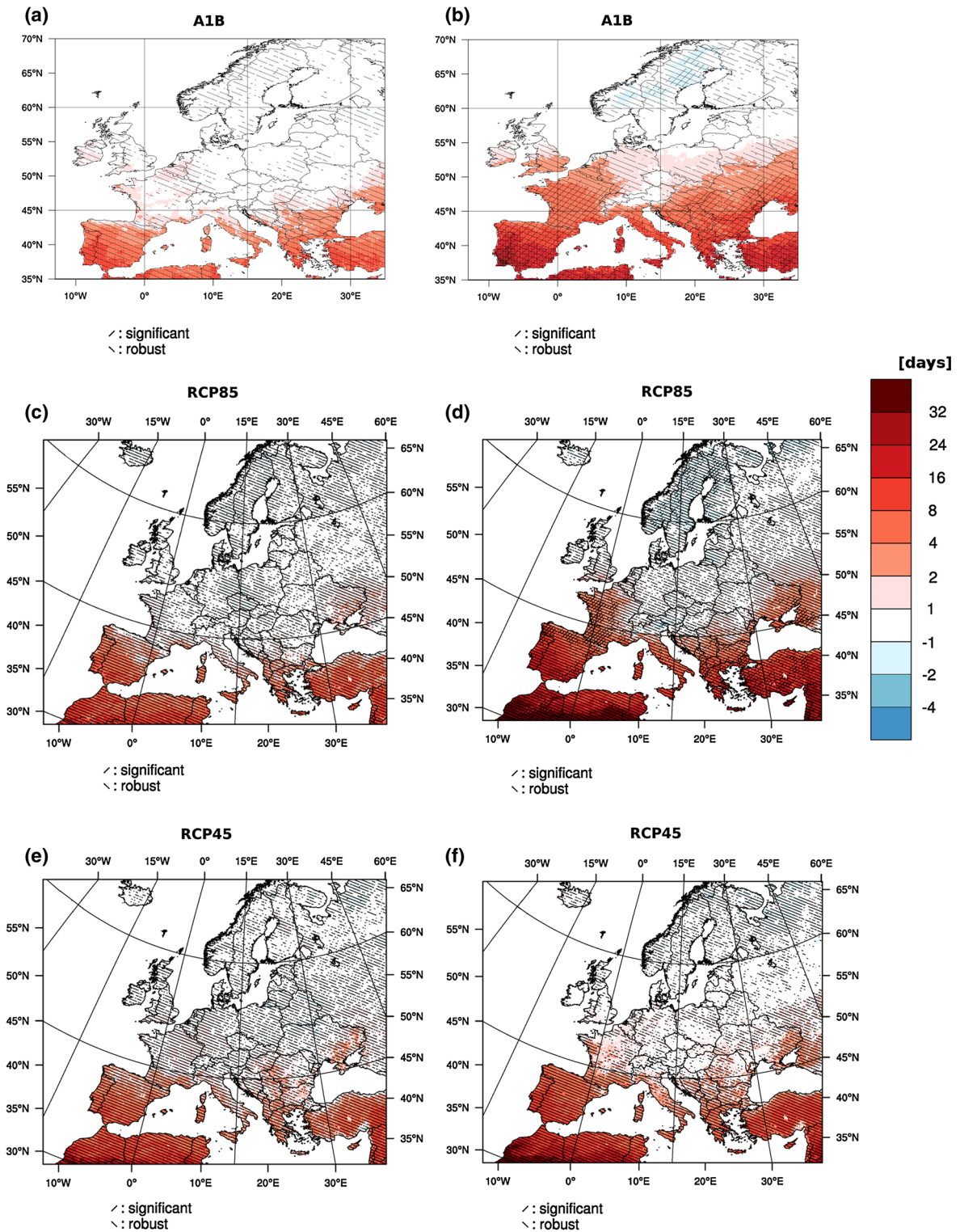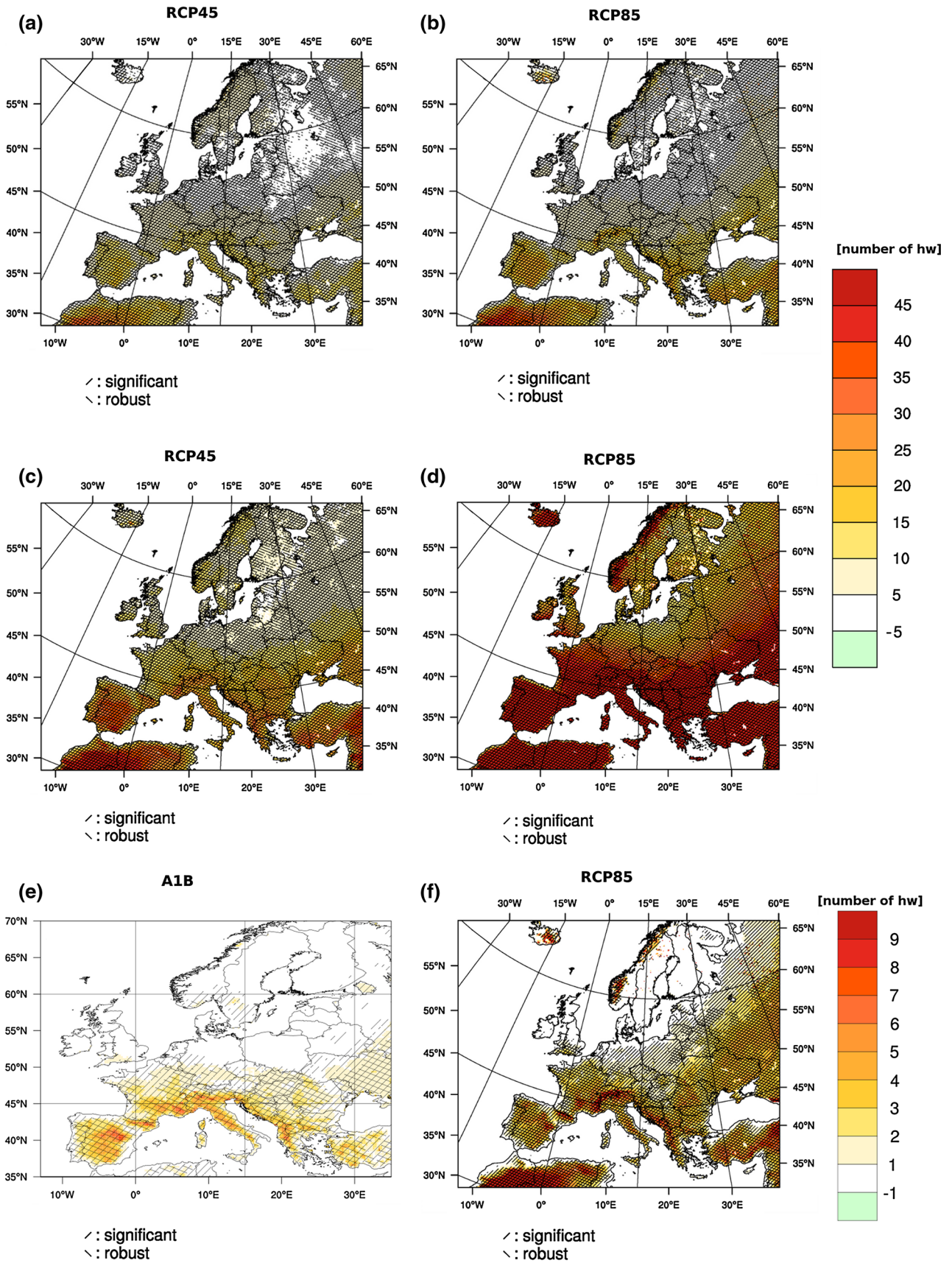
Fig. 7 Projected changes in the 95th percentile of the length of dry spells (days) for 2021–2050 compared to 1971-2000 (**a**, **c**, **e**) and 2071–2100 compared to 1971–2000 (**b**, **d**, **f**) for A1B (**a**, **b**), RCP8.5 (**c**, **d**) and RCP4.5 (**e**, **f**) scenarios. *Hatched areas* indicate regions with robust and/or statistical significant changes

(a) RCP45

: significant
: robust

(b) RCP85

: significant
: robust

(c) RCP45

: significant
: robust

(d) RCP85

: significant
: robust

(e) A1B

: significant
: robust

(f) RCP85

: significant
: robust

[number of hw]

45
40
35
30
25
20
15
10
5
−5

[number of hw]

9
8
7
6
5
4
3
2
1
−1

◀ **Fig. 8** Projected changes in the mean number of heat waves occurring in the months May–September for 2021–2050 compared to 1971–2000 (**a**, **b**) and 2071–2100 compared to 1971–2000 (**c** through **f**). Heat waves in figures **a** through **d** are defined as periods of more than 3 consecutive days exceeding the 99th percentile of the daily maximum temperature of the May to September season for the control period (1971–2000). Heat waves in figures **e** and **f** are defined as periods of more than 5 consecutive days with daily maximum temperature exceeding the mean maximum temperature of the May–September season for the control period (1971–2000) by at least 5 °C. They are based on A1B emission scenario (**e**), the RCP8.5 (**b**, **d**, **f**) and the RCP4.5 (**a**, **c**). *Hatched areas* indicate regions with robust and/or statistical significant change

detected. In this region, the decreasing length of extended dry spells is linked to an increasing number of all dry spells. This seems to be plausible, but follow-up studies are needed to understand the governing processes.

The investigation related to possible changes in heat waves shows clearly that the definition of the heat wave is dominating the regional change pattern more than the differences in scenarios or time periods. Therefore, it is of utmost importance to clearly define the index in the light of the specific study and most likely for the sector of interest.

Table 2 Projected changes of selected climate parameters and indices for 2071–2100 with respect to 1971–2000 spatially averaged for European sub-regions for A1B scenario

Scenario A1B	Climate parameters	Measure	Alpine	Atlantic	Continental	Northern	Southern
2071–2100 minus 1971–2000	Mean annual Temperature in K**	Median	3.4	2.5	3.3	3.8	3.6
		Min	2.8	1.9	2.1	3.2	2.3
		Likely in the range	3.1–4.5	2.1–3.5	2.8–4.5	3.5–5.0	3.3–4.1
		Max	5.4	4.7	5.7	5.8	5.5
	Frost days (1) per year*	Median	−50	−24	−44	−54	−24
		Min	−37	−13	−26	−38	−12
		Likely in the range	−38 to −57	−15 to −34	−27 to −53	−40 to −55	−12 to −31
		Max	−72	−39	−56	−71	−34
	Summer days (2) per year*	Median	14	21	32	7	48
		Min	4	9	21	3	33
		Likely in the range	11–20	16–32	22–41	5–14	33–51
		Max	21	34	43	27	51
	Tropical nights (4) per year*	Median	3	8	21	4	47
		Min	1	2	14	1	18
		Likely in the range	2–9	6–17	16–35	1–7	35–52
		Max	11	32	43	10	60
	Growing season length (5) days per growing season**	Median	47	41	52	41	36
		Min	27	23	20	25	14
		Likely in the range	34–56	33–51	33–62	27–46	27–41
		Max	75	55	81	61	51
	Warm spell duration index (14) days per year*	Median	57	44	42	67	91
		Min	46	29	26	37	67
		Likely in the range	51–84	35–72	37–69	47–96	85–112
		Max	126	125	94	119	144
	Cold spell duration index (15) days per year*	Median	−5	−5	−6	−6	−5
		Min	−4	−4	−4	−5	−3
		Likely in the range	−4 to −5	−4 to −6	−5 to −6	−5 to −8	−4 to −5
		Max	−8	−9	−9	−9	−8
	Annual total precipitation (27) in %**	Median	7	3	3	16	−15
		Min	1	9	−9	4	−7
		Likely in the range	5–12	−4 to 5	−1 to 5	13–21	−12 to −18
		Max	15	−11	12	29	−25
	Annual total precipitation where RR > 99p of 1971/2000 (26) in %**	Median	57	65	53	64	43
		Min	35	28	31	32	21
		Likely in the range	47–68	42–98	44–77	47–88	35–57
		Max	117	112	110	105	74

Numbers are based on 9 (indicated with*) and 20 (indicated with**) regional model simulations. The likely range defines the range of 66 % of all projected changes around the ensemble median

Table 3 Projected changes of selected climate parameters and indices for 2071–2100 with respect to 1971–2000 spatially averaged for European sub-regions based on RCP4.5 and RCP8.5

Scenarios	Climate parameters	Measure	Alpine RCP 4.5	Alpine RCP 8.5	Atlantic RCP 4.5	Atlantic RCP 8.5	Continental RCP 4.5	Continental RCP 8.5	Northern RCP 4.5	Northern RCP 8.5	Southern RCP 4.5	Southern RCP 8.5
2071–2100 minus 1971–2000	Mean annual temperature in K	Median	2.4	4.6	1.7	3.2	2.1	4.1	2.9	5.2	2.0	4.2
		Min	1.8	3.8	1.3	2.5	1.6	3.6	2.0	4.1	1.9	3.8
		Likely in the range	1.9–3.4	3.9–6.0	1.4–2.1	2.7–3.6	1.6–3.2	3.7–5.2	2.0–4.2	4.1–6.2	1.9–2.7	3.9–5.4
		Max	3.6	6.3	2.9	4.2	3.2	5.3	4.3	6.5	3.2	5.7
	Frost days (1) per year	Median	−40	−70	−28	−40	−34	−62	−40	−68	−22	−43
		Min	−25	−55	−12	−21	−16	−46	−24	−58	−10	−22
		Likely in the range	−26 to −41	−57 to −85	−15 to −30	−26 to −50	−18 to −40	−50 to −65	−26 to −43	−60 to −83	−11 to −30	−23 to −51
		Max	−47	−93	−33	−60	−41	−73	−52	−93	−31	−51
	Summer days (2) per year	Median	8	19	11	24	20	37	4	13	27	54
		Min	3	10	6	17	11	27	2	5	21	43
		Likely in the range	4–14	12–24	6–14	22–28	13–24	30–46	2–16	6–22	25–33	46–60
		Max	18	25	33	38	28	49	23	28	37	67
	Tropical nights (4) per year	Median	1	4	3	7	9	22	1	1	20	45
		Min	0	1	0	3	2	11	0	0	7	23
		Likely in the range	1–3	2–5	1–5	3–12	9–27	17–31	0–5	1–3	11–24	25–57
		Max	8	6	18	17	30	37	7	13	41	58
	Growing season length (5) days per growing season	Median	31	61	39	58	26	58	23	55	27	49
		Min	23	52	24	41	17	52	17	37	16	34
		Likely in the range	23–39	52–83	27–43	47–68	20–38	53–71	19–33	41–60	17–33	38–53
		Max	45	95	45	75	41	75	42	78	38	58
	Warm spell duration index (14) days per year	Median	34	96	20	65	23	73	35	82	34	124
		Min	26	73	17	46	16	52	22	64	28	90
		Likely in the range	29–55	77–136	20–31	49–87	18–42	58–93	23–42	75–113	32–69	98–177
		Max	69	162	55	102	54	106	63	130	83	186
	Cold spell duration index (15) days per year	Median	−5	−5	−5	−5	−6	−6	−7	−6	−5	−5
		Min	−3	−3	−2	−4	−3	−5	−4	−4	−3	−4
		Likely in the range	−4 to −7	−4 to −6	−3 to −6	−4 to −5	−4 to −7	−5 to −8	−6 to −8	−5 to −7	−3 to −5	−4 to −5
		Max	−7	−6	−6	−6	−7	−6	−8	−8	−6	−6
	Annual total precipitation (27) in %	Median	5	14	1	4	9	10	10	22	−6	−10
		Min	3	5	−1	−2	0	0	7	18	−11	0
		Likely in the range	4–8	7–16	−1 to 7	2–9	1–13	4–19	8–17	19–32	−10 to 0	−21 to −2
		Max	12	18	9	9	16	29	22	34	2	−27
	Annual total precipitation where RR > 99p of 1971/2000 (26) in %	Median	38	79	36	71	44	65	43	82	36	49
		Min	24	41	20	48	17	37	27	64	23	30
		Likely in the range	25–60	44–105	25–70	49–105	33–60	44–75	28–65	66–110	31–57	38–58
		Max	73	119	73	118	73	106	70	120	62	65

Numbers are based on 8 (RCP4.5) and 9 (RCP8.5) regional model simulations. The likely range defines the range of 66 % of all projected changes around the ensemble median

For example, the stricter (WMO) definition is more relevant to human health impacts. Any statement about possible changes in heat waves without this detailed information can easily be misleading regarding the risk of increases in future heat waves in Continental and Northern Europe.

The added value of regional climate simulations could be stated. RCMs provide higher daily precipitation intensities than GCMs, and they provide a significantly different climate change of daily precipitation intensities resulting in a smoother shift from weak to moderate and high intensities. Shifts of weak precipitation events are crucial for impact studies, in particular for hydrology, agriculture and air pollution.

The intention for this paper is not an in depth intercomparison of the time periods or scenarios, but the presentation of the new ensemble and a comparative analysis to currently used regional climate change information.

The EURO-CORDEX high-resolution data set will grow continuously and made available to the community. At the time of writing, more ensemble members are being conducted for both, the RCP4.5 and RCP8.5 scenarios. Most simulations will run until 2100 and will enlarge the ensemble. The EURO-CORDEX community agreed on common quality control and will provide a wealth of information for future climate research, impact assessment and adaptation.

References

Alcamo J, Moreno JM, Novaky B, Bindi M, Corobov R, Devoy RJN, Giannakopoulos C, Martin E, Olesen JE, Shvidenko A (2007) Climate change 2007: Impacts, Adaptation and Vulnerability. In: Parry ML, Canziani OF, Palutikof JP, van der Linden PJ, Hanson CE (eds) Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, pp 541–580

Christensen OB, Christensen JH, Machenhauer B, Botzet M (1998) Very high-resolution regional climate simulations over Scandinavia—Present climate. J Climate 11:3204–3229

Colin J, Déqué M, Radu R, Somot S (2010) Sensitivity study of heavy precipitations in Limited Area Model climate simulation: influence of the size of the domain and the use of the spectral nudging technique. Tellus A 62:591–604. doi:10.1111/j.1600-0870.2010.00467.x

Frich P, Alexander LV, Della-Marta P, Gleason B, Haylock M, Klein Tank AMG, Peterson T (2002) Observed coherent changes in climatic extremes during the second half of the twentieth century. Clim Res 19:193–212. doi:10.3354/cr019193

Giorgi F, Jones C, Asrar GR (2006) Addressing climate information needs at the regional level: the CORDEX framework. Bulletin World Meteorol Organ 58:175–183

Herrmann M, Somot S, Calmanti S, Dubois C, Sevault F (2011) Representation of daily wind speed spatial and temporal variability and intense wind events over the Mediterranean Sea using dynamical downscaling: impact of the regional climate model configuration. Nat Hazards Earth Syst Sci 11:1983–2001. doi:10.5194/nhess-11-1983-2011

Hewitt CD, Griggs DJ (2004) Ensembles-based predictions of climate changes and their impacts. Eos 85:566

IPCC SRES (2000) In: Nakićenović N, Swart R (eds) Special Report on Emissions Scenarios: a special report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge University Press, UK

Jacob D, Elizalde A, Haensler A, Hagemann S, Kumar P, Podzun R, Rechid D, Remedio AR, Saeed F, Sieck K, Teichmann C, Wilhelm C (2012) Assessing the transferability of the regional climate model REMO to different coordinated regional climate downscaling experiment (CORDEX) regions. Atmosphere 3:181–199. doi:10.3390/atmos3010181

Kotlarski S, Keuler K, Christensen OB, Déqué M, Gobiet A, Görgen K, Jacob D, Lüthi D, van Meijgaard E, Nikulin G, Suklitsch M, Teichmann C, Vautard R, Warrach-Sagi K (2013) Regional climate modelling on European scales: a joint standard evaluation of the Euro-CORDEX RCM ensemble. Climate Dynamics. In prep

Kupiainen M, Samuelsson P, Jones C, Jansson C, Willén U, Hansson U, Ullerstig A, Wang S, Döscher R (2011) Rossby Centre regional atmospheric model, RCA4. Rossby Centre Newsletter, June

Meijgaard E van, Van Ulft LH, Lenderink G, de Roode SR, Wipfler L, Boers R, Timmermans RMA (2012) Refinement and application of a regional atmospheric model for climate scenario calculations of Western Europe. Climate changes Spatial Planning publication: KvR 054/12, ISBN/EAN 978-90-8815-046-3, pp 44

Metzger MJ, Bunce RGH, Jongman RHG, Muecher CA, Watkins JW (2005) A climatic stratification of the environment of Europe. Glob Ecol Biogeogr 14:549–563

Moss RH, Edmonds JA, Hibbard KA, Manning MR, Rose SK, van Vuuren DP, Carter TR, Emori S, Kainuma M, Kram T, Meehl GA, Mitchell JFB, Nakicenovic N, Riahi K, Smith SJ, Stouffer RJ, Thomson AM, Weyant JP, Wilbanks TJ (2010) The next generation of scenarios for climate change research and assessment. Nature 463:747–756. doi:10.1038/nature08823

Pfeifer S, Gobiet A, Hänsler A, Mudelsee M, Rechid D, Ries H, Teichmann C, Weber B, Jacob D (2013) Mapping the robustness of regional climate change information. J App Meteorol Climatol. In review

Rockel B, Will A, Hense A (2008) Special issue regional climate modelling with COSMO-CLM (CCLM). Meteorol Z 17:347–348. doi:10.1127/0941-2948/2008/0309

Rogelj J, Meinshausen M, Knutti R (2012) Global warming under old and new scenarios using IPCC climate sensitivity range estimates. Nat Clim Change 2:248–253. doi:10.1038/NCLIMATE1385

Samuelsson P, Jones C, Willén U, Ullerstig A, Gollvik S, Hansson U, Jansson C, Kjellström E, Nikulin G, Wyser K (2011) The Rossby Centre Regional Climate Model RCA3: model description and performance. Tellus A 63:4–23. doi:10.1111/j.1600-0870.2010.00478.x

Skamarock WC, Klemp JB, Dudhia J, Gill DO, Duda DMBMG, Huang X-Y, Wang W, Powers JG (2008) A description of the advanced research WRF version 3. NCAR Technical note 475

Sun Y, Solomon S, Dai A, Portmann RW (2006) How often does it rain? J Clim 19: 916–934. doi:10.1175/JCLI3672.1

Taylor K, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Am Meteorol Soc 93:485–498. doi:10.1175/BAMS-D-11-00094.1

Van der Linden P, Mitchell JFB (eds) (2009) ENSEMBLES: climate change and its impacts: summary of research and results from ENSEMBLES project. Met Office Hadley Centre, Exeter

van Vuuren DP, Edmonds J, Kainuma M, Riahi K, Thomson A, Hibbard K, Hurtt GC, Kram T, Krey V, Lamarque J-F, Matsui T, Meinshausen M, Nakicenovic N, Smith SJ, Rose SK (2011) Representative concentration pathways: an overview. Climatic Change 109:5–31. doi:10.1007/s10584-011-0148-z

Vautard R, Gobiet A, Jacob D, Belda M, Colette A, Déqué M, Fernández J, García-Díez M, Goergen K, Güttler I, Halenka T, Karacostas T, Katragkou E, Keuler K, Kotlarski S, Mayer S, van Meijgaard E, Nikulin G, Patarčić M, Scinocca J, Sobolowski S, Suklitsch M, Teichmann C, Warrach-Sagi K, Wulfmeyer V, Yiou P (2013) The simulation of European heat waves from an ensemble of regional climate models within the EURO-CORDEX project. Climate Dynamics. doi:10.1007/s00382-013-1714-z

3. Article 2: Temporal and spatial scaling impacts on extreme precipitation

Atmospheric
Chemistry
and Physics

Open Access

Temporal and spatial scaling impacts on extreme precipitation

B. Eggert[1], **P. Berg**[2], **J. O. Haerter**[3], **D. Jacob**[1], **and C. Moseley**[4]

[1]Climate Service Center 2.0, Hamburg, Germany
[2]Hydrology Research unit, SMHI, Norrköping, Sweden
[3]Niels Bohr Institute, Copenhagen, Denmark
[4]Max Planck Institute for Meteorology, Hamburg, Germany

Correspondence to: B. Eggert (bastian.eggert@hzg.de)

Abstract. Convective and stratiform precipitation events have fundamentally different physical causes. Using a radar composite over Germany, this study separates these precipitation types and compares extremes at different spatial and temporal scales, ranging from 1 to 50 km and 5 min to 6 h, respectively. Four main objectives are addressed. First, we investigate extreme precipitation intensities for convective and stratiform precipitation events at different spatial and temporal resolutions to identify type-dependent space and time reduction factors and to analyze regional and seasonal differences over Germany. We find strong differences between the types, with up to 30 % higher reduction factors for convective compared to stratiform extremes, exceeding all other observed seasonal and regional differences within one type. Second, we investigate how the differences in reduction factors affect the contribution of each type to extreme events as a whole, again dependent on the scale and the threshold chosen. A clear shift occurs towards more convective extremes at higher resolution or higher percentiles. For horizontal resolutions of current climate model simulations, i.e., ~ 10 km, the temporal resolution of the data as well as the chosen threshold have profound influence on which type of extreme will be statistically dominant. Third, we compare the ratio of area to duration reduction factor for convective and stratiform events and find that convective events have lower effective advection velocities than stratiform events and are therefore more strongly affected by spatial than by temporal aggregation. Finally, we discuss the entire precipitation distribution regarding data aggregation and identify matching pairs of temporal and spatial resolutions where similar distributions are observed. The information is useful for planning observational
networks or storing model data at different temporal and spatial scales.

1 Introduction

The IPCC's fifth assessment report highlights an intensification of heavy precipitation events in North America and Europe (Hartmann et al., 2013) and projects further increase of extremes as global temperatures rise (Collins et al., 2013). The study of extreme events is complex due to a strong inhomogeneity of precipitation intensities in space and time. Assessment of precipitation extremes, e.g., as defined by an intensity threshold, is strongly scale dependent and therefore requires specification of the analyzed spatial and temporal resolution.

Even though spatial and temporal scales are far from independent (Taylor, 1938), it is often unclear how to compare data sets directly when their data are measured at differing resolutions. The data resolution needed by users, e.g., hydrologists or crop modelers, often differs from that at which observed or modeled data are recorded (Willems et al., 2012).

The primary societal interest in extreme precipitation lies in its hydrological implications, typically requiring statistics of precipitation extremes for the area of a given catchment or drainage system, which is not identical to that of model grid boxes or the observations.

Moreover, temporal scales relevant to flood risk vary enormously with area (Blöschl and Sivapalan, 1995; Westra et al., 2014): for catchments, hours to days are relevant (Mueller and Pfister, 2011), whereas urban drainage systems of ~ 10 km (Arnbjerg-Nielsen et al., 2013) are impacted at

timescales from minutes to hours (De Toffol et al., 2009), and soil erosion can occur at even smaller scales (Mueller and Pfister, 2011).

Areal reduction factors (ARFs) and intensity–duration functions have previously been used to describe the decrease of average precipitation intensity due to spatial and temporal aggregation (Bacchi and Ranzi, 1996; Smith et al., 1994). The capability of radar data to capture the spatial structure of storms was identified as a key factor in deriving the ARFs (Bacchi and Ranzi, 1996; Arnbjerg-Nielsen et al., 2013). A general outcome was that ARFs exhibit a decay with respect to the return period (Bacchi and Ranzi, 1996; Sivapalan and Blöschl, 1998) and a dependency on the observed region, resulting from different governing rainfall generation mechanisms (Sivapalan and Blöschl, 1998).

In the current study we separate the physically different processes leading to convective and stratiform type precipitation events. Using synoptic observation data, we classify precipitation events into these two types, allowing us to analyze their aggregated statistics individually across scales.

The two types physically differ in that convection is often initiated by local radiative surface heating, resulting in a buoyantly unstable atmosphere (Houze, 1997), whereas stratiform precipitation stems from large-scale frontal systems and relatively weak and uniform up-lifting. Analyzing these two types separately regarding their intensities at different scales can, e.g., be important when considering temperature changes, such as anthropogenic warming: over large scales, the changes were found to be moderate, whereas for very small scales it has been argued that the two processes may increase with warming (Trenberth, 1999; Trenberth et al., 2003; Trenberth, 2011; Lenderink and van Meijgaard, 2008), albeit at very differing rates (Berg et al., 2013). Using high-resolution model simulations, heavy precipitation at high temporal resolutions was suggested to increase strongly in a future climate and a dominant contribution to extreme events to stem from convective events (Kendon et al., 2014; Muller et al., 2011; Attema et al., 2014). In spite of their small horizontal and temporal range, convective events can cause substantial damage (Kunz, 2007; Kunz et al., 2009), e.g., through flash floods (Marchi et al., 2010).

Numerous studies have assessed the temporal and spatial characteristics of precipitation events using a storm centered, or *Lagrangian*, approach (Austin and Houze Jr., 1972; Houze Jr. and Hobbs, 1982; Moseley et al., 2013) which focuses on the storm dynamics, e.g., lifetime or history of its spatial extent. Moseley et al. (2013) showed that, for Lagrangian event histories of 30 min, the convective type can produce significantly higher intensities than the stratiform type. As we here focus on potential hydrological applications and those addressing possible impact of extremes, e.g., floods, defining events over a *fixed* surface area and time period is more appropriate (Berndtsson and Niemczynowicz, 1988; Onof et al., 1996; Bacchi and Ranzi, 1996; Michele et al., 2001; Marani, 2003, 2005). The statistics thereby constitute averages over a defined space–time window within which both dry and wet sub-intervals may occur.

In this study, we analyze at which fixed temporal and spatial scales convective precipitation dominates precipitation extremes. To this end, we analyze 2 years of mid-latitude high-resolution radar data (5 min temporally and 1 km spatially), classified by precipitation types and separated into seasons (summer vs. winter) and geographic areas (northern vs. southern Germany). Analysis of these data over large spatial and temporal periods characterizes the statistical aggregation behavior in space and time. It can quantify the requirements on minimal model resolution sufficient for the proper description of the respective extremes. Revisiting the Taylor hypothesis (Taylor, 1938), we contrast the two precipitation types as to how resolutions in space and time can be compared. Using a resulting effective advection velocity, we give a simple means of quantifying effective temporal averaging in models, resulting from a given spatial resolution.

The structure of the article is as follows: in Sect. 2 we describe the data and methods used. Section 3 presents the results for extremes at different resolutions (Sect. 3.1) and suggests a method to compare the corresponding probability density functions (PDFs) (Sect. 3.2). We close with discussions and conclusions (Sect. 4).

2 Data and methods

A Germany-wide radar composite (RADOLAN-RY) from the German Weather Service is used in this study. This data set is provided on an approximate 900 km × 900 km grid with a 1 km horizontal resolution and contains information derived from 17 radar measurement facilities (Fig. 1). The rainfall rates (R) were derived from raindrop reflectivities (Z) using the Z–R relationship (Steiner et al., 2004). The data are stored as discrete instantaneous intensities with an increasing bin size towards higher values. For the analysis, the 2-year time period covering 2007–2008 is considered. The data have been used (Moseley et al., 2013) and compared with gauge data previously (Berg et al., 2013).

For the current analysis, radar grid points are aggregated in time, i.e., $\Delta t \in \{5, 10, 15, 20, 30, 45, 60, 120, 180, 240, 360\}$ min, and in space over square grid box areas with linear dimensions $\Delta x \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 25, 50\}$ km. Aggregation includes all possible pairs $\{\Delta t, \Delta x\}$. Spatial aggregation is performed such that a coarser grid box starts at the bottom left corner of the domain and aggregates over the respective number of grid points towards the top right, with no overlap between the coarser grid boxes. As a consequence, the number of aggregated grid box scales $\sim 1/(\Delta t \Delta x^2)$. In cases where the original horizontal resolution cannot evenly be divided by the resolution of the coarser grid, the remaining grid points at the top and right border are not considered. This is the closest mimic of a gridded model.

Synoptic cloud observations, at 222 stations, obtained from the Met Office Integrated Data Archive System (MIDAS) data base (http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__dataent_ukmo-midas) are used to separate large-scale and convective precipitation following Berg et al. (2013). The locations of the stations used are shown in Fig. 1. The classification process is carried out such that first a classification is made for each station and each 3 hourly observation into *convective*, *stratiform*, *mixed* or *no observations*. Second, to ensure more stable conditions, the classifications are aggregated in space to quadrants over the region (see Fig. 1) such that each quadrant contains one single classification for each 3 hourly time period. The aggregated classification can only be *convective* (*stratiform*) if there are no simultaneous observations of *stratiform* (*convective*) in the quadrant, or else the classification will be considered to be of the *mixed* type.

For the aggregated time resolutions 5 to 180 min, the precipitation is flagged as *convective*, respectively *stratiform*, according to the corresponding 3 hourly time slice. For time resolutions longer than 3 hours, two 3 hourly time slices have to be considered. Here we classify the precipitation event as *stratiform* or *convective* only if the type is identified at least at one of the time slices and the other time slice was not identified as the opposite type of event. This procedure was found to be the best compromise between rigid classification and sufficient data availability at the coarsest averaging windows.

Next, for each averaging window, the total number of convective and stratiform *events*, i.e., single time steps with an intensity higher than $1\,\mathrm{mm\,day^{-1}}$, is counted. To ensure that enough events for statistical analysis are present, the analysis is restricted to resolutions where at least 500 convective and 500 stratiform events were detected. All other fields will be marked as insufficient (gray squares in the Figs. 3, 4 and 8).

3 Results

3.1 Quantifying the impact of spatial and temporal aggregation on convective and stratiform precipitation extremes

3.1.1 Differential impact on exceedance probabilities

We define the cumulative distribution function (CDF) as the probability of precipitation intensity exceeding a given intensity I:

$$\mathrm{CDF}(\Delta t, \Delta x, I) \equiv \frac{\int_{I}^{\infty} N(\Delta t, \Delta x, I')\mathrm{d}I'}{\int_{I_0}^{\infty} N(\Delta t, \Delta x, I')\mathrm{d}I'}, \quad (1)$$

where $N(\Delta t, \Delta x, I)$ is the number of data aggregates to resolution Δt and Δx with averaged precipitation intensity I, and I_0 is the lower measurement cutoff. In the following, we

Figure 1. Data used in the analysis. Map of Germany with the synoptic stations (red crosses) and the radar locations and approximate range (gray circles). Dashed black lines indicate the division of the domain into quadrants.

choose $I_0 = 1\,\mathrm{mm\,day^{-1}}$ throughout. $\mathrm{CDF}(\Delta t, \Delta x, I)$ thus describes the percentiles of precipitation intensity when conditioning on wet periods. Figure 2 shows $\mathrm{CDF}(\Delta t, \Delta x, I)$ for Germany for different Δt and Δx conditional on convective and stratiform events. Note the logarithmic representation of the data, i.e., the figure focuses on the high precipitation intensities between the 99.9th percentile (10^{-1}) and the 90th percentile (10^1) of the distribution.

It is important to realize the effect of aggregation at varying scales. Consider first spatial aggregation (see legend in Fig. 2). Convection forms patterns with intense and localized precipitation peaks, separated spatially by regions without precipitation (Austin and Houze Jr., 1972; Moseley et al., 2013; Berg et al., 2013). Performing averages over areas of increasing size therefore yields broad variation of averages at small spatial scales but rapid decrease of variation as data are aggregated over larger areas. Stratiform precipitation is more uniform in the sense that sampling over small areas yields a good description of the statistics also at larger areas of aggregation.

Consider now temporal aggregation from an interval well below the convective lifetime (e.g., $\ll 30\,\mathrm{min}$): the effect of temporal aggregation is to even out spatial variations due to the large-scale flow. This makes convection appear spatially
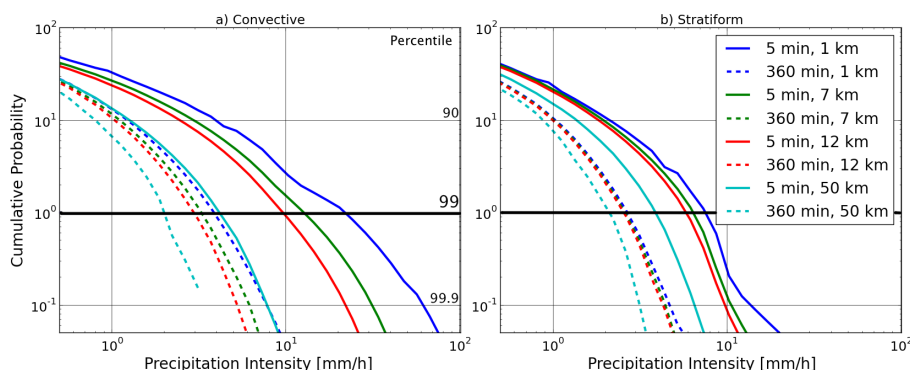
Figure 2. Cumulative probability density functions of precipitation intensities. All of Germany for the years 2007–2008, aggregated at different horizontal and temporal resolutions: **(a)** convective events and **(b)** stratiform events.

more uniform. For stratiform precipitation, patterns are already less localized in space and temporal aggregation will change the statistics less.

We make several observations in support of this assessment (Fig. 2): first, while convective precipitation can be much more intense (compare, e.g., the solid curves in Fig. 2a vs. b), the decrease of mean intensity due to aggregation is more pronounced than for stratiform precipitation. Second, we find that the relative differences in the CDFs between the 1 and 50 km data are stronger if the data are stored at 5 min resolution than for the 360 min data. For stratiform events we find almost no differences between precipitation intensities at resolutions below 12 km for a 360 min temporal resolution. Only at the largest regions, 50 km, do the spatial aggregations clearly modify the CDF as the non-precipitating region off the boundary of the event is included. This finding suggests that, for a given time resolution, there should be an associated horizontal resolution to store or collect data, i.e., a resolution that carries similar information about the distribution function.

More generally, this highlights the close match of the convective intensity CDFs when comparing two data sets of different resolution, namely 5 min and 50 km vs. 360 min and 1 km. For these pairs of resolutions time aggregation has a similar statistical effect on precipitation intensities as does spatial aggregation.

This latter observation can be appreciated when remembering the Taylor hypothesis of "frozen turbulence" (Taylor, 1938), which states that as the mean atmospheric flow advects eddies past a station, information about spatial variations can be gained as long as the properties of the eddies remain frozen in time. Consider, for example, an average convective event with constant precipitation intensity over its lifetime. According to Berg et al. (2013) and Moseley et al. (2013) the average convective event has a lifetime of approximately 30 min, a spatial extent of ~ 10 km and a propagation speed of ~ 10 m s^{-1}. When using a 50 km grid box and 5 min temporal resolution, the event will move about 3 km;

therefore it can be assumed that the event stays in one grid box. It will affect roughly $\frac{10 \times 10}{50 \times 50} \approx 4$ % of the cell at a time. When an event of ~ 10 km cross section moves over a location with ~ 10 m s^{-1}, its passage over the location would last ~ 1000 s, which is ~ 17 min and $\frac{17}{360} \approx 5$ % of the matching time interval of 6 h.

In the following we discuss how the actual observations depart from the approximation of the Taylor hypothesis and how this departure is influenced by the precipitation type. In reality, there are complications such that events change intensity also on short timescales, many events can be superimposed in space and time, and the large-scale flow is not constant.

To proceed, we now focus on intensity changes at a specific percentile, defined for a given combination of Δt and Δx by the inverse of Eq. (1), i.e., the intensity corresponding to a choice of exceedance probability. We will later return to the entire distribution functions in Sect. 3.2. Specifically, we now choose the 99th percentile of all detected precipitation events and refer to this percentile as *extreme precipitation*. This percentile was found to be a good compromise between the aim of focusing mainly on the high end of the intensity distributions and the need for sufficient data for the statistics. Using a percentile value as threshold to define precipitation extremes is a common practice.

For varying Δx and Δt, Figs. 3 and 4 show the 99th percentile of precipitation intensities for convective (termed $\hat{I}_{\mathrm{cv}}(\Delta t, \Delta x)$) and stratiform (termed $\hat{I}_{\mathrm{ls}}(\Delta t, \Delta x)$) events, respectively, for the entire region of Germany and separated into northern and southern Germany as well as for the whole year and separated into the summer (April–September) and winter (October–March) seasons. Note that we used a nonlinear scaling for the horizontal and vertical axes to better visualize the intensity changes at very high resolutions. The same plots as in Figs. 3 and 4 but with linear scales are shown in the Supplement. In the linear case the arcs, found when connecting fields with similarly extreme intensities, become almost straight lines. Straight lines mean that for any choice of
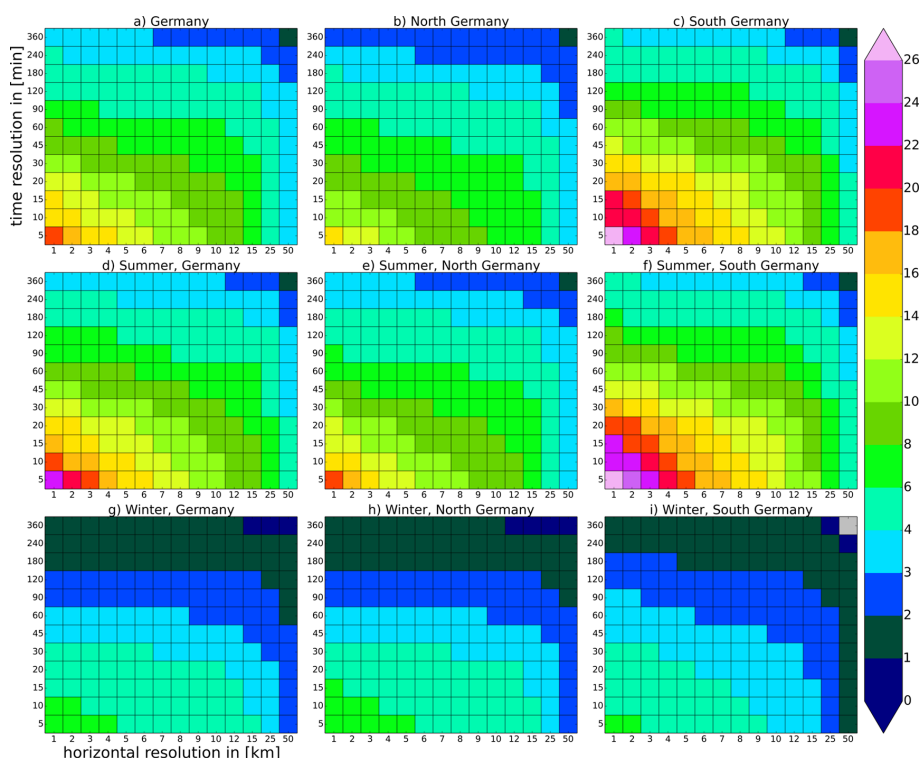
Figure 3. Convective extremes as a function of resolution. The 99th percentile of convective precipitation intensities, aggregated over different parts of Germany for the years 2007–2008, on different horizontal (horizontal axis) and temporal (vertical axis) resolutions: entire year **(a–c)**, summer season **(d–f)** and winter season **(g–i)**. All of Germany **(a, d, g)**, northern Germany **(b, e, h)**, southern Germany **(c, f, i)**; intensities given in $\mathrm{mm\,h^{-1}}$.

a resolution pair, equivalent resolutions, i.e., those of similar extremes, can be obtained by simple linear transformations.

When comparing $\hat{I}_{\mathrm{cv}}(\Delta t, \Delta x)$ (Fig. 3) to $\hat{I}_{\mathrm{ls}}(\Delta t, \Delta x)$ (Fig. 4), it is striking that at high temporal and spatial resolutions the intensity $\hat{I}_{\mathrm{ls}}$ is only about one-third of $\hat{I}_{\mathrm{cv}}$. However, $\hat{I}_{\mathrm{ls}}$ shows much less spatial and seasonal differences when compared to $\hat{I}_{\mathrm{cv}}$. For example, the increase in intensity at the highest resolution in summer vs. winter is about 220 % for $\hat{I}_{\mathrm{cv}}$ but only approximately 20 % for $\hat{I}_{\mathrm{ls}}$. This finding is in line with the relatively weak temperature response of stratiform precipitation intensities as reported recently (Berg et al., 2013).

Regionally, southern Germany exhibits higher $\hat{I}_{\mathrm{cv}}$ in summer as compared to the north. This may largely be due to not only complex orographic areas in southern Germany, e.g., the highly convectively active area of the Black Forest in southwestern Germany (Khodayar et al., 2013), but also latitudinal temperature differences and the more continental climate of the south.

The highest intensities of stratiform precipitation occur in northern Germany in the months May to September. We find that for time durations shorter than 3 h the highest intensities occur between June to August. For longer time durations, the highest intensities occur in the months September to November (see Supplement).

3.1.2 Scaling behavior of convective and stratiform precipitation events

To quantify the reduction due to spatial aggregation, we define the area reduction factor ARF(Δx) as the reduction of the 99th percentile at spatial resolution x relative to the percentile (here defined as $\hat{I}_{\mathrm{ori}}$) at the original resolution (5 min, 1 km). Varying now the spatial resolution while keeping the temporal resolution fixed (at 5 min), we define

$$\mathrm{ARF}(\Delta x) \equiv 1 - \frac{\hat{I}(\Delta x)}{\hat{I}_{\mathrm{ori}}}, \qquad (2)$$

where $\hat{I}$ and $\hat{I}_{\mathrm{ori}}$ is shorthand for either of the precipitation types. Analogously, we define the duration reduction factor DRF(Δt) as the intensity reduction due to temporal aggregation relative to $\hat{I}_{\mathrm{ori}}$, while keeping the spatial resolution at 1 km, i.e.,

$$\mathrm{DRF}(\Delta t) \equiv 1 - \frac{\hat{I}(\Delta t)}{\hat{I}_{\mathrm{ori}}}. \qquad (3)$$

ARFs and DRFs are shown in Fig. 5a and b, respectively, for both precipitation types and separately for the summer and winter seasons, as well as for northern and southern Germany. Considering $\hat{I}_{\mathrm{cv}}$, a strong intensity reduction can be
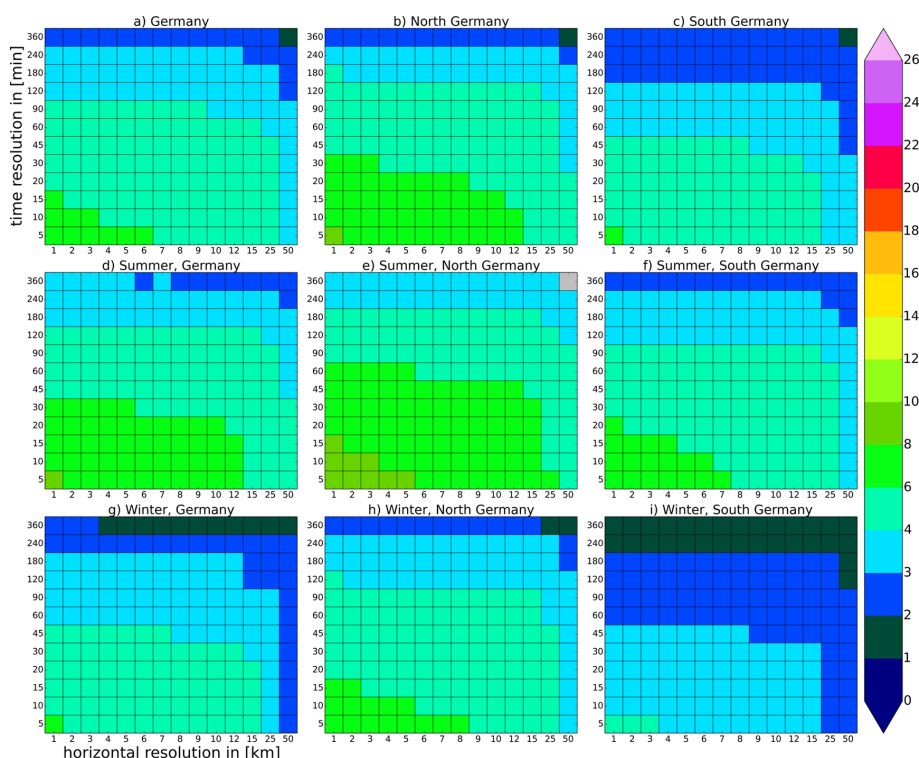
Figure 4. Stratiform extremes as a function of resolution; otherwise similar to Fig. 3.

seen when the spatial (Fig. 5a) or temporal (Fig. 5b) resolution is decreased. The reduction due to spatial aggregation shows clear seasonal and regional differences: the lowest ARFs occur in northern Germany in winter (68 % at 50 km grid size) and the highest in southern Germany in summer (84 % at 50 km grid size). Temporal aggregation is nearly independent of seasonal and regional distinctions and reaches values of about 80 to 85 % at a 6 hourly resolution. The differences found between $\hat{I}_{cv}$ and $\hat{I}_{ls}$ are hence larger than all other seasonal or regional differences within one type.

$\hat{I}_{ls}$ shows much less pronounced ARFs and DRFs. For the maximum spatial aggregation, only 52 % reduction is found in northern Germany in winter. The seasonal and regional differences are smaller than for $\hat{I}_{cv}$ and differ only by less than 10 percentage units. Temporal aggregation shows also only small regional and seasonal differences, causing DRFs of 60 to 70 % at a temporal resolution of 6 h.

3.1.3 Comparing the relevance of space compared to time aggregation

We can distinguish the behavior of spatial and temporal aggregation using two kinds of approaches (Deidda, 2000). The first approach would be to regard precipitation as a self-similar process (simple scaling). In this case the Taylor hypothesis (Taylor, 1938) would be obeyed, and temporal variations can be reinterpreted as spatial variations that are ad-

vected over a fixed location by a large-scale flow that is constant over the observed temporal and spatial scales.

Following the notion of "frozen turbulence", intensity change due to spatial aggregation can then be calculated from the intensity changes that result due to temporal aggregation multiplied by a constant velocity u, i.e., $\Delta x \approx \Delta t \cdot u$. This would only hold if precipitation extremes could be seen as objects of temporally constant characteristics that are translated by large-scale advection. If we also assume spatial inhomogeneity only to occur in the advection direction, a gauge station could be used to measure the precipitation intensities that fall over an area (Fig. 6a).

The second approach would assume that the spatial and temporal aggregation behavior of precipitation extremes would behave like a self-affine process (a process where the ratio of scales is changing as one of the scales changes). In this case, the simple linear relation that connects changes due to time aggregation with changes due to spatial aggregation through an advection velocity generally does not hold anymore (e.g., due to temporal (Fig. 6b) or spatial inhomogeneity (Fig. 6c)). A multifractal analysis is needed where, in short, the "velocity" itself would become a function of the respective spatial and temporal scales. If this function is known, it is possible also for self-affine processes to connect spatial and temporal scales. Proper understanding of the relationship between spatial and temporal aggregation is crucial,
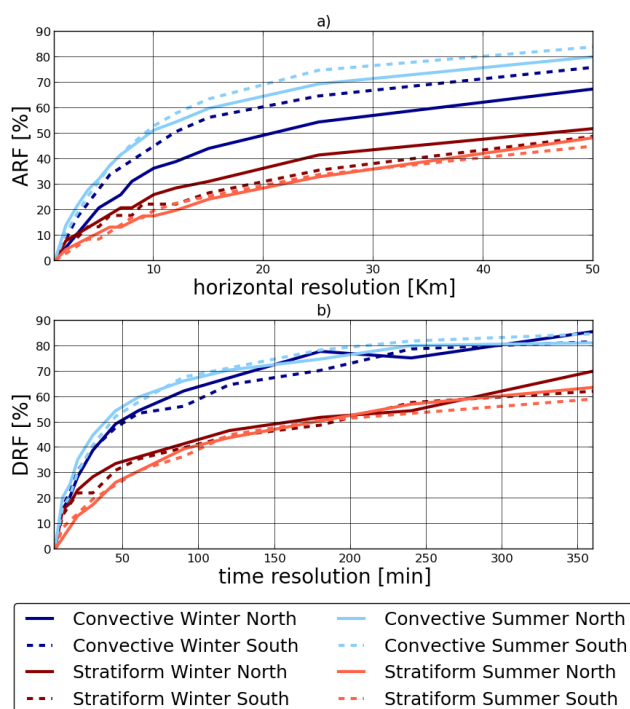
Figure 5. Area and duration reduction factors. **(a)** Area reduction factors at 5 min temporal resolution. **(b)** Duration reduction factors (DRFs) for 1 km × 1 km spatial resolution in percent for convective (blue) and stratiform (red) precipitation. Data shown for the summer and winter seasons and northern and southern Germany.
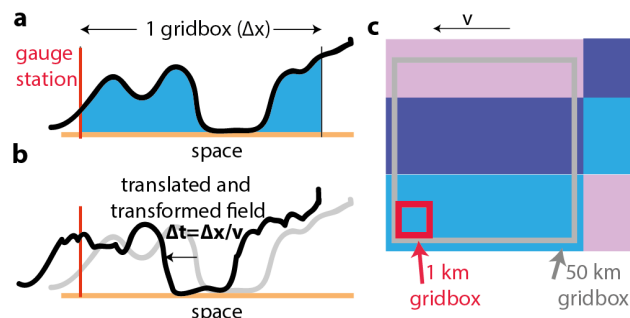

Figure 6. Schematic illustration of the Taylor hypothesis. **(a)** One-dimensional case showing space, grid box width and precipitation intensity (black curve); the location of a gauge station is marked in red. **(b)** Similar to **(a)** but illustrating how the curve may change due to small-scale dynamics after a time interval $\Delta t = \Delta x / v$, with v the atmospheric advection velocity. **(c)** Two-dimensional inhomogeneities (different colors indicate different intensities) perpendicular to the advection direction (direction indicated by the thin arrow). Small (red) and large (gray) grid boxes as marked.

e.g., for precipitation downscaling and bias correction methods (Wood et al., 2004; Piani et al., 2010a, b).

Our goal here is to characterize the differences in scaling of convective and stratiform extremes. Comparing the intensity reduction due to time aggregation for the 1 km data set (Fig. 3a, left column) with the intensity reduction that results from spatial aggregation at a temporal resolution of 5 min (bottom row), a 4 km spatial aggregation is comparable to that of spatial aggregation for roughly 15 min. Similarly, for stratiform precipitation (Fig. 4a) we find that 6 km spatial aggregation corresponds to 15 min temporally. There is hence a dependence on the precipitation type, a relation we now analyze.

Figure 7a shows for each horizontal resolution the matching temporal resolution that achieves similar intensity reduction. We describe the relation between temporal and spatial aggregation at a fixed Δx by

$$f_{\Delta x}(\Delta t) = |\hat{I}(\Delta t, \ 1\,\mathrm{km}) - \hat{I}(5\,\mathrm{min}, \ \Delta x)|. \tag{4}$$

We now define $\phi_{\Delta x}$ as the minimum value of $f_{\Delta x}$ w.r.t. Δt:

$$\phi_{\Delta x} = \min_{\Delta t} f_{\Delta x}(\Delta t). \tag{5}$$

The best matching time window Δt for a given Δx can be determined using the inverse function of $f_{\Delta x}$: $\Delta t = f^{-1}(\phi)$.

In practice, we determine Δt by an iterative numerical procedure, using first an interpolation between available resolutions for better approximation. The result for several high percentiles is shown for both precipitation types over Germany for the entire year on a log–log plot (Fig. 7a), i.e., straight lines represented power laws. If the Taylor hypothesis is obeyed, the exponent would equal unity.

Within the limitations of the relatively noisy data, we find that the data represent a straight line over most of the analyzed spatial range and can be fitted to a power law function $\Delta t = a \times \Delta x^b$ with fitting parameters a and b, or by using dimensionless variables (i.e., defining $\chi \equiv \Delta x / \Delta x_0$, $\tau \equiv \Delta t / \Delta t_0$ and $\widetilde{a} \equiv a\,\Delta x_0^b / \Delta t_0$) we have

$$\tau = \widetilde{a}\chi^b, \tag{6}$$

with fitting parameters $\widetilde{a}$ and b. The parameter $\widetilde{a}$ is a scaling parameter and describes the Δt_0 corresponding to Δx_0. χ^b describes how the relevance of space compared to the time aggregation changes with resolution.

In Fig. 7a and b, the best fit for the 99th intensity percentile is shown for convective and stratiform precipitation. We find that b is similar for both types (generally between 1.17 and 1.32), a departure from unity that should be confirmed by other data sources than the radar data at hand. An exponent $b > 1$ indicates that extreme precipitation is self-affine (self-similarity would require $b = 1$). The fractal properties of precipitation were already highlighted in earlier studies and are found to be a result of the hierarchical structure of precipitation fields (Schertzer and Lovejoy, 1987) with cells that are embedded in small mesoscale areas which in turn occur in clusters in large-scale synoptic areas (Austin and Houze Jr., 1972).

Table 1 displays $\widetilde{a}$ and b for the different percentiles shown in Fig. 7a (non-dimensional). We find that for convective pre-
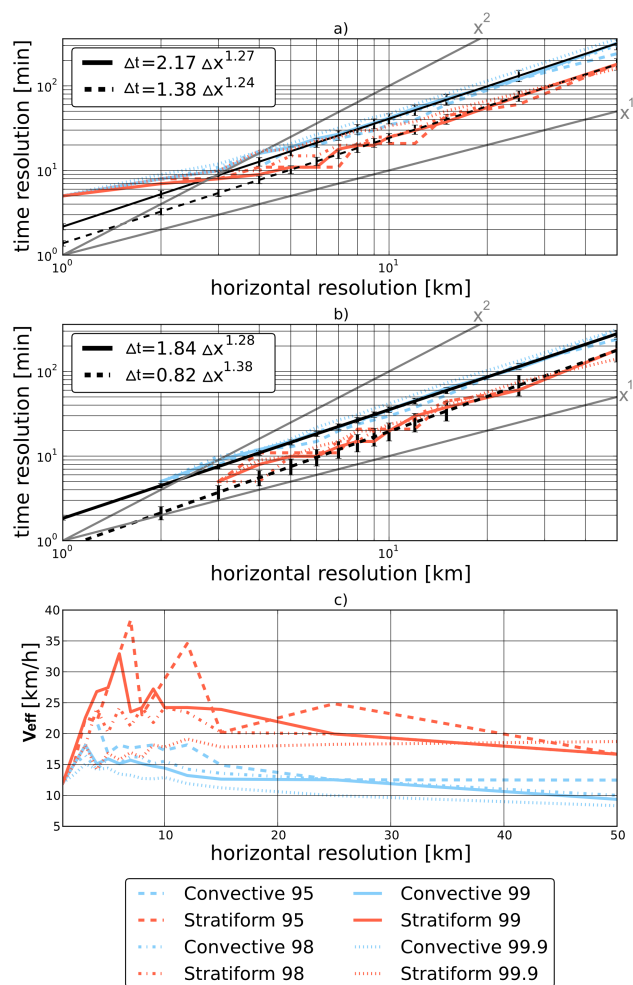
Figure 7. Consistent spatial and temporal resolutions. Δt derived using Eq. (5) for different values of Δx for convective (blue) and stratiform (red) precipitation extremes at the 95th, 98th, 99th and 99.9th percentiles. Black lines are least square fit of $\Delta t = a \times \Delta x^b$ with the fitting parameters a and b for the 99th percentile. Error bars indicate the standard deviation of parameter estimates. Gray lines show $\Delta t \sim \Delta x$ and $\Delta t \sim \Delta x^2$. **(a)** Initial resolutions $\Delta t_0 = 5\,\text{min}$, $\Delta x_0 = 1\,\text{km}$. **(b)** $\Delta t_0 = 5\,\text{min}$ and aggregated spatial resolutions $\Delta x_0 = 2\,\text{km}$ (convective) and $\Delta x_0 = 3\,\text{km}$ (stratiform). **(c)** v_{eff} (Eq. 7) for both precipitation types for Germany over the entire year.

cipitation $\widetilde{a}$ is near 0.5. Within the error bars there is no obvious dependence on percentile. This is also the case for the stratiform type but not for the 99.9th percentile, which has higher $\widetilde{a}$ and lower b values.

Since the values of b are similar for both precipitation types (Table 1), the difference between the matching temporal resolution of stratiform and convective events is kept constant over the entire range of Δx analyzed. We find that the different scaling between the two precipitation types mainly results from the varying $\widetilde{a}$.

Table 1. Estimation of the exponent b and the pre-factor $\widetilde{a}$ for the different precipitation types and percentiles together with the standard deviation of the parameter estimate.

Precipitation type	Percentile	$\widetilde{a}$	b
Convective	95th	0.51 ± 0.05	1.17 ± 0.03
	98th	0.45 ± 0.03	1.25 ± 0.02
	99th	0.43 ± 0.04	1.27 ± 0.02
	99.9th	0.55 ± 0.01	1.24 ± 0.01
	mean	0.49 ± 0.03	1.23 ± 0.02
Stratiform	95th	0.20 ± 0.04	1.32 ± 0.06
	98th	0.35 ± 0.03	1.18 ± 0.02
	99th	0.28 ± 0.02	1.24 ± 0.02
	99.9th	0.76 ± 0.03	0.96 ± 0.01
	mean*	0.28 ± 0.03	1.25 ± 0.03

* Excluding the 99.9th percentile.

Note also the kink in the observed curves in Fig. 7a at about 6 km, where a change of slope is observed. To show that this kink is a manifestation of the scale mismatch, we aggregate data spatially to 2 km (3 km for stratiform) horizontal resolution and re-plot (Fig. 7b). Due to this procedure the kink almost vanished. This test shows that aligning resolutions according to Eq. (6) allows smooth scaling.

For further analysis, and to make contact to the Taylor hypothesis, we use the ratio of the matching Δx and Δt to calculate the mean *effective* advection velocity, which we call v_{eff}. We define

$$v_{\text{eff}}(\chi) \equiv \chi/\tau = \chi^{1-b}/\widetilde{a}. \qquad (7)$$

This effective velocity is not obviously the same as the velocity obtained by tracking algorithms, such as in Moseley et al. (2013), as v_{eff} combines all reasons for changes caused by aggregation. The main sources for these changes are advection of the precipitation field out of the grid box, temporal inhomogeneity caused by the temporal evolution of the precipitation event (Fig. 6b) and horizontal inhomogeneities perpendicular to the advection direction, which will increase the area reduction factors (Fig. 6c).

Figure 7c shows v_{eff} calculated for different Δx for the 95th, 98th, 99th and 99.9th percentile, using data without seasonal distinctions over Germany. v_{eff} lies in the same range as the velocities calculated by Deidda (2000) and Moseley et al. (2013) who calculated the velocities using tracking techniques. This shows that advection is likely the major source for changes due to temporal and horizontal aggregation. Low v_{eff} for horizontal resolutions below about 2 to 4 km are again a result of the mismatch of the 5 min temporal resolution and the 1 km spatial resolution explained above.

Note the deviating value of $\widetilde{a}$ for the 99.9th percentile of stratiform precipitation. This could be explained by mesoscale stratiform systems with embedded convection,

i.e., systems that are somewhat intermediate between stratiform and convective events. The corresponding graph (Fig. 7c) shows intermediary behavior, connecting the curves of convective precipitation (low Δx) to those of stratiform precipitation at high Δx. Due to substantial noise at high spatial resolution it is not possible to identify whether v_{eff} shows a constant behavior ($b = 1$) at the high resolutions, therefore the results in Zawadzki (1973) and Waymire et al. (1984) that indicate the Taylor hypothesis holds for timescales less than 40 min can neither be confirmed nor rejected.

Realizing that v_{eff} combines all sources for changes caused by aggregation enables a simplified view on the aggregation process. In a similar way as in Deidda (2000) we can use v_{eff} to generalize the Taylor hypothesis for a self-affine process by using v_{eff} instead of a constant velocity to describe the relation between space and time. Following the Taylor hypothesis we can now interpret the matching temporal and spatial scales from Fig. 7a as the mean time that is needed to advect the information about the precipitation field over the matching horizontal scale (implicitly including all other sources of aggregation changes as described above). For example the typical timescale for a convective precipitation area to cross a grid box with a 10 km grid size, a typical resolution of state-of-the-art climate models, would be about 40 min. For a stratiform precipitation event the information about the precipitation field is already captured after about 20 to 25 min. Reasons for the lower effective advection velocity might be that stratiform events are statistically more homogeneous than convective events which results in a shorter period to capture the structure of the event. Also, convective events often occur at high-pressure weather conditions where low wind velocities might entail lower advection velocities.

Aggregation effects at a specific resolution will always be a combination of duration and area reduction factors. Connecting space and timescales using v_{eff} allows the association of temporal and spatial scales, shown in Fig. 7a. If, for a given spatial resolution, a larger temporal output period is used as indicated by Fig. 7a, the event will on average be advected beyond the grid box area, leading to high duration reduction factors (a "smearing out").

3.1.4 Dominance of convective vs. stratiform extremes including event occurrences

So far we have only illustrated differences in the 99th percentiles of detected convective and stratiform events with precipitation intensities above 1 mm day^{-1}, i.e., conditional probability density functions. The sample size therefore depends on the number of detections of the specific precipitation type, the resolution of the data set and the area fraction in the detected quadrants with precipitation intensities higher than the specified threshold. Including the events without precipitation in the statistics will have a major impact on the percentile values; therefore a sensitivity analysis per-

Table 2. Occurrence of convective and stratiform events. Number of quadrants of Germany classified as convective (C) or stratiform (S) in the 3 hourly synoptic observations. The maximum possible values for the 2 years and for all four quadrants is 23 360. This number reduces by about half for the seasonal data and again by half for the sub-regions of Germany.

Area	Type	Year	Summer	Winter
All	S	1358	206	1152
All	C	1537	1270	267
North	S	761	103	658
North	C	741	590	151
South	S	597	103	494
South	C	796	680	116

forming the same analyses shown in Figs. 3 and 4 but with non-conditional probability density functions was done (not shown). This demonstrated that v_{eff} is not strongly affected by this threshold. Naturally, due to the high number of non-precipitation values, the high percentiles show correspondingly lower intensities. Table 2 indicates the event occurrences classified as convective or stratiform in the 3 hourly synoptic observations.

To consider the strong variation in occurrences, e.g., concerning season, we find that also the relative occurrence frequency of the two types of events has to be accounted for. We again use the 99th percentile for all data above 1 mm day^{-1}, but now without distinction of precipitation type, for each aggregation interval as well as for each region and season. In the following we redefine $\hat{I}$ as the corresponding intensity (see Supplement for $\hat{I}$ values).

To assess the relative likelihood of a certain precipitation type to cause extreme precipitation, Fig. 8 shows the ratio of the number of convective events exceeding the intensity $\hat{I}$ vs. the total number (convective + stratiform) of events exceeding $\hat{I}$, i.e., $N_{\mathrm{cv}}(I > \hat{I})/(N_{\mathrm{cv}}(I > \hat{I}) + N_{\mathrm{ls}}(I > \hat{I}))$.

However, dominance again depends on resolution: e.g., in southern Germany (all year) 80–90 % of precipitation extremes are of the convective type for the higher resolutions. Only when the data are aggregated to resolutions with grid spacings of 25 km and more does the percentage of stratiform events become appreciable. Even stronger differences occur between seasons: in summer, convection dominates extremes but is of less importance in winter (less than 10 % for the aggregated data sets and less than 35 % even at the very high-resolution data sets).

It is important to note that we used a percentile threshold for this analysis and the corresponding intensity threshold fluctuates with seasons. To test whether our findings simply are a consequence of overall higher intensities in summer we also compare similar intensities for summer and winter (using the 98th percentile for summer and the 99th percentile in winter, see Fig. 8g–i and Supplement). This revealed that seasonal differences nonetheless prevail.
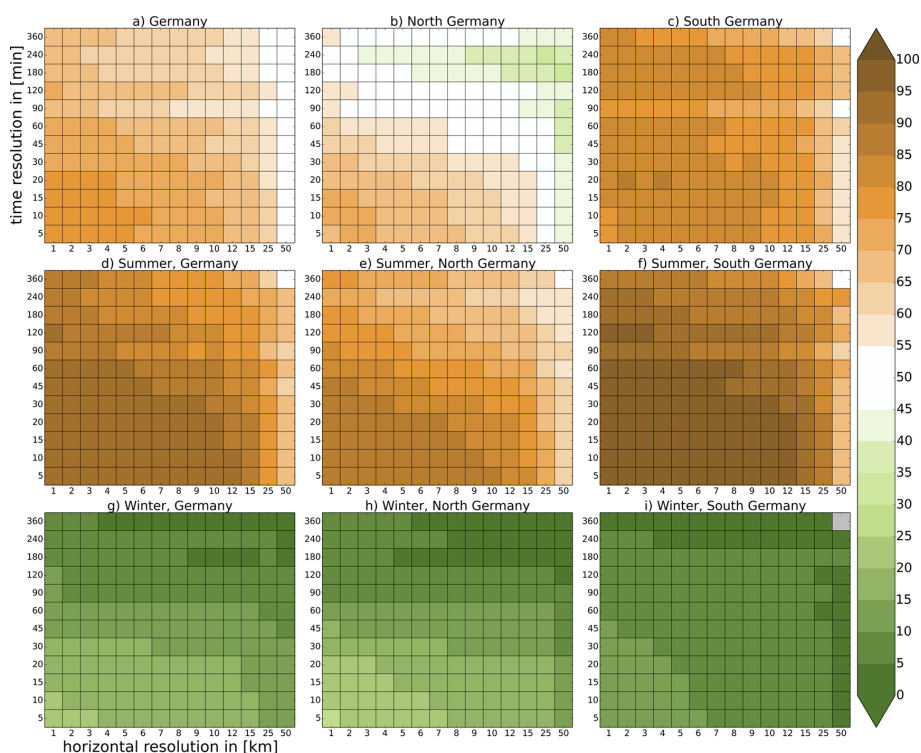
Figure 8. Convective dominance as a function of resolution including dry periods. The ratio of the number of convective precipitation events with precipitation intensities greater than or equal to the threshold intensity. Threshold intensity is defined as the 99th percentile of total precipitation intensities over the different parts of Germany for the years 2007–2008. Panels otherwise as in Fig. 3.

Figure 9 shows the convective dominance as a function of the horizontal resolution for the 95th, 98th, 99th and 99.9th percentiles. The role of convective precipitation in the extremes increases with higher percentiles, and convective precipitation becomes more relevant also over larger aggregated areas and time steps (see Supplement). At relatively low percentiles convective and stratiform events have the same exceedance probability, but with increasing percentile convection dominates, especially at high spatial resolution.

3.2 Assessing PDF changes due to data aggregation

The results of Sect. 3.1 highlight the interdependence of spatial and temporal scales and their impact on extreme precipitation. Changing resolutions, however, modifies the entire distribution function. To give an estimate of the information loss due to the aggregation process, we adopt a measure similar to that of the Perkins skill score (Perkins et al., 2007), originally designed to validate a model against observations by assigning a skill score. Here, we use it to quantify the overlap between two intensity PDFs at different horizontal and temporal resolutions. We define the *PDF overlap* as
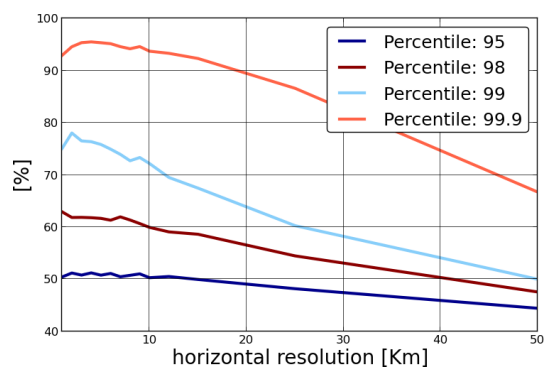
Figure 9. Convective dominance vs. horizontal resolution. The ratio of the number of convective precipitation events with precipitation intensities greater than or equal to the labeled percentile of total precipitation intensities over entire Germany for the years 2007–2008. The data are aggregated to 5 min temporal and different horizontal resolutions.

$$S(\Delta t_1, \Delta x_1; \Delta t_2, \Delta x_2) \equiv \int\limits_{I_0}^{\infty} \min\left(\rho_{\Delta t_1, \Delta x_1}(I), \rho_{\Delta t_2, \Delta x_2}(I)\right) \mathrm{d}I, \quad (8)$$

where I is precipitation intensity, I_0 is the measurement cutoff, $\rho_{\Delta t, \Delta x}(I)$ is the normalized PDF as in Eq. (1) and $\min(\cdot, \cdot)$ gives the minimum of the two arguments. Hence,
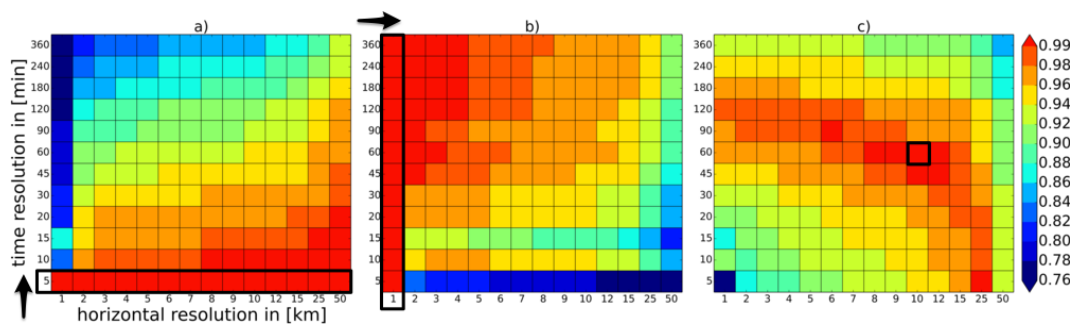
Figure 10. PDF overlap for convective precipitation intensity. All of Germany for the years 2007–2008, aggregated to different horizontal (horizontal axis) and temporal (vertical axis) resolutions. **(a)** PDF overlap of each horizontal resolution between every temporal resolution and the 5 min data. **(b)** PDF overlap of each temporal resolution between every horizontal resolution and the 1 km data. **(c)** PDF overlap of each horizontal and temporal resolution compared to the 10 km, 60 min data.

$S(\Delta t_1, \Delta x_1; \Delta t_2, \Delta x_2)$ quantifies the overlap between PDFs of aggregated data at the spatiotemporal resolutions (Δt_1, Δx_1) and (Δt_2, Δx_2). If the two PDFs are identical, the overlap value is 1; if there is no overlap at all, it is 0. The PDF overlap is a means of comparing not only a fixed percentile of precipitation intensity but measuring the similarity of entire distribution functions. It is hence a way to quantify our initially qualitative discussion regarding Fig. 2.

We aggregate convective precipitation intensities over Germany and present the PDF overlap in three different ways: Fig. 10a shows the PDF overlap between the aggregated time resolution with the corresponding 5 min data but at fixed horizontal resolution, i.e., $S(5\,\text{min}, \Delta x; \Delta t, \Delta x)$ at matrix element position ($\Delta t, \Delta x$). For the spatially highly resolved data ($\Delta x < 7\,\text{km}$), the PDF overlap degrades quickly when temporal resolution is reduced, while degradation is much slower at lower spatial resolution. In practice, if a defined spatial area, say a metropolitan region of 25 km, is of interest, performing measurements at 60 min resolution may lead to a tolerable margin of error while a smaller region of 2 km would require 5 or 10 min temporal resolution for the same margin of error. The chart could hence be used to estimate the error when data are available at one resolution but another is of interest. In Fig. 10b we present an analogous analysis, but we have now fixed the temporal resolution and compare to the 1 km data sets, i.e., $S(\Delta t, 1\,\text{km}; \Delta t, \Delta x)$ at matrix element position ($\Delta t, \Delta x$). A similar pattern emerges with degradation now occurring for decreased spatial resolution.

In a third analysis (Fig. 10c) we calculate the overlap $S(60\,\text{min}, 10\,\text{km}; \Delta t, \Delta x)$ among aggregated data of spatiotemporal resolution (t, x) and the data set at 60 min temporal resolution and 10 km spatial resolution. This reference point was chosen because it is close to current state-of-the-art regional climate model simulation over Europe. The plot shows a ridge with values close to 1, ranging from 5 min and 25 km to 120 min and 1 km resolution. Apparently all spatiotemporal resolutions along this curve produce PDFs which

differ only slightly from the 5 min, 10 km aggregation. PDF overlap values quickly decrease when departing from this ridge. Comparing this ridge with the intensity decrease in the 99th percentile as illustrated in Fig. 3a, we find that the PDF overlap mirrors the changes found in the 99th percentile. Using cumulative PDF measures as the Kolmogorov–Smirnov statistics is an alternative way of comparing PDFs. Figure 10c shows that different pairs of resolution give very similar PDFs. This can be used when comparing data sets of different resolution. This information also proved to be useful for statistical bias correction, further analyzed in the paper by Haerter et al. (2015).

For stratiform precipitation (Fig. 11), the analogous PDF overlap degrades more slowly compared to convective precipitation. For example, at a 50 km grid size we find that twice the temporal aggregation can be tolerated as compared to convective precipitation when a given PDF overlap is demanded (Fig. 11a). Similar conclusions hold for the degradation as function of horizontal resolution (Fig. 11b). Starting at about 20 min we again find that the Δx can be increased to about twice the value for convective events to achieve the same PDF overlap value. For the overlap $S(60\,\text{min}, 10\,\text{km}; \Delta t, \Delta x)$, shown in Fig. 11c, the lower sensitivity to resolution changes for stratiform precipitation translates to a substantial widening of the red-shaded area near the ridge, indicating much lower errors of estimating extremes at unavailable resolutions when stratiform precipitation is concerned compared to the case for convective precipitation (Fig. 10c). Performing measurements over extended regions can already serve as a reasonable predictor of more local extremes. We also find that due to the different area and duration reduction factors of stratiform and convective type events, the ridge with values close to 1 is shifting. For the stratiform type we find that this ridge ranges from 5 min and 25 km to 90 min and 1 km resolution.
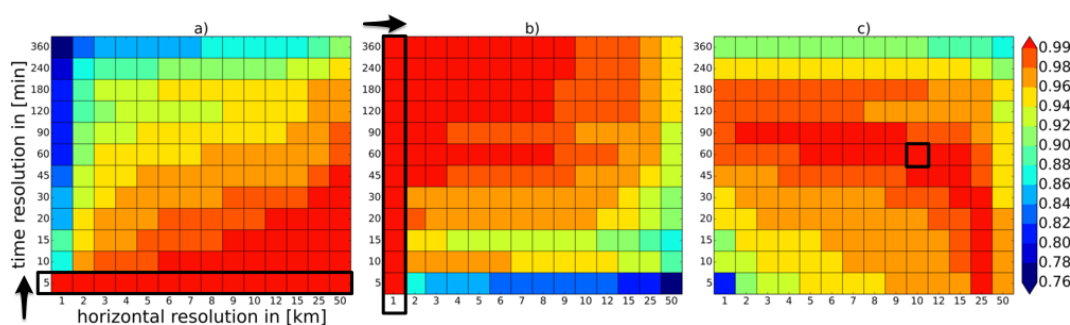
Figure 11. PDF overlap of stratiform precipitation intensity; otherwise similar to Fig. 10.

4   Discussion and conclusions

Precipitation is strongly inhomogeneous in time and space. Averaging over a specific temporal or spatial interval therefore transforms the distribution function. The resulting smoothing especially affects the extreme values, as it narrows the distribution function while preserving the mean. In this study, the focus is on how such averaging affects the two synoptically identifiable precipitation types, namely stratiform and convective extreme precipitation events. Convective events are known to produce strong, short-duration and localized precipitation while stratiform events are less bursty and cover larger areas. Using synoptic observations we separate radar-derived high-resolution precipitation intensities conditional on events of either of these two types. Unlike other studies, we here concentrate on the different aggregation behavior of the two precipitation types at different seasons and regions of Germany. Although we have not analyzed this behavior in other regions and climate zones, we expect that the findings will depend on the mean advection velocity and also the orography might have an impact on results.

4.1   Space–time dependency of intensity distributions

We found that convective extremes were considerably stronger in the south than in the north of Germany and also showed clear seasonal differences with the highest extremes occurring in summer. Stratiform extremes showed much more moderate differences over seasons and regions.

When aggregating data temporally or spatially, we find much stronger reduction for convective than for stratiform events (about 20 to 30 % higher). These differences are larger than seasonal or regional differences that were observed within one type. This highlights the importance of distinguishing between these two types of events, for example for statistical downscaling exercises. After the type separation, only the convective extremes show clear regional and seasonal differences and only in the area reduction factors. For the convective type, the strongest intensity reductions with

spatial scale were found in southern Germany in summer and the lowest in northern Germany in winter.

4.2   Temporal and spatial scales at which shifts occur between dominantly convective and dominantly stratiform extreme events

Depending on the spatial and temporal resolution, different meteorological events will be considered extreme. We point out that this makes it difficult to compare different studies of extremes in which these extremes were defined at different scales. To demonstrate this we present the contribution of convective events to the total, as a function of data aggregation, for the 99th percentile of all precipitation events.

This information is needed to identify which space–time resolutions contain comparable information about the distribution function, including the extremes. It will further help to identify at which resolution and percentile one can expect to obtain information about convective extreme precipitation events. Besides expected seasonal and regional differences with higher contribution of convective events in summer and over southern Germany, we also found a clear dependency on the scale and the threshold used. Over northern Germany, stratiform events contribute to the 99th percentile extremes only at horizontal resolutions coarser than 12 km when the duration interval is kept constant to 5 min. For a higher threshold (99.9th percentile), convective events dominate even more strongly and convective extremes consequently prevail over even larger areas and durations.

4.3   Pairs of temporal and spatial resolutions with similar aggregation effects on the extremes

For proper choice of model output resolution, precipitation downscaling as well as bias correction, the relation between the DRFs as compared to ARFs is important. Originating from the radar data resolution of 5 min temporally and 1 km spatially, we produced sequences of aggregation, both in space and time, yielding (i) temporally aggregated intensities for spatial scales held fixed and (ii) spatially aggregated intensity for temporal scales held fixed. Associating the respective aggregation resolution by matching identical pre-

cipitation extremes, we yield pairs of temporal and spatial resolutions, which define a curve.

The results allow us, e.g., to identify pairs (Δx, Δt) of spatial and temporal resolutions for which the decrease in extreme precipitation intensities due to temporal aggregation matches that due to horizontal aggregation. In terms of the Taylor hypothesis, the timescales can roughly be viewed as the mean duration needed to advect the precipitation pattern by the width of a grid box (Fig. 6).

For example, if for a given horizontal grid size a larger temporal output interval is used the event will likely be advected further than the size of the grid box, leading to strong duration reduction factors. We find that for state-of-the-art regional climate simulations, performed at a 11 km horizontal resolution, the temporal resolution needed in order to avoid stronger duration than area reduction effects would be approximately 20 to 25 min.

In practice, in regional climate models the temporal output is often lower than the resolution computed here. It should therefore be reconsidered why many regional models do not output at sub-hourly frequency and why often only daily averages are stored.

If a model can resolve some small-scale features, e.g., convective extremes, information can only be preserved by outputting at the appropriate temporal resolution, while information gets lost when using lower temporal resolutions (Fig. 8). High temporal resolution is accessible by most models already (most models have computing time steps $\sim$ seconds–minutes) but is not routinely output at such short periods. Recording at higher frequency would mainly affect storage space and not simulation run time (assuming efficient I/O handling).

The pairs of corresponding grid sizes and durations define a velocity v_{eff}, which can be used to generalize the Taylor hypothesis to the situation where temporal scales change disproportionately compared to spatial scales (self-affinity; Deidda, 2000). For constant v_{eff} as function of spatial scale, the Taylor hypothesis would be obeyed. However, v_{eff} of convective and stratiform extreme precipitation algebraically decreases with increasing Δx with similar exponents for both precipitation types. The main scaling difference between convective and stratiform events can be described by a constant scaling factor. This scaling factor leads to about 1.75 times higher advection velocities for stratiform than for convective events.

4.4 PDF overlap

Changes caused by temporal aggregation depend on the spatial scale of the data and vice versa. We examine these dependencies by comparing pairs of PDFs derived for different aggregation resolutions using a method developed by Perkins et al. (2007), here defined as PDF overlap.

We find that PDF changes that were observed when decreasing the temporal resolution from 5 min to 2 h at 50 km

horizontal resolution are quantitatively comparable with PDF changes when going from 5 to 30 min at 10 km horizontal resolution or from 5 to 10 min at 2 km horizontal resolution.

Furthermore, we show that the PDF overlap of a certain reference resolution (we chose as an example 60 min, 10 km) compared to all other aggregated resolutions shows a ridge with values close to 1. This ridge ranges from 5 min and 25 km to 120 min at 1 km resolution for convective type events (Fig. 10c) and from 5 min and 25 km to 90 min at 1 km resolution for stratiform events (Fig. 10c). These differences can be explained by the strong area reduction factors found for the convective type. The patterns found in this analysis are very similar to the patterns found in Figs. 3 and 4, highlighting that most of the differences found in the PDF overlap result from changes in the extremes.

The Supplement related to this article is available online at doi:10.5194/acp-15-5957-2015-supplement.

References

Arnbjerg-Nielsen, K., Willems, P., Olsson, J., Beecham, S., Pathirana, A., Bülow Gregersen, I., Madsen, H., and Nguyen, V.-T.-V.: Impacts of climate change on rainfall extremes and urban drainage systems: a review, Water Sci. Technol., 68, 16–28, doi:10.2166/wst.2013.251, 2013.

Attema, J. J., Loriaux, J. M., and Lenderink, G.: Extreme precipitation response to climate perturbations in an atmospheric mesoscale model, Environ. Res. Lett., 9, 014003, doi:10.1088/1748-9326/9/1/014003, 2014.

Austin, P. M., and Houze Jr., R. A.: Analysis of the structure of precipitation patterns in New England, J. Appl. Meteorol., 11, 926–935, doi:10.1175/1520-0450(1972)011<0926:AOTSOP>2.0.CO;2, 1972.

Bacchi, B. and Ranzi, R.: On the derivation of the areal reduction factor of storms, Atmos. Res., 42, 123–135, doi:10.1016/0169-8095(95)00058-5, 1996.

Berg, P., Moseley, C., and Haerter, J. O.: Strong increase in convective precipitation in response to higher temperatures, Nat. Geosci., 6, 181–185, doi:10.1038/ngeo1731, 2013.

Berndtsson, R. and Niemczynowicz, J.: Spatial and temporal scales in rainfall analysis – Some aspects and future perspectives, J. Hydrol., 100, 293–313, 1988.

Blöschl, G. and Sivapalan, M.: Scale issues in hydrological modelling: a review, Hydrol. Process., 9, 251–290, doi:10.1002/hyp.3360090305, 1995.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A., and Wehner, M.: Long-term climate change: projections, commitments and irreversibility, in: Climate Change 2013: the Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, UK and New York, NY, USA, 2013.

Deidda, R.: Rainfall downscaling in a space-time multifractal framework, Water Resour. Res., 36, 1779–1794, doi:10.1029/2000WR900038, 2000.

De Toffol, S., Laghari, A. N., and Rauch, W.: Are extreme rainfall intensities more frequent? Analysis of trends in rainfall patterns relevant to urban drainage systems, Water Sci. Technol., 59, 1769–1776, doi:10.2166/wst.2009.182, 2009.

Haerter, J. O., Eggert, B., Moseley, C., Piani, C., and Berg, P.: Statistical precipitation bias correction of gridded model data using point measurements, Geophys. Res. Lett., 42, 1919–1929, doi:10.1002/2015GL063188, 2015.

Hartmann, D., Tank, A. K., Rusticucci, M., Alexander, L., Brönnimann, S., Charabi, Y., Dentener, F., Dlugokencky, E., Easterling, D., Kaplan, A., Soden, B., Thorne, P., Wild, M., and Zhai, P.: Observations: atmosphere and surface, in: Climate Change 2013: the Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK and New York, NY, USA, 2013.

Houze, R. A.: Stratiform precipitation in regions of convection: a meteorological paradox?, B. Am. Meteorol. Soc., 78, 2179–2196, doi:10.1175/1520-0477(1997)078<2179:SPIROC>2.0.CO;2, 1997.

Houze Jr., R. A., and Hobbs, P. V.: Organization and structure of precipitating cloud systems, Adv. Geophys., 24, 225–315, 1982.

Kendon, E. J., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., and Senior, C. A.: Heavier summer downpours with climate change revealed by weather forecast resolution model, Nat. Clim. Change, 4, 1–7, doi:10.1038/NCLIMATE2258, 2014.

Khodayar, S., Kalthoff, N., and Scha, G.: The impact of soil moisture variability on seasonal convective precipitation simulations, Part I: Validation, feedbacks, and realistic initialisation, Meteorol. Z., 22, 489–505, doi:10.1127/0941-2948/2013/0403, 2013.

Kunz, M.: Von Wettersystemen zu Extremereignissen: Gefährdungsanalyse über orografisch strukturiertem Gelände, in: Verständnis, Vorsorge und Bewältigung von Naturkatastrophen, Abschlusssymposium 2007, Graduiertenkolleg "Naturkatastrophen", 24/25 July 2007, edited by: Senitz, S., Universitätsverlag, Karlsruhe, 195–203, 2007.

Kunz, M., Sander, J., and Kottmeier, C.: Recent trends of thunderstorm and hailstorm frequency and their relation to atmospheric characteristics in southwest Germany, Int. J. Climatol., 29, 2283–2297, doi:10.1002/joc.1865, 2009.

Lenderink, G. and van Meijgaard, E.: Increase in hourly precipitation extremes beyond expectations from temperature changes, Nat. Geosci., 1, 511–514, doi:10.1038/ngeo262, 2008.

Marani, M.: On the correlation structure of continuous and discrete point rainfall, Water Resour. Res., 39, 1128, doi:10.1029/2002WR001456, 2003.

Marani, M.: Non-power-law-scale properties of rainfall in space and time, Water Resour. Res., 41, W08413, doi:10.1029/2004WR003822, 2005.

Marchi, L., Borga, M., Preciso, E., and Gaume, E.: Characterisation of selected extreme flash floods in Europe and implications for flood risk management, J. Hydrol., 394, 118–133, doi:10.1016/j.jhydrol.2010.07.017, 2010.

Michele, C. D., Kottegoda, N. T., Rosso, R., and De Michele, C.: The derivation of areal reduction factor of storm rainfall from its scaling properties, Water Resour. Res., 37, 3247–3252, doi:10.1029/2001WR000346, 2001.

Moseley, C., Berg, P., and Haerter, J. O.: Probing the precipitation life cycle by iterative rain cell tracking, J. Geophys. Res.-Atmos., 118, 13361–13370, doi:10.1002/2013JD020868, 2013.

Mueller, E. N. and Pfister, A.: Increasing occurrence of high-intensity rainstorm events relevant for the generation of soil erosion in a temperate lowland region in Central Europe, J. Hydrol., 411, 266–278, doi:10.1016/j.jhydrol.2011.10.005, 2011.

Muller, C. J., O'Gorman, P. A., and Back, L. E.: Intensification of Precipitation Extremes with Warming in a Cloud-Resolving Model, J. Climate, 24, 2784–2800, doi:10.1175/2011JCLI3876.1, 2011.

Onof, C., Northrop, P., Wheater, H. S., and Isham, V.: Spatiotemporal storm structure and scaling property analysis for modeling, J. Geophys. Res., 101, 26415, doi:10.1029/96JD01378, 1996.

Perkins, S. E., Pitman, A. J., Holbrook, N. J., and McAneney, J.: Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions, J. Climate, 20, 4356–4376, doi:10.1175/JCLI4253.1, 2007.

Piani, C., Haerter, J., and Coppola, E.: Statistical bias correction for daily precipitation in regional climate models over Europe, Theor. Appl. Climatol., 99, 187–192, 2010a.

Piani, C., Weedon, G., Best, M., Gomes, S., Viterbo, P., Hagemann, S., and Haerter, J.: Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models, J. Hydrol., 395, 199–215, 2010b.

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

Sivapalan, M. and Blöschl, G.: Transformation of point rainfall to areal rainfall: intensity-duration-frequency curves, J. Hydrol., 204, 150–167, 1998.

Smith, J., Bradley, A., and Baeck, M.: The space-time structure of extreme storm rainfall in the southern plains, J. Appl. Meteorol., 33, 1402–1417, doi:10.1175/1520-0450(1994)033<1402:TSSOES>2.0.CO;2, 1994.

Steiner, M., Smith, J. A., and Uijlenhoet, R.: A microphysical interpretation of radar reflectivity-rain rate relationships, J. Atmos. Sci., 61, 1114–1131, 2004.

Taylor, G.: The spectrum of turbulence, P. Roy. Soc. Lond., 164, 476–490, 1938.

Trenberth, K. E.: Conceptual Framework for Changes of Extremes of the Hydrological Cycle with Climate Change, Springer Netherlands, 327–339, doi:10.1007/978-94-015-9265-9_18, 1999.

Trenberth, K. E.: Changes in precipitation with climate change, Clim. Res., 47, 123–138, doi:10.3354/cr00953, 2011.

Trenberth, K. E., Dai, A., Rasmussen, R. M., and Parsons, D. B.: The changing character of precipitation, B. Am. Meteorol. Soc., 84, 1205–1217, doi:10.1175/BAMS-84-9-1205, 2003.

Waymire, E., Gupta, V. K., and Rodriguez-Iturbe, I.: A spectral theory of rainfall intensity at the meso-β scale, Water Resour. Res., 20, 1453–1465, doi:10.1029/WR020i010p01453, 1984.

Westra, S., Fowler, H. J., Evans, J. P., Alexander, L. V., Berg, P., Johnson, F., Kendon, E. J., Lenderink, G., and Roberts, N. M.: Future changes to the intensity and frequency of short-duration extreme rainfall, Rev. Geophys., 52, 2014RG000464, doi:10.1002/2014RG000464, 2014.

Willems, P., Arnbjerg-Nielsen, K., Olsson, J., and Nguyen, V.: Climate change impact assessment on urban rainfall extremes and urban drainage: methods and shortcomings, Atmos. Res., 103, 106–118, doi:10.1016/j.atmosres.2011.04.003, 2012.

Wood, A. W., Leung, L. R., Sridhar, V., and Lettenmaier, D.: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, Climatic Change, 62, 189–216, 2004.

Zawadzki, I.: Statistical properties of precipitation patterns, J. Appl. Meteorol., 12, 459–472, doi:10.1175/1520-0450(1973)012<0459:SPOPP>2.0.CO;2, 1973.

4. Article 3: Statistical precipitation bias correction of gridded model data using point measurements

Statistical precipitation bias correction of gridded model data using point measurements

Jan O. Haerter[1], Bastian Eggert[2], Christopher Moseley[3], Claudio Piani[4], and Peter Berg[5]

[1]Niels Bohr Institute, Copenhagen, Denmark, [2]Helmholtz Zentrum Geesthacht, Climate Service Center 2.0, Hamburg, Germany, [3]Max Planck Institute for Meteorology, Hamburg, Germany, [4]Computer Science, Mathematics and Environmental Sciences, American University of Paris, Paris, France, [5]Hydrology Research Unit, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

Abstract It is well known that climate model output data cannot be used directly as input to impact models, e.g., hydrology models, due to climate model errors. Recently, it has become customary to apply statistical bias correction to achieve better statistical correspondence to observational data. As climate model output should be interpreted as the space-time average over a given model grid box and output time step, the status quo in bias correction is to employ matching gridded observational data to yield optimal results. Here we show that when gridded observational data are not available, statistical bias correction can be carried out using point measurements, e.g., rain gauges. Our nonparametric method, which we call scale-adapted statistical bias correction (SABC), is achieved by data aggregation of either the available modeled or gauge data. SABC is a straightforward application of the well-known Taylor hypothesis of frozen turbulence. Using climate model and rain gauge data, we show that SABC performs significantly better than equal-time period statistical bias correction.

1. Introduction

Statistical bias correction, as a field, has received substantial attention in recent years as it is a simple tool that makes impact studies possible in situations where climate model data are available but are subject to inherent biases. Since the early approaches [*Wood et al.*, 2004], statistical bias correction techniques have now diversified considerably [*Maraun et al.*, 2010; *Piani et al.*, 2010a, 2010b; *Haerter et al.*, 2011; *Teutschbein and Seibert*, 2012; *Piani and Haerter*, 2012] and have been widely applied to a range of global [*Piani et al.*, 2010b; *Li et al.*, 2010] and regional climate model (RCM) data sets [*Berg et al.*, 2012a; *Gudmundsson et al.*, 2012; *Teutschbein and Seibert*, 2012].

Most current statistical bias correction techniques have in common that some form of quantile mapping is applied to match the probability distribution function of climate model output to that of observed climate data. It is generally necessary that the spatial and temporal resolution of modeled and observed data match as closely as possible—e.g., in order to avoid the so-called inflation or deflation issue [*von Storch*, 1999; *Maraun*, 2013]. However, in many practical situations, observed spatial data may not be available. Some type of derived data set, such as reanalysis data [*Dee et al.*, 2011] or data products combining data from multiple sources [*Weedon et al.*, 2011; *Berg et al.*, 2015], must be used. The former, as a model interpolation technique, still suffers from error and bias, while in the latter, due to the blend of data sources, little is sometimes known on the exact details of the underlying data resolution in specific regions or time periods. Further, both model and observations must typically be regridded to obtain matching resolutions. For station data (point measurements), regridding to a common grid, i.e., finding an area representation of observations, is impossible from the outset—especially when few stations are available in the region of interest. One possible option are then stochastic methods to account for variability at small scales [*Eden et al.*, 2014; *Wong et al.*, 2014], which, however, entail assumptions on the distribution functions.

In this study we take a different approach that may be suitable when gridded observations of dynamical variables are not available, but reliable station measurements are. The approach is a simple application of the Taylor hypothesis of frozen turbulence. The original Taylor hypothesis states that as the mean atmospheric flow advects eddies past a station, the properties of the eddies remain unaltered [*Taylor*, 1938]. The hypothesis has previously been used for precipitation disaggregation [*Deidda*, 2000]. Recognizing this, mapping to a common grid is no longer necessary as lacking spatial information can be compensated by increased temporal information.
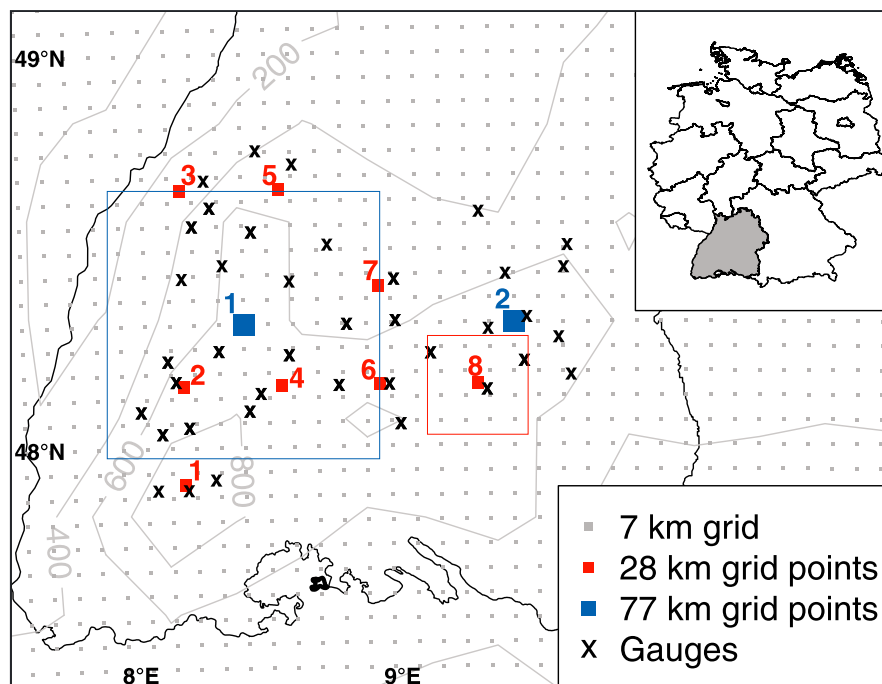
Figure 1. Map of the data used. Figure shows available rain gauge stations, the model grid at 7 km, and grid points (red and blue numbers) used for the 28 km and 77 km correction. Red and blue boxes exemplify areas of respective grid boxes. Also shown: topography in meters, coordinates, and state boundary (black) of Baden-Württemberg, Germany.

Here we will consider simulated precipitation versus rain gauge station observations. Each value of simulated precipitation represents a space and time average determined by the grid size $\delta_{x,\text{mod}}$ and time step $\delta_{t,\text{mod}}$ of the output. By contrast, station precipitation data represent a spatial point measurement ($\delta_{x,\text{obs}} = 0$) and a time average determined by the measurement time step ($\delta_{t,\text{obs}}$). In general, we will have $\delta_{t,\text{mod}} \neq \delta_{t,\text{obs}}$. In the case of high-frequency observations, one might be tempted to simply aggregate station data to a coarser time interval so that $\delta_{t,\text{mod}} = \delta_{t,\text{obs}}$ and use the resulting data set to perform a statistical bias correction of the simulated precipitation. For example, if correcting simulated daily values of precipitation with hourly station data, one might simply derive daily station data and proceed as usual. We will show that further aggregation onto longer time intervals of the observed station data may lead to better quality bias correction, while also avoiding the inflation issue [*Maraun*, 2013]. We also show the inverse, when model data have comparably high resolution. In that case, moderate coarse graining of the model data improves the statistical bias correction.

2. Data

We use a fine-resolution (1 km and 5 min) composite of radar images from the RADOLAN-RY product of the German Weather Service (DWD) for Southern Germany, which was aggregated for varying resolutions [*Eggert et al.*, 2015]. Rainfall rates were calculated from radar echoes with the *Z-R* relationship [*Steiner et al.*, 2004] and are available for the 2 years 2007–2008. Additionally, a set of 1 h resolution rain gauge data from Baden-Württemberg in southwestern Germany was used for the bias correction experiments. The stations constitute a relatively dense network covering both the Black Forest mountain range and the topographically less variable Rhine Valley (Figure 1). The station network is most dense in the period 1997–2004, which is used here. Stations with more than 10% missing data were discarded from the analysis, and for the remaining stations NaNs were set to zero precipitation to simplify. This has little impact on the results presented here.

The model data are taken from a simulation with the COSMO-CLM (COnsortium for Small scale Modelling-CLimate Mode) [*Doms and Schättler*, 2002]. The current simulation uses ERA-Interim reanalysis [*Dee et al.*, 2011] as driving data in a double nesting setup with a second 7 km nest domain covering all

of Germany, with the lateral boundaries well outside Germany's borders. The output time step is 1 h for precipitation. (Simulation details: *Berg et al.* [2012b].) For the current investigations, the model results are applied at the original 7 km resolution, and also as representatives for coarser model resolution by remapping to 28 km and 77 km. As shown in *Berg et al.* [2012b] and *Fosser et al.* [2014], the model has a strong bias in mean precipitation amounts but performs well regarding the intensity distribution for both daily and hourly timescales. Equally long periods of 8 years are used for the model and station data.

3. Methodology

The Taylor hypothesis [*Taylor*, 1938] states that observations at a given spatial and temporal resolution should be similar to those at lower spatial but higher temporal resolution. Our method is hence based on the idea that lacking spatial resolution can be compensated by increased temporal resolution. Point measurements, i.e., data at high spatial resolution, should therefore be compared to spatially averaged data at correspondingly higher temporal resolution.

3.1. Frozen Turbulence and Comparison of PDFs

The probability density function (PDF) of precipitation intensity I depends on the resolution in space and time, expressed by the size of the grid box or time step (δ_x or δ_t), of the precipitation data. Our methodology is based on the assumption that PDFs of precipitation intensity obtained with different spatial and temporal data resolutions (δ_x, δ_t) can have similar, albeit not identical, features. For example, should the Taylor hypothesis hold perfectly, then PDF($\delta_x, 0, I$) = PDF($0, \delta_t = \delta_x/v, I$), where v is an advection speed.

Using observed or modeled precipitation data (section 2), we produce PDFs of precipitation intensity at varying spatial and temporal resolutions. We then compare the PDFs at different resolutions (δ_x, δ_t): The *PDF agreement*

$$S\left(\delta_x, \delta_t; \delta_x', \delta_t'\right) \equiv 1 - \int_{I_0}^{\infty} dI \, \left| w(\delta_x, \delta_t, I) - w\left(\delta_x', \delta_t', I\right) \right| \tag{1}$$

measures the similarity of the PDFs corresponding to different resolutions [compare *Perkins et al.*, 2007; *Eggert et al.*, 2015]. In equation (1), $w(\delta_x, \delta_t, I)$ is some function of the probability density function of precipitation intensity and I_0 defines a possible low-intensity cutoff. Note that for most regions of the globe, zero intensity constitutes the bulk of the probability weight in the PDF, i.e., dry periods outnumber the wet. Intensity weighting gives more emphasis to nonzero intensities. For this study, we therefore choose $w(\delta_x, \delta_t, I)$ as the intensity-weighted PDF of precipitation, i.e., $w(\delta_x, \delta_t, I) \equiv I \cdot \text{PDF}(\delta_x, \delta_t, I)/ \int_0^{\infty} dI' \, I' \cdot \text{PDF}(\delta_x, \delta_t, I')$ and $I_0 = 0$. Other reasonable choices are the Kolmogoroff-Smirnov statistics or the use of the bare PDF with a nonzero intensity cutoff $I_0 > 0$. We found all to give similar results, albeit with varying degree of noise (not shown).

3.1.1. Observed Data

For observed high-resolution data (section 2) and for a given reference resolution (e.g., $\delta_{x0} = 25$ km, $\delta_{t0} = 5$ min, shown as a large white circle in Figures 2a and 2b), we now determine all equivalent resolutions. This can be done in two equivalent ways leading to similar, albeit not identical, results:

1. Consider Figure 2a, where we directly compare PDFs of all resolutions with that of the reference resolution. The PDF agreement is therefore maximal, i.e., unity, not only for $(\delta_x, \delta_t) = (\delta_{x0}, \delta_{t0})$ but also for other pairs of resolutions the figure shows high-agreement values along a "ridge" with increasing space and decreasing time resolution. To the extent that the ridge top can be approximated by a line, this indicates that PDFs for resolutions $\left(\delta_x, \delta_t^* - r\delta_x\right)$, with δ_t^* the ridge intersect with the vertical axis, are all very similar. Defining, analogously, δ_x^* as the ridge intersect with the horizontal axis and $r^* \equiv \delta_x^*/\delta_t^*$, $v^* = 1/r^*$ has units of a velocity and gives the effective speed of advection [*Taylor*, 1938]. We note that as r^* is the slope of the ridge, it can also be estimated when any two points (δ_t, δ_x) and $\left(\delta_t', \delta_x'\right)$ on this line are known, i.e., $r^* \approx \left(\delta_t' - \delta_t\right) / \left(\delta_x' - \delta_x\right)$. This will then also deliver δ_t^*, e.g.,

$$\delta_t^* = \delta_t + r^*\delta_x \left(= \delta_t' + r^*\delta_x'\right) \ . \tag{2}$$

We will refer to the line defined by r^* and δ_t^* as the relevant ridge. In short, scale-adapted statistical bias correction (SABC) demands finding this relevant ridge. This ridge will then be used to identify the proper scale adaptation.
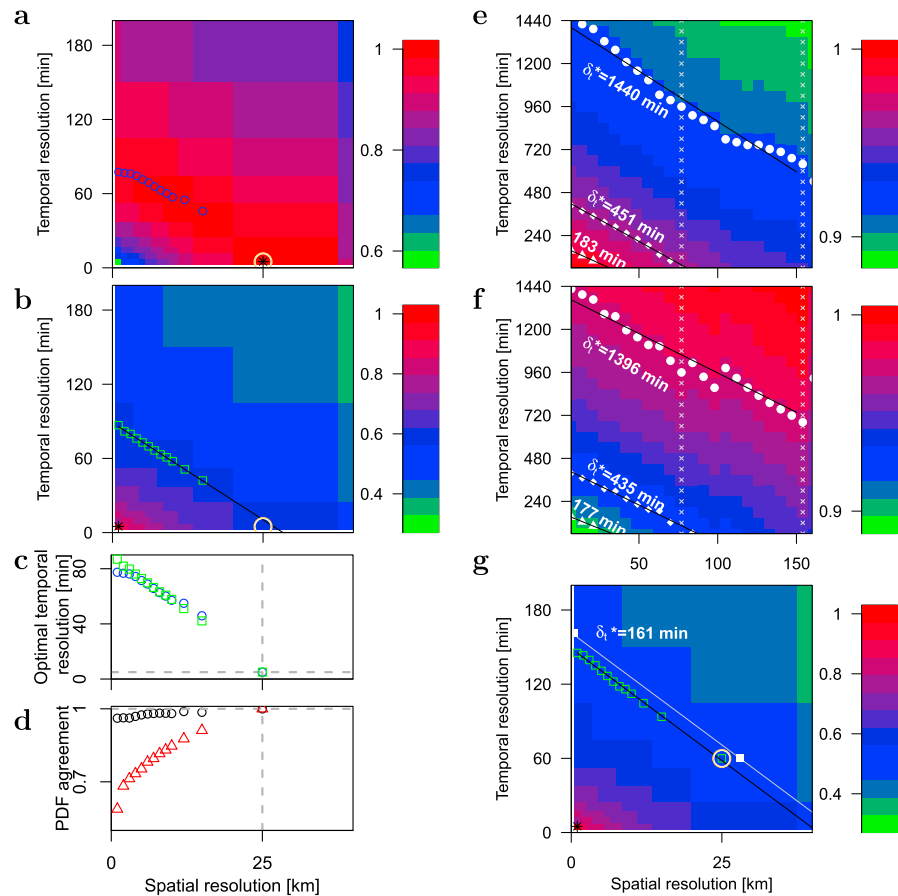
Figure 2. Mapping of distribution functions. (a) PDF agreement (equation (1)) of radar precipitation PDFs of different spatial and temporal resolution to the origin radar precipitation PDF with 25 km and 5 min resolution (shown as black star symbol). Resolutions equivalent to the reference resolution (white circle) shown as blue circles. (b) Similar to Figure 2a but using contours of similar PDF agreement for extrapolation. The origin distribution is now chosen at 1 km and 5 min (star symbol). Optimal resolutions shown as green squares. (c) Comparison of equivalent temporal resolutions from Figures 2a and 2b as function of spatial resolution. Dashed lines indicate reference resolution. (d) PDF agreement as function of spatial resolution: black circles indicate maximal PDF agreement derived in Figure 2a; red triangles denote PDF agreement when keeping temporal resolution fixed to 5 min. (e) PDF agreement with an origin resolution chosen at (7 km; 60 min) for model data. White triangles/diamonds/circles are computed ridges for reference resolutions (28 km; 60 min) and (77 km; 60 min), as well as $\delta_t^* = 1440$ min, respectively. (f) Similar to Figure 2c but for an origin resolution chosen at (154 km; 1440 min). Note the similar results for the ridges in Figures 2c and 2d. (g) Comparison to radar data. Color bars denote respective PDF agreement.

2. Another approach is to use a contour line of PDF agreement, not the maximum, to obtain the optimal resolutions. To obtain contours of PDF agreement, we specify a different origin resolution, that is, a resolution that serves as a comparison. In (1), this origin had been set to the same value as the reference resolution, but we now show that it can also be moved to another point in the plane, i.e., away from the reference resolution. Figure 2b exemplifies this for the case where the origin is set to (1 km; 5 min). Using now the PDF agreement corresponding to the reference resolution (i.e., 25 km and 5 min), i.e., comparing PDF(1 km, 5 min) to PDF(25 km, 5 min), we find the line of similar resolutions, this time by comparing with the respective PDF agreement (compare Figure 2b, green line).

Not surprisingly, the contour line (symbols in Figure 2c) differs very little from the ridge found in method (1). This is clear when noting that all PDFs along the ridge are very similar. Hence, also the agreement of any of them with an origin resolution away from the ridge should yield comparable values. Methods (1) and (2) are two equivalent ways to obtain the value of δ_t^*, i.e., the desired temporal resolution of the rain gauge. In practice (section 3.3) we use method (2).

We can compare the PDF agreement $S\left(0, \delta_t^*, \delta_{x0}; \delta_{t0}\right)$ of PDF $\left(0, \delta_t^*\right)$, obtained with method (1) with the standard approach of simply taking the time resolutions to match, i.e., maintaining a temporal gauge resolution of 5 min (Figure 2d). The values obtained for the ridge top are reasonably good (black circles), and much better than when using unchanged temporal resolution (red triangles). From the example we see that when the gauge data are first aggregated to approximately 80 min, it constitutes a much better representation of the distribution function of spatial data, i.e., PDF(0, 80 min) ≈ PDF(25 km, 5 min), while PDF(0, 5 min) ≠ PDF(25 km, 5 min).

3.1.2. Climate Model Data

In Figures 2e and 2f we repeat the analysis of PDF agreement for high-resolution climate model data (7 km spatially and 60 min temporally). Setting the origin resolution first to the highest model resolution (lower left corner in Figure 2e), we compute the contour lines for several choices of reference resolution (solid white symbols). These choices are made to reflect the cases of (i) typical high-resolution RCM output resolution (28 km, 60 min), (ii) high-resolution global climate model output (77 km, 60 min), and (iii) daily gauge temporal resolution (1440 min). Fitting linear functions, we extrapolate the corresponding values of δ_t^* (compare equation (2)). The fits show the following: For (i), the equivalent gauge resolution would be approximately 3 h. For (ii), gauge resolution should be chosen at 7.5 h. For (iii), daily gauge resolution requires model resolution given by any of the symbols (white circles) in Figure 2e. One possible choice is (150 km, ≈ 12 h). That is, if a model of 150 km spatial resolution was available, optimal temporal resolution would be approximately 12 hourly.

We repeat the analysis for another origin resolution (154 km spatially, and 1440 min, temporally, upper right corner in Figure 2f). Producing again the contour lines and corresponding fits, we obtain very similar values of δ_t^*, confirming that the choice of origin resolution has little impact on the resulting contour lines.

To compare these results with the observational data, we also extrapolate the contour line corresponding to one of the reference resolutions (28 km, 60 min) in the radar data (Figure 2g). In the observational data, the closest corresponding spatial resolution is 25 km. We use this to obtain the slope r. Together with the reference resolution (28 km, 60 min) we yield the extrapolated value of $\delta_t^* \approx 161$ min. This value lies somewhat lower than those from the model data (there, $\delta_t^* \approx 180$ *min*). However, given the shortcomings of both observational and modeled data and limitations of available resolutions, the agreement of the results is remarkably good.

The previous analysis shows that the patterns obtained with observational data can approximately be reproduced using model data. However, typically, the model output resolution will lie at lower resolutions than the one used for Figures 2e and 2f. Consider, e.g., the resolution (77 km; 1 h); available spatial and temporal model resolutions are hence 77, 154, 231, … , km, respectively 60, 120, 180, … , min (shown as white crosses in Figures 2e and 2f). Desired contour lines must hence be determined using exclusively these discrete combinations of resolutions. We will show how to obtain these in the following.

3.2. Bias Correction Methodology

We distinguish the two cases resulting from comparably low model resolution (model limited) and comparably low gauge resolution (gauge limited).

3.2.1. Model-Limited Correction

Consider a climate model with a given output resolution (blue cross, Figure 3a). We assume that observations are only available through a rain gauge (red cross, Figure 3a), with a relatively high gauge resolution. The goal is now to estimate the contour line that is defined by the model output resolution, i.e., the blue cross symbol must lie on the contour line. The offset δ_t^* for this line will define the required coarsening, i.e., scale adaptation, of the gauge data. Once the scale adaptation has been produced, standard statistical bias correction [e.g., *Piani et al.*, 2010b] can be performed.

3.2.2. Gauge-Limited Correction

In many practical situations, it is also possible that rain gauge resolution is poor compared to the model resolution (Figure 3b, *gauge-limited* case). This may especially be the case in very data sparse regions, such as areas of the globe with little infrastructure (deserts and glaciated regions), but even in developed areas, complete spatial coverage by subdaily precipitation gauges is by no means standard and daily temporal resolution is usually the best available data. Under such circumstances, scale adaptation can only be performed by aggregating the *model* data. It must thereby be assured that the available gauge resolution lies on top of the contour line. Available model resolutions are shown in Figure 3b (small cross symbols). It is
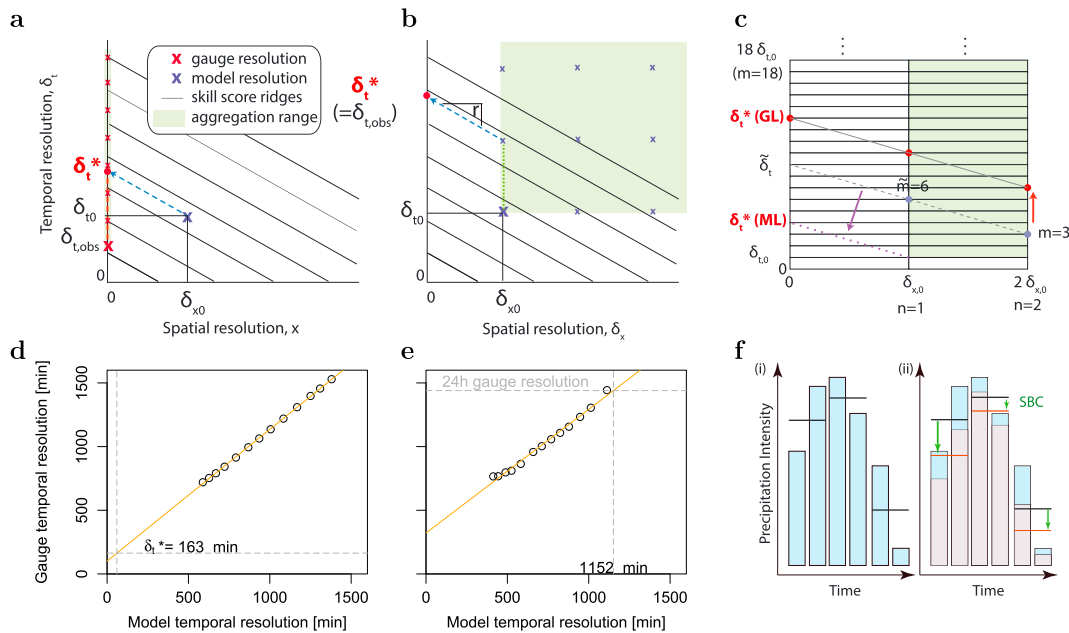
Figure 3. Scale-adapted statistical bias correction. (a) Model-limited correction. Two-dimensional schematic of spatial and temporal averaging intervals and PDF agreement ridges (black lines). Red and blue cross: symbols indicate available gauge and model resolutions, respectively. (b) Gauge-limited correction. Similar to Figure 3a but also including the range of resolutions where aggregation could be performed (green shades). Dotted green line indicates a possible aggregation choice. (c) Schematic showing the construction of equivalent resolutions. Corners of green-shaded region show available aggregated model resolutions. Dashed gray line indicates choice of $(n=2, m=3)$ and $(n=1, \tilde{m}=6)$, yielding $r=3\delta_{t0}/\delta_{x0}$ and $\tilde{\delta}_t = \tilde{m}\delta_{t0} + r\delta_{x0} = 9\delta_{t0}$. Solid gray line shows equivalent resolutions yielding $\tilde{\delta}_t = \delta_t^*$ for the gauge-limited case (GL). Dotted purple line shows equivalent resolutions yielding $\tilde{\delta}_t = \delta_t^*$ for a model-limited case (ML); this line must be extrapolated (purple arrow indicates extrapolation). (d) Model-limited, extrapolation of $\tilde{t}$ as $\delta_x \to 0$ for an assumed model of resolution (28 km; 1 h). (e) Gauge-limited case, interpolation of $\tilde{\delta}_t$ as $\delta_x \to 0$ for an assumed model of resolution (77 km; 1 h) and gauge of 24 h resolution. (f) Schematic of data aggregation and disaggregation: (i) aggregation of pairs of two subsequent precipitation measurements (blue bars) to respective averages (black lines). (ii) Statistical bias correction (SBC) of aggregated data (black lines converted to red lines) and subsequent disaggregation of corrected data (gray bars).

now possible that multiple model resolutions can in principle be used for scale-adapted correction, i.e., any that are compatible with a contour line through $(0, \delta_t^*)$. Note that once the correction has been performed, the model data can be disaggregated to its original resolution. For each aggregated data point, we simply separate the data by applying the appropriate correction to each of its individual contributions (Figure 3f).

3.3. Practical Implementation

In practice, the resolution of the model output defines the reference resolution $(\delta_{x0}, \delta_{t0})$ and we set the origin resolution equal to this. Available model resolutions are all combinations of integer multiples $(n\,\delta_{x0}; m\,\delta_{t0})$ of the reference resolution, with n and m integers (Figure 3c). Implementation of scale-adapted bias correction follows four steps:

1. Set δ_t to δ_{t0}, i.e., $m=1$. Observe all PDFs for several accessible $n\,\delta_{x0}$, e.g., δ_{x0} and $2\delta_{x0}$.
2. Fix a specific PDF in (1) by choosing a multiplication $\tilde{n}$, e.g., choose the resolution $(2\delta_{x0}, \delta_{t0})$. Now obtain the value $\tilde{m}$ where $\tilde{m} \cdot \delta_{t0}$ yields maximum PDF agreement between PDF$(\tilde{n}\delta_{x0}, \delta_{t0})$ and PDF$(\delta_{x0}, \tilde{m}\delta_{t0})$, i.e., both will lie on the same contour line. The pairs $(\tilde{n}\delta_{x0}, \delta_{t0})$ and $(\delta_{x0}, \tilde{m}\delta_{t0})$ define lines which can be used to extrapolate to $\delta_x = 0$, yielding a value $\tilde{\delta}_t$ (compare equation (2) and Figure 3c).
3. Repeat steps (1) and (2) starting with $m>1$ in (1).
4. Extrapolate or interpolate the resolution for required δ_t^*, i.e., achieving $\tilde{\delta}_t \approx \delta_t^*$.

Note: For better results, in step (2) a fit with respect to PDF agreement should be used to obtain $\tilde{m}$, thereby allowing noninteger $\tilde{m}$ and yielding more exact estimation of $\tilde{\delta}_t$.

Once δ_t^* has been obtained, in the ML case the gauge data will be aggregated to the resolution δ_t^*. In the GL case, the model data will be aggregated to a resolution closest to the line defined by the contour through δ_t^*. Once the data have been aggregated, standard statistical bias correction will be performed.
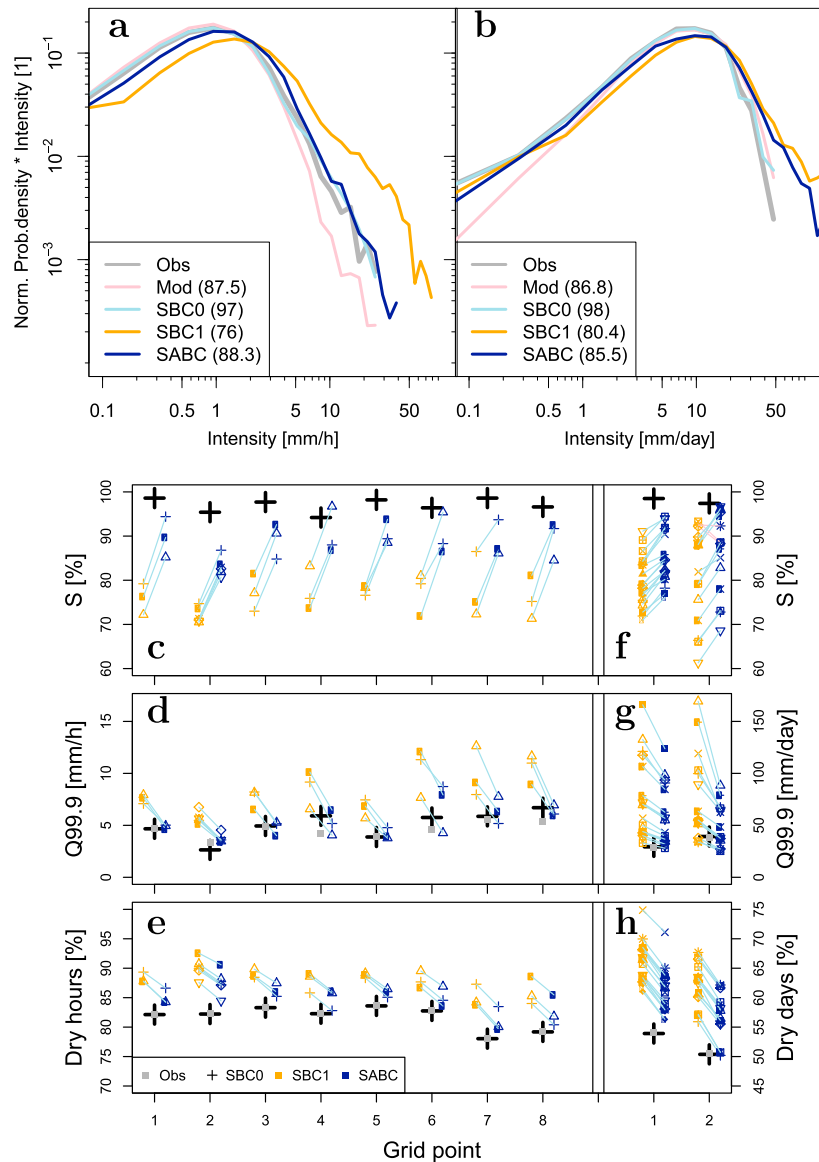
Figure 4. Bias correction results. (a) Model-limited bias correction; intensity-weighted PDFs for the estimated observations (Obs, gray), i.e., the average of all stations in a grid box, 28 km model data (Mod, pink), SBC1 using a single gauge at a 1 h resolution (orange), and SABC using a single gauge at 3 h temporal data aggregation (dark blue). PDFs shown are aggregate distributions for all stations and grid boxes. Values marked in parentheses in legend are PDF agreement values with ρ_{obs} in percent. (b) Gauge-limited bias correction for 77 km model; curves and colors analogous to Figure 4a. SBC1 is now for a single station and model data aggregated to 24 h; SABC uses 19 h temporal aggregation of model data. Note the different horizontal axes and units in Figures 4a and 4b. Results for individual grid boxes and stations for the ML case shown in Figure 4a: (c) PDF agreement , (d) 99.9th intensity percentile, and (e) dry period fraction. Colors as in Figures 4a and 4b. Symbols for each grid box are the different available stations (Figure 1), thin blue (pink) lines are guides to the eye, linking same stations for increased (decreased) performance. (f–h) Similar to Figures 4c–4e but corresponding to the GL case shown in Figure 4b.

4. Results

Does scale-adapted bias correction yield measurable improvement in practice? We carry out both types of corrections (Figures 3d, 3e, and 4) and compare each to the correction where temporal resolutions are simply matched for model and observations. Specifically, the tests are as follows.

(ML) Model-limited correction. We employ the same model used in Figures 2e and 2f but start from a minimal resolution of $\delta_{x0} = 28$ km spatially and 60 min temporally. This is to mimic output resolution typical of current RCMs. We assume station data to be available at 1 h resolution.

(GL) Gauge-limited correction. Again, using the same model and hourly temporal resolution, we now use coarse spatial resolution (77 km), mimicking state-of-the-art global climate model output. The station resolution is assumed to be daily.

Following the steps in section 3.3, we now obtain an estimate for the corresponding contour line (white triangles and circles in Figure 2f for the ML and GL cases) without the knowledge of the data corresponding to spatial scales below δ_{x0}. Specifically, we compute various curves (such as those shown in Figure 3c) and for each determine the corresponding $\tilde{\delta}_t$. This procedure yields a set of corresponding temporal resolutions ($\tilde{m}\delta_{t0}, \tilde{\delta}_t$), which is plotted in Figures 3d and 3e. For the ML case, these points allow us to extrapolate the δ_t^* corresponding to model temporal resolution δ_{t0}, yielding $\delta_t^* \approx 163$ min. The best equivalent gauge aggregation (integer multiple of 1 h) is hence 3 h. We proceed analogously for the GL case, where we determine the value of $\tilde{m}$ that corresponds to $\delta_t^* = 1440$ min, yielding $\tilde{m} \approx 19$, hence a model aggregation to 19 h (compare Figure 2e, where the value is ≈ 18 h).

To obtain a proxy for spatially averaged observations, we first group rain gauge stations into 28 km (respectively 77 km) grid boxes and simply average their precipitation intensities. We discard grid boxes with fewer than three stations. For the ML case, we maintain a temporal resolution of 1 h, and for the GL case, we aggregate temporally to 24 h. The resulting spatially averaged signals will be less variable than that of any station by itself, but, due to the finite number of stations per grid box, somewhat more variable than the actual spatial average (i.e., that corresponding to an infinite number of stations per grid box). For each grid box, the averaged station data serve as the "ground truth" for our procedure (we call its probability density function ρ_{obs} in the following).

We now use any single station within the grid box as constituting the only available station for a given practical situation. Our hypothesis is that SABC yields better agreement with ρ_{obs} than the standard correction with matching time resolutions (SBC1). To exemplify the bias correction procedure, we use simple empirical quantile mapping by estimating a regularly spaced quantile distribution, following *Gudmundsson et al.* [2012]. Our approach is, however, generic and should also apply for more sophisticated bias correction techniques [e.g., *Piani et al.*, 2010b; *Mehrotra and Sharma*, 2012; *Rocheta et al.*, 2014]. As a benchmark, we first perform a correction with ρ_{obs} (SBC0). This represents the *ideal* correction obtainable when sufficient data are at hand. As expected, the resulting corrected model PDF is very close to ρ_{obs}, as seen in Figures 4a and 4b.

For the ML case, for each available grid box and for each possible choice of associated station, we now repeat the procedure using the hourly data—assuming that only this *single* station is available for the bias correction. For each combination of a grid box and station, this yields a histogram for the corrected data. SBC1 significantly shifts the original distributions to more extreme intensities (Figure 4a); affecting the overall PDF agreement and high percentiles. Repeating for SABC (gauge data coarsened to 3 h), markedly closer agreement with ρ_{obs} is reached. Besides the overall skill, also extremes and dry period fraction are consistently improved for each single station used as reference (Figures 4c–4e). Also, the average intensities are corrected well to that of the relevant reference station, which trivially produces a spread of results around the average of ρ_{obs} with only small differences between SBC0 and SABC.

An analogous comparison is performed for the GL case; now SBC1 employs temporal coarsening of model data to match the assumed daily resolution of the station, leading to generally too heavy intensities compared to ρ_{obs} (Figure 4b). Again, SABC overall yields substantial improvement for PDF, extremes and dry period fraction. Disaggregation back to hourly data (Figure 3f) retains the increased skill seen in the daily statistics but does not, e.g., affect other bias such as in the diurnal cycle.

We end with a comment on GL corrections: When multiple options are available for coarsening, one relevant criterion for optimal choice might be loss of data samples in the statistics. Reduction of sample size is greatest for intermediate options of spatial coarsening, but low for both the original spatial resolution and very high values of the spatial coarsening. We consider the latter only a theoretical option, as it may require substantial domain sizes and removes all spatial information, e.g., orographic effects (details on sample reduction: Appendix A).

5. Discussion and Conclusion

Statistical bias correction of precipitation has emerged as an indispensible tool at the intersection between climate and impact modeling—but hinges on the availability of adequate observational data. Especially in data sparse regions of the globe, successful bias correction is hampered by the lack of agreement of model precipitation output resolution and the resolution observed. We have shown that even with observational data from a single rain gauge station, considerable improvement can be achieved when aggregating the available data to yield maximal PDF agreement between model and observations. We suggest the use of this method, which we have called scale-adapted statistical bias correction (SABC), in situations where bias correction is required but only limited point measurements are available.

Crucially, SABC capitalizes on the ability of the model to capture the space-time dynamics of precipitation. The model is hence employed to associate adequate scales. In SABC, the variability of precipitation intensity—measured at a station—is compared to model data measured over a spatial domain but accordingly chosen, shorter time intervals. Our results point to the utility of outputting temporally highly resolved model data, even (or especially *if*) spatial resolution is low—the information on fluctuations, encoded in the higher temporal output rate, will then be preserved and can be used for SABC. In this way, the method can help remedy the inflation issue [*von Storch*, 1999; *Maraun*, 2013]. SABC can naturally not impact on inherent model shortcomings in simulating temporal variability [*Maurer and Pierce*, 2014]. Also, when atmospheric advection is not described adequately, the association of scales needed for SABC may itself be biased.

The methodology is straightforward and can be applied for any type of climate model data. It requires only to compute probability density functions for several choices of coarsened resolutions. This is simply done by data aggregation, e.g., by doubling or tripling of spatial and temporal scales. Pairs of matching resolutions then define appropriate gauge resolutions. Our results are encouraging in that quantitative improvement of intensity distributions is reached for the overall histogram, extreme precipitation as well as dry periods.

SABC should not be limited to precipitation correction but may also be relevant for other meteorological variables where the hypothesis of frozen turbulence applies, i.e., where advection of the quantity is the dominant cause of local fluctuations. While our method works well even in regions of moderate topographic variation (Southern Germany), stronger variation, e.g., mountainous regions, may introduce features that naturally require fine-scale knowledge of local climate, not captured by single gauges. Especially when gauge density is low, as is the case in vast parts of the globe, our method may allow for substantially improved bias correction at essentially no cost in terms of model output, data storage, or mathematical complexity.

Studying precipitation intensity at resolutions finer than that of convective systems has recently become of widespread interest [e.g., *Lenderink and Van Meijgaard*, 2008; *Berg et al.*, 2013; *Eggert et al.*, 2015]. Observationally, data describing such extremes is often limited to individual gauges: Measurements from gauges are usually considered more reliable than other sources of data, i.e., constitute the preferred source of information. Yet modeling of station scale characteristics will—for the time being—not be feasible. Our study speaks to an alternative solution, where low spatial resolution could again be compensated by high temporal resolution. A station with hourly temporal resolution could then be modeled by a regional climate model with approximately 12 km spatial and 5 min temporal output. Again, simply using the same temporal output rate as available from the station would lead to sizeable error in the comparison (compare Figures 2a and 2b).

As an alternative to recent stochastic methods [*Eden et al.*, 2014; *Wong et al.*, 2014], SABC could also be used for simple, direct, downscaling of model data to the point scale—possibly circumventing the need for a statistical model. In the example of the previous paragraph, each value of (12 km, 5 min) model output would then be a proxy for an hourly average for a point measurement. An analysis of such downscaling is, however, left to a future study.

Finally, it may be promising to combine SABC with two-dimensional bias correction [*Piani and Haerter*, 2012; *Mehrotra and Sharma*, 2015]. Given the large amount of data necessary to populate 2-D histograms of dynamical variables, for example, temperature and precipitation, gridded data sets with sufficiently long time series are hard to come by. Using SABC, climate impact modelers can access the information in station

data directly without prior gridding. This has the potential to make 2-D bias correction a standard procedure and will be the focus of future work.

Appendix A: Sample Size Reduction

We comment on the possible choices for gauge-limited corrections: When multiple choices are available for possible coarsening, one relevant criterion for optimal choice might be loss of data samples in the statistics. When coarsening, using equivalent resolutions along the relevant ridge, the reduction of sample size is $R(\tilde{n}) \equiv \tilde{n}^2 \cdot \tilde{m}(\tilde{n})$. Dropping the tilde for simplicity, and noting that $m = \delta_t/\delta_{t0}$, $\delta_t = \delta_t^* - r\delta_x$, and $n = \delta_x/\delta_{x0}$, we have

$$R(n) = \frac{\delta_t^*}{\delta_{t0}} n^2 - \frac{r}{r_0} n^3, \qquad (A1)$$

where we have defined $r_0 \equiv \delta_{t0}/\delta_{x0}$. Equation (A1) states that R initially increases as a function of n but decays for large n. Noting that $n, m \geq 1$, we have the bounds

$$n_{\min} \equiv 1 \leq n \leq \frac{\delta_t^* - \delta_{t0}}{r\delta_{x0}} \equiv n_{\max}. \qquad (A2)$$

Intermediate choices of n yield extreme reductions of data, for the case studied in Figure 4b, $R_{\mathrm{extr}} > 100$ is possible. Generally, $n = 1$ may be a reasonable choice (in our example $R(1) = 19$) but in some cases also $n = n_{\max}$ should be considered, especially when high temporal resolution, i.e., small δ_{t0}, is available. In those cases, $R(n_{\max}) < R(1)$ is possible, allowing a larger sample size to be preserved. In the plot, n ranges from unity to $n_{\max}$. Extremal R_{extr} occurs at intermediate values of n and can substantially exceed $R(1)$ and $R(n_{\max})$. Dependencies on system parameters: $R(1) = \delta_t^*/\delta_{t0} - r/r_0$. The value of n where extremal R is reached: $n_{\mathrm{extr}} = 2r_0\delta_t^*/3r\delta_{t0}$; extremal R: $R_{\mathrm{extr}} = 4r_0^2\delta_t^{*3}/27r^2\delta_{t0}^3$. In practice, however, possible degradation of PDF agreement, even along the ridge, should be evaluated when using large n. Assessment of this question should be left for future work.

References

Berg, P., H. Feldmann, and H.-J. Panitz (2012a), Bias correction of high resolution regional climate model data, *J. Hydrol.*, 448, 80–92.

Berg, P., S. Wagner, H. Kunstmann, and G. Schädler (2012b), High resolution regional climate model simulations for Germany: Part I. Validation, *Clim. Dyn.*, 40, 401–414, doi:10.1007/s00382-012-1508-8.

Berg, P., C. Moseley, and J. O. Haerter (2013), Strong increase in convective precipitation in response to higher temperatures, *Nat. Geosci.*, 6, 181–185.

Berg, P., T. Bosshard, and W. Yang (2015), Model consistent pseudo-observations of precipitation and their use for bias correcting regional climate models, 118–132, 3, doi:10.3390/cli3010118.

Dee, D. P., et al. (2011), The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, 137, 553–597, doi:10.1002/qj.828.

Deidda, R. (2000), Rainfall downscaling in a space-time multifractal framework, *Water Resour. Res.*, 36(7), 1779–1794.

Doms, G., and U. Schättler (2002), A description of the nonhydrostatic regional model LM. Part I: Dynamics and numerics, *COSMO Newsletter*, 2, 225–235.

Eden, J. M., M. Widmann, D. Maraun, and M. Vrac (2014), Comparison of GCM- and RCM-simulated precipitation following stochastic postprocessing, *J. Geophys. Res. Atmos.*, 119, 11,040–11,053, doi:10.1002/2014JD021732.

Eggert, B., P. Berg, J. Haerter, D. Jacob, and C. Moseley (2015), Temporal and spatial scaling impacts on extreme precipitation, *Atmos. Chem. Phys. Discuss.*, 15(2), 2157–2196.

Fosser, G., S. Khodayar, and P. Berg (2014), Benefit of convection permitting climate model simulations in the representation of convective precipitation, *Clim. Dyn.*, 44, 45–60, doi:10.1007/s00382-014-2242-1.

Gudmundsson, L., J. Bremnes, J. Haugen, and T. Engen-Skaugen (2012), Technical note: Downscaling RCM precipitation to the station scale using statistical transformations—A comparison of methods, *Hydrol. Earth Syst. Sci.*, 16(9), 3383–3390.

Haerter, J., S. Hagemann, C. Moseley, and C. Piani (2011), Climate model bias correction and the role of timescales, *Hydrol. Earth Syst. Sci.*, 15(3), 1065–1079.

Lenderink, G., and E. Van Meijgaard (2008), Increase in hourly precipitation extremes beyond expectations from temperature changes, *Nat. Geosci.*, 1(8), 511–514.

Li, H., J. Sheffield, and E. F. Wood (2010), Bias correction of monthly precipitation and temperature fields from intergovernmental panel on climate change ar4 models using equidistant quantile matching, *J. Geophys. Res.*, 115, D10101, doi:10.1029/2009JD012882.

Maraun, D. (2013), Bias correction, quantile mapping, and downscaling: Revisiting the inflation issue, *J. Clim.*, 26(6), 2137–2143.

Maraun, D., et al. (2010), Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, *Rev. Geophys.*, 48, RG3003, doi:10.1029/2009RG000314.

Maurer, E. P., and D. W. Pierce (2014), Bias correction can modify climate model simulated precipitation changes without adverse effect on the ensemble mean, *Hydrol. Earth Syst. Sci.*, 18, 915–925.

Mehrotra, R., and A. Sharma (2012), An improved standardization procedure to remove systematic low frequency variability biases in GCM simulations, *Water Resour. Res.*, 48, W12601, doi:10.1029/2012WR012446.

Mehrotra, R., and A. Sharma (2015), Correcting for systematic biases in multiple raw GCM variables across a range of timescales, *J. Hydrol.*, *520*, 214–223.

Perkins, S. E., A. J. Pitman, N. J. Holbrook, and J. McAneney (2007), Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions, *J. Clim.*, *20*(17), 4356–4376, doi:10.1175/JCLI4253.1.

Piani, C., and J. Haerter (2012), Two dimensional bias correction of temperature and precipitation copulas in climate models, *Geophys. Res. Lett.*, *39*, L20401, doi:10.1029/2012GL053839.

Piani, C., J. Haerter, and E. Coppola (2010a), Statistical bias correction for daily precipitation in regional climate models over Europe, *Theor. Appl. Climatol.*, *99*(1–2), 187–192.

Piani, C., G. Weedon, M. Best, S. Gomes, P. Viterbo, S. Hagemann, and J. Haerter (2010b), Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models, *J. Hydrol.*, *395*(3), 199–215.

Rocheta, E., J. Evans, and A. Sharma (2014), Assessing atmospheric bias correction for dynamical consistency using potential vorticity, *Environ. Res. Lett.*, *9*(12), 124010.

Steiner, M., J. A. Smith, and R. Uijlenhoet (2004), A microphysical interpretation of radar reflectivity-rain rate relationships, *J. Atmos. Sci.*, *61*(1), 1114–1131.

Taylor, G (1938), The spectrum of turbulence, *Proc. R. Soc. A, 164*, 476–490.

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

von Storch, H (1999), On the use of "inflation" in statistical downscaling, *J. Clim.*, *12*(12), 3505–3506.

Weedon, G., S. Gomes, P. Viterbo, W. J. Shuttleworth, E. Blyth, H. Österle, J. C. Adam, N. Bellouin, O. Boucher, and M. Best (2011), Creation of the watch forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century, *J. Hydrometeorol.*, *12*(5), 823–848.

Wong, G., D. Maraun, M. Vrac, M. Widmann, J. M. Eden, and T. Kent (2014), Stochastic model output statistics for bias correcting and downscaling precipitation including extremes, *J. Clim.*, *27*, 6940–6959.

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

5. Article 4: Uncertainties in snow and precipitation projections in the northern Alps: the role of model biases

Uncertainties in snow and precipitation projections in the northern Alps: the role of model biases

Bastian Eggert[1,2], Andreas Haensler[2], Daniela Jacob[1,2]

[1]Faculty of Sustainability, Leuphana University Lueneburg

[2]Climate Service Center (GERICS), Helmholtz Zentrum Geesthacht

Abstract

Precipitation, especially when falling as snow, is a key element of the Alpine climate system. Within this region, regional climate models are known to contain common biases in the simulated surface-temperature and precipitation fields. While climate simulations tend to overestimate precipitation observations likely underestimate precipitation in particular when falling as snow. Based on EURO-CORDEX data, we analyzes how these errors impact the temperature distribution and the temperature dependent occurrence distribution of snowfall- and precipitation-frequencies. Differences are found between the shape of the observed and the simulated distributions. To analyzed the impact of these biases on the climate change signal (CCS) a probabilistic framework is developed in which the CCS is decomposed into two parts. Changes in the temperature distribution and changes in the temperature dependency of the event occurrence. Past changes are compared against observations allowing to quantify the impact of model biases on the CCS. The observations show an increase in (heavy) precipitation-frequency above the model spread at most altitudes. The main reason is a common cold bias in the surface-temperature of all ensemble members. A second reason is a bias in the temperature dependency of these events, that also impacts snowfall-frequency changes. We find, that mean snowfall-fractions at a specific temperature depend on the regional climate model and discrimination method. These differences impact the temperature dependency of snowfall-frequencies and hence influence the CCS. Using the RCP4.5 scenario, future changes are analyzed with and without bias-adjustment. Differences are found and related to biases identified in the past period.

1 Introduction

Projected future changes in precipitation and especially snowfall are expected to have a high environmental and economic impact in the Alpine region (*Beniston* [2003]; 200 [2007]). Snowfall plays a major role for the winter tourism sector as well as for hydro-power production and water management systems (*Schaefli et al.* [2007]; *Coppola et al.* [2016]). Heavy-snowfall events may lead to avalanches, cause severe damage on buildings and lead to traffic breakdowns (200 [2007]), whereas heavy-rain events can cause floods or landslides (*Ruiz-Villanueva et al.* [2012]).

Indications for a changing climate have already been observed in the last decades, which makes the region particularly interesting to validate simulated past changes against observations. In the 20th century an increase in surface-minimum-temperature of over 2 °C

has been observed (*Beniston* [2006]). Half of this increase is belived to be caused by high values in the North Atlantic circulation (NAO) index, that occurred towards the end of the 20th century (*Beniston et al.* [2010]). Especially for lower elevations, where the temperature trends were found to be stronger (*Ceppi et al.* [2010]), a decrease in the proportion of snowfall days relative to precipitation days is observed (*Serquet et al.* [2011]; *Scherrer et al.* [2004]). In contrast to changes in snow and temperature, observed long-term trends for precipitation show strong regional differences. An increase in the total precipitation amount was found north of the Alpine chain, that is mainly due to positive precipitation trends in the winter and spring season ([*Brunetti et al.*, 2006]), while a decrease was observed over the central Alps (*Brugnara et al.* [2011]), and south of the Alps (*Brunetti et al.* [2006]).

Projections using the new greenhouse gas concentration pathway RCP4.5 (*Moss et al.* [2010]), indicate a mean-surface-temperature increase of 2.5 °C over the alpine region by the end of the century, with a slightly stronger increase in winter compared to the summer season (*Smiatek et al.* [2016]; *Jacob et al.* [2013]). For RCP8.5, substantially larger changes are projected with a temperature increase of more than 4.5 °C (*Jacob et al.* [2013]). At the same time, winter precipitation is expected to increase under the RCP4.5 scenario by about 12 %, while summer precipitation is projected to decrease by 1.7% (*Smiatek et al.* [2016]). Projected changes of total annual precipitation show that the Alps are in a transition zone between a precipitation increase in the north of Europe, and a decrease in the south (*Jacob et al.* [2013]). In winter, this transition zone is expected to move further south, suggesting a slight increase of total- and heavy-precipitation events over the alpine region (*Jacob et al.* [2013]). These findings are in line with the already observed precipitation trends found by *Brunetti et al.* [2006]. *Frei et al.* [2017], analyze snowfall changes over the alpine region and find reductions in September to May snowfall from -25% in the RCP4.5 scenario, to -45%, using an ensemble of RCP8.5 projections.

In order to make informed decisions, stakeholders need to know about uncertainties that are associated with the model projections. Most studies that analyze uncertainties use large model ensembles and consider different greenhouse gas emission scenarios (e.g. *Jacob et al.* [2013]; *Gobiet et al.* [2014]). However, knowledge gaps, and insufficient temporal-spatial resolution, could lead to common errors in all models. Uncertainties related to model biases can therefore not always be captured by the ensemble spread (*Foley* [2010]; *Eggert et al.* [2015]). This could be particularly relevant for the analysis of projected changes in the next few decades (*Hawkins and Sutton* [2009]), or when changes under fixed global warm-

ing thresholds (e.g. 1.5 or 2 °C global warming) are compared. For these kind of studies, scenario uncertainties are reduced and model uncertainty as well internal variability gain of importance.

To determine if the models are able to correctly represent the relevant processes in a region, models are validated against observations. In the Alpine region, it is known that there are significant biases in the simulations of current and past climate (*Gobiet et al.* [2014]). *Smiatek et al.* [2016], analyzed an ensemble of EURO-CORDEX simulations and found a mean-surface-temperature bias in the range from -1.9 to -0.8 K and a wet precipitation bias ranging from 14.8 % in summer to 41.5 % in the winter season. They also showed that the overestimation of precipitation is found for the frequency of days with precipitation sums above 1 mm/day, as well as for the mean precipitation intensity. Similar model biases for temperature and precipitation are also found in other studies (*Kotlarski et al.* [2014]; *Frei et al.* [2017]; *Rajczak et al.* [2013]). However it has to be kept in mind that also the observations are prone to errors that influence the biases mentioned above (e.g. *Kotlarski et al.* [2014]; *Prein and Gobiet* [2016]).

In this study we analyze past and future precipitation- and snowfall-frequency changes over the northern alpine region using a 1 mm/day and a 25 mm/day threshold at different altitudes. For these variables it is well known that not just the variables show a temperature dependency (*Allen and Ingram* [2002]; *Trenberth* [2011]; *O'Gorman* [2014b]), but also the CCS of these variables is found to be dependent on the mean temperature in the control period (*Piazza et al.* [2013]; *de Vries et al.* [2014]; *O'Gorman* [2014a]). The snow-albedo feedback (SAF), has been identified as an important mechanism, that may cause nonlinear snow and temperature changes (*Kotlarski et al.* [2011]; *Giorgi et al.* [1997]; *Winter et al.* [2016]). These findings suggest that the temperature bias found for the alpine region will have an impact on the CCS.

To date, little research effort has been spend on the question of how these errors in the models and the observations affect the climate change signal (CCS). Therefore we develop a probabilistic decomposition-framework to test the sensitivity of past and projected changes to biases in the temperature distribution and in the temperature dependency of precipitation and snowfall events.

The framework is used to analyze the following research questions.

- How do the known temperature and precipitation biases modify the temperature distribution and in the temperature dependency of total-precipitation- and snowfall-frequencies at different altitudes in the northern alpine region.
- How strong can different snowfall-fraction estimates influence snowfall-frequency changes.
- How strong are the observed model biases influencing the past and projected CCS of total-precipitation- and snowfall-frequencies at different altitude.

We show, that the framework can also be used to perform a temperature dependent bias-adjustment of the analyzed parameters. After we quantify the impact of biases in the temperature distribution, and in the temperature dependent precipitation distributions on the CCS, we use the bias-adjusted data to identify the combined effect.

The paper is structured as follows: in Sect. 2 we describe the data and methods used. Sect. 3 presents the results from the snow discrimination analysis (Sect. 3.1), the validation experiment (Sect. 3.2), the future projections (Sect. 3.3) and the bias-adjustment (Sect. 3.4). We close with a summary and conclusions (Sect. 4).

2 Data and Methods

2.1 Research area and model selection
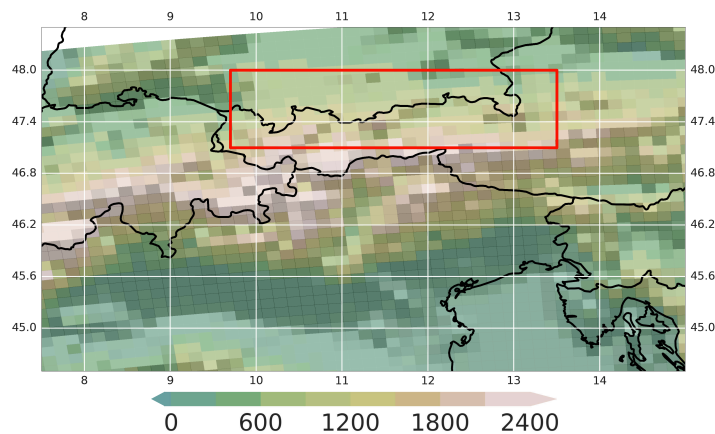
Figure 1. Orography of the alpine region at $0.11°$ resolution as used in the regional climate model REMO. The red box indicates the study area "northern Alpine region".

GCM	RCM
MPI-M-MPI-ESM-LR	MPI-CSC-REMO2009
MPI-M-MPI-ESM-LR	SMHI-RCA4
ICHEC-EC-EARTH	SMHI-RCA4
ICHEC-EC-EARTH	DMI-HIRHAM5
ICHEC-EC-EARTH	KNMI-RACMO22E
IPSL-IPSL-CM5A-MR	IPSL-INERIS-WRF331F

Table 1. Overview of global and regional climate models used in the present study

The domain of interest for this study is the northern Alpine region (9.7 - 13.5 °E, 47.1 - 48 °N) illustrated by the red box in Fig. 1. The analysis is performed on a 0.11° grid. All observation data sets were interpolated to this grid. We use simulations from the EUROpean branch of the COordinated Regional Downscaling EXperiment initiative (EURO-CORDEX) (*Jacob et al.* [2013]). As snowfall is a mandatory output variable for this study we can only use a subset of all available simulations Tab. 1. From each RCM that outputs snowfall, we selected at least one simulation. Only SMHI-RCA4 is used twice to establish a connection between the MPI-M-MPI-ESM-LR and the ICHEC-EC-EARTH driven simulations.

2.2 Observation Data

We use the Alpine precipitation gridded data sets (EURO4M-APGD) version 1.0 (*Isotta et al.* [2013]), hereafter referred to as EURO4M data. This data set covers the European Alps and is available at a horizontal resolution of 5 km. Especially over the northern part of the Alps the station density is high with more than 1 station / 100 km^2 in Germany, Austria and Switzerland (*Isotta et al.* [2013]). For precipitation observations only EURO4M data is used. The high resolution and the high station density is a key feature to capture the high spatial variability of precipitation. EURO4M was the only data available that can be expected to have at least a similar effective resolution as the EURO-CORDEX simulation for the domain.

We do not have temperature observations on the same resolution as the precipitation data, therefore we use two different temperature data sets to account for uncertainties resulting from the scale mismatch. One of these data sets is E-OBS version 14.0. E-OBS provides daily surface-air-temperature observations at a 0.25° horizontal resolution (*Haylock et al.* [2008]). The second set of temperature observations came from the HISTALP data set (*Chi-*

144 *mani et al.* [2011]). HISTALP provides monthly information at a high horizontal resolution

145 of 5min degree. Because the analysis is performed on daily data, we need to add daily devia-

146 tions to the monthly mean data. This information is taken from the E-OBS data set.

2.2.1 Creation of snowfall data sets

148 Gridded snowfall observations at a daily temporal resolution are not available for the

149 studied region. Therefore snowfall is estimated from EURO4M precipitation observations,

150 using both temperature data sets described above. These estimates affect the snowfall amounts

151 that will fall at temperatures close to 0 °C. This can have an impact on the correlation be-

152 tween snowfall and temperature.

153 To investigate this impact, we use two different snowfall discrimination methods. These

154 methods will be called SFE1 and SFE2. HISTALP also provides monthly snowfall informa-

155 tion. In the supplementary material a comparison of the yearly cycle of total precipitation

156 and snowfall between HISTALP and EURO4M is given.

157 The air-temperature thresholds (ATT scheme) SFE1 uses a simple temperature de-

158 pendent function, that was derived from snow flux observations over Sweden (*Feiccabrino*

159 *et al.* [2013]), to obtain a snow-precipitation ratio (PrsnFrac). This function is also used

160 by *O'Gorman* [2014a] who finds that the function can be used over large parts of northern

161 Europe. For surface-temperature values above 7 °C it is assumed that no snowfall will oc-

162 cur and the snowfall-fraction is set to 0. At surface-temperature values below -4 °C the snow

163 fraction is set to 1, all precipitation will be seen as snow. Between -4 to 7 °C equation 1 is

164 used to determine the snow-fraction (PrsnFrac).

$$PrsnFrac = \exp(-8.58\text{E}{-}5 * (T + 7.5)^{4.12}) \tag{1}$$

165 For the second snowfall discrimination approach SFE2, we decided to use snowfall

166 statistics from the EURO-CORDEX ensemble rather than to use a second conventional method.

167 This has the advantage that differences in snowfall between models and observations can di-

168 rectly be related to differences in precipitation and temperature, as the snowfall-fractions

169 at a specific temperature are statistically the same. Specifically we create a lookup-table

170 (LUT) from the model ensemble. This LUT provides ensemble mean snowfall-frequency-

171 and snowfall-intensity-fractions, dependent on temperature, precipitation intensity and alti-

172 tude. The same dependencies that will later be used for further analysis. To create the LUT,

173 the information is binned in the following way: 1 Kelvin bins for temperature, 100 m bins for

174 orthographic height and 1 mm/day bins for precipitation intensity. For orthographic height

175 and precipitation intensity a smoothing is applied as centered mean over neighbored bins.

176 Figure 2 (a), shows the dependency of mean snowfall-frequency-fractions (number of

177 snowfall events vs. number of precipitation events) and snowfall-intensity-fractions (mean

178 snowfall-intensity vs. mean total precipitation-intensity) on total precipitation at 0 °C. For

179 temperatures below -5 °C the snowfall-fraction is set to 1. For temperatures above 10 °C the

180 snowfall-fraction is set to 0. In the temperature range between -5 to 10 °C we use the calcu-

181 lated snowfall-frequencies from the LUT to randomly declare precipitation days as snow days

182 (Fig. 2 a: event frequency). Compared to SFE1 the temperature range where the snowfall-

183 fractions can deviate from zero is extendet. However, between 7 and 10 °C the models show

184 also snow-fraction values ~0 °C (Fig. 2 b).

185 Solid and liquid precipitation can occur on the same day, therefore not all precipitation

186 on a snowfall day can be counted as snow. In order to match the snowfall-fractions to the

187 values from the LUT, we reduce the snowfall intensity of each event until the mean snowfall-

188 fractions at each temperature bin are equal (Fig. 2 a: intensity adjustment). Since this step is

189 done after we declared only a certain percentage of all precipitation days as snowfall days,

190 the intensity reduction for the remaining events is lower compared to SFE1.

191 Figure 2 b shows the temperature dependency of SFE1 and SFE2 compared to snow-

192 fractions simulated by the EURO-CORDEX models. Between about -5 to 5 °C snow frac-

193 tion statistics strongly depend on the climate model. Also SFE1 and SFE2 show clear dif-

194 ferences in this temperature range. For temperatures higher than 2 °C SFE1 estimates are

195 above the model spread. SFE2 uses statistical information from the model ensemble and rep-

196 resents the ensemble mean. The difference between different ensemble members can be up to

197 55%.

198 With two different temperature data sets and two different functions to estimate snow-

199 fall, we obtain four snowfall data sets based on EURO4M total precipitation data: SFE1-

200 EOBS, SFE2-EOBS, SFE1-HISTALP and SFE2-HISTALP. In Sect. 3.1 we show a compar-

201 ison of these data stets. Hereafter we will call snowfall estimates that are based on observa-

202 tional data "snowfall observations", in order to distinguish them from snowfall estimates that

203 are based on EURO-CORDEX simulations.
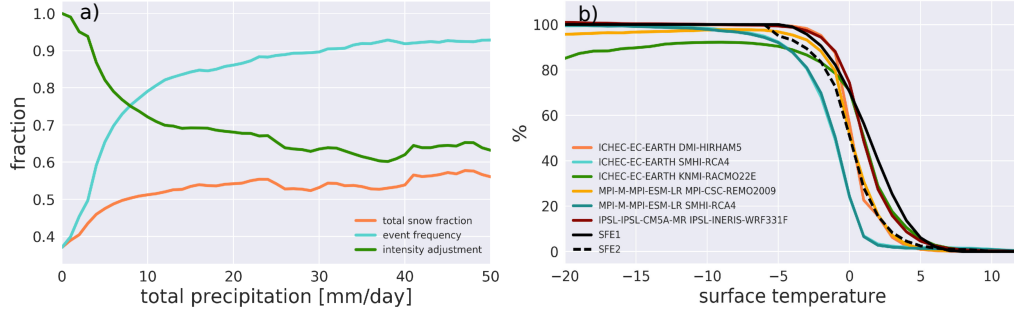
204 **Figure 2.** a) Ensemble mean statistics used in the SFE2 look-up table. Dependency of the total snowfall-

205 fraction, the snow event frequency and the intensity adjustment on total precipitation intensity at 0 °C. b)

206 Snowfall-fractions as a function of surface-air-temperature in EURO-CORDEX simulations compared to the

207 estimations SFE1 and SFE2 for the period 1971-2000.

208 ## 2.3 Introduction to the concept of the probabilistic decomposition-framework

209 ### 2.3.1 Distinguishing occurrence probabilities by mean daily surface air-temperatures

210 We separate temperature related changes from the total CCS of precipitation- and

211 snowfall-frequencies using the law of total probability (*Fahrmeir et al.* [2016]).

212 According to this law we can split the total probability that an event will occur ($P(I)$),

213 into separate probabilities that an event will occur at a specific temperature range ($P(I|T_i)$).

214 The total probability can than be calculated by weighting the probabilities $P(I|T_i)$ with the

215 probability that a day will have a mean daily air-temperature in the temperature range $P(T_i)$

216 and integrate over the entire sample space of temperatures. In practice we sum from T_{min} to

217 T_{max} using 1 K temperature bins (eq. 2).

$$P(I) = \sum_{T=T_{min}}^{T_{max}} P(I|T_i) * P(T_i) \tag{2}$$

218 The letter I stands for indices, the letter T for temperature. $P(I)$ and $P(I|T_i)$ are used as gen-

219 eral terms to indicate the total probability and the temperature dependent probability. In this

220 study these terms will be use for all data sets and indices.

²²¹   *2.3.2 Using the probability separation to analyze the climate change signal*

²²²   The chance, that an event will occur in the future period (P_f), can be thought of as the

²²³ sum of the chance that an event occurs today (P_h) and a term that represents changes in the

²²⁴ occurrence probability (P_{ccs}). Adding the terms for future changes to equation 2, we get the

²²⁵ following equation for the total occurrence probability in the future period ($P_f(I)$).

$$P_f(I) = \sum_{T=T_{min}}^{T_{max}} [P_h(I|T_i) + P_{ccs}(I|T_i)] * [P_h(T_i) + P_{ccs}(T_i)] \tag{3}$$

²²⁶   After expanding the product, we obtain the following terms:

²²⁷   1. $P_h(I|T_i) * P_h(T_i)$: The total occurrence probability in the control period ($P_h(I)$).

²²⁸   2. Three terms that together form the CCS.

²²⁹   (a) $P_h(T_i) * P_{ccs}(I|T_i)$: CCS caused by changes in the temperature dependent oc-

²³⁰   currence probability $P(I|T_i)$. Changes in this term will change the temperature-

²³¹   dependency of the events. A main cause for changes in this term are changes in the

²³²   large scale circulation. This term will further be referred to as $P(I|T_i)$ dependent

²³³   CCS.

²³⁴   (b) $P_h(I|T_i) * P_{ccs}(T_i)$: CCS caused by changes in the temperature distribution $P(T_i)$.

²³⁵   These changes are mainly related to thermodynamic effects. This term will further

²³⁶   be referred to as $P(T_i)$ dependent CCS.

²³⁷   (c) $P_{ccs}(I|T_i) * P_{ccs}(T_i)$: CCS caused by correlations between changes in the $P(T_i)$

²³⁸   and the $P(I|T_i)$ distribution. This term will further be referred to as correlation

²³⁹   dependent CCS.

²⁴⁰   *2.3.3 Decomposition of model biases in the climate change signal*

²⁴¹   In the first part of the analysis we compare the two 20 year time periods 1971-1989

²⁴² and 1990-2008. For both periods, observation data is available, allowing a comparison of the

²⁴³ simulated changes against observations.

²⁴⁴   From equation 3 we derive three terms that describe changes in the event occurrence

²⁴⁵ (2: a-c). Comparing the model simulations with observational data, we find differences in

²⁴⁶ the historic distributions and in the CCS. To analyze the impact of these biases individually,

²⁴⁷ a δ term is added to each probability distribution (2: a-c), that represents a possible model

²⁴⁸ bias. We obtain the following equation for the observed distribution changes:

$$P_{ccsobs}(Index) = [P_h(T_i) + \delta P_h(T_i)] * [P_{ccs}(I|T_i) + \delta P_{ccs}(I|T_i)] +$$

$$[P_h(I|T_i) + \delta P_h(I|T_i)] * [P_{ccs}(T_i) + \delta P_{ccs}(T_i)] +$$

$$[P_{ccs}(I|T_i) + \delta P_{ccs}(I|T_i)] * [P_{ccs}(T_i) + \delta P_{ccs}(T_i)]$$

(4)

After expanding the products, we obtain the CCS as simulated by the model:

$$P_h(T_i) * P_{ccs}(I|T_i) + P_h(I|T_i) * P_{ccs}(T_i) + P_{ccs}(I|T_i) * P_{ccs}(T_i)$$

additionally we obtain nine terms, that add up to the total bias in the climate change signal see table 2. If percentage changes are analyzed we receive two additional terms P1 and P2 (for more details see the supplementary material). These terms become important if there is a mismatch in the magnitude of the historic distribution between models and observations. We will refer to the terms P1 and P2 as scaling terms. In section 3.2.3 the terms from Tab. 2 are used to analyze the origin of differences we find between observed and simulated past climate changes.

2.3.4 Adjustment of the temperature and precipitation distribution

Decomposing the CCS bias according to Tab. 2 enables us to quantify how strong biases in the temperature distribution and biases in the temperature dependency of the events influence the climate signal. Under consideration of this knowledge, we analyze the combined impact from biases found in the historic distributions P(T) and P(I|T) on the simulated CCS. In order to calculate the combined impact, we adjust the simulated distributions to the observed data. For the historic distributions, we take the observed data instead of data from the simulations. The future distributions are calculated by adding the simulated changes from the model ensemble to the observed data from the control period. For the analysis of percentage changes, the percentage CCS is added to the observed distributions instead of absolut values. This adjustment is similar to the delta change method which is often used for bias-adjustments (*Schmucki et al.* [2014]; *Bosshard et al.* [2011]). Here we use the delta change method on the P(T) and on the P(I|T) distributions. Superimposing changes in the P(I|T) distribution upon the observed P(I|T) distribution in the control period, leads to a temperature dependent bias-adjustment of precipitation.

distribution	nr.	term	description and comments
	1	$P_h(T_i) * \delta P_{ccs}(I|T_i)$	bias caused by a bias in the CCS of $P_{ccs}(I|T_i)$. These differences are caused by all non-temperature related changes in precipitation e.g. caused by circulation changes
Model bias in the CCS, caused by changes in $P(I|T)$	2	$\delta P_h(T_i) * P_{ccs}(I|T_i)$	bias caused by a bias in the historic temperature distribution $P_h(T)$
	3	$\delta P_h(T_i) * \delta P_{ccs}(I|T_i)$	bias caused by correlations between the bias in the historic temperature distribution $P_h(T)$ and the bias in the CCS of the temperature dependent frequency-distribution
	P1.	$100 * \frac{P_h(T_i)*P_{ccs}(I|T_i)}{P_h(I)} -$ $100 * \frac{P_h(T_i)*P_{ccs}(I|T_i)}{P_{obs}(I)}$	additional term for percentage changes as result of differences in the base period $P_h(I)$ and $P_{obs}(I)$ (see supplementary material for details)
	4	$P_h(I|T_i) * \delta P_{ccs}(T_i)$	bias caused by a bias in the CCS of $P(T)$
Model bias in the CCS caused by changes in $P(T)$	5	$\delta P_h(I|T_i) * P_{ccs}(T_i)$	bias caused by a bias in the historic distribution $P_h(I|T)$
	6	$\delta P_h(I|T_i) * \delta P_{ccs}(T_i)$	bias caused by correlations between the bias in the historic distribution $P_h(I|T)$ and the bias in the climate change signal of $P(T)$
	P2.	$100 * \frac{P_h(I|T_i)*P_{ccs}(T_i)}{P_h(I)} -$ $100 * \frac{P_h(I|T_i)*P_{ccs}(T_i)}{P_{obs}(I)}$	additional term for percentage changes as result of differences in the base period $P_h(I)$ and $P_{obs}(I)$ (see supplementary material for details)
Model bias in the CCS caused by changes in the correlation between $P(T)$ and $P(I|T)$	7, 8, 9	$P_{ccs}(I|T_i) * \delta P_{ccs}(T_i)$ $\delta P_{ccs}(I|T_i) * P_{ccs}(T_i)$ $\delta P_{ccs}(I|T_i) * \delta P_{ccs}(T_i)$	We find the three terms (7, 8, 9) to be small and therefore combine these terms in the result section
SUMS	SUM1	SUM of the terms with the number: 2, 5, P1, P2	SUM1: sum of all products that do not include a δP_{ccs} term. These terms only need observation information from the control period and can be calculated for the future periods.
	SUM2	SUM of the terms with the numbers: 1, 3, 4, 6, 7, 8, 9	SUM2: sum of all products that include a δP_{ccs} term. These terms need an observed CCS and can not be calculated for future periods.

–12–

Table 2. Overview of the nine terms (plus P1 and P2 for percentage changes) that add up to the total bias in

the climate change signal

3 Results

3.1 Comparison of the snowfall discrimination methods SFE1 and SFE2 with model simulated snow fractions.

Two snowfall discrimination methods are used in order to analyze the sensitivity of observed snow frequency changes to different snowfall estimations. To understand discrepancies in the CCS caused by differences in snowfall-fraction, we need to analyze how strong the temperature dependency of snowfall is influenced by the discrimination methods and how the snowfall estimates compare to the simulations. For this comparison we apply the discrimination methods SFE1 and SFE2 (Sect. 2.3.1) to temperature and precipitation data from each ensemble member. We obtain one simulated and two estimated snowfall ensembles.

We start with a validation of these three data sets. Fig. 3 shows temperature dependent snowfall statistics for the three ensembles. SFE2 matches the simulated ensemble mean snowfall values closely over the entire temperature range (Fig. 3 a). As a consequence of SFE1 being at the upper limit of the simulated snow fraction range (Fig. 2 b), also the mean snowfall values are at the upper limit. Figure 3 b shows the standard deviation of snowfall intensities. SFE2 is again able to capture the simulated curve progression. SFE1 agrees better with the models in the standard deviation than in the mean values. For higher percentiles SFE1 seems to better match the model ensemble. A likely explanation for this observation is, that the simulated snow fraction statistically increases for more intense events (Fig. 2 a). The strongest discrepancy between SFE1 and the simulated snowfall is found for snowfall-frequencies using a 1 mm/day threshold (Fig. 3 c). The assumption that every precipitation event at temperatures below 7 °C will have a certain snowfall-fraction, leads to an increased number of snowfall events between about -6 to 7 °C. This moves the temperature range where snowfall events are most likely to occur, towards higher temperatures. Intense snowfall events are not as strongly affected by this (Fig. 3 d).

Analyzing the curve progression we find that the most snowfall is simulated at surface air-temperatures of about -5 °C with values between 4 to 7 mm/day. The highest snowfall-frequency is located at temperatures close to -6 °C. For heavy-snowfall events the distribution becomes narrower and shifts slightly to higher temperatures.
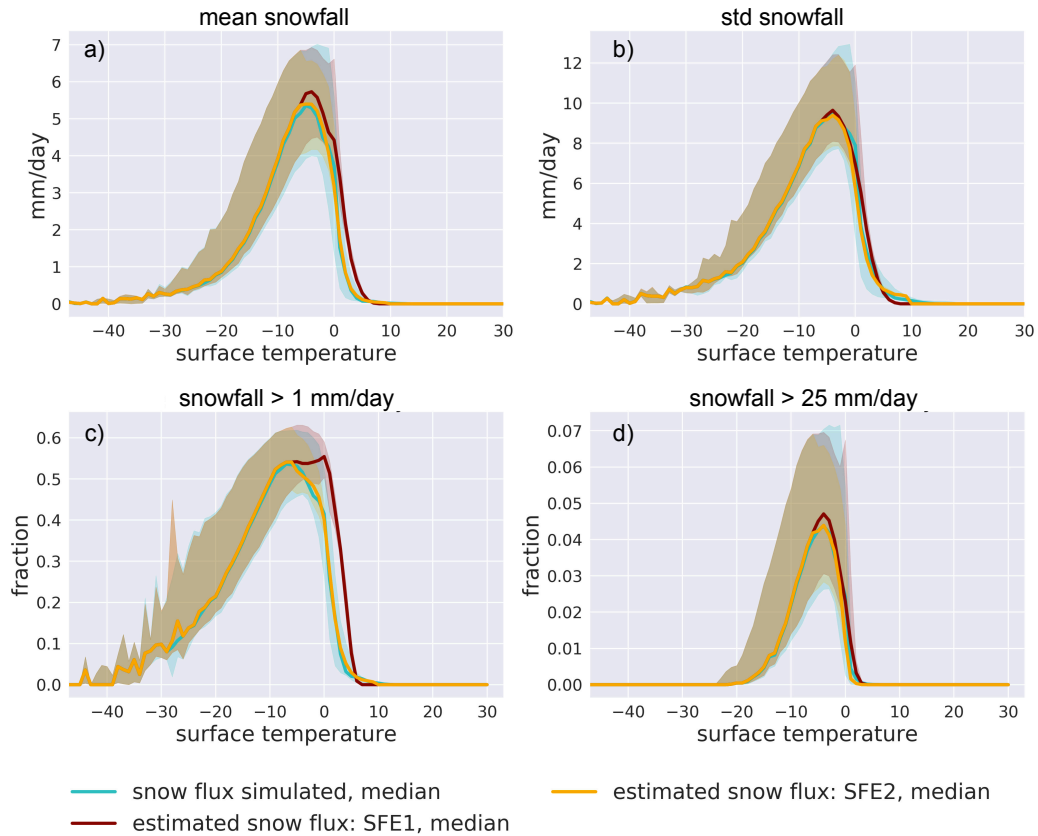
Figure 3. Comparison of estimated and simulated snow indices using snow, precipitation and surface air-temperature data from the EURO-CORDEX ensemble. a) mean snowfall, b) standard deviation of snowfall, c) snowfall-frequency, d) heavy-snowfall-frequency. Shaded areas indicate the ensemble spread (5th to 95th percentile).

3.2 Observed climate changes 1971-2008

To analyze the sensitivity of projected precipitation- and snowfall-frequencies to model biases, we compare simulated and observed historic changes (CCS_h), between the periods 1971-1989 and 1990-2008. To give an overview about the observed CCS_h and to indicate differences between the observation data sets, we show timelines of the observational data sets described in Sect. 2.2 as field-mean values over the entire northern Alpine region (Sect. 2.1). The straight lines indicate temporal mean values over the periods 1971-1989 and 1990-2008. Between these periods we find a temperature increase of 0.7 to 0.9 °C, depending on the temperature data set chosen (Fig. 4 a). The difference in the temperature change signal between EOBS and HISTALP, is due to slightly colder temperatures in HISTALP over the period

318   1971 to 1990. We also see an increase in the frequency and in the mean intensity of precipi-

319   tation (Fig. 4 b, d). At the same time mean snowfall amounts and frequencies are decreasing

320   (Fig. 4 c, e, f). These findings are in line with earlier studies that analized station data in the

321   Alpine region (*Brunetti et al.* [2006]; *Beniston* [2006]; *Beniston and Rebetez* [1996]; *Scher-*

322   *rer et al.* [2004]).
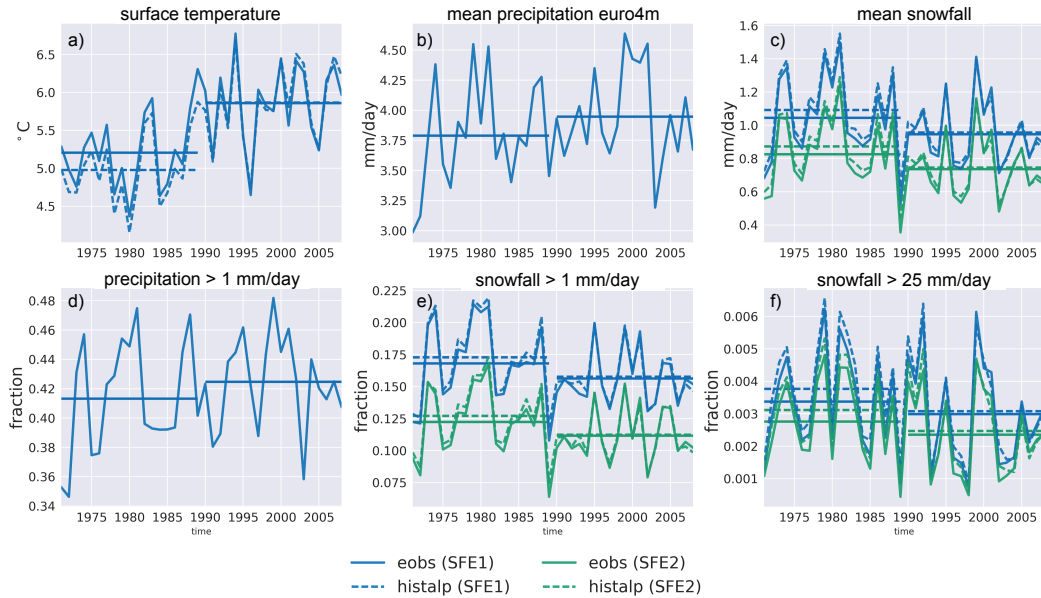
Figure 4.   Yearly mean timelines of the different observation data sets described in Sect. 2. The data rep-

resents a field-mean over the northern Alpine region. Straight lines indicate 20 year mean values over the

periods 1971 to 1989 and 1990 to 2008.

3.2.1 Comparing simulated vs. observed probability distributions

327   Earlier studies revealed significant biases in the historic model simulations over the

328   Alpine region (*Kotlarski et al.* [2014]; *Frei et al.* [2017]; *Rajczak et al.* [2013]; *Smiatek et al.*

329   [2016]). We analyze how these biases modify the temperature distribution as well as the tem-

330   perature dependency of (heavy) precipitation- and snowfall-frequencies at different altitudes.

331   Assuming a change in the temperature distribution, biases in these distributions will alter

332   the CCS. In the following section we will use this information to explain differences we find,

333   between the observed and the simulated CCS in the past period.

Figure 5 a, shows the temperature probability distribution P(T) for the periods 1971-1989 and 1990-2000 as well as changes that occurred between these periods, for the model ensemble and the observations. Changes in the temperature distribution will have an impact on all indices. The strength of this impact depends on the temperature dependency of the event, indicated by the temperature dependent probability distribution P(I|T) (Fig. 5 b, d, f, h). As explained in Sect. 2.3 the total occurrence probability of an event (P(I)) is the product of the probability P(T) and P(I|T) (Fig. 5 c, e, g, f).

Analyzing changes in the temperature distributions we find that compared to the simulations the center of the observed distribution is shifted towards higher values (Fig. 5 a). Also the shapes of the observed and simulated distributions differ. The simulated distributions are stronger skewed to the right with a more pronounced peak at 0 °C. The cold bias is also a result of the shape mismatch and not just a result from a shift towards higher temperatures in the observations.

Analyzing the temperature distributions at different heights (see supplementary material) we find, a shift from a bimodal distribution, to a unimodal distribution with altitude. At altitudes below 1000 m two modes close to 0 °C and ~12 °C are visible. At high altitudes above 1500 m only the single peak at 0 °C is found. This shift is not as pronounced in the observation data set explaining the weaker skewness of the temperature distribution when using data from all altitudes.

For total-precipitation-frequencies we make the following observations. Figure 5 b, shows the temperature dependent frequency-distribution (P(I|T)) for precipitation events above 1 mm/day. Observations and the simulations show a substantial different shape of this distribution. The observed distribution is skewed left with a maximum at ~12 °C, whereas the simulated distributions are skewed right and have a peak at ~1 °C. Because of these differences, the simulations show a higher chance for precipitation events at temperatures below ~12 °C than observed and a lower chance above ~12 °C. For heavy-precipitation events (Fig. 5 f), this pattern is more pronounced and shifted towards lower temperatures. The switch from a wet to a dry bias is located at ~ 8 °C. At 0 °C the chance for a heavy-precipitation event to occur, is 2 to 5 times larger in the model ensemble than in the observation data.

The chance that a (heavy) precipitation day with a certain mean-surface-temperature will occure (P(I)), is shown in figure 5 c (g). The cold temperature bias, in combination with

the overestimation of precipitation-frequencies at cold temperatures, leads to an overestimation of precipitation days at temperatures below ~5 °C. The largest bias is found at temperature close to 0 °C. In this temperature range the models simulate a ~50 % higher occurrence probability for a precipitation event.

Analyzing altitudinal aspects we find a lower altitude dependency in the precipitation-frequency in the observations than in the models (see supplementary material). The lower dependency on altitude seems to be a result of a latitude dependency that is specific to the northern Alpine region (*Isotta et al.* [2013]; *Masson and Frei* [2014]). Other reasons could be measurement errors, that are found to be especially large for high elevations in the winter season (*Prein and Gobiet* [2016]).

Figure 5 d, e show the P(I|T) distribution for snow and heavy-snowfall-frequencies. Precipitation-frequencies are overestimated below 0 °C, at these temperatures most precipitation will fall as snow, therefore also the snowfall-frequencies are overestimated. Differences between precipitation and snowfall are described by the snowfall-fraction. Especially in the range between -5 to 5 °C the different snowfall-fraction statistics illustrated in Fig. 2 introduce additional errors. Depending on the model or discrimination method, a different fraction of the total precipitation is identified as snow. To analyze this effect, we use the two discrimination methods SFE1 and SFE2 from section 2.3.1. The largest differences between SFE1 and SFE2 are found for snowfall-frequencies using the 1 mm/day threshold (Fig. 3). For this threshold SFE1 shows higher snowfall-frequencies for all temperatures above 0 °C. This is shifting the maximum in the observed P(I|T) distribution from about -2 °C using SFE2, to 0 °C using SFE1 (Fig. 5 d blue shade). The maximum P(I|T) values in the model ensemble are located at ~-5 °C with probabilities between 45 to 62 %. Because of the different snowfall-fraction statistics, the observations show a higher probability for a snowfall day at temperature above 0 °C although total precipitation-frequency is overestimated in the models.

The total probability P(I) for snowfall events is shown in Fig. 5 e (Fig. 5 f, heavy-snowfall events). Because of the peak in the temperature distribution at 0 °C, the maximum is located at higher temperatures than for P(I|T). The peak in the temperature distribution is more pronounced in the simulations than in the observations, causing a stronger shift towards higher values. We find that models and observations both show the most snowfall days at temperatures between -1 to 0 °C.

A comparison of the snowfall-frequency-distributions at different altitudes (see supplementary material) shows a shift in the maximum of P(I|T) from about 0 °C at altitudes below 1000 m to ~-6 °C at altitudes above 1500 m, that can not be found for heavy-snowfall-frequencies. We also find that the differences between models and observations become larger at high altitudes.
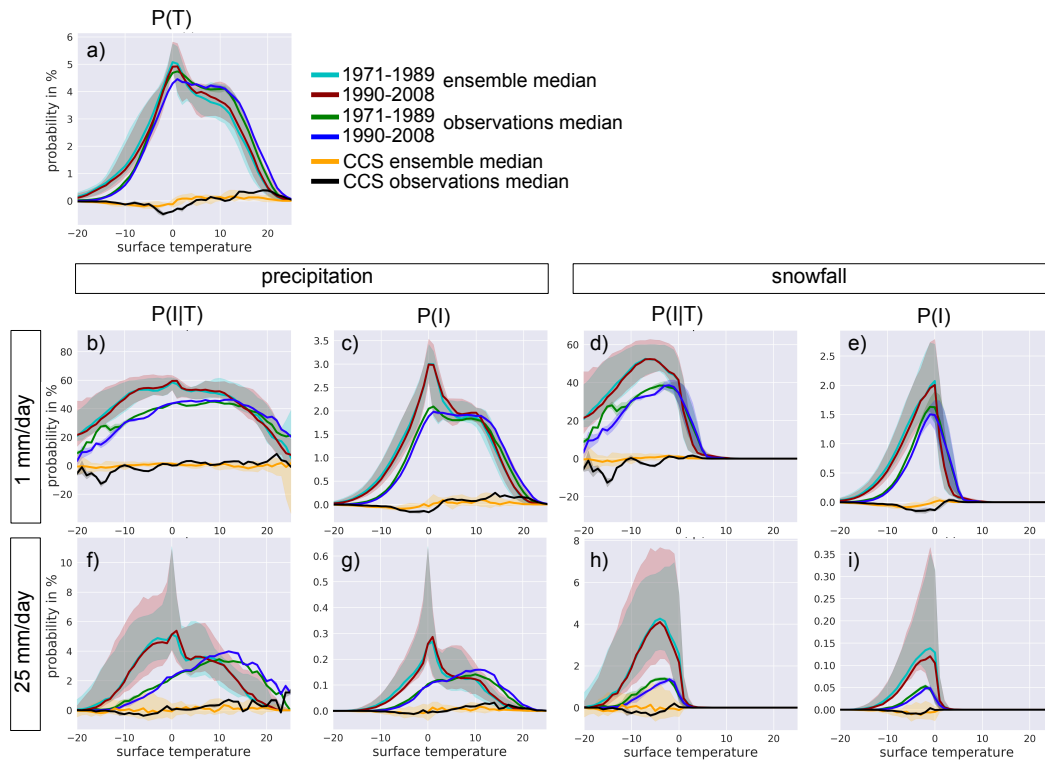
Figure 5. Temperature dependent probability distributions P(T), P(I|T) and P(I) for precipitation- and snowfall-frequencies. Shaded areas indicate the ensemble spread (5th to 95th percentile).

3.2.2 Past snowfall and precipitation-frequency changes

We compare the simulated CCS$_h$ between the two selected past periods 1971-1989 and 1990-2008 against observations. This enables us to compare not just the P(T) and P(I|T) distributions themself but also the sensitivity to changes in these distributions.

Analyzing observed changes in the surface-temperature distribution (Fig. 5 a), a strong reduction in the probability distribution for temperatures close to 0 °C and an increase for

temperatures above 12 °C is found. In the ensemble mean of the simulations, P(T) is increasing for most temperatures above about 1 °C and decreasing for temperatures below this threshold. The observed reduction between -5 to 3 °C is less pronounced in the model ensemble. The number of observed precipitation events increased at temperatures above 2 °C and decreased below (Fig. 5 c), causing a decrease in snowfall-frequency (Fig. 5 e).

In Fig. 6 we analyze the difference between the observations and the simulations integrated over the entire temperature range. For this comparisons we decompose the climate change signal according to equation 3. Fig. 6 shows the total CCS_h as well as the decomposed terms. These terms show the CCS_h if only one of the two distributions (P(T), P(I|T)) would have changed. The CCS_h of the model ensemble is illustrated by boxplots. The observed CCS_h is indicated by orange lines.

Analyzing past precipitation-frequency changes (Fig. 6 a, c), the observations show an increase of ~3 % (P(I) CCS) and ~9 % for heavy-precipitation-frequencies. The mean increase in precipitation-frequency of the model ensemble is close to 1 % for both intensity thresholds. Comparing the CCS_h we obtain by changing only one of the two distributions (P(T) dep. CCS, P(I|T) dep. CCS) we find, that changes in the temperature distribution as well as the correlation terms played only a minor role for precipitation- and heavy-precipitation-frequency changes.

For snowfall-frequencies (Fig. 6 e, g) the different snow estimation methods (SFE1, SFE2), in combination with the different temperature data sets, result in considerably different climate change signals which increases the spread in the obsesrvations. SFE1-EOBS is always the top orange line and SFE2-HISTALP always the bottom line (see supplementary material for details). Unlike precipitation-frequency changes, snowfall-frequency changes were mainly driven by temperature changes. For heavy-snowfall-frequencies, however, changes in both distributions are important. Comparing the simulations against the observations we see that the observations show a stronger decrease in snowfall-frequency and that these differences are mainly caused by changes in the temperature distribution.

Analyzing altitudinal aspects we find, that the observed increase in (heavy) precipitation-frequency is (100 %) 50 % stronger at altitudes above 1000 m. This behavior is not found in the simulations. The observed decrease in the snowfall-frequency is about 50 % stronger at low altitudes. For heavy-snowfall events the altitude dependency becomes even more significant. The decrease in snowfall-frequency is mostly underestimated in the simulations at

all altitudes. At low altitudes this is mainly caused by underestimating P(I|T) dep. changes whereas at high altitudes an underestimation of P(T) dep. changes is found to be the main reason.
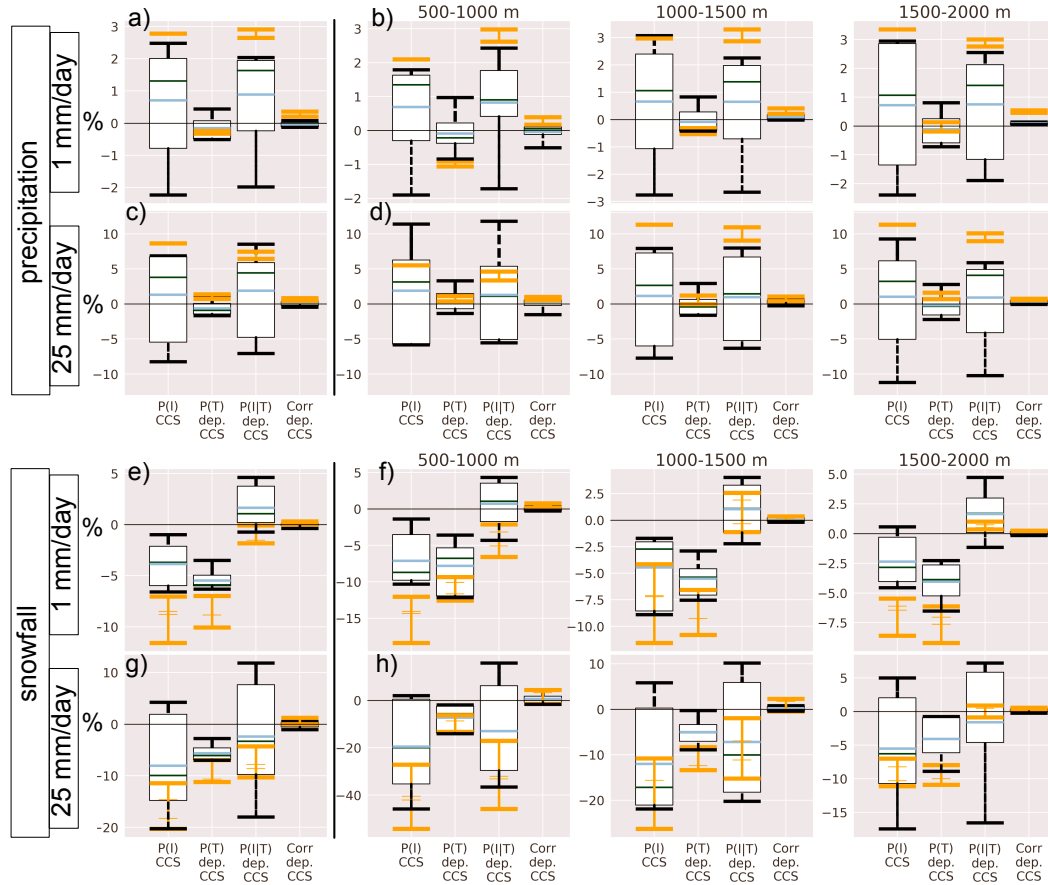
Figure 6. CCS$_h$ of precipitation-freq. (a, b), heavy-precipitation-freq. (c, d) snowfall-freq. (e,f) and heavy-snowfall-freq. (g, h), in total (P(I) CCS) and separated by P(T) dependent CCS, P(I|T) dependent CCS and correlation dependent CCS, according to equation 3. Observed changes are indicated by orange lines. The boxplots show the interquartile range of the model ensemble. The whiskers indicate the ensemble spread (5th to 95th percentile). The ensemble (mean) median is illustrated by the dark green (blue) line.

3.2.3 Decomposition of model biases in the historical climate change signal

In Fig. 6 differences between the observed and simulated CCS$_h$ become visible. Decomposing the CCS$_h$ in P(T) and P(I|T) dependent changes gives first insights about the causes for the observed changes. To relate differences in the CCS$_h$ to specific model biases, a further step is needed. The observed impact of temperature changes on snowfall-

frequencies (P(T) dep. CCS) could be stronger than simulated because of the following three reasons. First, the observations could experience stronger temperature changes as simulated. Second, a bias in the temperature dependency of snowfall-frequencies within the control period (model-bias) could result in a lower CCS_h. A third reason could be correlations between these biases (correlation-Bias). Similar considerations can be made for P(I‖T) dependent changes and the correlation dependent changes. In total we obtain nine terms that explain why the observed and the simulated CCS_h differ. Each of these terms points to different biases. Detailed knowledge about the main reasons that cause the bias in CCS_h , can help to detect shortcomings in the model simulations. It also enables us to explain why bias-adjustment is changing the CCS_h. As the observations are not free from errors, the detailed information also helps to identify, if the CCS_h bias may partly be explained by observation errors. In the following section we decompose the biases in the CCS_h according to Sect. 2.3.3.

In Fig. 7 (a-d) the term 1 describes differences in the CCS_h of the respective indices, that are caused by different changes in the temperature-dependency. For this term a large spread in the model ensemble is visible. This indicates disagreements between the ensemble members in terms of circulation changes. For precipitation- and heavy-precipitation-frequencies (Fig. 7 a, b) the ensemble mean value of term 1 is close to the observations, whereas snowfall-frequencies show a positive bias. The positive bias means that the models show a lower reduction in the frequency of snowfall than the observations

Term 4 represents differences in the CCS_h of the respective indices, caused by different changes in the temperature distribution. For this term a better model agreement is found for all indices (Fig. 7 a-d). However, for all four indices a positive bias is shown. Considering only the different changes in the temperature distribution between models and observations, these differences would cause a stronger increase in the precipitation-frequency which leads to a reduced decrease in the snowfall-frequency.

The terms 3, 6, 7, 8, 9 all present correlation related biases in the CCS_h. Terms 7, 8, 9 show small values and are therefore combined in one boxplot. The terms 3 and 6 indicate differences in the CCS_h of the respective indices, that are caused by correlations between a model-bias and a different change of either the temperature distribution or the temperature-dependency. Because these terms include a model-bias, they will be influenced by bias-adjustment approaches. The bias described by term 3 reduces the simulated precipitation-

frequency increase by ~2 %. The models simulate fewer days with mean daily surface-temperatures above ~3 °C than observed. Especially for these days the observations show a stronger increase in the precipitation-frequency (Fig. 5 a, b).

The remaining terms 2, 5, P1 and P2 all describe an influence of a specific model-bias on the CCS_h. To calculate these terms, observations are needed only for the control period. Information about future changes in the temperature distribution or in the temperature-dependency of the event are not needed. Therefore we can also calculate these terms for future climate scenarios (Sec. 3.3). In order to analyze the combined impact of these terms and to identify the relative impact compared to all other terms, we create two sums. The first sum (SUM1) includes the terms [2, 5, P1, P5] and the second sum (SUM2) includes all remaining terms [1, 3, 4, 6, 7, 8, 9].

For (heavy) precipitation-frequency we find that the mean simulated increase is (~7 %) ~2 % weaker in the simulations than in the observations. For the 1 mm/day threshold term 3 is identified as the major reason for discrepancies between observations and simulations. For heavy-precipitation SUM1 explains most part of the bias. In both cases the identified model biases strongly influence the climate signal.

Analyzing snowfall-frequencies we find that the observed decrease is underestimated in the simulations by about 5 %. This underestimation is mainly caused by the model biases described in the terms 1 and 4. The observations show a stronger decrease of days with surface-temperatures close near 0 °C (term 4) as well as a stronger decrease in the snowfall probability at temperatures below 0 °C (term 1) (compare to Fig. 5). However, the cold temperature bias and the temperature dependency of these events lead to a stronger increase in snowfall-frequency (SUM1). Unlike for precipitation-frequency changes, SUM1 and SUM2 point in different directions, therefore partly compensate each other.

3.3 Future climate change

After the validation of precipitation- and snowfall-frequency changes within the past period, we analyze the effect of temperature changes as well as the effect of the model biases on climate projections, using the RCP4.5 scenario. For this study we compare the period 1971-2000 to 2071-2100. Figure 8 (a-d) shows the projected CCS for snow- and precipitation-frequencies. In addition the expected biases in the projected CCS that are caused by the terms (2, 5, P1, P5) are illustrated for the future period.
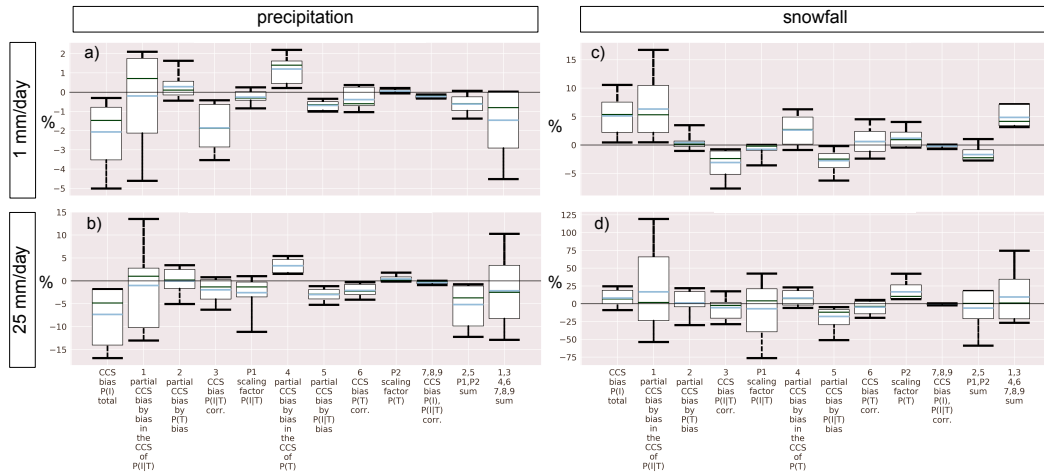
Figure 7. Difference in the percentage CCS$_h$ between each ensemble member and the mean of the observations for (a) precipitation-freq., (b) heavy-precipitation-freq., (c) snowfall-freq., (d) heavy-snowfall-freq., in total and separated according to equation 4. For a description of the terms see Tab. 2. Boxplots as in Fig. 6.

Analyzing precipitation-frequency changes we find, that the ensemble mean shows a slight positive trend. Two out of six models, however, show a decrease. For heavy-precipitation-frequencies all models point in the same direction, with a 10 % increase in the ensemble mean. It shows that the increase in the precipitation-frequency is mainly a result of changes in the P(I|T) distribution. Changes in the P(T) distribution caused a decrease in precipitation- and in heavy-precipitation-frequency.

Snowfall-frequencies show a decrease of ∼19 % in the ensemble mean. These changes are almost completely explained by changes in the temperature distribution. Changes in heavy-snowfall-frequencies, caused by changes in the temperature distribution are also close to 19%. However, changes in the P(I|T) distribution considered by itself, would lead to an increase in heavy-snowfall-frequency of about 7%. Therefore the total decrease in heavy-snowfall-frequency is only ∼13% (∼1 % decrease because of the correlation term).

For precipitation-frequencies as well as for snowfall-frequencies, we find that the bias in the CCS, that is caused by model biases in the P(T) and P(I|T) distributions, is increasing compared to the validation study. In the case of precipitation, these biases likely reduce the expected increase in precipitation-frequency. The projected increase should therefore be

higher for bias-adjusted data. For snowfall-frequencies the biases are also negative. The projected decrease in snowfall-frequency could be overestimated due to these biases.
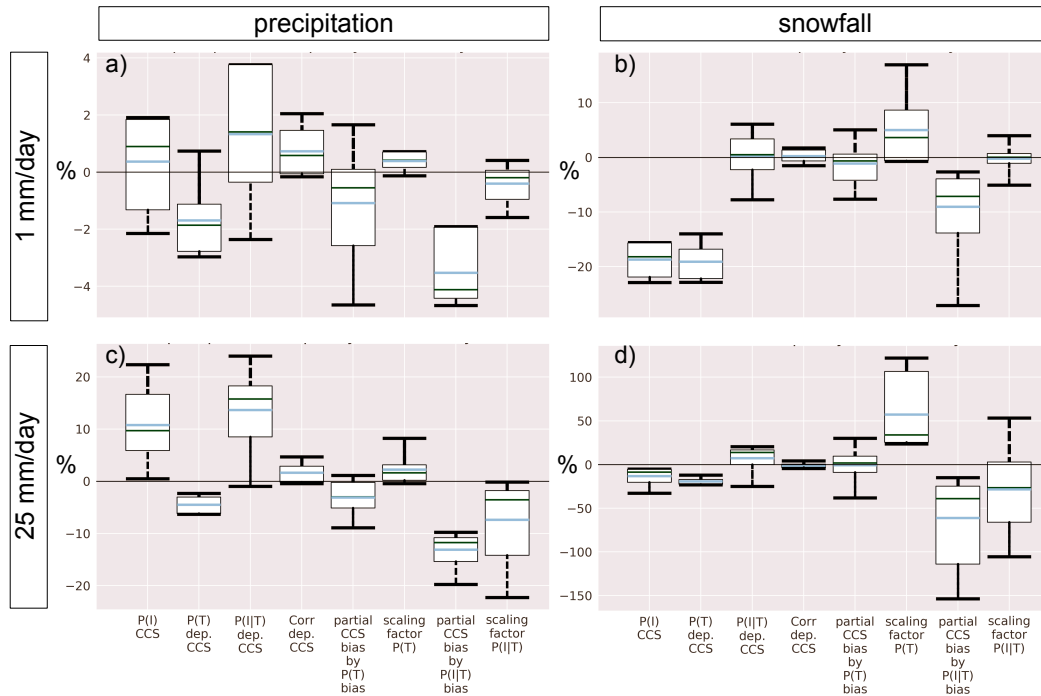
Figure 8. Future climate change signal for (a) precipitation-freq., (b) snowfall-freq., (c) heavy-precipitation-freq., (d) heavy-snowfall-freq., as simulated by the model ensemble. The first four boxplots in each graph show the climate change signal in total, and separated by P(T) and P(I|T) dependent changes according to equation 4. The other four boxplots illustrate a possible bias in the climate change signal that results from a bias in the P(T) and P(I|T) distributions in the control period.

3.4 Bias-adjustment

From the decomposition analysis we learn, that the different CCS_h between models and observations can partly be explained by model biases in the P(T) and the P(I|T) distributions. After we gained understanding about these processes and quantified the impact of the relevant terms, we analyze the total effect of these changes. The historical distributions P(T) and P(I|T) are adjusted to the observations (Sect. 2.3.4). The distributions for the future period are obtained by adding the projected CCS to the adjusted distributions. Figure 9 shows the CCS for the adjusted data. Comparing Fig. 9 a-d, against the non-bias-

adjusted CCS_h from Fig. 6, we find a closer agreement to the observed total CCS_h for both precipitation-frequencies. Looking at the dependent climate signals (P(T), P(I|T)), we see only an improvement for heavy-precipitation-frequencies. For the 1 mm/day threshold, the bias in the CCS_h is mainly caused by a bias in the CCS_h of P(T) and P(I|T), that is not corrected by the adjustment. For snowfall-frequencies we show, that the model biases in P(T) and P(I|T) partly compensate the underestimated decrease in the simulations (Fig. 7 (c, d)). In this case the bias-adjustment leads to a larger difference between the simulated and the observed CCS_h.

For the future projections (Fig. 9 e-h) the changes caused by the bias-adjustment point in the same direction as for the control period but prove to be stronger. Because of the large differences between models and observations in P(I|T), the temperature increase leads to a clear increase in precipitation-frequency in the bias corrected case and to a decrease if the data is not bias corrected (P(T) dep. CCS) . We also find a stronger increase for P(I|T) dependent changes in the precipitation-frequency. This differences are mainly explained by the cold temperature bias. The models show lower occurrence probabilities of days with high surface-temperatures where the probability of precipitation events is increasing the most.
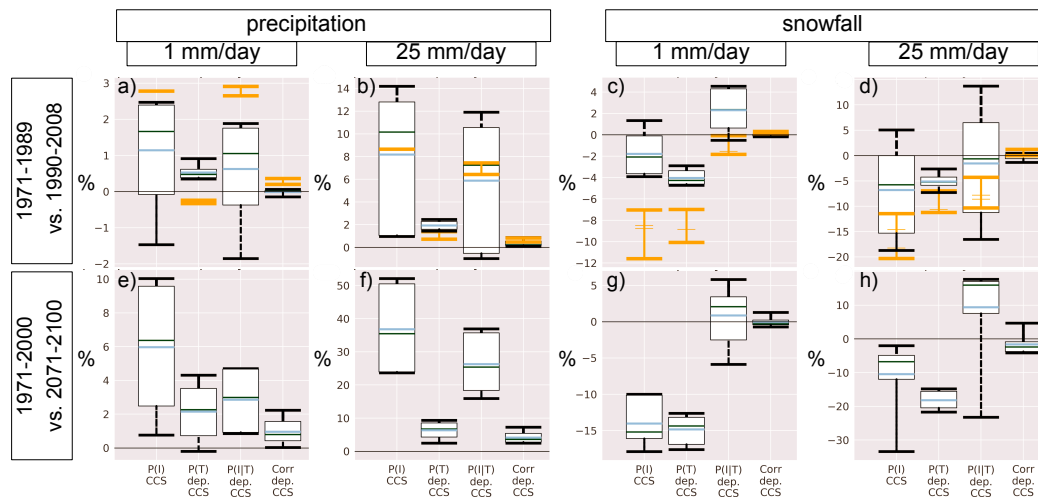
Figure 9. Bias-adjusted climate change signal (a-d) between the periods 1971-1989 and 1990-2008, (e-h) between the periods 1971-2000 and 2071-2100. Boxplots as in Fig. 6.

4 Summary and Conclusions

Probability decomposition: In this study we split the occurrence probability of a (snowfall) precipitation event into the two probabilities P(T) and P(I|T). The separation allows us to analyze future changes caused by a change in each of these distributions separately. Comparing the simulations against observations we find profound differences in the shape of their distributions. These differences can be connected to known model- as well as observation errors. The mismatch in the shape of these distributions has important impacts on bias-adjustment methods. Adding a constant factor to the temperature data set in order to correct the cold bias will only shift the distribution. This shift could even enhance the mismatch compared to the observations. The peak in the temperature distribution at 0 °C is physically explainable by snow-cover and surface-temperature interactions and should not be shifted by the adjustment procedure. However, the peak is likely overestimated in the simulations. This overestimation could be caused by too many snowfall days as well as the overall cold bias in the region, that may leads to an enhanced snow-cover.

For the temperature dependency of precipitation-frequencies we make three important observations. First, the slopes of the P(I|T) distribution in the models and the observations are opposite within the temperature range ~0 °C to 12 °C. The observations show a statistical increase with temperature, while the models show a decrease. The observed increase may be overestimated, because the observations suffer from undercatch errors that are more pronounced for snowfall (particularly snowfall that falls at low temperatures) than for rain (*Prein and Gobiet* [2016]). Especially for heavy-precipitation however, theoretical explanations from *Trenberth* [2011] suggest, that this relation should be positive. Second, the slope of the P(I|T) curve is not constant over the entire temperature range. A shift in the temperature distribution will therefore have a different impact depending on the temperature ranges that are subject to changes. In the case of a simple shift of the distribution towards higher temperatures, these changes would mainly occur at the tails. For temperature it can be expected, that the location of the tails will be closely related to the location of the mean value. We find, that this statistical effect offers an additional explanation to the observed dependency of snowfall changes on the long term mean-surface-temperature (*Ceppi et al.* [2010]). Other reasons for this dependency are temperature dependent feedback processes like the snow-albedo effect (*Winter et al.* [2016]). A major implication of this finding is, that a model bias in the temperature distribution will impact future changes in snow and precipitation-frequencies. Third, the P(I|T) distribution of the model ensemble shows a kink between 0 to

1 °C that is not existent in the observed distribution. This kink could be caused by parameterisations that describe interactions between liquid- and ice-cloud-content and should be further investigated.

Estimation of snowfall-fractions: To investigate the impact of different snowfall-fraction discrimination methods on the CCS$_h$ of precipitation- and snowfall-frequencies, we use two different methods in this study. We find, that the different methods clearly modify the temperature dependency of snowfall. This modification is strongest for temperatures close to 0 °C were the chance for snowfall and heavy-snowfall events is high. Differences in the snowfall-fraction therefore show a large impact. Depending on the discrimination method, the temperature range with the highest snowfall occurrence shifts from -2 °C using the ensemble snowfall-fraction statistics to 0 °C when the ATT-sheme is applied. The shift towards higher temperatures reduces the projected decrease in snowfall. Comparisons to the HISTALP data set suggest, that all models underestimate snowfall-fractions in a temperature range close to 0 °C. One reason for this finding may be, that the model grid is to coarse to represent the topography. Sub-grid orography information may be a key to correctly represent snowfall-fractions. Increasing the horizontal resolution may increase the snow-fraction at temperatures close to 0 °C. This would shift the peak in the snowfall probability towards higher temperatures and therefore reduces the projected decrease in snowfall-frequency. These findings should be the object of further analysis.

The clear effect the different snowfall-fractions have on the projected changes in snowfall-frequency also highlight the importance of a clear definition when precipitation is counted as snow. For the tourism industry e.g. it could be beneficial if a day where only sleet is falling is not counted as snow day. Excluding sleet would however shift the temperature range where the most snowfall occurs to lower values.

Validation of the past CCS$_h$: Analyzing the climate change signal between the periods 1971-1989 and 1990 -2008 we find, that the different snow discrimination methods, result in a clear spread in the CCS$_h$ of snowfall-frequencies for both intensity thresholds. The spread in the CCS$_h$ between the different estimates, in combination with the use of the two different temperature data sets, is in the same range as the CCS$_h$ itself. The upper bound is always SFE1-EOBS and the lower bound always SFE2-HISTALP (for details see supplementary material). The lack of trustful information on snowfall-fractions as well as the spread between the RCM's leads to uncertainties in the CCS$_h$. Comparing the differences in the CCS$_h$

between models and observations we make the following findings. In the past period the observations show an increase in the precipitation- and heavy-precipitation-frequencies above the model spread at most altitudes. We find that the cold temperature bias is the main cause for the lower precipitation-frequency increase in the simulations. The strongest increase in precipitation is found at days with surface-temperatures above ~10 °C. Because of the cold bias these temperatures are less likely to occur in the simulations. A reason for the enhanced precipitation increase at temperatures above ~10 °C may be an increase in convective events (*Giorgi et al.* [2016]).

The strongest cause for differences between the ensemble members are differences in the CCS_h of the P(I|T) distribution. These differences are likely connected to differences in the large scale circulation. For the CCS_h of snowfall-frequencies the higher decrease found in the observations, is essentially a result of stronger changes in the P(T) and in the P(I|T) distribution. However, model biases in the P(T) and in the P(I|T) distribution reduce the lower decrease in the simulations, as the different shape of the simulated distributions amplifies the decrease in snowfall-frequencies.

Future climate change: Future climate projections (RCP4.5) indicate a mean-surface-temperature increase of 2.5 °C in the alpine region by the end of the century (*Smiatek et al.* [2016]). This increase will have a strong impact on the P(T) distribution. Analyzing the projected changes we find, that the ensemble mean shows an increase of precipitation-frequencies of less than 1 % with some models projecting a decrease. For heavy-precipitation-frequencies we find a clear increase of about 10 % in the ensemble mean. In both cases changes in the temperature distribution cause a decrease. This can be explained by the negative temperature dependency found in the ensemble data, that is likely caused by model biases. Snowfall-frequencies are changing stronger than heavy-snowfall-frequencies. This confirms findings from earlier studies (*O'Gorman* [2014b]; *Frei et al.* [2017]). In contrast to *O'Gorman* [2014b] our findings suggest, that these differences are mainly the result of changes in the P(I|T) distribution. These differences cause an increase for heavy-snowfall-frequencies which counteracts with the decrease caused by changes in the temperature distribution.

Bias-adjustment: After we analyzed the biases in the P(T) and P(I|T) distribution and quantified their impact on the past and future CCS, we investigate the combined effect of biases in both distributions. For this analysis the distributions are bias-adjusted. After the adjustment we find a stronger increase for precipitation and heavy-precipitation-frequencies.

Because of the change in the P(I|T) distribution the temperature increase now causes an increase for precipitation and heavy-precipitation-frequencies. Especially for the heavy-precipitation-frequency the increase with temperature is in line with theoretical concepts *Trenberth* [2011]. For snowfall-frequencies the spread in the CCS_h between the models and the observation increased. The differences are not reduced, because the main cause for the spread is, that the simulations show compared to the observations a weaker decrease in the probability of days close to 0 °C.

In contrast to other studies that compared bias-adjusted and non bias-adjusted data ([*Frei et al.*, 2017]), we find clear changes in the results. Our results indicate that a temperature dependent bias-adjustment is needed. Adjusting only the mean precipitation and temperature fields will not account for precipitation changes, that result because of the temperature adjustment.

In this study we identified important biases in the model ensemble and found that these biases are modifying the simulated past CCSh as well as the projected CCS by the end of the century. We find a temperature dependent precipitation bias and temperature bias, that make it questionable if raw climate model output can be used for the analysis of snowfall. The framework we present helps to detect these biases and to quantify possible impacts. It shows that when thoughtful used, bias-adjustment can be a valid tool to detect important processes and dependencies, that need to be captured in order to project future changes. This makes it interesting for model development. The systematic assessment also enables us to explain the changes caused by the bias-adjustment used in this study. Our results however show clear limitations of bias-adjustment approaches. Feedback processes like the snow-albedo feedback are difficult to correct by statistic approaches. Reducing model biases caused by these feedback mechanism is likely only possible by continuous development of dynamical models.

5 Acknowledgements

References

(2007), *Climate Change in the European Alps*, OECD Publishing, doi:10.1787/ 9789264031692-en.

Allen, M. R., and W. J. Ingram (2002), Constraints on future changes in climate and the hydrologic cycle, *Nature*, *419*(6903), 224–232, doi:10.1038/nature01092.

Beniston, M. (2003), Climatic Change in Mountain Regions: A Review of Possible Impacts, in *Advances in Global Change Research*, pp. 5–31, Springer Netherlands, doi:10.1007/ 978-94-015-1252-7_2.

Beniston, M. (2006), Mountain Weather and Climate: A General Overview and a Focus on Climatic Change in the Alps, *Hydrobiologia*, *562*(1), 3–16, doi:10.1007/ s10750-005-1802-0.

Beniston, M., and M. Rebetez (1996), Regional behavior of minimum temperatures in Switzerland for the period 1979?1993, *Theoretical and Applied Climatology*, *53*(4), 231–243, doi:10.1007/bf00871739.

Beniston, M., B. Uhlmann, S. Goyette, and J. I. Lopez-Moreno (2010), Will snow-abundant winters still exist in the Swiss Alps in an enhanced greenhouse climate?, *International Journal of Climatology*, *31*(9), 1257–1263, doi:10.1002/joc.2151.

Bosshard, T., S. Kotlarski, T. Ewen, and C. Schär (2011), Spectral representation of the annual cycle in the climate change signal, *Hydrology and Earth System Sciences*, *15*(9), 2777–2788, doi:10.5194/hess-15-2777-2011.

Brugnara, Y., M. Brunetti, M. Maugeri, T. Nanni, and C. Simolo (2011), High-resolution analysis of daily precipitation trends in the central Alps over the last century, *International Journal of Climatology*, *32*(9), 1406–1422, doi:10.1002/joc.2363.

Brunetti, M., M. Maugeri, T. Nanni, I. Auer, R. Böhm, and W. Schöner (2006), Precipitation variability and changes in the greater Alpine region over the 1800–2003 period, *Journal of Geophysical Research*, *111*(D11), doi:10.1029/2005jd006674.

Ceppi, P., S. C. Scherrer, A. M. Fischer, and C. Appenzeller (2010), Revisiting Swiss temperature trends 1959-2008, *International Journal of Climatology*, *32*(2), 203–213, doi: 10.1002/joc.2260.

Chimani, B., R. Böhm, C. Matulla, and M. Ganekind (2011), Development of a longterm dataset of solid/liquid precipitation, *Advances in Science and Research*, *6*, 39–43, doi:10. 5194/asr-6-39-2011.

Coppola, E., F. Raffaele, and F. Giorgi (2016), Impact of climate change on snow melt driven runoff timing over the Alpine region, *Climate Dynamics*, doi:10.1007/s00382-016-3331-0.

de Vries, H., G. Lenderink, and E. van Meijgaard (2014), Future snowfall in western and central Europe projected with a high-resolution regional climate model ensemble, *Geophysical Research Letters*, *41*(12), 4294–4299, doi:10.1002/2014gl059724.

Eggert, B., P. Berg, J. O. Haerter, D. Jacob, and C. Moseley (2015), Temporal and spatial scaling impacts on extreme precipitation, *Atmospheric Chemistry and Physics*, *15*(10), 5957–5971, doi:10.5194/acp-15-5957-2015.

Fahrmeir, L., C. Heumann, R. Künstler, I. Pigeot, and G. Tutz (2016), *Statistik*, Springer Berlin Heidelberg, doi:10.1007/978-3-662-50372-0.

Feiccabrino, J., D. Gustafsson, and A. Lundberg (2013), Surface-based precipitation phase determination methods in hydrological models, *Hydrology Research*, *44*(1), 44, doi:10.2166/nh.2012.158.

Foley, A. M. (2010), Uncertainty in regional climate modelling: A review, *Progress in Physical Geography*, *34*(5), 647–670, doi:10.1177/0309133310375654.

Frei, P., S. Kotlarski, M. A. Liniger, and C. Schär (2017), Snowfall in the Alps: Evaluation and projections based on the EURO-CORDEX regional climate models, *The Cryosphere Discussions*, pp. 1–38, doi:10.5194/tc-2017-7.

Giorgi, F., J. W. Hurrell, M. R. Marinucci, and M. Beniston (1997), Elevation Dependency of the Surface Climate Change Signal: A Model Study, *Journal of Climate*, *10*(2), 288–296, doi:10.1175/1520-0442(1997)010<0288:edotsc>2.0.co;2.

Giorgi, F., C. Torma, E. Coppola, N. Ban, C. Schär, and S. Somot (2016), Enhanced summer convective rainfall at Alpine high elevations in response to climate warming, *Nature Geoscience*, *9*(8), 584–589, doi:10.1038/ngeo2761.

Gobiet, A., S. Kotlarski, M. Beniston, G. Heinrich, J. Rajczak, and M. Stoffel (2014), 21st century climate change in the European Alps—A review, *Science of The Total Environment*, *493*, 1138–1151, doi:10.1016/j.scitotenv.2013.07.050.

Hawkins, E., and R. Sutton (2009), The Potential to Narrow Uncertainty in Regional Climate Predictions, *Bulletin of the American Meteorological Society*, *90*(8), 1095–1107, doi:10.1175/2009bams2607.1.

Haylock, M. R., N. Hofstra, A. M. G. K. Tank, E. J. Klok, P. D. Jones, and M. New (2008), A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006, *Journal of Geophysical Research*, *113*(D20), doi:10.1029/2008jd010201.

Isotta, F. A., C. Frei, V. Weilguni, M. P. Tadić, P. Lassègues, B. Rudolf, V. Pavan, C. Cacciamani, G. Antolini, S. M. Ratto, M. Munari, S. Micheletti, V. Bonati, C. Lussana, C. Ronchi, E. Panettieri, G. Marigo, and G. Vertačnik (2013), 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, doi: 10.1002/joc.3794.

Jacob, D., J. Petersen, B. Eggert, A. Alias, O. B. Christensen, L. M. Bouwer, A. Braun, A. Colette, M. Déqué, G. Georgievski, E. Georgopoulou, A. Gobiet, L. Menut, G. Nikulin, A. Haensler, N. Hempelmann, C. Jones, K. Keuler, S. Kovats, N. Kröner, S. Kotlarski, A. Kriegsmann, E. Martin, E. Meijgaard, C. Moseley, S. Pfeifer, S. Preuschmann, C. Radermacher, K. Radtke, D. Rechid, M. Rounsevell, P. Samuelsson, S. Somot, J.-F. F. Soussana, C. Teichmann, R. Valentini, R. Vautard, B. Weber, P. Yiou, E. van Meijgaard, C. Moseley, S. Pfeifer, S. Preuschmann, C. Radermacher, K. Radtke, D. Rechid, M. Rounsevell, P. Samuelsson, S. Somot, J.-F. F. Soussana, C. Teichmann, R. Valentini, R. Vautard, B. Weber, and P. Yiou (2013), EURO-CORDEX: new high-resolution climate change projections for European impact research, *Regional Environmental Change*, *14*(2), 563–578, doi:10.1007/s10113-013-0499-2.

Kotlarski, S., T. Bosshard, D. Lüthi, P. Pall, and C. Schär (2011), Elevation gradients of European climate change in the regional climate model COSMO-CLM, *Climatic Change*, *112*(2), 189–215, doi:10.1007/s10584-011-0195-5.

Kotlarski, S., K. Keuler, O. B. Christensen, A. Colette, M. Déqué, A. Gobiet, K. Goergen, D. Jacob, D. Lüthi, E. van Meijgaard, G. Nikulin, C. Schär, C. Teichmann, R. Vautard, K. Warrach-Sagi, and V. Wulfmeyer (2014), Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble, *Geoscientific Model Development*, *7*(4), 1297–1333, doi:10.5194/gmd-7-1297-2014.

Masson, D., and C. Frei (2014), Spatial analysis of precipitation in a high-mountain region: exploring methods with multi-scale topographic predictors and circulation types, *Hydrology and Earth System Sciences*, *18*(11), 4543–4563, doi:10.5194/hess-18-4543-2014.

Moss, R. H., J. A. Edmonds, K. A. Hibbard, M. R. Manning, S. K. Rose, D. P. van Vuuren, T. R. Carter, S. Emori, M. Kainuma, T. Kram, G. A. Meehl, J. F. B. Mitchell, N. Nakicenovic, K. Riahi, S. J. Smith, R. J. Stouffer, A. M. Thomson, J. P. Weyant, and T. J. Wilbanks (2010), The next generation of scenarios for climate change research and assessment, *Nature*, *463*(7282), 747–756, doi:10.1038/nature08823.

O'Gorman, P. (2014a), Contrasting responses of mean and extreme snowfall to climate change., *Nature*, *512*, 416–8.

O'Gorman, P. A. (2014b), Contrasting responses of mean and extreme snowfall to climate change, *Nature*, *512*(7515), 416–418, doi:10.1038/nature13625.

Piazza, M., J. Boé, L. Terray, C. Pagé, E. Sanchez-Gomez, and M. Déqué (2013), Projected 21st century snowfall changes over the French Alps and related uncertainties, *Climatic Change*, *122*(4), 583–594, doi:10.1007/s10584-013-1017-8.

Prein, A. F., and A. Gobiet (2016), Impacts of uncertainties in European gridded precipitation observations on regional climate analysis, *International Journal of Climatology*, *37*(1), 305–327, doi:10.1002/joc.4706.

Rajczak, J., P. Pall, and C. Schär (2013), Projections of extreme precipitation events in regional climate simulations for Europe and the Alpine Region, *Journal of Geophysical Research: Atmospheres*, *118*(9), 3610–3626, doi:10.1002/jgrd.50297.

Ruiz-Villanueva, V., M. Borga, D. Zoccatelli, L. Marchi, E. Gaume, and U. Ehret (2012), Extreme flood response to short-duration convective rainfall in South-West Germany, *Hydrology and Earth System Sciences*, *16*(5), 1543–1559, doi:10.5194/hess-16-1543-2012.

Schaefli, B., B. Hingray, and A. Musy (2007), Climate change and hydropower production in the Swiss Alps: quantification of potential impacts and related modelling uncertainties, *Hydrology and Earth System Sciences*, *11*(3), 1191–1205, doi:10.5194/hess-11-1191-2007.

Scherrer, S. C., C. Appenzeller, and M. Laternser (2004), Trends in Swiss Alpine snow days: The role of local- and large-scale climate variability, *Geophysical Research Letters*, *31*(13), n/a–n/a, doi:10.1029/2004gl020255.

Schmucki, E., C. Marty, C. Fierz, and M. Lehning (2014), Simulations of 21st century snow response to climate change in Switzerland from a set of RCMs, *International Journal of Climatology*, *35*(11), 3262–3273, doi:10.1002/joc.4205.

Serquet, G., C. Marty, J.-P. Dulex, and M. Rebetez (2011), Seasonal trends and temperature dependence of the snowfall/precipitation-day ratio in Switzerland, *Geophysical Research Letters*, *38*(7), n/a–n/a, doi:10.1029/2011gl046976.

Smiatek, G., H. Kunstmann, and A. Senatore (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, doi:10.1002/2015jd024727.

839   Trenberth, K. (2011), Changes in precipitation with climate change, *Climate Research*, *47*(1),

840   123–138, doi:10.3354/cr00953.

841   Winter, K. J. P. M., S. Kotlarski, S. C. Scherrer, and C. Schär (2016), The Alpine snow-

842   albedo feedback in regional climate models, *Climate Dynamics*, *48*(3-4), 1109–1124, doi:

843   10.1007/s00382-016-3130-7.

Appendices

A. Declarations

Bastian Eggert

Mühlenstraße 1A

21423 Winsen (Luhe)

017634976432

Bastian-e@gmx.net

Hiermit erkläre ich, dass ich mich noch keiner Doktorprüfung unterzogen oder mich um Zulassung zu einer solchen beworben habe.

Ich versichere, dass die Dissertation mit dem Titel "Improvements, Uncertainties and Scale Dependencies of High Resolution Climate Simulations" noch keiner Fachvertreterin bzw. Fachvertreter vorgelegen hat, ich die Dissertation nur in diesem und keinem anderen Promotionsverfahren eingereicht habe und, dass diesem Promotionsverfahren keine endgültig gescheiterten Promotionsverfahren vorausgegangen sind.

Ich versichere, dass ich die eingereichte Dissertation "Improvements, Uncertainties and Scale Dependencies of High Resolution Climate Simulations" selbstständig und ohne unerlaubte Hilfsmittel verfasst habe. Anderer als der von mir angegebenen Hilfsmittel und Schriften habe ich mich nicht bedient. Alle wörtlich oder sinngemäß anderen Schriften entnommenen Stellen habe ich kenntlich gemacht.

Winsen (Luhe)

18.10.2017

Bastian Eggert

B. Authors contribution and articles publication status

Article #	Title	Specific contribution of all authors	Author status	Weighting factor	Puplication status* of contributions
[1]	EURO-CORDEX: new high-resolution climate change projections for European impact research	**BE:** Data evaluation and analysis, development of research methods, discussion of results, comments on the manuscript DJ: Concept, discussion of results, manuscript preparation, development of research methods, JP: Concept, discussion of results, manuscript preparation, development of research methods, AK: Data evaluation and analysis, development of research methods, discussion of results, CR: Data evaluation and analysis, discussion of results, AH, CM, SPf, SPr, DR, BW, LMB, EG, SK, EM, MR, JFS, RV: Discussion of results, development of research methods RVau, KK, SKot, SS: Providing model data, discussion of results, manuscript preparation AA, OBC, AB, AC MD, GG, AG, LM, GN, NH, CJ, NK, AM, AVM, KR, PS, CT, PY: Providing model data, discussion of results	Co-authorship: Important contribution [Wichtiger Anteil]	0.5	**Published** Reg Environ Change, 14, 563–578, doi:10.1007/s10113-013-0499-2 **2013** [IF = 3.15 (5-year)]
[2]	Temporal and spatial scaling impacts on extreme precipitation	**BE:** Concept, development of research methods, data evaluation and analysis, discussion of results, manuscript preparation PB: Concept, discussion of results, development of research methods, comments on the manuscript JOH: Concept, development of research methods, discussion of results, comments on the manuscript DJ: Discussion of results, comments on the manuscript CM: Concept, discussion of results, development of research methods, comments on the manuscript	Co-authorship: Predominant contribution [Überwiegender Anteil]	1.0	**Published** Atmos. Chem. Phys., 15, 5957-5971, doi:10.5194/acp-15-5957-2015, **2015** [IF = 5.90 (5-year)]

Article #	Title	Specific contribution of all authors	Author status	Weighting factor	Puplication status* of contributions
[3]	Statistical precipitation bias correction of gridded model data using point measurements	**BE:** Concept, development of research methods, data evaluation and analysis, discussion of results, comments on the manuscript, JOH: Concept, development of research methods, data evaluation and analysis, discussion of results, manuscript preparation CM: Concept, development of research methods, discussion of results, comments on the manuscript CP: development of research methods, discussion of results, comments on the manuscript PB: Concept, development of research methods, data evaluation and analysis, discussion of results, comments on the manuscript	Co-authorship: Important contribution [Wichtiger Anteil]	0.5	**Published** Geophys. Res. Lett., 42, 1–11, doi:10.1002/ 2015GL063188 **2015** [IF = 4.41 (5-year)]
[4]	Uncertainties in snow and precipitation projections in the northern Alps: the role of model biases	**BE:** Concept, development of research methods, data evaluation and analysis, discussion of results, manuscript preparation DJ:discussion of results, AH: discussion of results	Co-authorship: Predominant contribution [Überwiegender Anteil]	1.0	**Submittet** J. Geophys. Res. Atmos.

Acronyms of the individual authors:

AA: Antoinette Alias; OBC: Ole Bøssing Christensen; LMB: Laurens M. Bouwer; PB: Peter Berg; AB: Alain Braun; AC: Augustin Colette; MD: Michel Déqué; **BE** : Bastian Eggert; GG: Goran Georgievski; EG: Elena Georgopoulou; AG: Andreas Gobiet; LM: Laurent Menut; GN: Grigory Nikulin; AH: Andreas Haensler; JOH: Jan O. Haerter; NH: Nils Hempelmann; CJ: Colin Jones; DJ: Daniela Jacob; KK: Klaus Keuler; SK: Sari Kovats; NK: Nico Kröner; SKot: Sven Kotlarski; AK: Arne Kriegsmann; EM: Eric Martin; EVM: Erik van Meijgaard; CM: Christopher Moseley SPf: Susanne Pfeifer; CP: Claudio Piani; SPr: Swantje Preuschmann; JP: Juliane Petersen; CR: Christine Radermacher; KR: Kai Radtke; DR: Diana Rechid; MR: Mark Rounsevell; PS: Patrick Samuelsson; SS: Samuel Somot; JFS: Jean-Francois Soussana; CT: Claas Teichmann; RV: Riccardo Valentini; RVau: Robert Vautard; BW: Björn Weber; PY: Pascal Yiou

Ich versichere, dass alle in diesem Anhang gemachten Angaben jeweils einzeln und insgesamt vollständig der Wahrheit entsprechen

Bastian Eggert

d

C. Curriculum Vitae

Contact

Bastian Eggert
Climate Service Center Germany (GERICS)
Helmholtz-Zentrum Geesthacht
Fischertwiete 1
20095 Hamburg
Tel: +49 (0) 40 226 338 434
Fax: +49 (0) 40 226 338 163

Work experience and education

05/2012 to 11/2017	Scientific associate Climate Service Center Germany (GERICS), Helmholtz-Zentrum Geesthacht, Fischertwiete 1, 20095 Hamburg, Germany
09/2011 to 04/2012	Scientific associate Max Planck Institute for Meteorology, Bundesstraße 55, 20146 Hamburg, Germany
2011	Graduated from the University of Hamburg, Germany in Meteorology (Diploma)
08/2010 to 08/2011	Diploma student Climate Service Center Germany (GERICS), Helmholtz-Zentrum Geesthacht, Fischertwiete 1, 20095 Hamburg, Germany
12/2008 to 09/2010	Student research assistant Max Planck Institute for Meteorology, Bundesstraße 55, 20146 Hamburg, Germany

Research Interests

Temporal and spatial scale dependencies of climate models
Uncertainties of precipitation projections
Ensemble analysis of climate models
Physical parameterizations in climate models

Research projects and initiatives involved

Since 2012	**CORDEX** -- COordinated Regional climate Downscaling Experiment Contribution: Regional climate change assessment
2012 to 2014	**ECLISE** -- Enabling CLimate Information Services for Europe -- collaborative FP7 research project under the Environment Programme of the European Commission Contribution: Convection-resolving (2 km) Regional Climate simulations for application to selected ECLISE case studies.
2012 to 2013	**Congo Project** -- Climate Change Scenarios for the River Congo Basin – Funded by GIZ and BMU Contribution: Regional climate change assessment; Capacity building activities
2011 to 2012	**GENUS** -- **G**eochemistry and **E**cology of the **N**amibian **U**pwelling **S**ystem Contribution: Regional climate change assessment

Publications
Journal articles

Eggert, B., Berg, P., Haerter, J. O., Jacob, D., and Moseley, C.: Temporal and spatial scaling impacts on extreme precipitation, Atmos. Chem. Phys., 15, 5957-5971, doi:10.5194/acp-15-5957-2015, 2015.

Haerter, J. O., **B. Eggert**, C. Moseley, C. Piani and P. Berg (2015), Statistical precipitation bias correction of gridded model data using point measurements, Geophys. Res. Lett., 42, 1–11, doi:10.1002/2015GL063188.

Jacob D, Petersen J, **Eggert B**, Alias A, Christensen OB, Bouwer LM, Braun A et al. (2013) EURO-CORDEX: new high-resolution climate change projections for European impact research. Reg Environ Change. doi:10.1007/s10113-013-0499-2

Teichmann C, **Eggert B**, Elizalde A, Haensler A, Jacob D, Kumar P, Moseley C,

Pfeifer S, Rechid D, Remedio AR, Ries H, Petersen J, Preuschmann S, Raub T, Saeed F, Sieck K, Weber T. How Does a Regional Climate Model Modify the Projected Climate Change Signal of the Driving GCM: A Study over Different CORDEX Regions Using REMO. *Atmosphere*. 2013; 4(2):214-236.

Non peer-reviewed:

Weber, T., A. Kriegsmann, **B. Eggert** and D. Jacob (2015): Analysis of Climate Change Projections for the Okavango River Basin. Published in the Okavango Basin Information System (OBIS)

Stefan Sobolowski (UNI), Petter Lind (SMHI), Aristeidis Koutroulis (TUC), Lennart Marien (CSC), Erik Kjellström (SMHI), Bert van Ulft (KNMI), **Bastian Eggert** (CSC), Youmin Chen (UNI), Geert Lenderink (KNMI) and David Lindstedt (SMHI) and Ioannis Tsanis (TUC) (2013), Simulating extreme precipitation in the island of Crete with non- hydrostatic high-resolution RCMs. ECLISE Project Report 2.4, http://www.eclise- project.eu

Conferences

Poster presentation at the EGU General Assembly 2015, Vienna, Austria, 12-17 April, 2015. **Eggert B**, Berg P, Haerter J, Moseley M, and Jacob D. Finding pairs of optimal temporal and spatial resolution for precipitation analyses

Presentation at the International REKLIM conference: "Our Climate – Our Future, Regional perspectives on a global challenge", Berlin 6-9 October 2014. **Eggert, B.**, P. Berg, J. O. Haerter, C. Moseley, D. Jacob: Impact of temporal and spatial resolution on extreme event statistics and its implications on climate modelling.

Poster presentation at the 3rd Lund Regional-scale Climate Modelling Workshop. Lund, Sweden, 16-19 June 2014. **Eggert B**, Marien L., Jacob D. High-resolution sensitivity studies over Crete with the non-hydrostatic regional climate model REMO-NH

Poster presentation at the EGU General Assembly 2013, Vienna, Austria, 07-12 April. **Eggert B**, Berg P, Haerter J, Moseley M, and Jacob D. Impact of temporal and spatial resolution on extreme rainfall event statistics

Poster presentation, International Conference on Regional Climate CORDEX 2013; 4-7 November 2013 Brussels, Belgium. **Eggert B**, Berg P, Haerter J, Moseley M, and Jacob D. Impact of temporal and spatial resolution on extrem rainfall event statistics

Diploma Thesis

Eggert B (2011): "Auswirkungen der Oberflächeneigenschaften in REMO auf die Simulation der unteren Atmosphäre", University of Hamburg

Lectures

Workshop: Regional climate projections and accosiated uncertainties. 5 day cours, BMBF project SASSCAL in Lusaka, Sambia, Mai 2016

Handling and Processing of Climate Model raw Data: 5 day course at a capacity building workshop in Yaoundé, Cameroon on behalf of GIZ. Feb 2013.

Data Processing and Visualization with Python: 2 day course at a capacity building workshop in Douala, Cameroon on behalf of GIZ. Nov 2012.

Awards

Best poster presentation, International Conference on Regional Climate CORDEX 2013; 4-7 November 2013 Brussels, Belgium

h

D. Supplementary material

D.1 Article 1: EURO-CORDEX: new high-resolution climate change projections for European impact research

	REMO	RCA4	CCLM	RACMO2	WRF	ALADIN5.1	HIRHAM
Institution	CSC	SMHI	CLMCOM[1]	KNMI	IPSL-INERIS	Météo-France	DMI
Grid resolution	0.11 ° x0.11°	0.11 ° x0.11°	0.11 ° x0.11°	0.11 ° x0.11°	0.11 ° x0.11°	0.11 ° x0.11° [2]	0.11 ° x0.11°
Grid (lat*lon)	433*433	438* 456	450*438	444*456	442 * 454	432*288 (total grid number)	452*432
Rotation	lon -162° lat 39.25	lon -162° lat 39.25°	lon -162° lat 39.25°	lon -162° lat 39.25°	CORDEX specifications	0°E, 90°S	lon -162° lat 39.25°
Vertical levels	27	40	40	40	32	31	31
Boundary layer scheme	Louis 1979	Cuxart et al 2000	Louis 1979	Lenderink and Holtslag 2004; Siebesma et al. 2007	YSU, Hong et al. 2006	Ricard and Royer 1993	Louis 1979
Number of points (sponge zone)	8	10	12	8 (16)[3]		8	
Convection	Mass flux Tiedtke 1989; Nordeng 1994 for CAPE closure; Pfeifer 2006	Kain and Fritsch 1990, 1993; Kain 2004; Jones and Sanchez 2002	Tiedtke 1989	Tiedtke 1989; Nordeng 1994; Neggers et al 2009	Grell and Devenyi 2002	Mass flux Bougeault 1985	Tiedtke 1989
Microphysics	Lohmann and Roeckner 1996	Rasch and Kristjánsson 1998	Doms et al. 2007; Baldauf and Schulz 2004	Tiedtke 1993; Tompkins et al, 2007; ECMWF-IFS 2007;Neggers 2009	Hong et al. 2004	Ricard and Royer 1993	Lohmann and Roeckner 1996
Radiation	Morcrette 1986; Giorgetta and Wild 1995	Savijärvi 1990; Sass et al. 1994	Ritter and Geleyn 1992	Fouquart and Bonnel 1980; Mlawer et al.1997	RRTMG, Lacono et al, 2008	Morcrette 1990	Morcrette et al 1986; Giorgetta and Wild 1995
Land surface scheme	Hagemann 2002; Rechid et al. 2009	Samuelsson et al. 2006	TERRA-ML Doms et al. 2007	Van den Hurk et al 2000; Balsamo et al. 2009	NOAH	Douville et al 2000	Hagemann 2002
Soil thermal layers	5	5	10	4	4	4	5
Soil moisture layers	1	3	8	4	4	2	1
Main references	Jacob et al. 2012	Samuelsson et al. 2011; Kupiainen et al. 2011	Rockel et al. 2008; http://www.cosmo-model.or	Meijgaard van et al. 2012	Version 3.3.1, Skamarock et al. 2008	Colin et al. 2010; Herrmann et al. 2011	Christensen et al. 1998

[1] The Climate Limited-area Modeling Community (http://www.clm-community.eu)
[2] Simulation on MED-CORDEX domain
[3] 8 for temperature, humidity, surface pressure, 16 for horizontal wind components

Table s1. Summary of grid configurations and parameterisations for models used in the present study

References

Balsamo G, Viterbo P, Beljaars A, van den Hurk BJJM, Hirschi M, Betts A, Scipal K (2009) A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System. J Hydrometeor 10: 623-643

Baldauf M, Schulz JP (2004) Prognostic precipitation in the Lokal-Modell (LM) of DWD. COSMO Newsletter 4: 177–180

Bougeault P (1985) A simple parameterization of the large-scale effects of cumulus convection. Mon. Weather Rev 113: 2108- 2121

Christensen OB, Christensen JH, Machenhauer B, Botzet, M (1998): Very high-resolution regional climate simulations over Scandinavia - Present climate. J Climate 11:3204-3229

Colin J, Déqué M, Radu R, Somot S (2010) Sensitivity study of heavy precipitations in Limited Area Model climate simulation: influence of the size of the domain and the use of the spectral nudging technique. Tellus-A, DOI: 10.1111/j.1600-0870.2010.00467.x (published on-line, May 2010)

Cuxart J, Bougeault P, Redelsperger J-L (2000) A turbulence scheme allowing for mesoscale and large-eddy simulations. Quart J Roy Meteor Soc 126: 1–30

Doms G, Förstner J, Heise E, Herzog HJ, Raschendorfer M, Schrodin R, Reinhardt T, Vogel G (2007): A description of the non-hydrostatic regional model LM. Part II: Physical parameterization. [Available online at http://www.cosmomodel.org/content/model/documentation/core/cosmoPhysParamtr.pdf].

Douville H, Planton S, Royer JF, Stephenson DB, Tyteca S, Kergoat L, Lafont S, Betts RA (2000) The importance of vegetation feedbacks in doubled-CO2 time-slice experiments. J Geophys Res 105: 14841-14861

ECMWF-IFS 2007 IFS documentation-Cy31r1. PART IV: Physical Processes. http://www.ecmwf.int/research/ifsdocs/CY31r1/PHYSICS/IFSPart4.pdf. Assessed 28 September 2011

Fouquart Y, Bonnel B (1980) Computations of solar heating of the earth's atmosphere : A new parameterization. Beitr Phys Atmos 53: 35-62

Giorgetta M, Wild M (1995) The water vapour continuum and its representation in ECHAM4. Max Planck Institut für Meteorologie Report 162: 38

Grell GA, Devenyi D (2002) A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. Geophysical Research Letters 29

Hagemann S (2002) An improved land surface parameter dataset for global and regional climate models. Max Planck Institut für Meteorologie Report 336

Herrmann M, Somot S, Calmanti S, Dubois C, Sevault F (2011) Representation of daily wind speed spatial and temporal variability and intense wind events over the Mediterranean Sea using dynamical downscaling: impact of the regional climate model configuration. Nat Hazards Earth Syst Sci 11: 1983-2001. doi:10.5194/nhess-11-1983-2011

Hong S-Y, Lim J-O J (2006) The WRF single-moment 6-class microphysics scheme (WSM6). J Korean Meteor Soc 42: 129-151

Hong S-Y, Noh Y, Dudhia J (2006) A new vertical diffusion package with an explicit treatment of entrainment processes. Mon Wea Rev 134: 2318–2341

Hong S-Y, Dudhia J, Chen S-H (2004) A revised approach to microphysical processes for the bulk parameterization of cloud and precipitation Mon Wea Rev 132: 103-120

Jacob D, Elizalde A, Haensler A, Hagemann S, Kumar P, Podzun R, Rechid D, Remedio AR, Saeed F, Sieck K, Teichmann C, Wilhelm C (2012) Assessing the transferability of the regional climate model REMO to different coordinated regional climate downscaling experiment (CORDEX) regions. Atmosphere 3: 181-199. doi:10.3390/atmos3010181.

Jones CG, Sanchez E (2002) The representation of shallow cumulus convection and associated cloud fields in the Rossby Centre atmospheric model. HIRLAM Newsletter 41

Kain JS (2004) The Kain–Fritsch Convective Parameterization: An Update. J Appl Meteor 43: 170–181

Kain JS, Fritsch JM (1993) Convective parameterizations for mesoscale models: The Kain-Fritsch scheme. In: Emanuel KA, Raymond DJ (ed) The representation of cumulus convection in numerical models. American Meteorological Society, Monograph, Boston, USA

Kain JS, Fritsch JM (1990) A one-dimensional entraining/detraining plume model and its application in convective parameterization. J Atmos Sci 47: 2784–2802

Kupiainen M, Samuelsson P, Jones C, Jansson C, Willén U, Hansson U, Ullerstig A, Wang S, Döscher R (2011) Rossby Centre regional atmospheric model, RCA4. Rossby Centre Newsletter

Iacono MJ, Delamere JS, Mlawer EJ, Shephard MW, Clough SA, Collins WD (2008) Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. Journal of Geophysical Research 113

Lenderink G, Holtslag AAM (2004) An updated length-scale formulation for turbulent mixing in clear and cloudy boundary layers. QJR Meteorol Soc 130: 3405–3427

Lohmann U, Roeckner E (1996) Design and performance of a new cloud microphysics scheme developed for the ECHAM4 general circulation model. Clim Dyn 12: 557–572

Louis JF (1979) A parametric model of vertical eddy fluxes in the atmosphere. Bound Layer Meteorol 17: 187–202

Meijgaard E van, Van Ulft LH, Lenderink G, de Roode SR, Wipfler L, Boers R, Timmermans RMA (2012) Refinement and application of a regional atmospheric model for climate scenario calculations of Western Europe. Climate changes Spatial Planning publication: KvR 054/12, ISBN/EAN 978-90-8815-046-3, pp 44

Mlawer EJ, Taubman S J, Brown PD, Iacono MJ, Clough SA (1997) Radiative transfer for inhomogeneous atmospheres : RRTM, a validated correlated-k model for the longwave. J Geophys Res 102: 16663-16682. doi: 10.1029/97JD00237

Morcrette JJ (1990) Impact of changes to the radiation transfer parameterizations plus cloud optical properties in the ECMWF model. Mon Weather Rev 118: 847-873

Morcrette JJ (1989) Description of the radiation scheme in the ECMWF model. ECMWF technical memorandum 165

Morcrette J-J, Smith L, Fouquart Y (1986) Pressure and temperature dependence of the absorption in longwave radiation parameterizations. Contributions to Atmospheric Physics 59: 455-469

Neggers RAJ (2009) A dual mass flux framework for boundary layer convection. Part II: Clouds. J Atmos Sci 66: 1489-1506. doi:10.1175/2008JAS2636.1

Neggers RAJ , Koehler M, Beljaars ACM (2009) A dual mass flux framework for boundary layer convection. Part I: Transport. J Atmos Sci 66: 1465-1487. doi:10.1175/2008JAS2635.1

Nordeng TE (1994) Extended versions of the convective parametrization scheme at ECMWF and their impact on the mean and transient activity of the model in the tropics. Technical memorandum 206: 41

Pfeifer S (2006) Modeling cold cloud processes with the regional climate model remo. Reports on Earth System Science 23

Rasch PJ, Kristjánsson JE (1998) A comparison of the CCM3 model climate using diagnosed and predicted condensate parameterizations. J Climate 11: 1587–1614

Rechid D, Raddatz T J, Jacob D (2009) Parameterization of snow-free land surface albedo as a function of vegetation phenology based on MODIS data and applied in climate modeling. Theor Appl Climatol 95: 245–255

Ricard JL, Royer JF(1993) A statistical cloud scheme for use in an AGCM. Ann Geophysicae 11: 1095-1115

Ritter B, Geleyn J-F (1992) A comprehensive radiation scheme of numerical weather prediction with potential application to climate simulations. Mon Wea Rev, 120: 303-325

Rockel B, Will A, Hense A (ed) (2008) Special issue Regional climate modelling with COSMO-CLM (CCLM). Meteorologische Zeitschrift 17

Samuelsson P, Jones C, Willén U, Ullerstig A, Gollvik S, Hansson U, Jansson C, Kjellström E, Nikulin G and Wyser K (2011) The Rossby Centre Regional Climate Model RCA3: model description and performance. Tellus 63A. doi: 10.1111/j.1600-0870.2010.00478.x

Samuelsson P, Gollvik S, Ullerstig A (2006) The land-surface scheme of the Rossby Centre regional atmospheric climate model (RCA3). Report in Meteorology 122

Sass BH, Rontu L, Savijärvi H and Räisänen P (1994) HIRLAM-2 Radiation scheme: Documentation and tests. Hirlam technical report No 16

Savijärvi H (1990) A fast radiation scheme for mesoscale model and short-range forecast models. J Appl Met 29: 437-447

Siebesma AP, Soares PMM, Teixeira J (2007) A Combined Eddy-Diffusivity Mass-Flux Approach for the Convective Boundary Layer. J Atmos Sci 64: 1230-1248. doi:10.1175/JAS3888.1

Skamarock WC, Klemp JB, Dudhia J, Gill DO, Duda DMBMG, Huang X-Y, Wang W, Powers JG (2008) A description of the advanced research WRF version 3. NCAR Technical note 475

Tompkins A, Gierens K, Rädel G (2007) Ice supersaturation in the ECMWF Integrated Forecast System. QJR Meteorol Soc 133: 53-63

Tiedtke M (1993) Representation of clouds in large-scale models. Mon Wea Rev 121: 3040-3061

Tiedtke M (1989) A comprehensive mass flux scheme for cumulus parameterization in large-scale models. Mon. Weather Rev. 117: 1779- 1800

Van den Hurk BJJM, Viterbo P, Beljaars ACM, Betts AK (2000) Offline validation of the ERA40 surface scheme. ECMWF Tech Report 75

GCM	GCM Member	RCM	Scenarios	Time
MPI-ESM-LR	r1i1p1	CCLM	RCP4.5, RCP8.5	until 2100
CNRM-CM5-LR	r1i1p1	CCLM	RCP4.5	until 2100
EC-EARTH	r12i1p1	CCLM	RCP4.5	until 2100
HadGEM2-ES	r1i1p1	CCLM	RCP4.5	until 2100
CNRM-CM5-LR	r8i1p1	ALADIN V5.2	RCP4.5, RCP8.5	until 2100
MPI-ESM-LR	r1i1p1	REMO	RCP4.5, RCP8.5	until 2100
IPSL-CM5A-MR	r1i1p1	WRF331	RCP4.5	until 2100
EC-EARTH	r1i1p1	RACMO2	RCP4.5, RCP8.5	until 2100
EC-EARTH	r12i1p1	RCA4	RCP8.5	until 2100
CNRM-CM5-LR	r1i1p1	RCA4	RCP8.5	until 2100
HadGEM2-ES	r1i1p1	RCA4	RCP8.5	until 2100
MPI-ESM-LR	r1i1p1	RCA4	RCP8.5	until 2100
IPSL-CM5A-MR	r1i1p1	RCA4	RCP8.5	until 2100
EC-EARTH	r3i1p1	DMI-HIRHAM	RCP4.5 RCP8.5	until 2050

Table s2a: Overview of the global and regional climate models for RCP4.5 and RCP8.5 used in the present study

GCM	RCM	Resolution	Time
HadCM3Q16	C4IRCA3*	25 km	until 2100
ARPEGE	CNRM-RM5.1	25 km	until 2100
ARPEGE	HIRHAM5*	25 km	until 2100
ECHAM5	HIRHAM5	25 km	until 2100
BCM	HIRHAM5	25 km	until 2100
HadCM3Q0	CLM*	25 km	until 2100
ECHAM5	REGCM3	25 km	until 2100
ECHAM5	RACMO2	25 km	until 2100
HadCM3Q0	HadRM3Q0*	25 km	until 2100
HadCM3Q16	HadRM3Q16	25 km	until 2100
HadCM3Q3	HadRM3Q3	25 km	until 2100
ECHAM5	REMO*	25 km	until 2100
IPSL	REMO	25 km	until 2100
BCM	RCA*	25 km	until 2100
ECHAM5	RCA*	25 km	until 2100
HadCM3Q3	RCA	25 km	until 2100
ECHAM5-r1	RACMO2	50 km	until 2100
ECHAM5-r2	RACMO2*	50 km	until 2100
MIROC3.2	RACMO2*	50 km	until 2100
ECHAM5	RCA	50 km	until 2100

Table s2b: Overview of the global and regional climate models for emission scenario A1B used in the present study. The RCMs indicated with * are used for the calculation of parameters based on 9 simulations
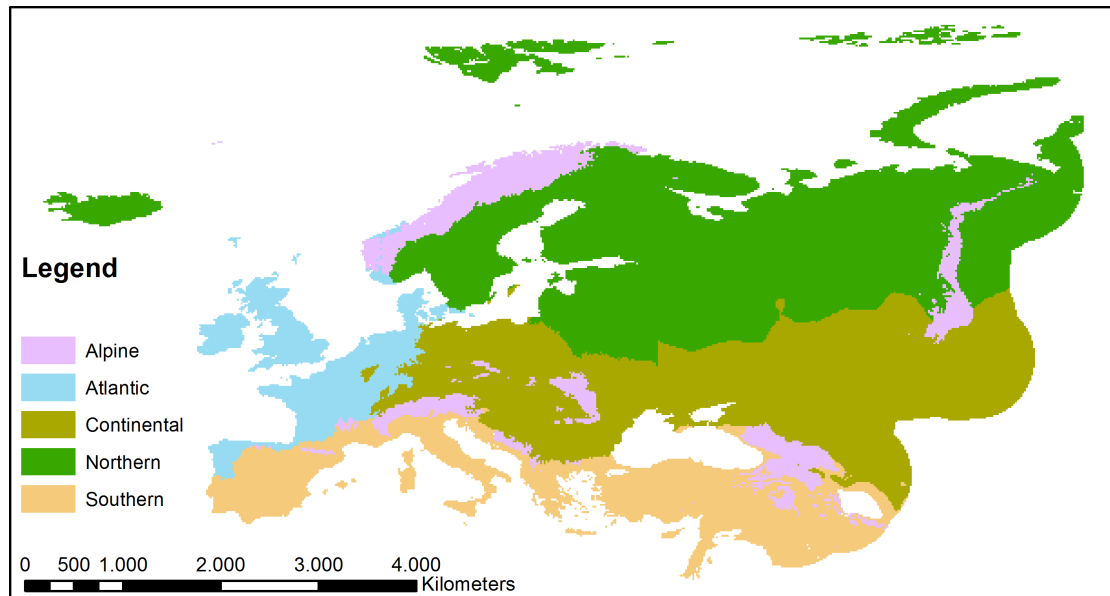
Figure s1: Aggregated IPCC zones for Europe based on the classification of Metzger et al. (2005). The zones are used to calculate area means of several climate impact indices and their potential change.

Reference

Metzger M J, Bunce R G H, Jongman R H G, Muecher C A and Watkins J W (2005) A climatic stratification of the environment of Europe. Global Ecology and Biogeography 14:549-563

Seasonal mean temperature and precipitation climate change signals
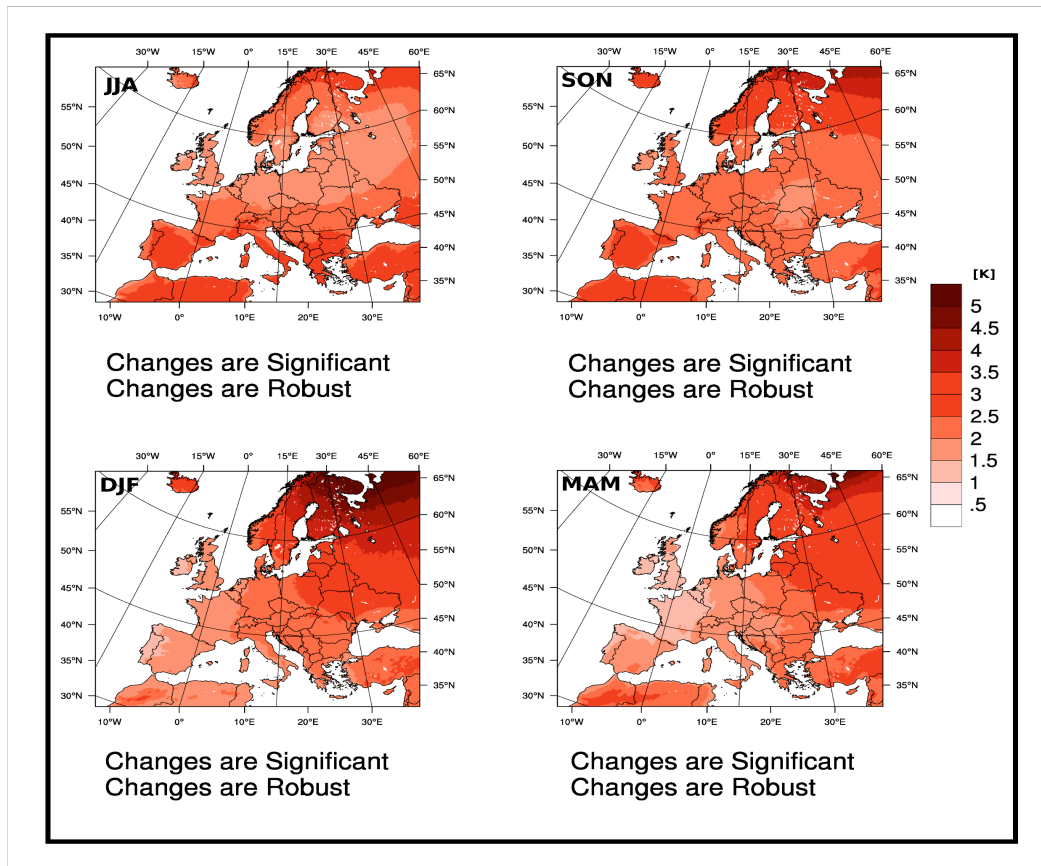
Figure s2: Projected seasonal changes of temperature [K] based on the RCP4.5 scenario for the period 2071-2100 compared to 1971-2000. Changes are robust and significant across the entire European continent.
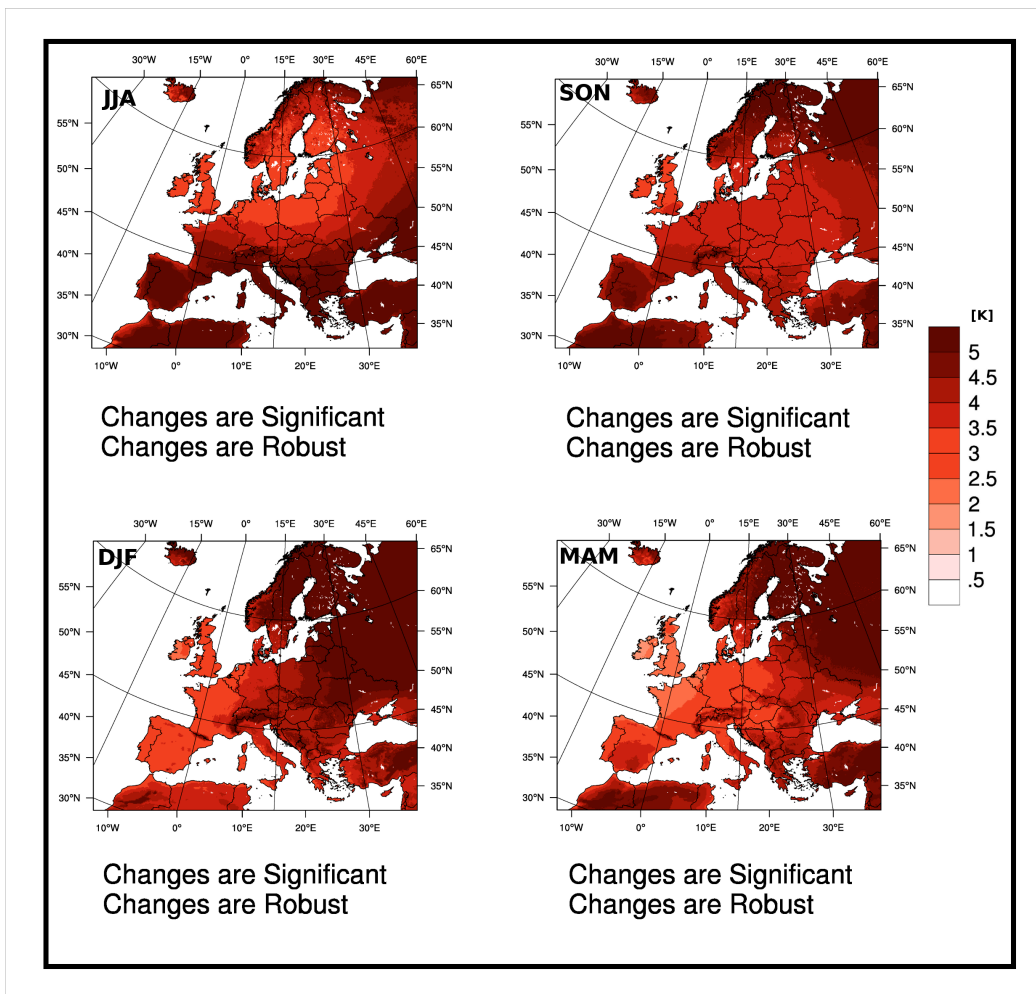
Figure s3: Projected seasonal changes of temperature [K] based on the RCP8.5 scenario for the period 2071-2100 compared to 1971-2000. Changes are robust and significant across the entire European continent.
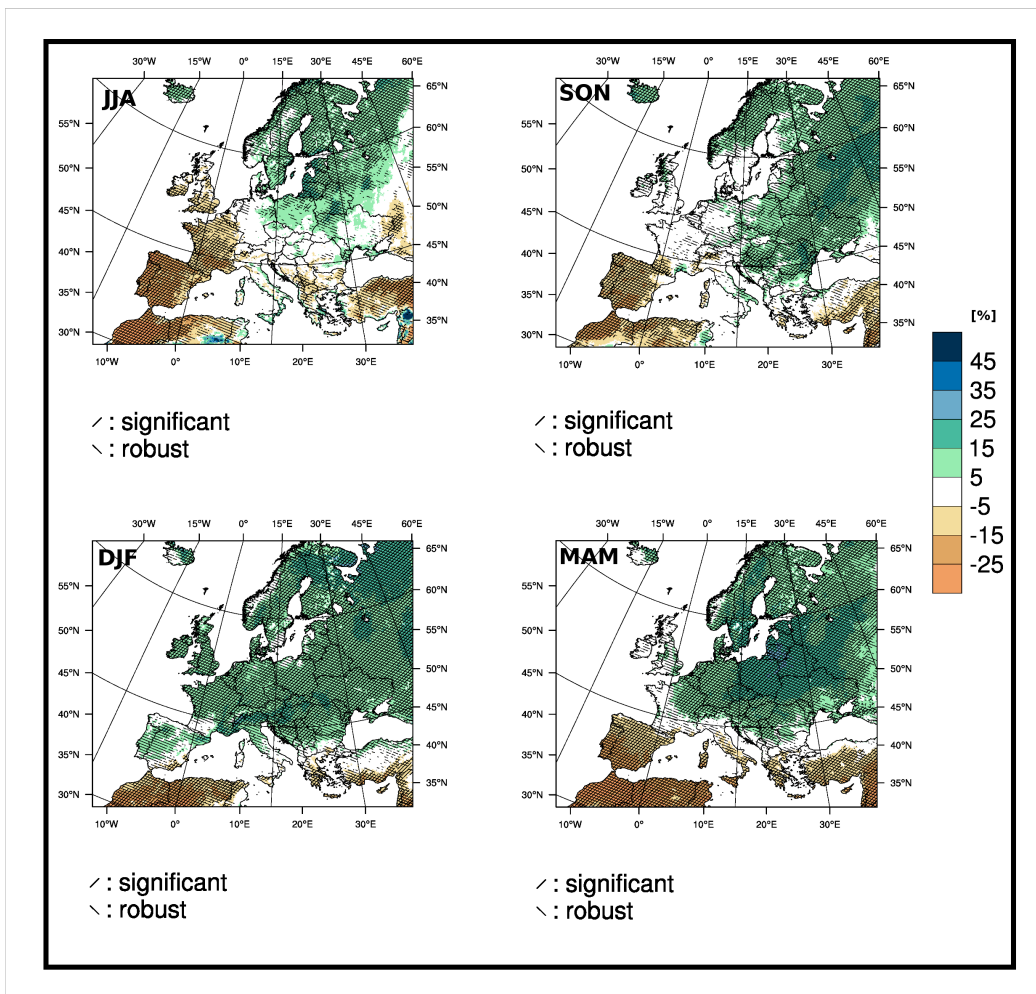
Figure s4: Projected seasonal changes of precipitation [%] based on the RCP4.5 scenario for the period 2071-2100 compared to 1971-2000. Hatched areas indicate regions with robust and/or statistical significant change.
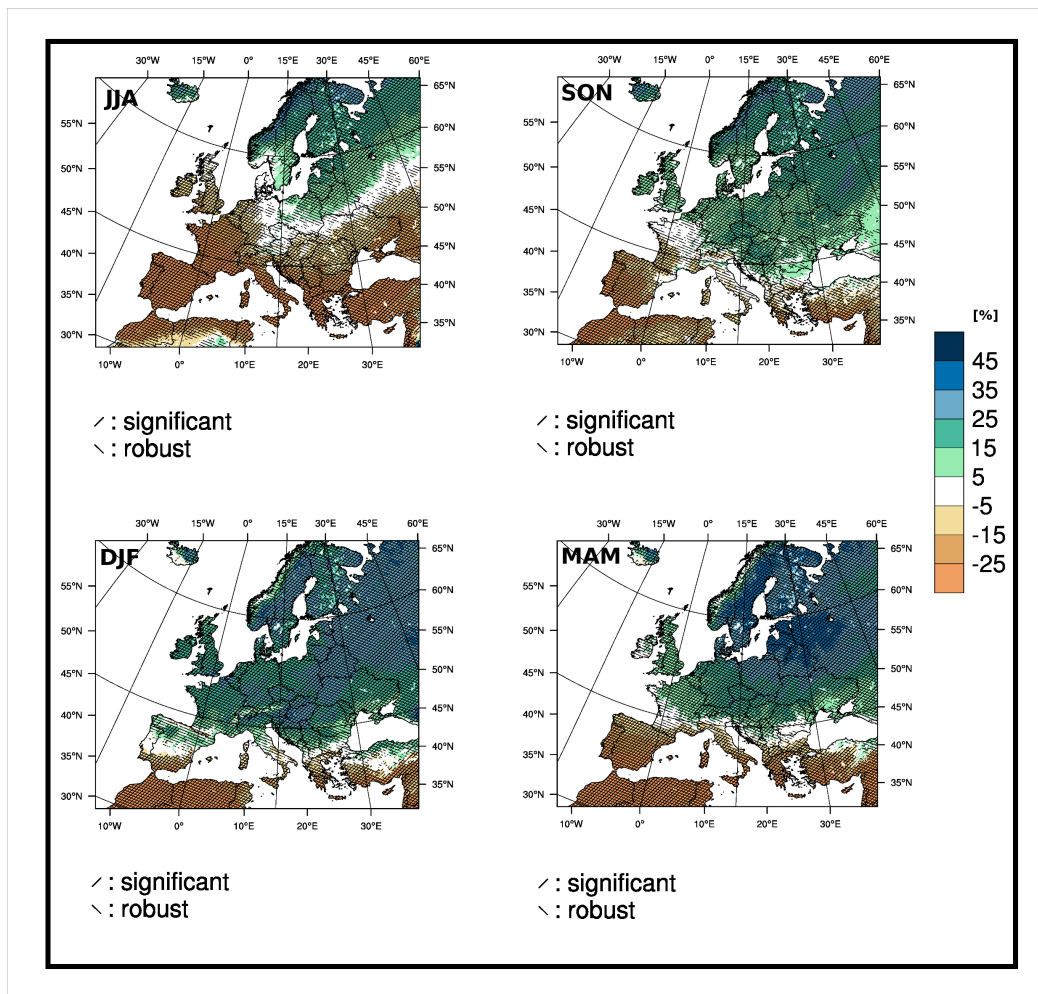
Figure s5: Projected seasonal changes of precipitation [%] based on the RCP8.5 scenario for the period 2071-2100 compared to 1971-2000. Hatched areas indicate regions with robust and/or statistical significant change.

Figures s2 to s5 show the seasonal climate change signals of mean temperature and precipitation for the end of the 21st century, in the ensemble mean for the scenarios RCP4.5 and RCP8.5. Temperature change signals are significant and robust in all seasons, but they show more regional heterogeneity than the yearly mean. The horizontal patterns look very similar for the two scenarios, with a smaller magnitude for RCP4.5. In winter and spring, there is a less pronounced warming in Western Europe compared to the yearly mean, with a strong increasing gradient towards the north east. In contrast, in summer, and to a less extent in autumn, there is a north-south gradient with the strongest warming in the Mediterranean and Black Sea region. However, a particularly strong warming in the Sub-Arctic region (over 5 °C for RCP8.5 and over 3.5 °C for RCP4.5) is seen throughout the year.

Same as for the yearly means, seasonal precipitation changes show a separating band between significant increases in the north and significant decreases in the south. However, this separation band shifts seasonally. In winter, only the most southern parts of Spain, Italy, Greece, and Turkey experience a decrease, while there is an increase for the rest of Europe. The separation band moves

northward in spring, until in summer an increase is only seen in Scandinavia and North Eastern Europe, while there is a drying in Western, South-Eastern, and parts of Central Europe. In autumn, the band moves southward again. Same as for temperature, the precipitation change patterns are very similar for both scenarios, but for RCP8.5 they are larger in magnitude, and the separating band with insignificant changes is narrower.
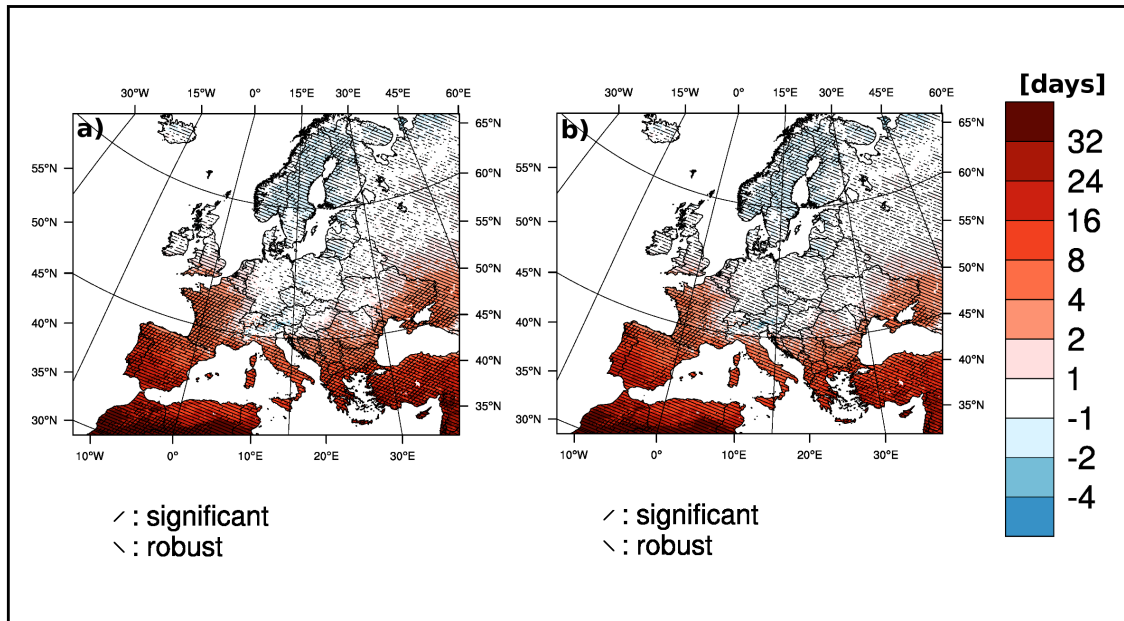
Figure s6: Sensitivity test for the change in the 95th percentile of the length of dry spells [days] for the RCP8.5 ensemble for the period 2071-2100 compared to 1971-2000. a) Ensemble mean without the RCM Aladin simulation for Med-CORDEX region. b) Ensemble mean including the Aladin simulation. Hatched areas indicate regions with robust and/or statistical significant change.
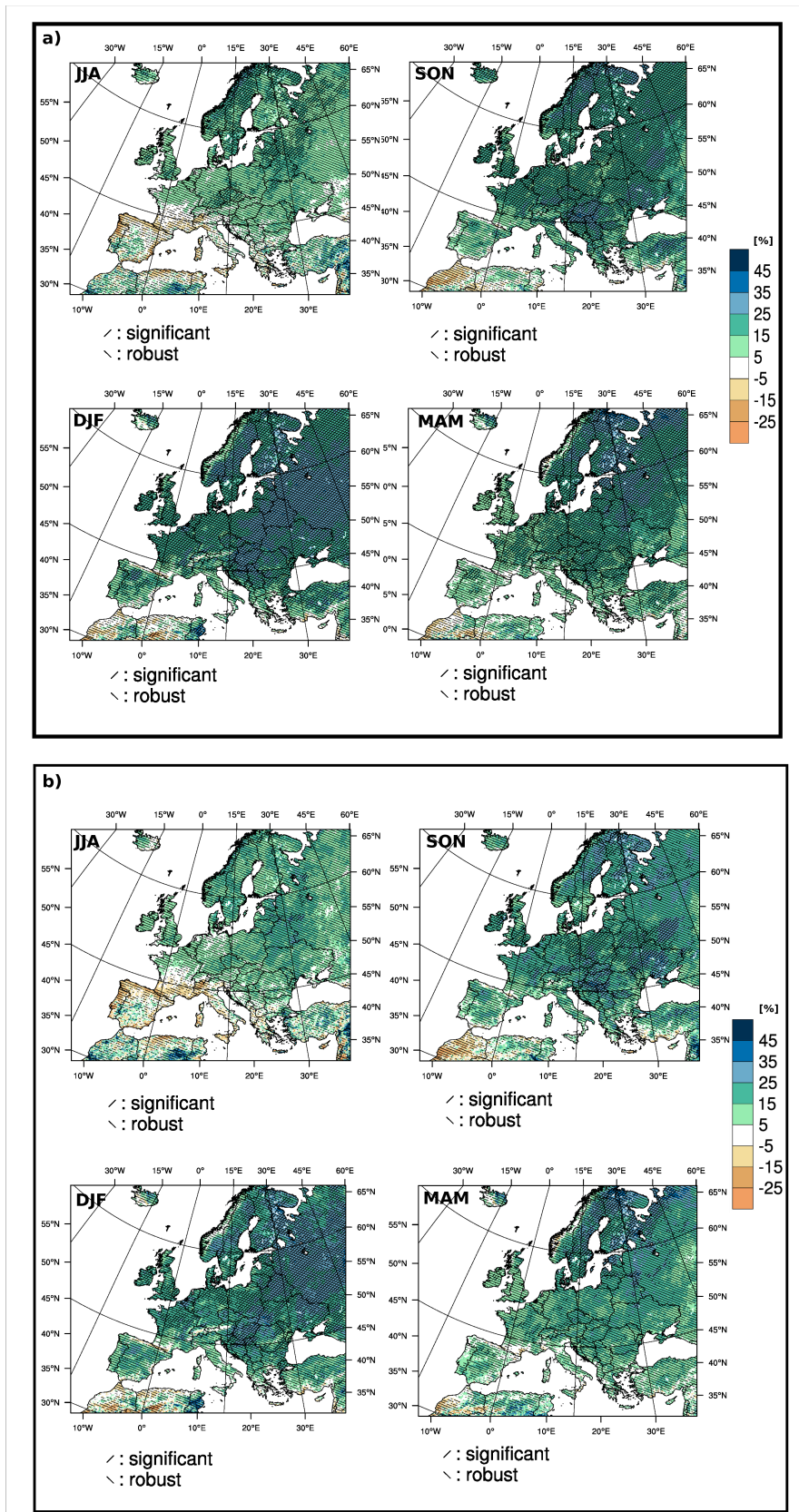
Figure s7: Sensitivity test for projected seasonal changes of heavy precipitation [%] based on the

RCP8.5 scenario for the period 2071-2100 compared to 1971-2000. a) Full ensemble (see Supplementary Material Table S2) b) small ensemble (GCM: MPI-ESM-LR, RCM: CCLM; GCM: CNRM-CM5-LR, RCM: ALADIN V5.2; GCM: MPI-ESM-LR, RCM: REMO; GCM: EC-EARTH, RCM: RACMO2; GCM: IPSL-CM5A-MR, RCM: RCA4). Hatched areas indicate regions with robust and/or statistical significant change.

Figure s6 shows the sensitivity of the change signal of dry spells to the changes in the size of the ensemble, for the end of the 21st century under RCP8.5: If the simulation by ALADIN, which, in contrast to the other simulations, covers only the Mediterranean (Med-CORDEX http://www.medcordex.eu/) region, is removed from the ensemble, only small changes in robustness in Central Europe can be seen. The magnitude of the changes and significance are hardly affected. Figure s7 shows a similar sensitivity test for the change in heavy precipitation on a seasonal basis, for RCP8.5. For the end of the 21st century, the result from the full ensemble of 9 simulations is compared with the results from a reduced ensemble of 5 simulations. Here, the robustness and, again, the magnitude of the changes is hardly affected, only the significance is weaker for the smaller ensemble.

D.2 Article 2: Temporal and spatial scaling impacts on extreme precipitation

Supplement of

Temporal and spatial scaling impacts on extreme precipitation

B. Eggert et al.

Correspondence to: B. Eggert (bastian.eggert@hzg.de)
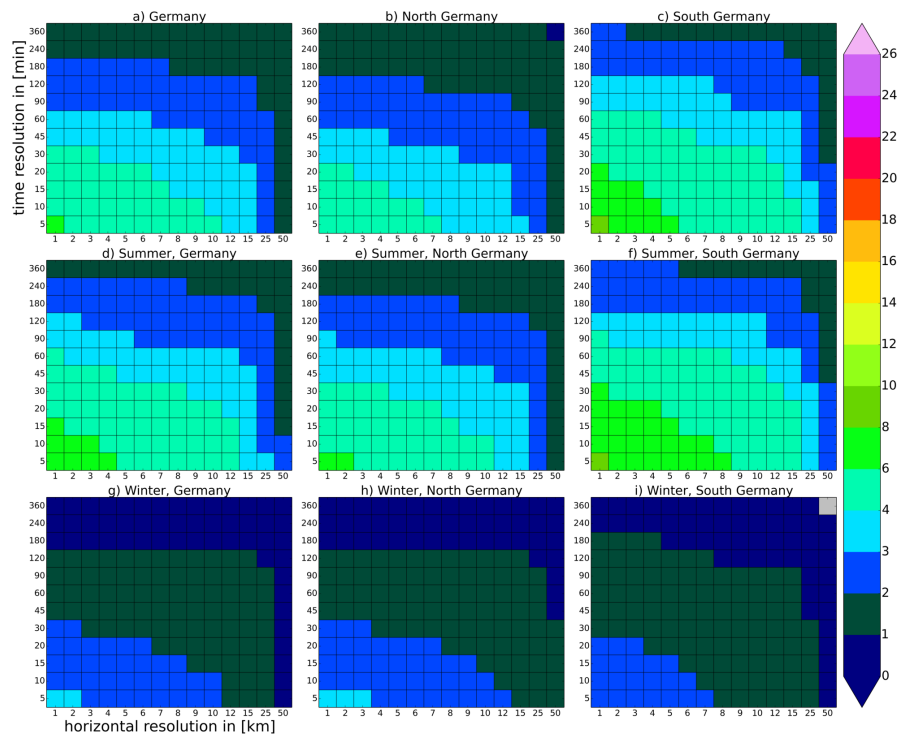
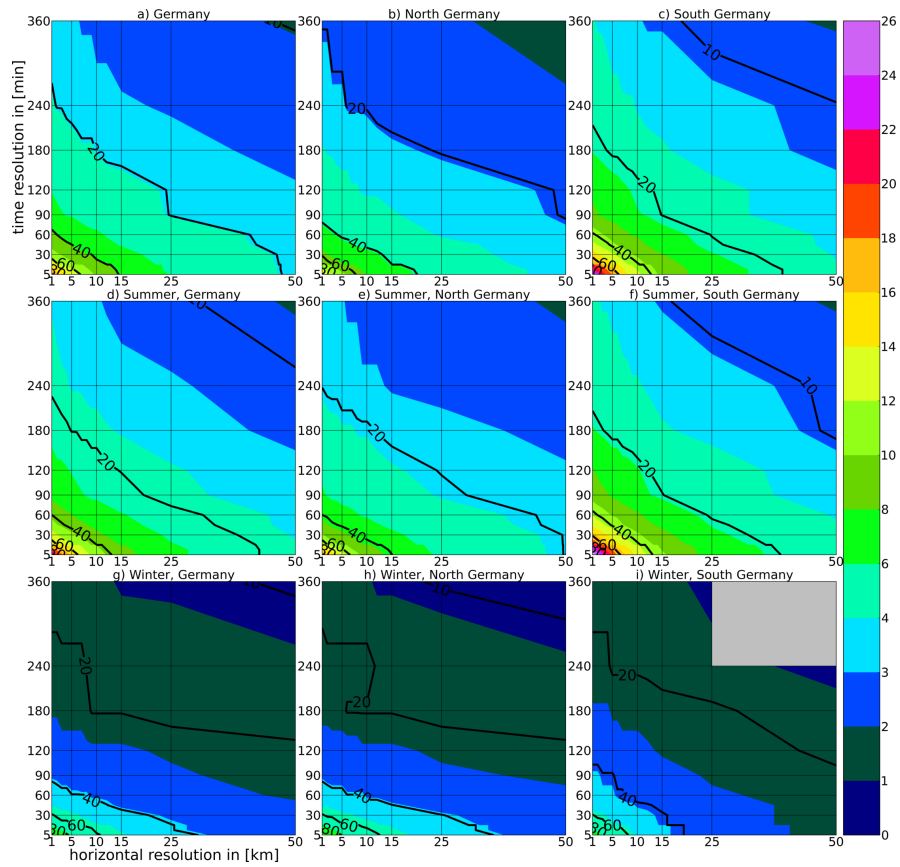Figure 1: **Convective extremes as function of resolution.** The **95th** percentile of convective precipitation intensities, aggregated over different parts of Germany for the years 2007-2008, on different horizontal (horizontal axis) and temporal (vertical axis) resolutions: Entire year (a,b,c), summer season (d,e,f) and winter season (g,h,i). All of Germany (a,d,g), North Germany (b,e,h), South Germany (c,f,i); Intensities given in $mm\ h^{-1}$.
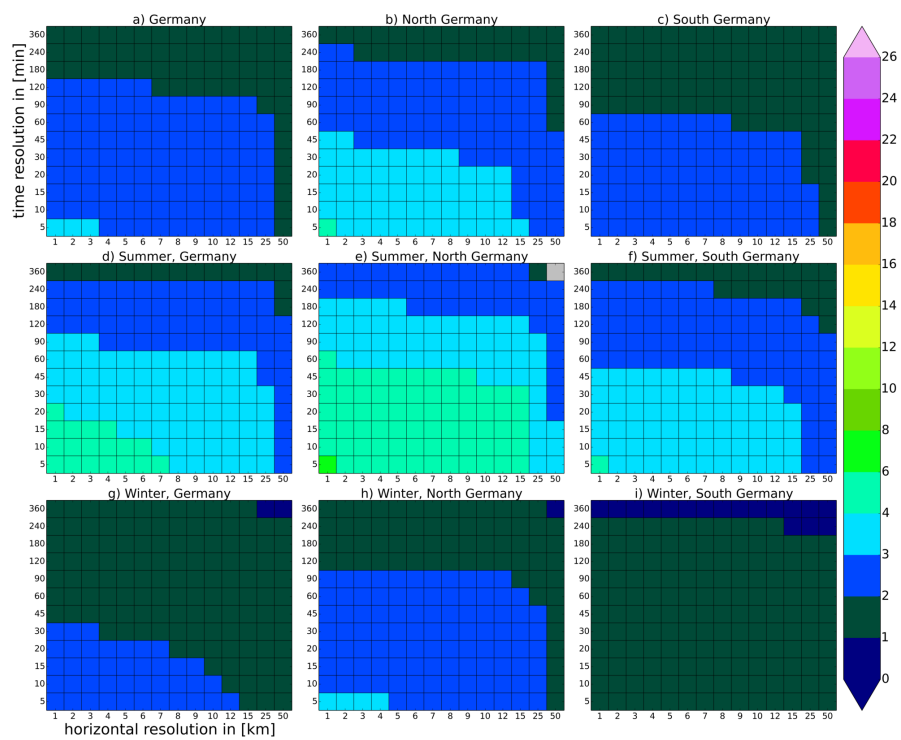
Figure 2: **Stratiform extremes as function of resolution.** Otherwise similar to Fig. 1.
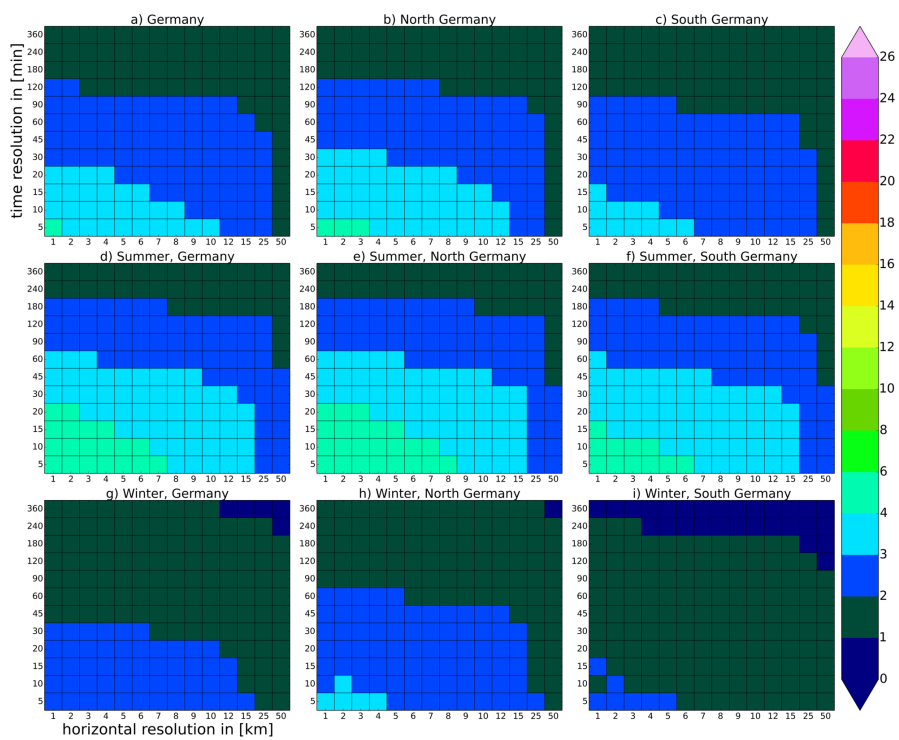
Figure 3: **Precipitation extremes as function of resolution.** Otherwise similar to Fig. 1.
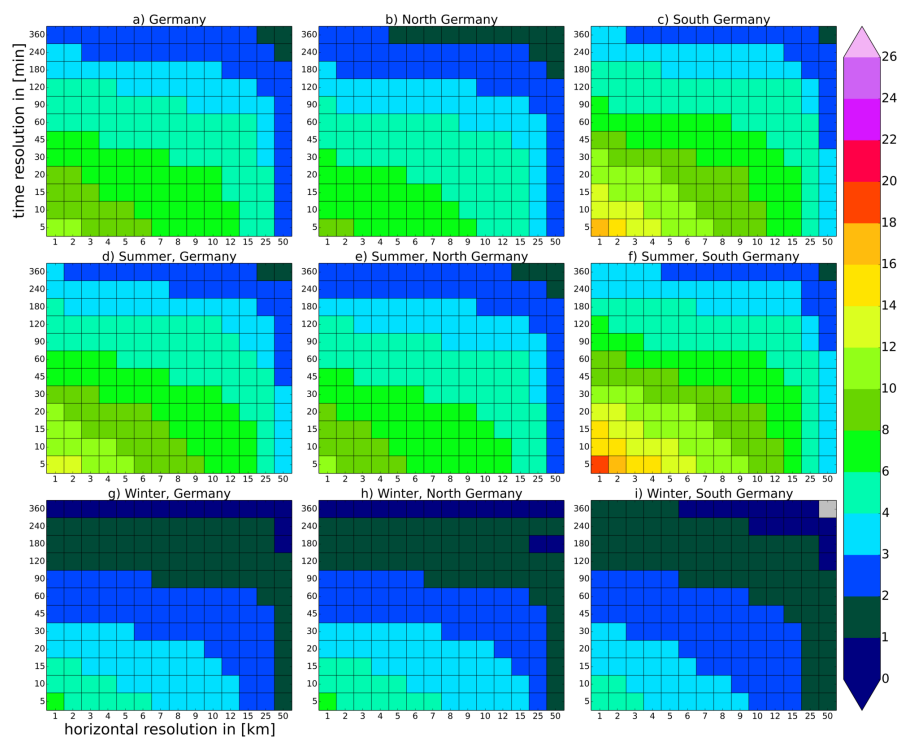
Figure 4: **Convective extremes as function of resolution.** The **98th** percentile of convective precipitation intensities, Otherwise similar to Fig. 1.
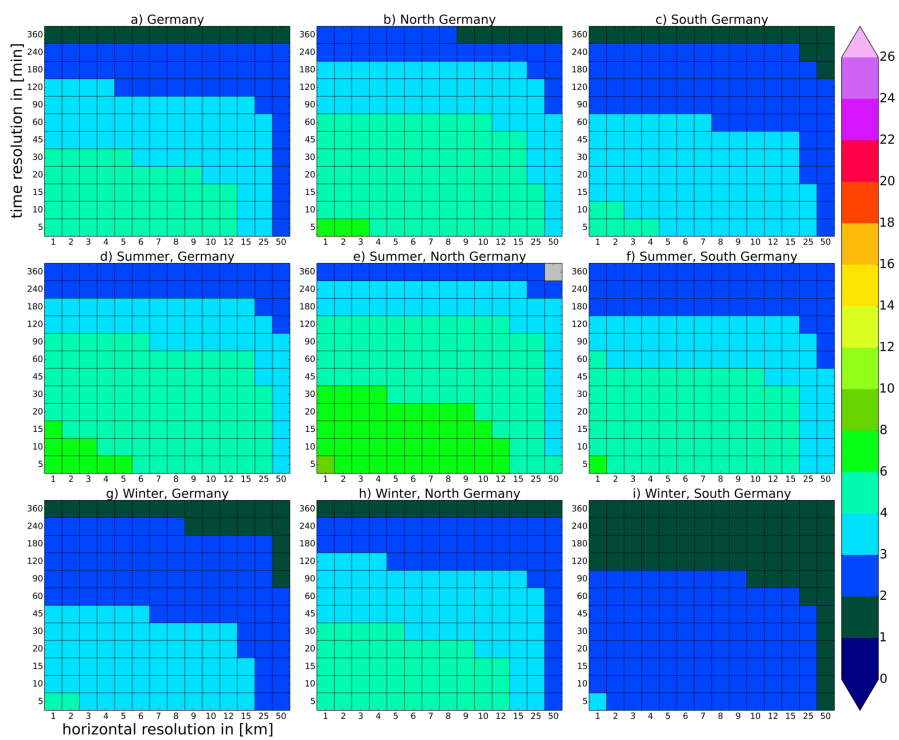
Figure 5: **Stratiform extremes as function of resolution.** The **98th** percentile of stratiform precipitation intensities, Otherwise similar to Fig. 1.
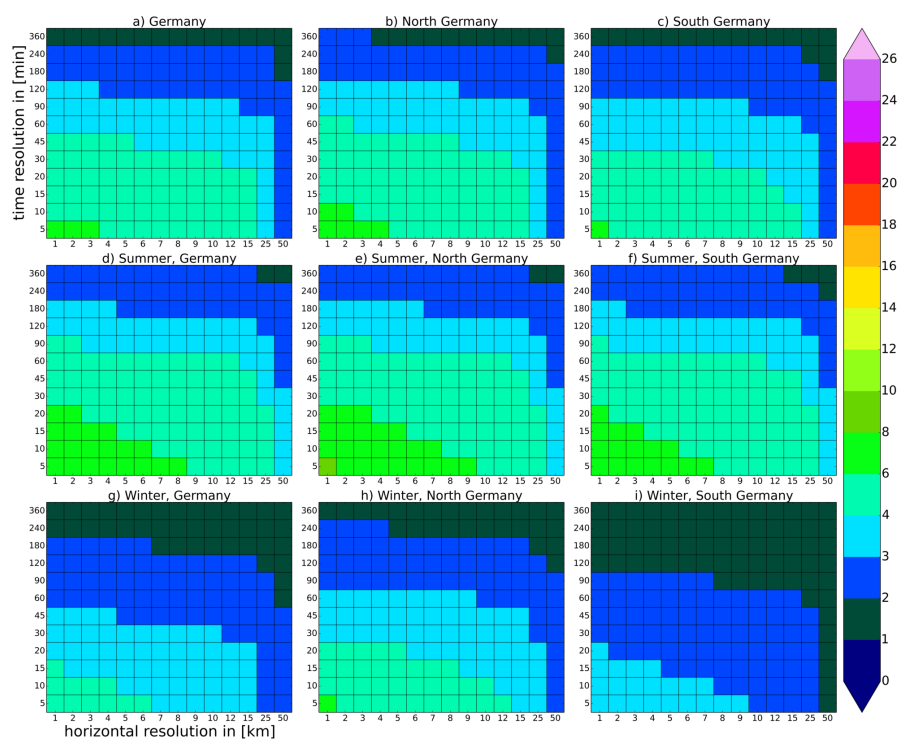
Figure 6: **Precipitation extremes as function of resolution.** The **98th** percentile of precipitation intensities, Otherwise similar to Fig. 1.
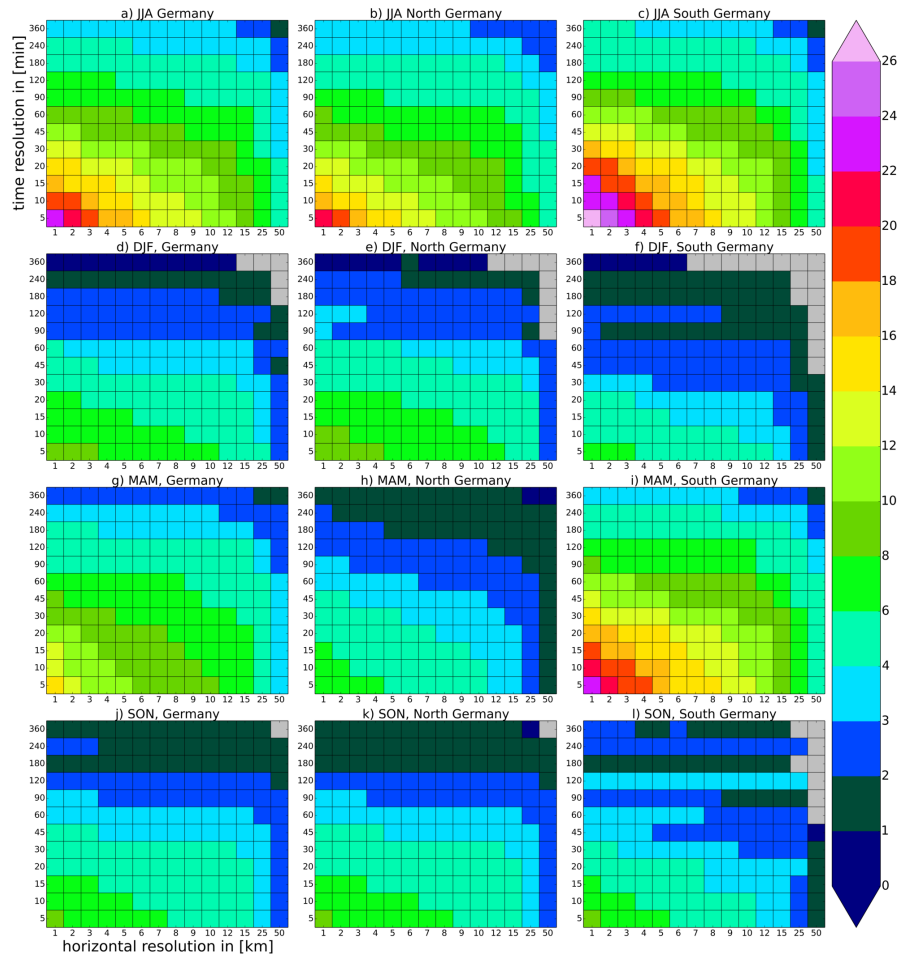
Figure 7: **Convective extremes as function of resolution.** The **99th** percentile of convective precipitation intensities, aggregated over different parts of Germany for the years 2007-2008, on different horizontal (horizontal axis) and temporal (vertical axis) resolutions: Seasons: JJA (a,b,c), DJF (d,e,f), MAM (g,h,i), and SON (j,k,l). All of Germany (a,d,g), North Germany (b,e,h), South Germany (c,f,i); Intensities given in $mm\ h^{-1}$.
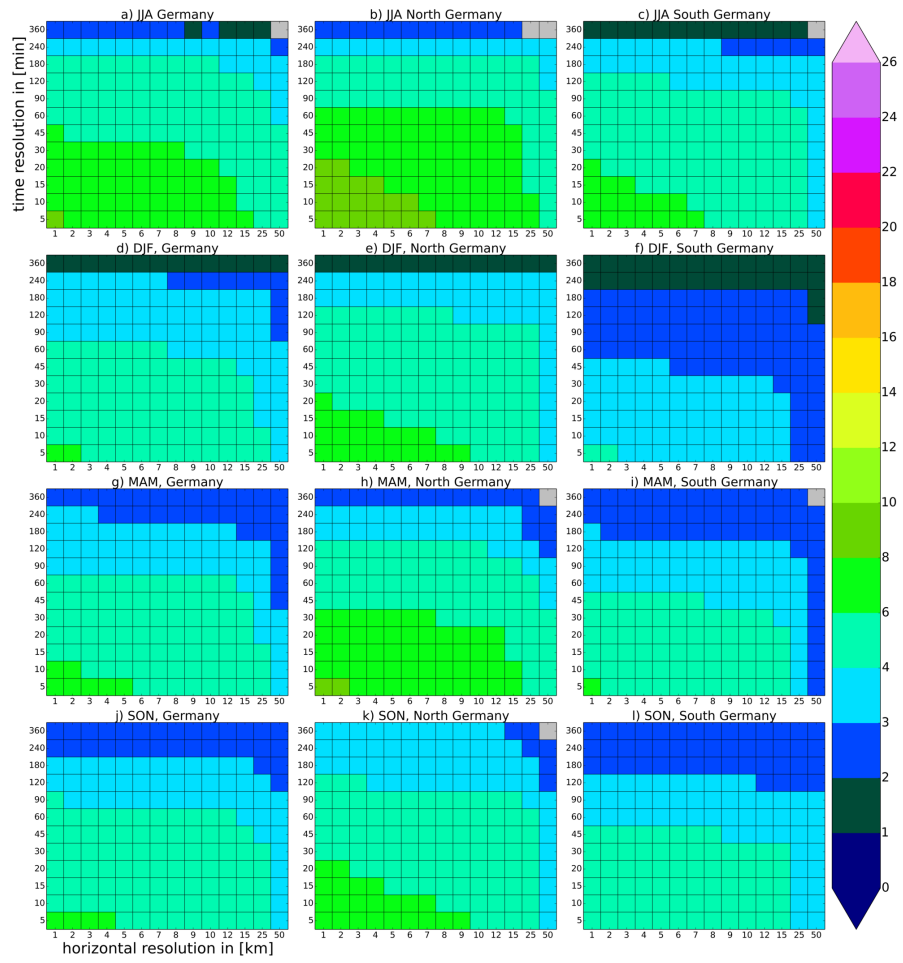
Figure 8: **Stratiform extremes as function of resolution.** Otherwise similar to Fig. 7.
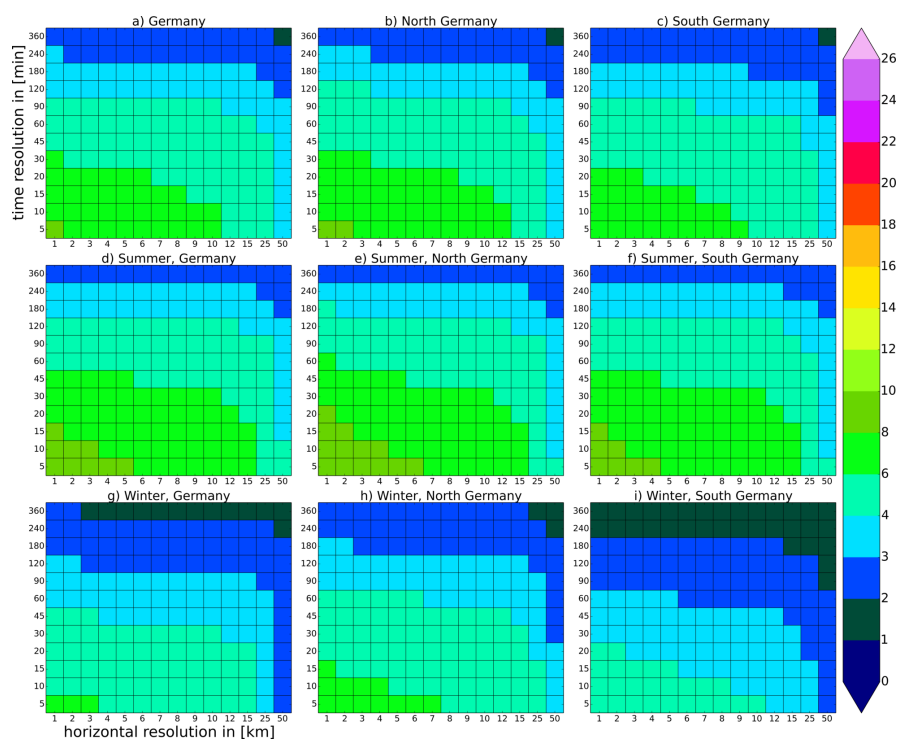
Figure 9: **Precipitation extremes as function of resolution.** The **99th** percentile of precipitation intensities. Otherwise similar to Fig. 1.
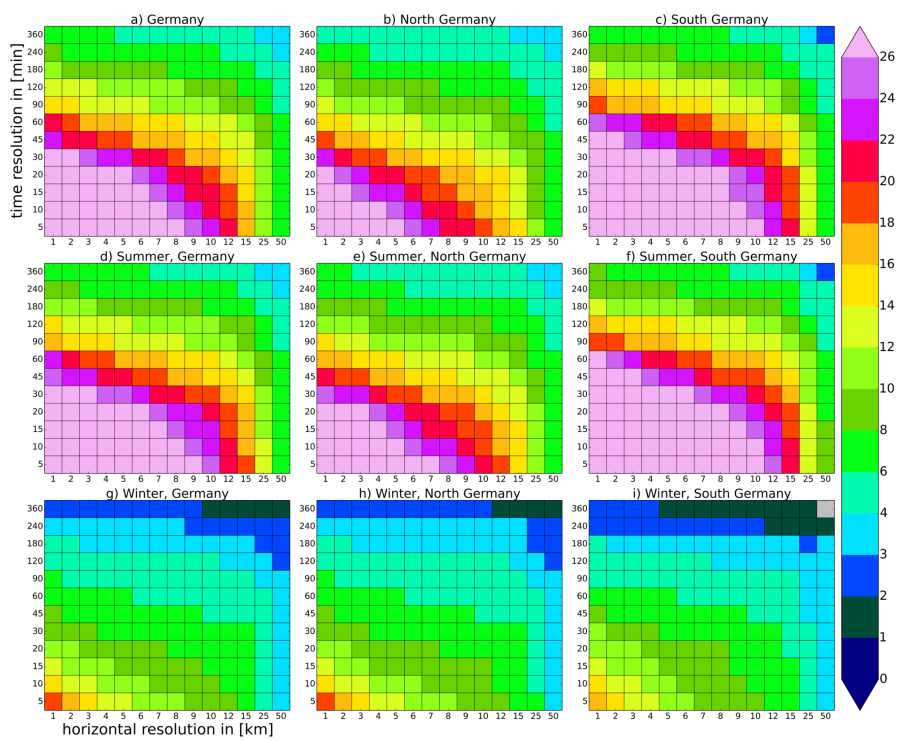
Figure 10: **Convective extremes as function of resolution.** The **99.9th** percentile of convective precipitation intensities. Otherwise similar to Fig. 1.
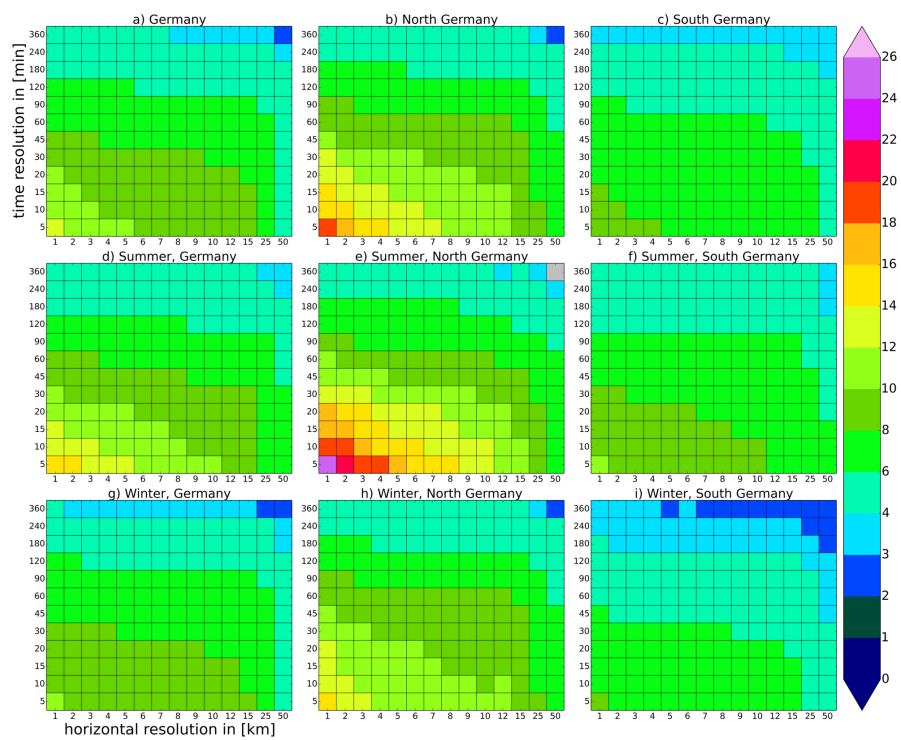
Figure 11: **Stratiform extremes as function of resolution.** The **99.9th** percentile of stratiform precipitation intensities. Otherwise similar to Fig. 1.
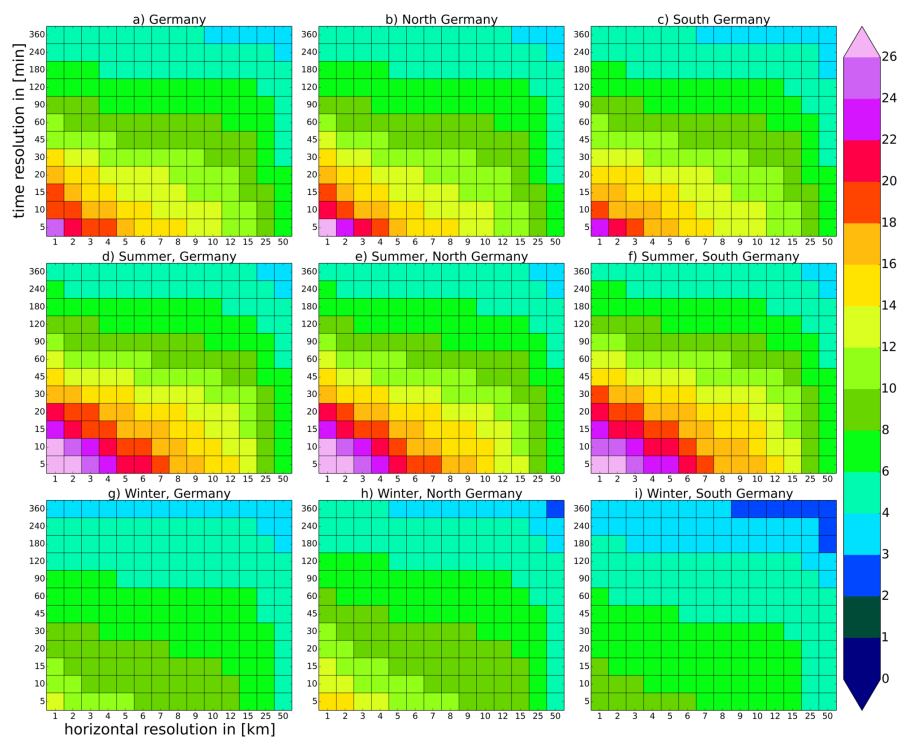
Figure 12: **Precipitation extremes as function of resolution.** The **99.9th** percentile of precipitation intensities. Otherwise similar to Fig. 1.

Figure 13: **Convective extremes as function of resolution.** The **99th** percentile of convective precipitation intensities with a linear scaling at the x and y-axis. The contour lines indicate the intensity decrease in %. Otherwise similar to Fig. 1.
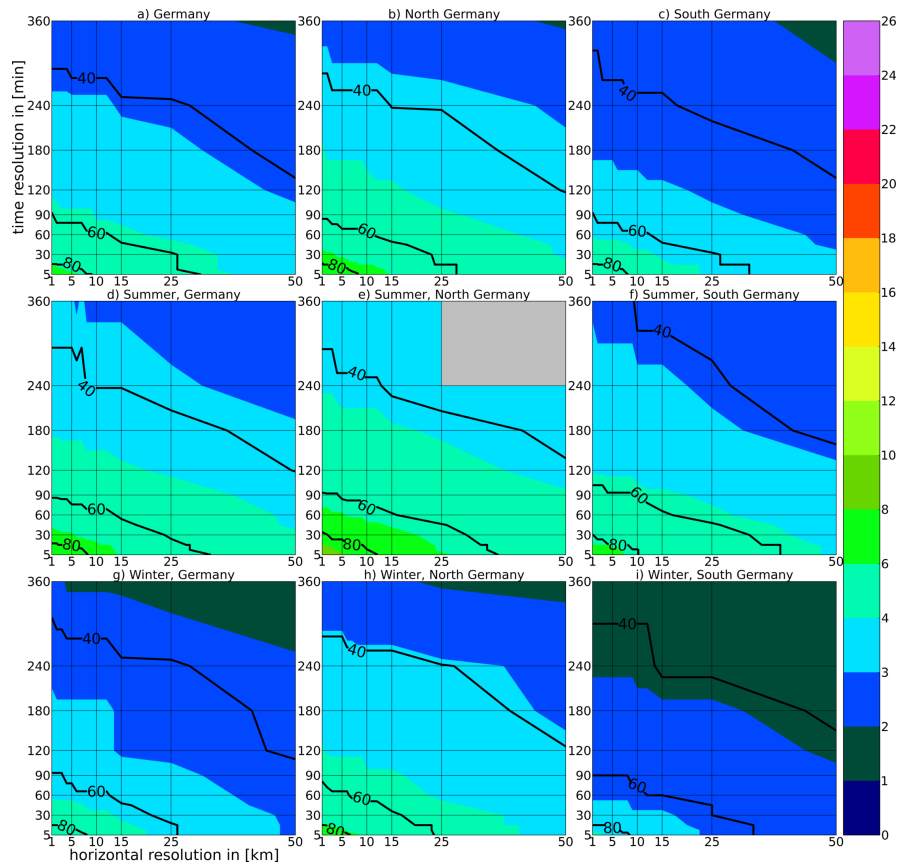
Figure 14: **Stratiform extremes as function of resolution.** The **99th** percentile of stratiform precipitation intensities with a linear scaling at the x and y-axis. The contour lines indicate the intensity decrease in %. Otherwise similar to Fig. 1.
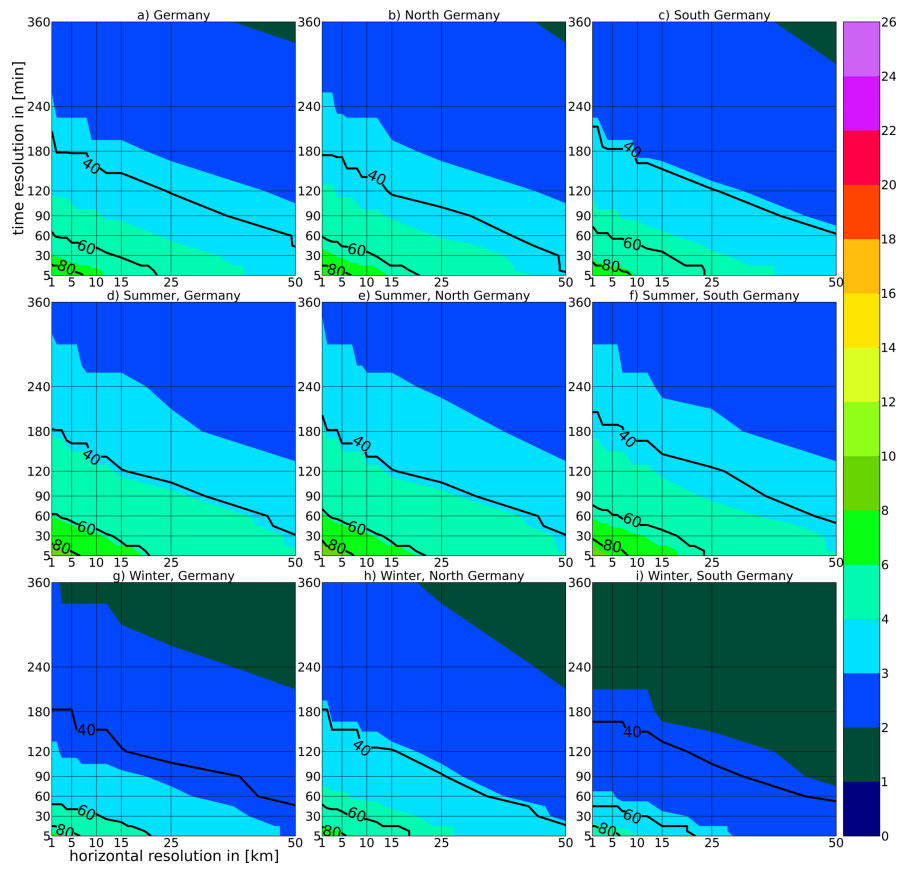
Figure 15: **Precipitation extremes as function of resolution.** The **99th** percentile of precipitation intensities with a linear scaling at the x and y-axis. Otherwise similar to Fig. 1.
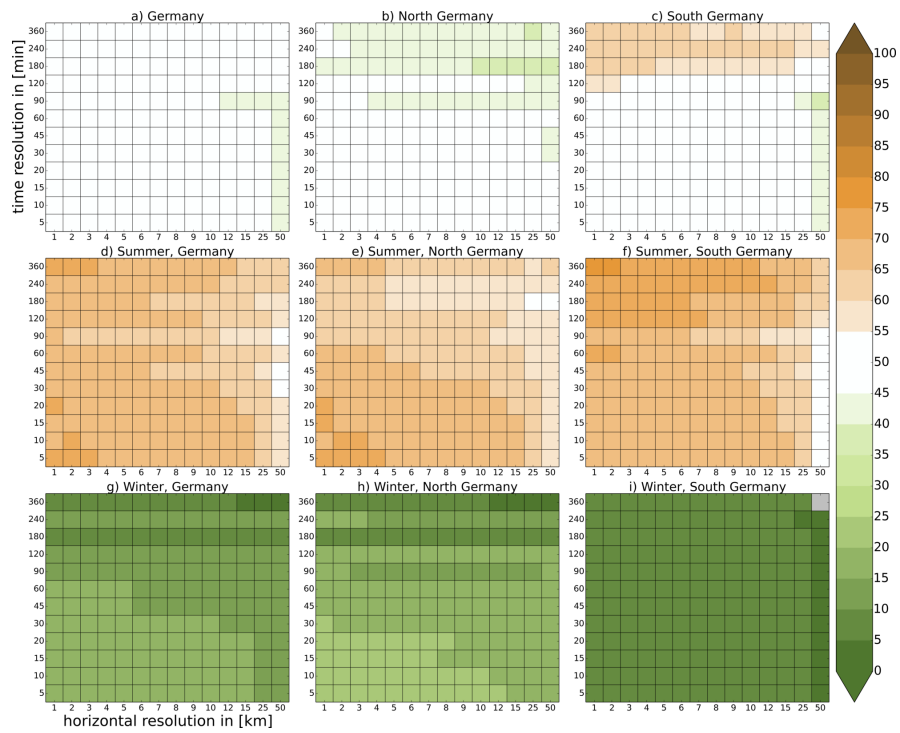
Figure 16: **Convective dominance as function of resolution including dry periods.** The ratio of the number of convective precipitation events with precipitation intensities larger or equal threshold intensity. Threshold intensity is defined as the **95th** percentile of total precipitation intensities over the different parts of Germany for the years 2007-2008. Panels otherwise as in Fig. 1.
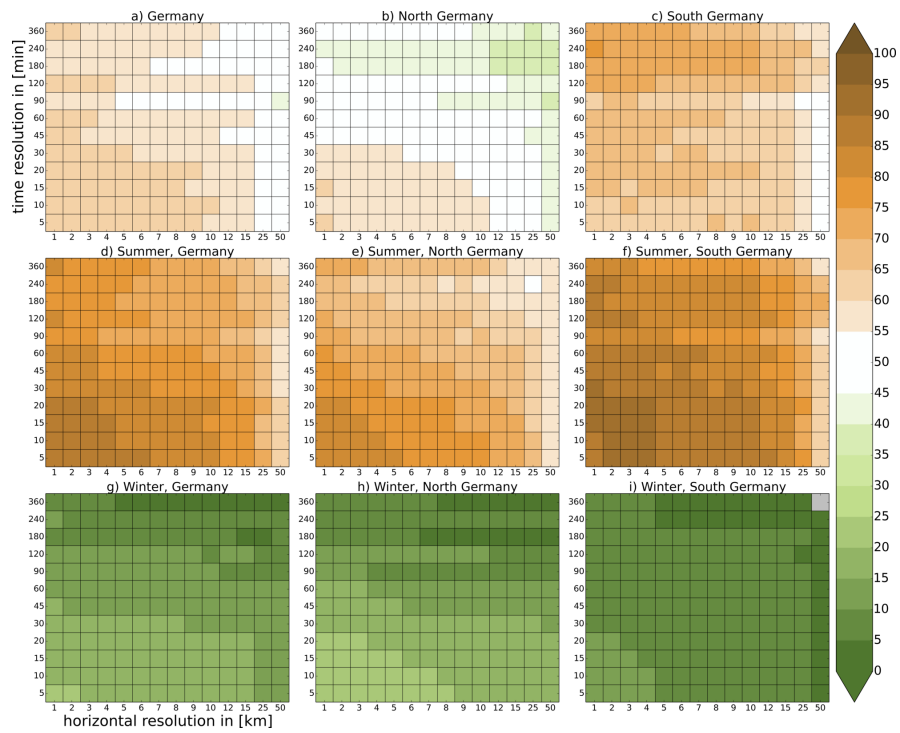
Figure 17: **Convective dominance as function of resolution including dry periods.** The ratio of the number of convective precipitation events with precipitation intensities larger or equal threshold intensity. Threshold intensity is defined as the **98th** percentile of total precipitation intensities over the different parts of Germany for the years 2007-2008. Panels otherwise as in Fig. 1.
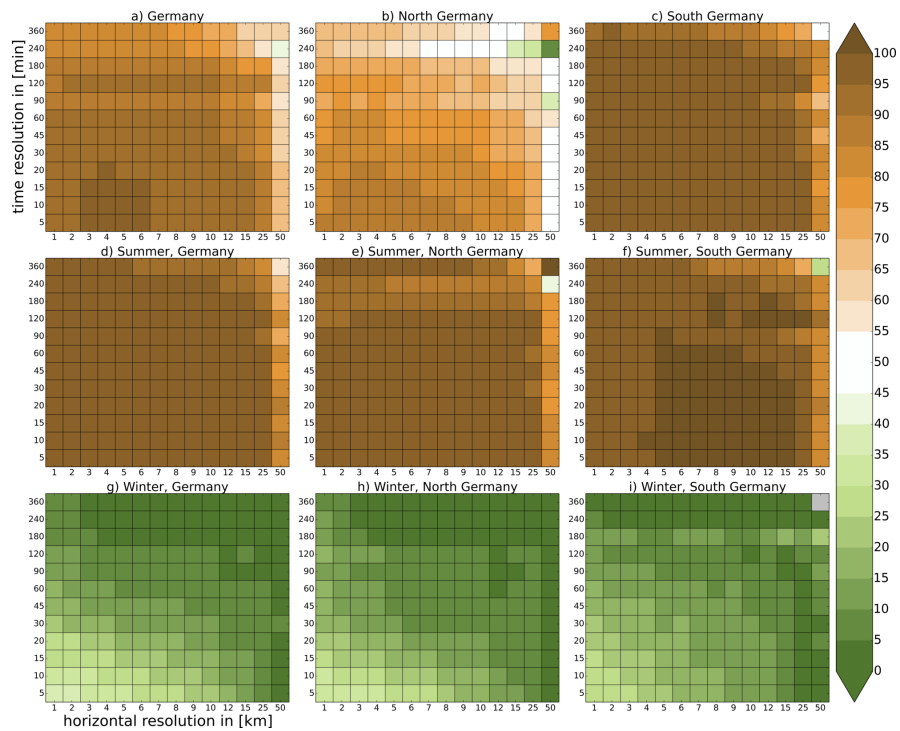
Figure 18: **Convective dominance as function of resolution including dry periods.** The ratio of the number of convective precipitation events with precipitation intensities larger or equal threshold intensity. Threshold intensity is defined as the **99.9th** percentile of total precipitation intensities over the different parts of Germany for the years 2007-2008. Panels otherwise as in Fig. 1.
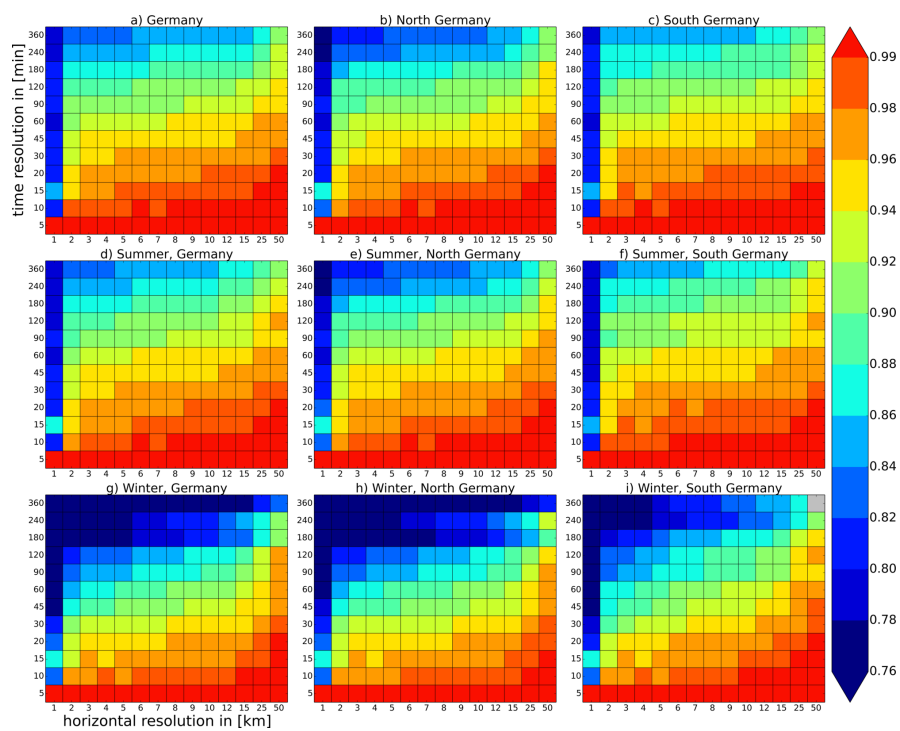
Figure 19: **PDF overlap for convective precipitation intensity.** PDF overlap of each horizontal resolution between every temporal resolution and the 5 minute data. Panels otherwise as in Fig. 1.
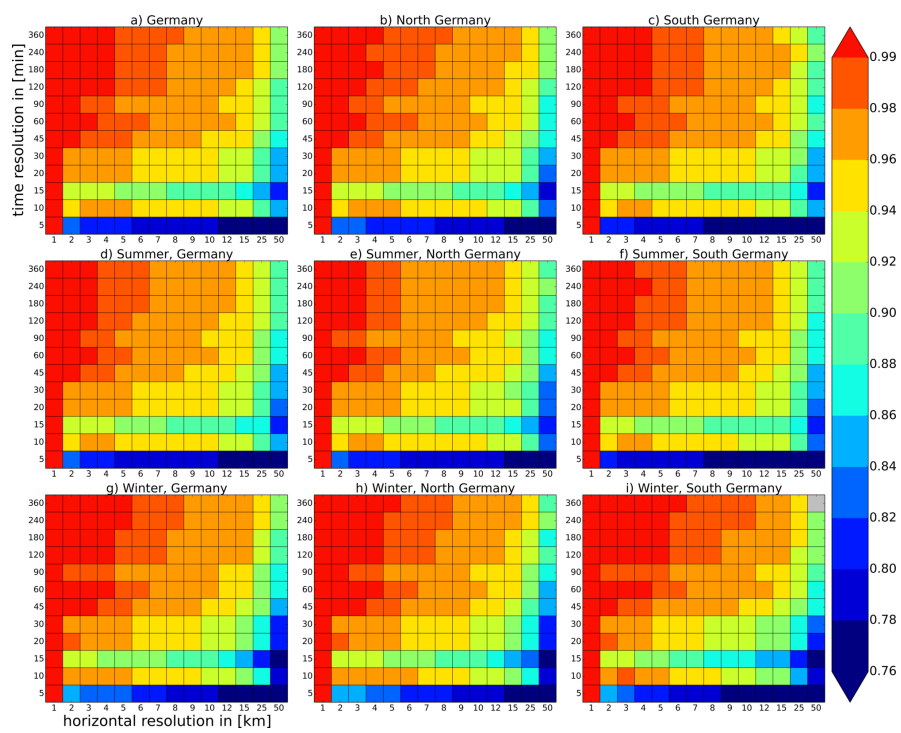
Figure 20: **PDF overlap for convective precipitation intensity.** PDF overlap of each temporal resolution between every horizontal resolution and the 1 *km* data. Panels otherwise as in Fig. 1.
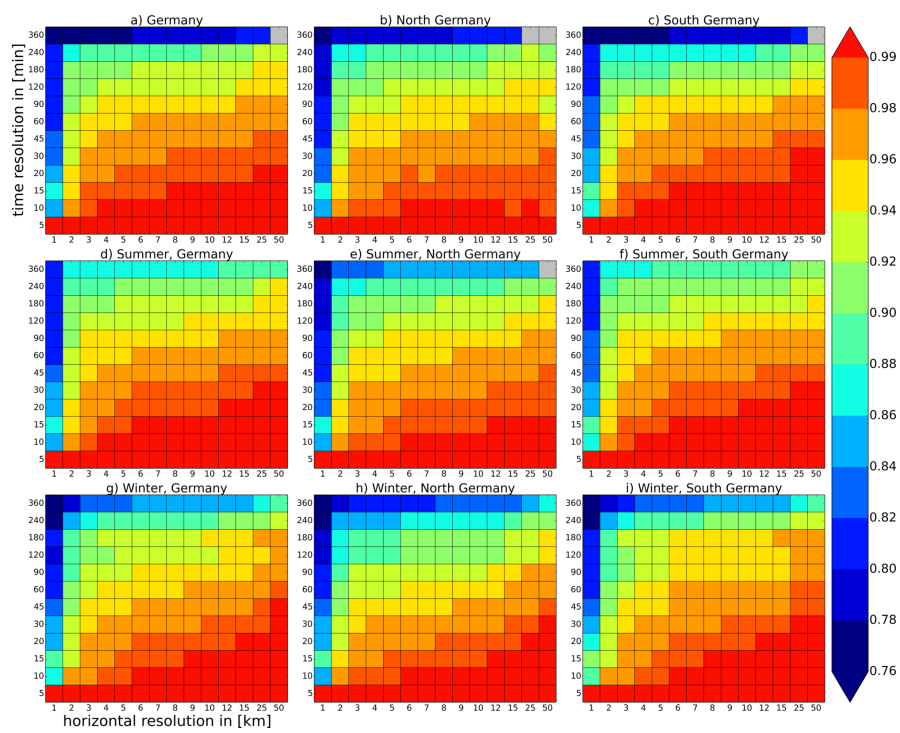
Figure 21: **PDF overlap for stratiform precipitation intensity.** PDF overlap of each horizontal resolution between every temporal resolution and the 5 minute data. Panels otherwise as in Fig. 1.
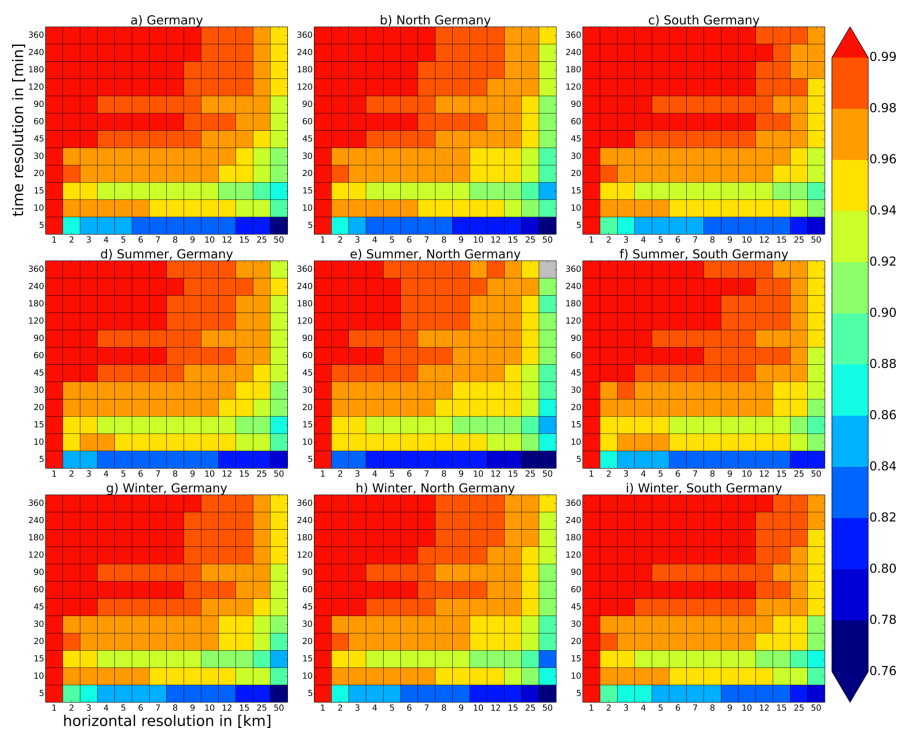
Figure 22: **PDF overlap for stratiform precipitation intensity.** PDF overlap of each temporal resolution between every horizontal resolution and the 1 *km* data. Panels otherwise as in Fig. 1.
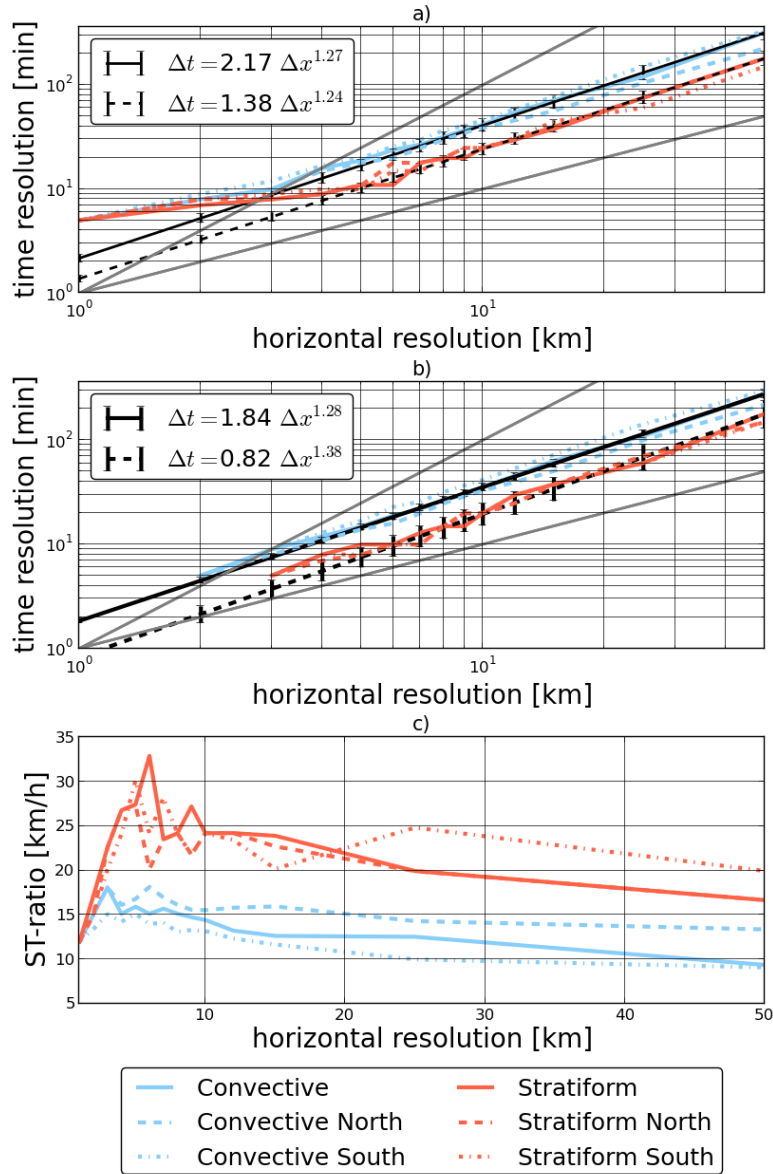
Figure 23: **Consistent spatial and temporal resolutions**. Δt derived using Eq. 5 for different values of Δx for convective (blue) and stratiform (red) precipitation extremes at the 99th percentile for entire germany, north germany and south germany. Black lines are least square fit of $\Delta t = a \times \Delta x^b$ with the fitting parameters a and b for entire germany. Errorbars indicate the standard deviation of parameter estimates. Gray lines show $\Delta t \sim \Delta x$ and $\Delta t \sim \Delta x^2$, respectively. **a**, Initial resolutions $\Delta t_0 = 5\ min$, $\Delta x_0 = 1\ km$. **b**, $\Delta t_0 = 5\ min$, and aggregated spatial resolutions $\Delta x_0 = 2\ km$ (convective) and $\Delta x_0 = 3\ km$ (stratiform). **c**, ST-ratio (Eq. 7) for both precipitation types for different parts of Germany over the entire year.

D.3 Article 4: Uncertainties in snow and precipitation projections in the northern Alps: the role of model biases

Supplement material to: Uncertainties in snow and precipitation projections in the northern Alps: the role of model biases

Bastian Eggert[1,2], Andreas Haensler[2], and Daniela Jacob[1,2]

[1]Faculty of Sustainability, Leuphana University Lueneburg
[2]Climate Service Center (GERICS), Helmholtz Zentrum Geesthacht

September 19, 2017

Percentage change

For percentage change we receive additional terms. As we will see below, the climate change signal from the model is not canceled out, because of the different denominators $P_h(I)$ and $P_{obs}(I)$.

The difference in the percentage climate change signal between model and observations can be written like this: $\delta P_{ccs_p}(I) = 100 * \frac{P_{ccs}(I)}{P_h(I)} - 100 * \frac{P_{ccs\_obs}(I)}{P_{obs}(I)}$

For percentage change, the first part of the analysis from above (terms 1-3) now have the following form:

1. $100 * \frac{P_h(T_i) * \delta P_{ccs}(I|T_i)}{P_{obs}(I)}$:

2. $100 * \frac{\delta P_h(T_i) * P_{ccs}(I|T_i)}{P_{obs}(I)}$:

3. $100 * \frac{\delta P(T_i) * \delta P_{ccs}(I|T_i)}{P_{obs}(I)}$:

Additionaly to these three terms we obtain term P1. This term describes differences that result from percentage changes using model data as reference and percentage changes using observation data as reference.

P1. $100 * \frac{P_h(T_i) * P_{ccs}(I|T_i)}{P_h(I)} - 100 * \frac{P_h(T_i) * P_{ccs}(I|T_i)}{P_{obs}(I)}$

Terms 4-6 can be decomposed in the same way. We obtain the extra term P2.

P2. $100 * \frac{P_h(I|T_i) * P_{ccs}(T_i)}{P_h(I)} - 100 * \frac{P_h(I|T_i) * P_{ccs}(T_i)}{P_{obs}(I)}$

In the result section we will refer to the terms P1 and P2 as scaling terms. These terms become important if there is a mismatch in the magnitude of the historic distributions between models and observations.
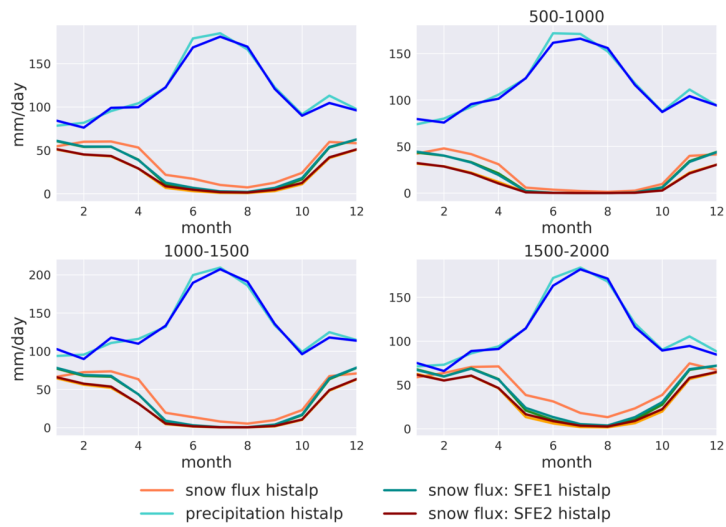
Figure S1: Yearly cycle of precipitation and snowfall amounts over the entire northern Alpine region and for selected altitude ranges.
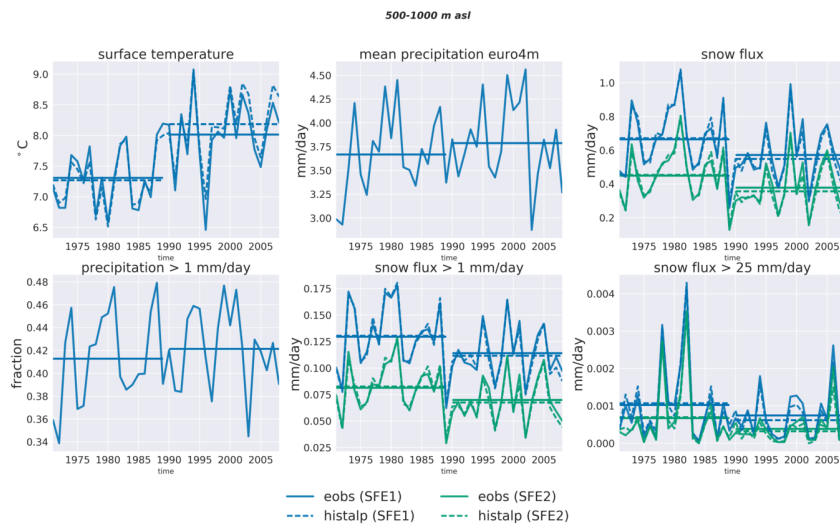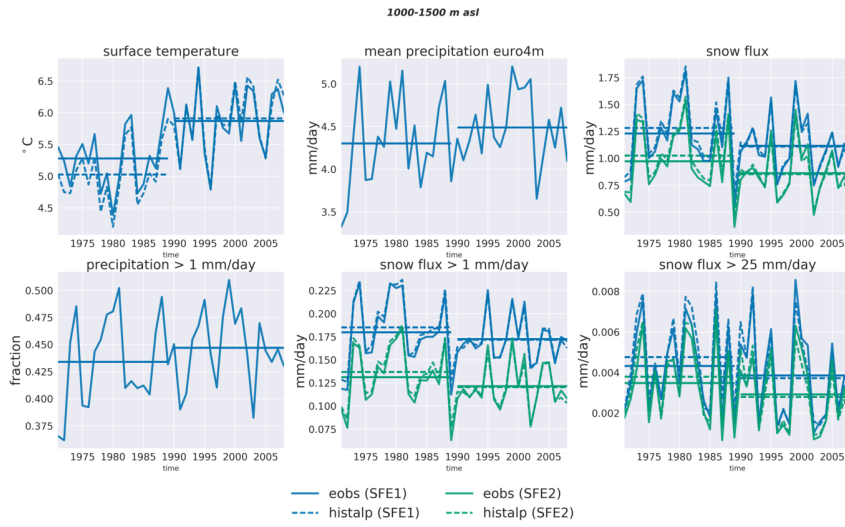

Figure S2: Yearly mean timelines of different observation data sets. Data is an average over the northern Alpine region, considering only grid boxes with a mean altitude between 500 to 1000 m asl. Straight lines indicate 20 year mean values over the periods 1971-1989 and 1990-2008.

all altitudes	1971-1989	1990-2008	difference	difference in %
surface temperature eobs [$°C$]	5.200	5.900	0.700	NaN
surface temperature histalp [$°C$]	5.000	5.900	0.900	NaN
mean precipitation euro4m [mm/day]	3.800	3.900	0.200	4.2
snow flux SFE1: eobs [mm/day]	1.000	0.900	-0.100	-9.4
snow flux SFE1: histalp [mm/day]	1.100	1.000	-0.100	-12.4
snow flux SFE2: eobs [mm/day]	0.800	0.700	-0.100	-10.9
snow flux SFE2: histalp [mm/day]	0.900	0.700	-0.100	-14.5
precipitation frequency [%]	41.300	42.500	1.200	2.8
snowfall frequency SFE1: eobs [%]	16.800	15.600	-1.200	-7.0
snowfall frequency SFE1: histalp [%]	17.300	15.800	-1.500	-8.8
snowfall frequency SFE2: eobs [%]	12.200	11.200	-1.000	-8.8
snowfall frequency SFE2: histalp [%]	12.700	11.200	-1.500	-11.6
heavy snowfall frequency SFE1: eobs [%]	0.338	0.299	-0.039	-11.4
heavy snowfall frequency SFE1: histalp [%]	0.377	0.308	-0.069	-18.3
heavy snowfall frequency SFE2: eobs [%]	0.276	0.236	-0.040	-14.7
heavy snowfall frequency SFE2: histalp [%]	0.311	0.247	-0.064	-20.7

Table S1: Observed time averages over the 20 year time periods 1971-1989 and 1990-2008. Data is an average over the northern Alpine region.

500-1000 m asl	1971-1989	1990-2008	difference	difference in %
surface temperature eobs [$°C$]	7.300	8.000	0.7	NaN
surface temperature histalp [$°C$]	7.300	8.200	0.9	NaN
mean precipitation euro4m [mm/day]	3.700	3.800	0.1	3.3
mean snowfall SFE1: eobs [mm/day]	0.700	0.600	-0.1	-13.9
mean snowfall SFE1: histalp [mm/day]	0.700	0.500	-0.1	-18.3
mean snowfall SFE2: eobs [mm/day]	0.400	0.400	-0.1	-15.9
mean snowfall SFE2: histalp [mm/day]	0.500	0.400	-0.1	-21.5
precipitation frequency [%]	41.300	42.200	0.9	2.1
snowfall frequency SFE1: eobs [%]	13.000	11.400	-1.6	-12.1
snowfall frequency SFE1: histalp [%]	13.100	11.200	-1.9	-14.4
snowfall frequency SFE2: eobs [%]	8.100	7.000	-1.1	-14.2
snowfall frequency SFE2: histalp [%]	8.300	6.800	-1.5	-18.2
heavy snowfall frequency SFE1: eobs [%]	0.103	0.075	-0.028	-27.1
heavy snowfall frequency SFE1: histalp [%]	0.107	0.062	-0.045	-42.1
heavy snowfall frequency SFE2: eobs [%]	0.067	0.039	-0.028	-42.4
heavy snowfall frequency SFE2: histalp [%]	0.070	0.032	-0.038	-53.8

Table S2: Observed time averages over the 20 year time periods 1971-1989 and 1990-2008. Data is an average over the northern Alpine region, concidering only altitudes between 500 to 1000 m asl.
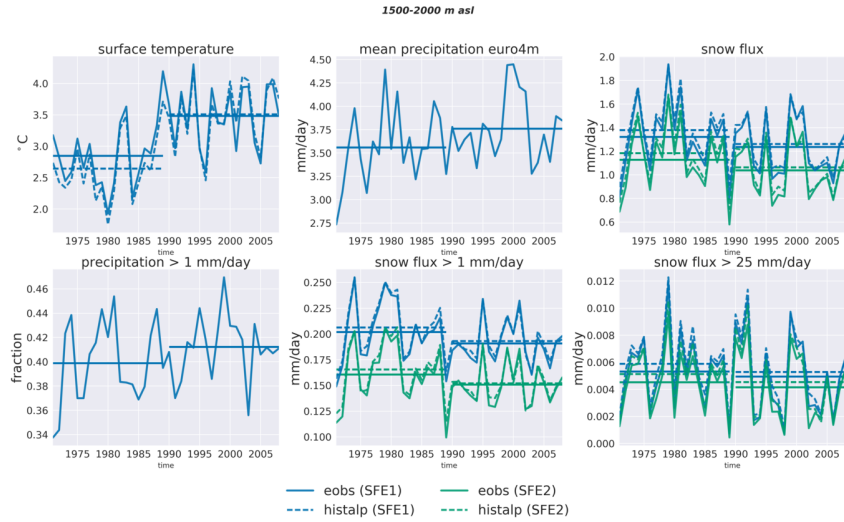
Figure S3: Yearly mean timelines of different observation data sets. Data is an average over the northern Alpine region, considering only grid boxes with a mean altitude between 1000 to 1500 m asl. Straight lines indicate 20 year mean values over the periods 1971-1989 and 1990-2008.

1000-1500 m asl	1971-1989	1990-2008	difference	difference in %
surface temperature eobs [$°C$]	5.300	5.900	0.600	NaN
surface temperature histalp [$°C$]	5.000	5.900	0.900	NaN
mean precipitation euro4m [mm/day]	4.300	4.500	0.200	4.3
mean snowfall SFE1: eobs [mm/day]	1.200	1.100	-0.100	-9.4
mean snowfall SFE1: histalp [mm/day]	1.300	1.100	-0.200	-13.5
mean snowfall SFE2: eobs [mm/day]	1.000	0.900	-0.100	-11.3
mean snowfall SFE2: histalp [mm/day]	1.000	0.900	-0.200	-16.9
precipitation frequency [%]	43.400	44.700	1.3	3.0
snowfall frequency SFE1: eobs [%]	18.000	17.200	-0.8	-4.1
snowfall frequency SFE1: histalp [%]	18.500	17.200	-1.3	-7.2
snowfall frequency SFE2: eobs [%]	13.100	12.200	-0.9	-7.4
snowfall frequency SFE2: histalp [%]	13.700	12.100	-1.6	-12.0
heavy snowfall frequency SFE1: eobs [%]	0.433	0.386	-0.047	-10.8
heavy snowfall frequency SFE1: histalp [%]	0.477	0.373	-0.104	-21.8
heavy snowfall frequency SFE2: eobs [%]	0.349	0.293	-0.056	-16.1
heavy snowfall frequency SFE2: histalp [%]	0.380	0.280	-0.100	-26.2

Table S3: Observed time averages over the 20 year time periods 1971-1989 and 1990-2008. Data is an average over the northern Alpine region, concidering only altitudes between 1000 to 1500 m asl.
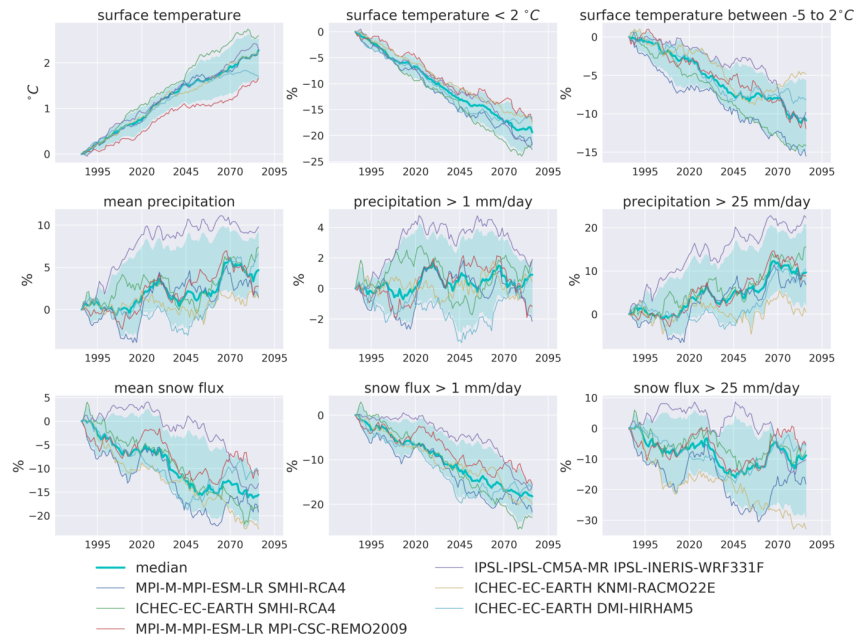

Figure S4: Yearly mean timelines of different observation data sets. Data is an average over the northern Alpine region, considering only grid boxes with a mean altitude between 1500 to 2000 m asl. Straight lines indicate 20 year mean values over the periods 1971-1989 and 1990-2008.

1500-2000 m asl	1971-1989	1990-2008	difference	difference in %
surface temperature eobs [$°C$]	2.800	3.500	0.600	NaN
surface temperature histalp [$°C$]	2.600	3.500	0.900	NaN
mean precipitation euro4m [mm/day]	3.600	3.800	0.200	5.6
mean snowfall SFE1: eobs [mm/day]	1.300	1.200	-0.100	-6.5
mean snowfall SFE1: histalp [mm/day]	1.400	1.300	-0.100	-8.5
mean snowfall SFE2: eobs [mm/day]	1.100	1.000	-0.100	-7.9
mean snowfall SFE2: histalp [mm/day]	1.200	1.100	-0.100	-10.2
precipitation frequency [%]	39.900	41.200	1.3	3.3
snowfall frequency SFE1: eobs [%]	20.200	19.100	-1.1	-5.5
snowfall frequency SFE1: histalp [%]	20.600	19.300	-1.3	-6.4
snowfall frequency SFE2: eobs [%]	16.100	15.100	-1.0	-6.2
snowfall frequency SFE2: histalp [%]	16.600	15.200	-1.4	-8.1
heavy snowfall frequency SFE1: eobs [%]	0.533	0.495	-0.038	-7.0
heavy snowfall frequency SFE1: histalp [%]	0.590	0.529	-0.061	-10.3
heavy snowfall frequency SFE2: eobs [%]	0.452	0.415	-0.037	-8.2
heavy snowfall frequency SFE2: histalp [%]	0.513	0.454	-0.059	-11.5

Table S4: Observed time averages over the 20 year time periods 1971-1989 and 1990-2008. Data is an average over the northern Alpine region, concidering only altitudes between 1500 to 2000 m asl.

Figure S5: Yearly mean timelines of different indices. Data is an average over the northern Alpine region. The shaded areas indicate the ensemble spread from the 5th to the 95th percentile.
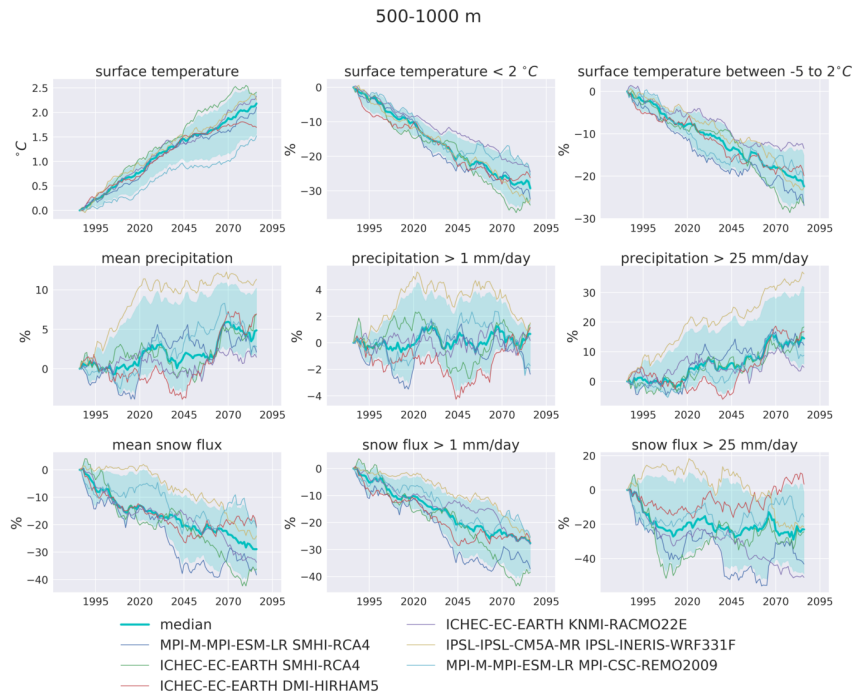
Figure S6: As Fig. S5, but averaged using gridboxes with an average altitude between 500 to 1000 m asl.

1000-1500 m
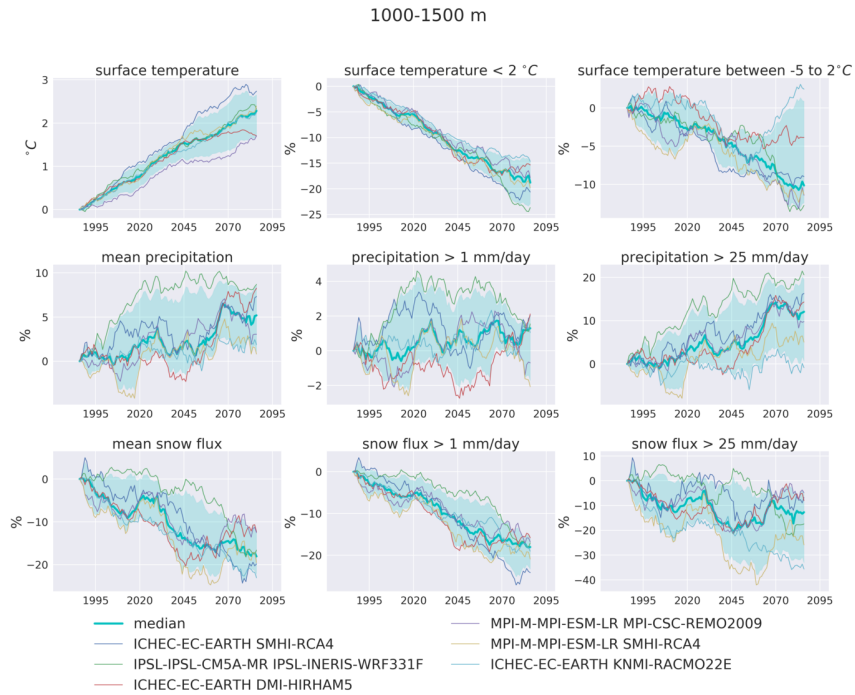
Figure S7: As Fig. 5, but averaged using gridboxes with an average altitude between 1000 to 1500 m asl.

1500-2000 m
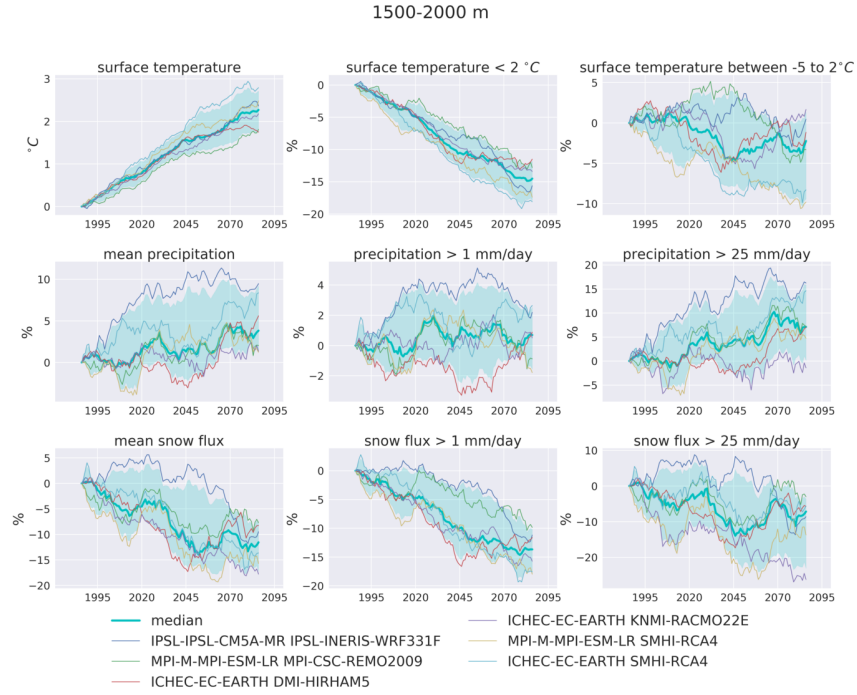
Figure S8: As Fig. 5, but averaged using gridboxes with an average altitude between 1500 to 2000 m asl.

		P(I) hist	P(I) future	P(I) CCS	P(T) dep. CCS	P(I|T) dep. CCS	corr. dep. CCS
EOBS	all-all	0.41	0.42	2.78	-0.33	2.91	0.20
	500-1000	0.41	0.42	2.10	-1.06	2.98	0.17
	1000-1500	0.43	0.45	2.97	-0.53	3.30	0.21
	1500-2000	0.40	0.41	3.35	-0.19	3.00	0.54
HISTALP	all-all	0.41	0.42	2.78	-0.24	2.65	0.36
	500-1000	0.41	0.42	2.10	-0.91	2.62	0.39
	1000-1500	0.43	0.45	2.97	-0.31	2.87	0.41
	1500-2000	0.40	0.41	3.35	0.13	2.76	0.46
ENSEMBLE mean	all-all	0.51	0.51	0.71	-0.15	0.89	-0.03
	500-1000	0.48	0.49	0.69	-0.09	0.83	-0.04
	1000-1500	0.52	0.52	0.66	-0.08	0.66	0.08
	1500-2000	0.54	0.54	0.73	-0.12	0.75	0.09

Table S5: Precipitation frequency changes within the historic period

		P(I) hist	P(I) future	P(I) CCS	P(T) dep. CCS	P(I|T) dep. CCS	corr. dep. CCS
EOBS	all-all	0.03	0.03	8.65	0.73	7.45	0.47
	500-1000	0.02	0.02	5.50	0.37	4.62	0.51
	1000-1500	0.03	0.03	11.30	-0.02	10.93	0.39
	1500-2000	0.02	0.03	11.32	0.69	10.06	0.57
HISTALP	all-all	0.03	0.03	8.65	1.39	6.43	0.83
	500-1000	0.02	0.02	5.50	1.12	3.37	1.00
	1000-1500	0.03	0.03	11.30	1.21	9.05	1.05
	1500-2000	0.02	0.03	11.32	1.60	8.96	0.76
ENSEMBLE mean	all-all	0.04	0.04	1.33	-0.64	1.89	0.09
	500-1000	0.02	0.02	1.91	0.48	1.28	0.15
	1000-1500	0.04	0.04	1.16	-0.12	0.99	0.30
	1500-2000	0.06	0.06	1.02	-0.16	0.92	0.27

Table S6: Heavy precipitation frequency changes within the historic period

precipitation frequency > 1 mm/day
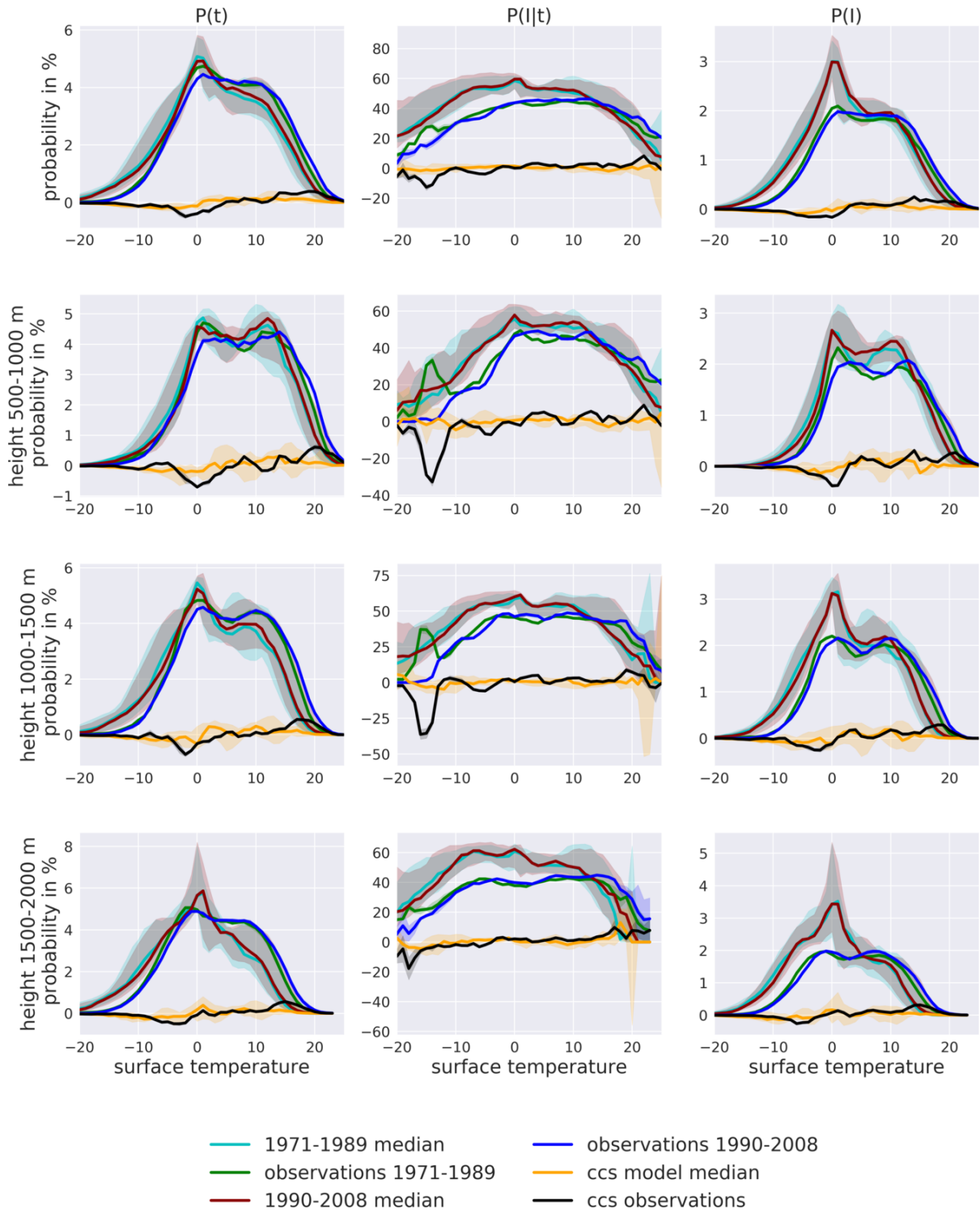
Figure S9: Temperature dependent probability distributions, P(I), P(I|T) and P(I) for precipitation frequencies, at different altitudes. For the periods 1971-1989 and 1990-2008.
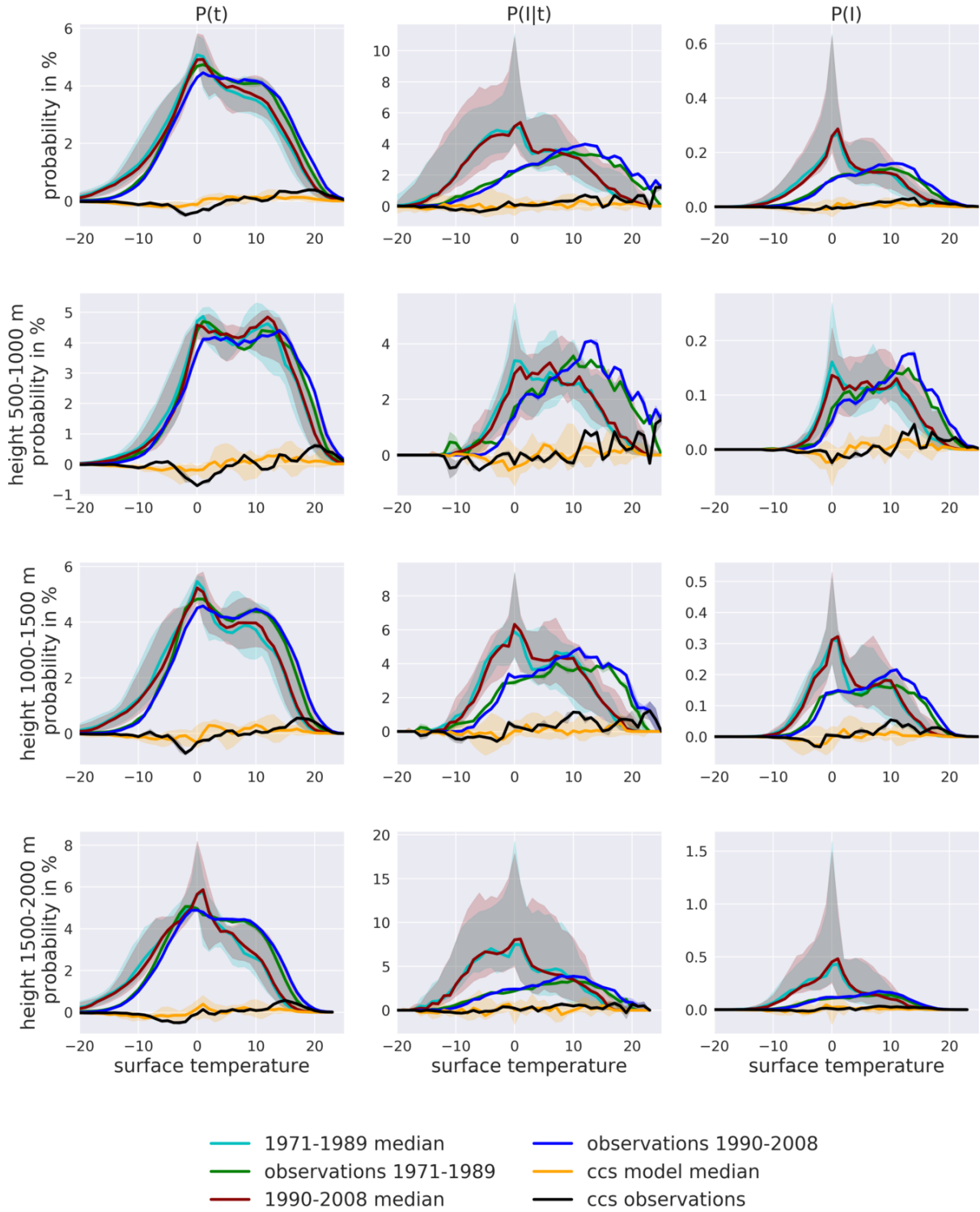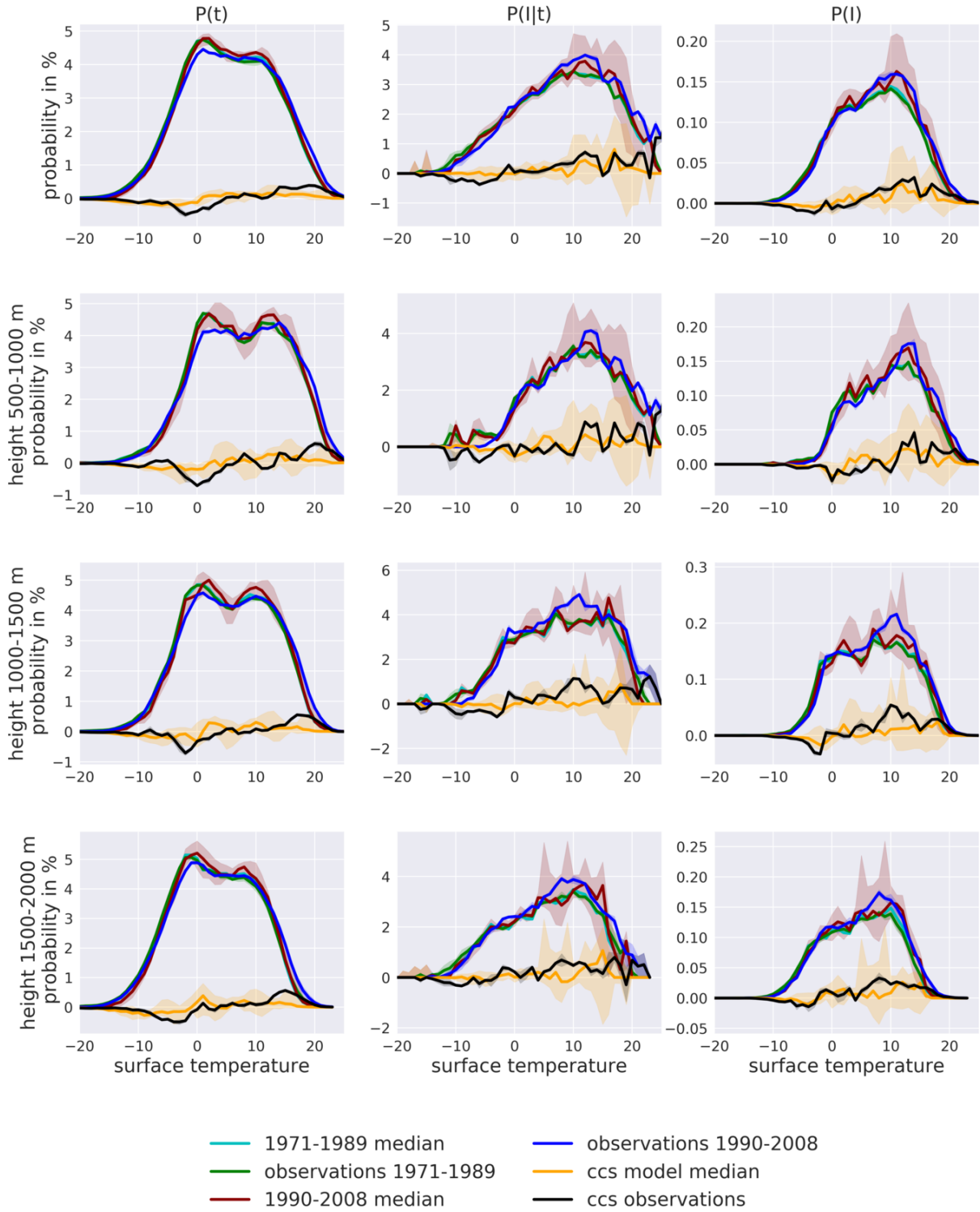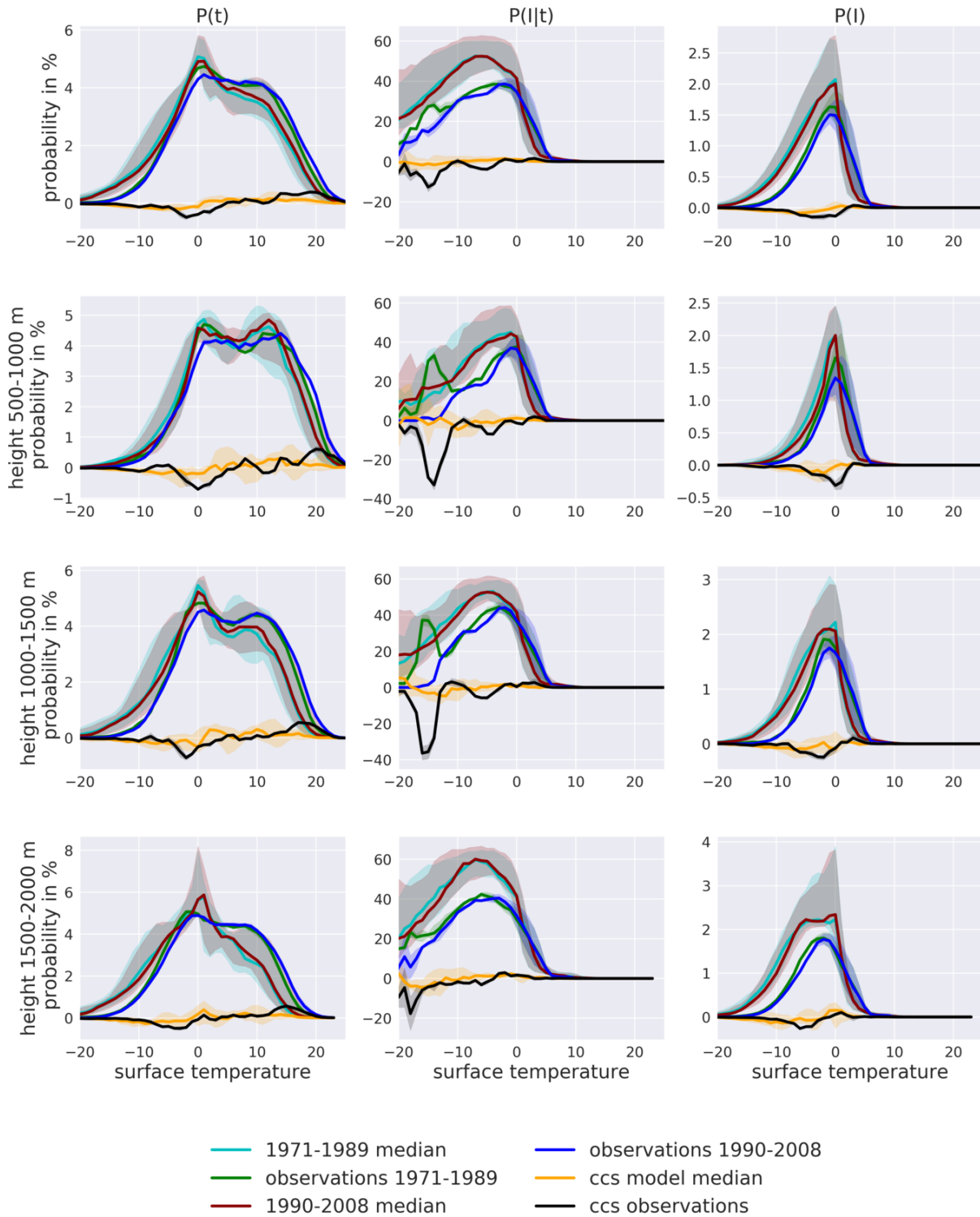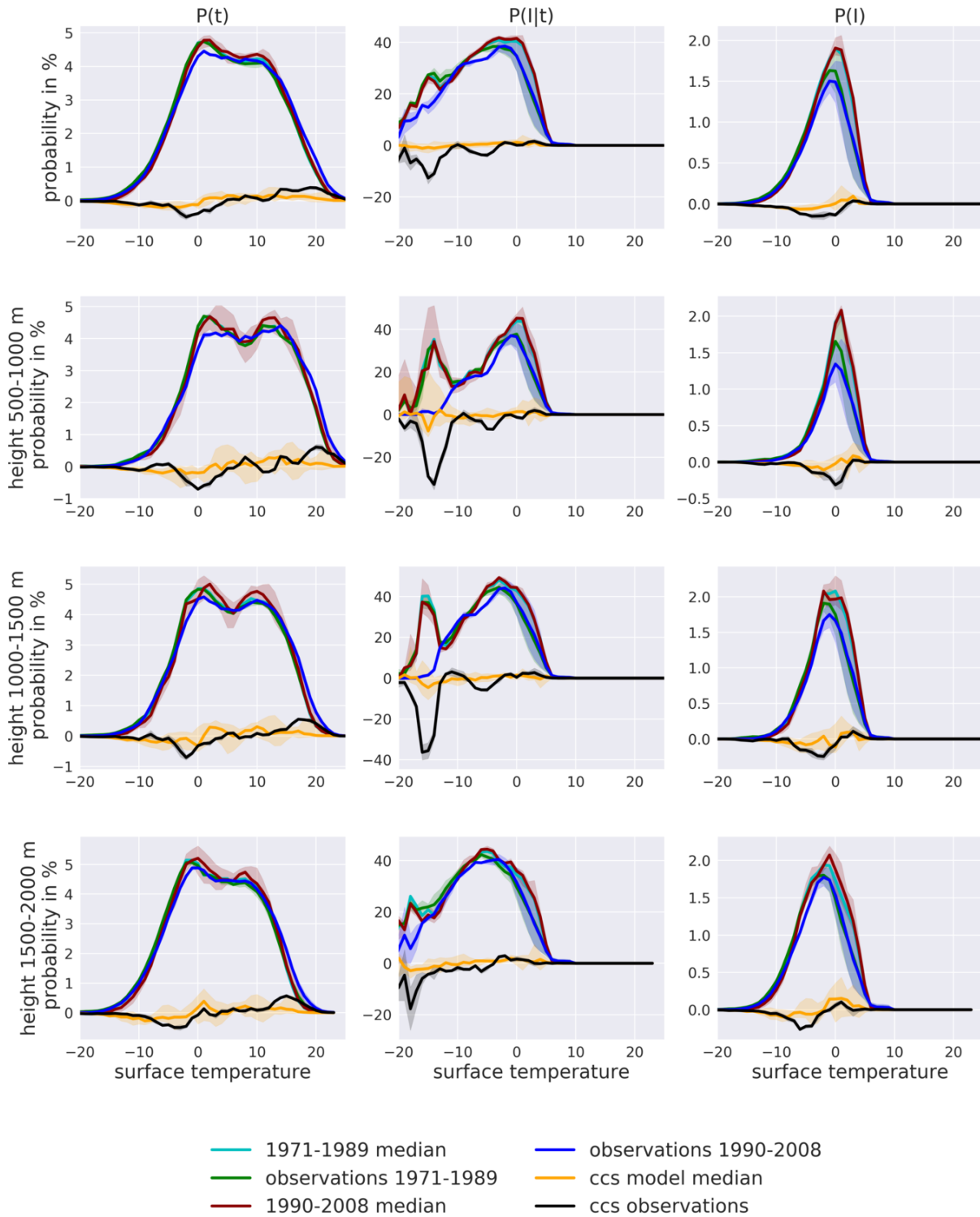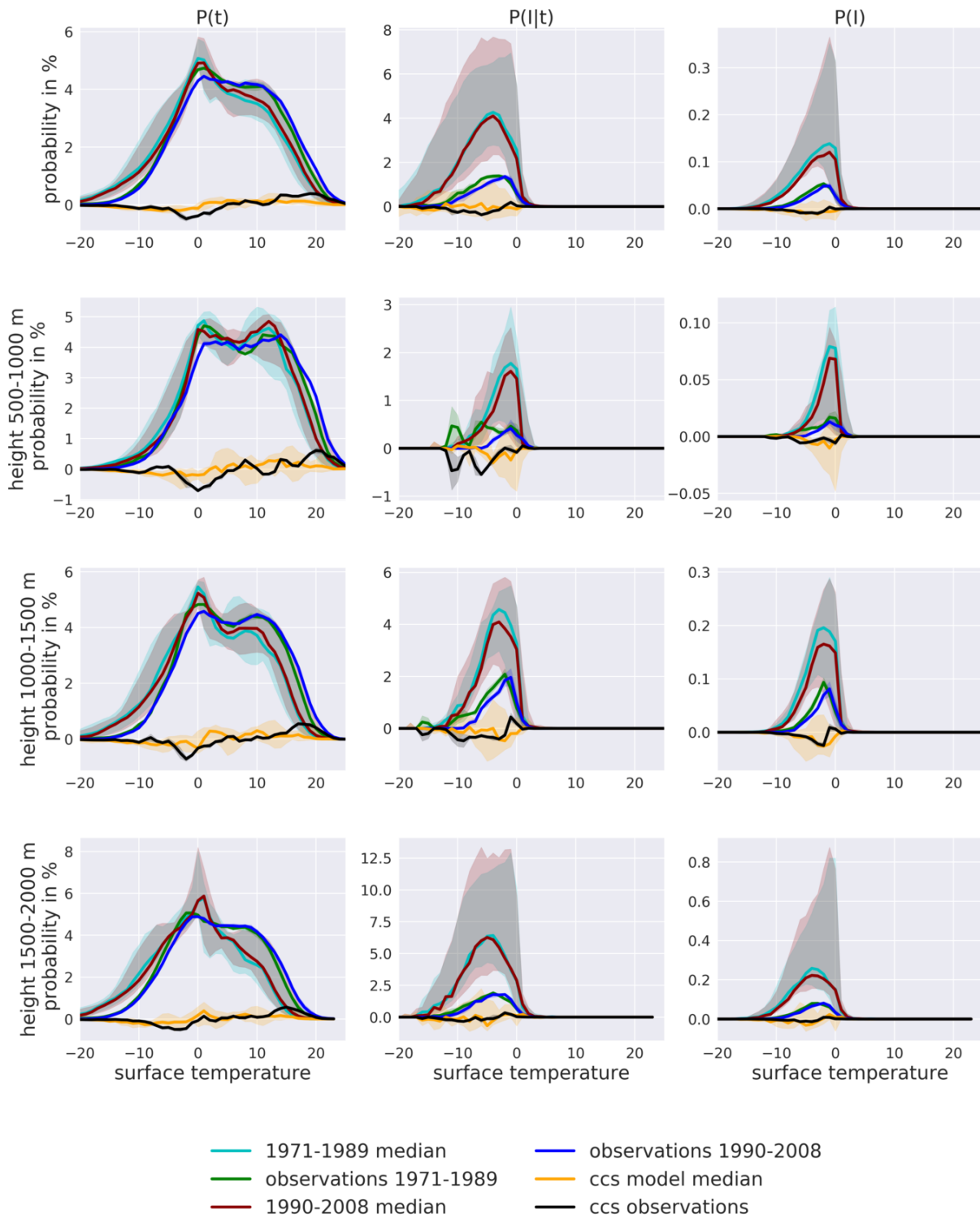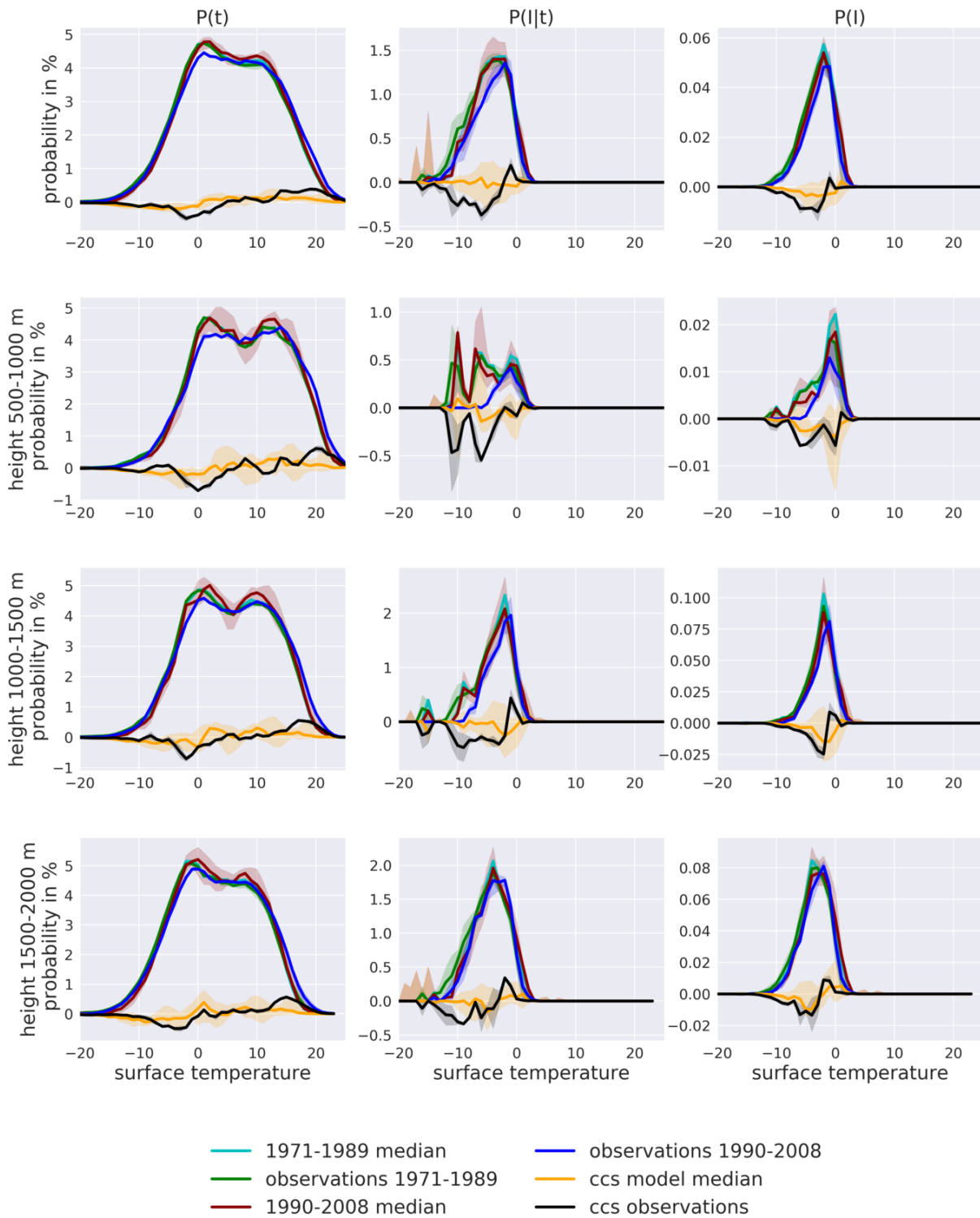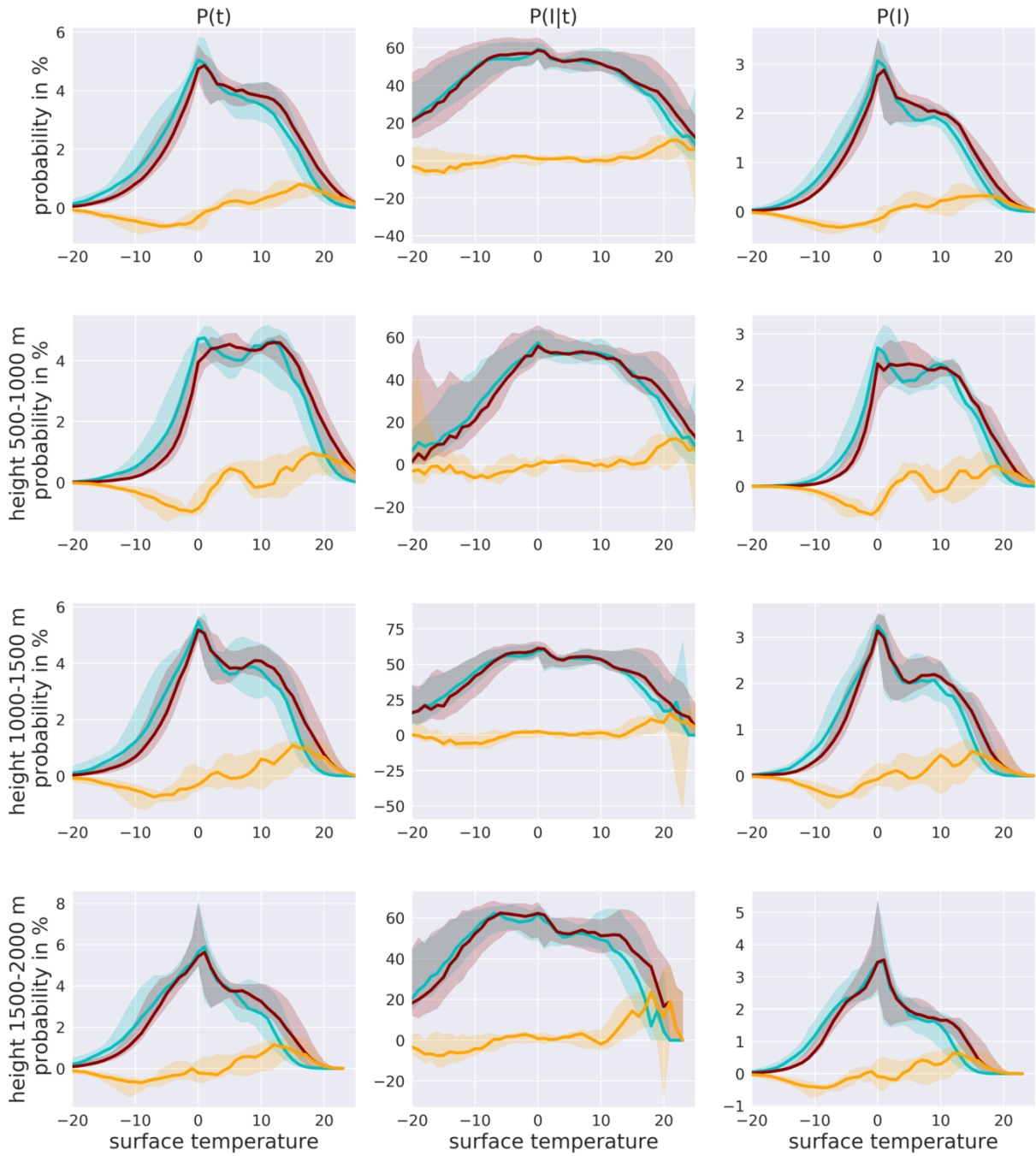
precipitation frequency > 1 mm/day
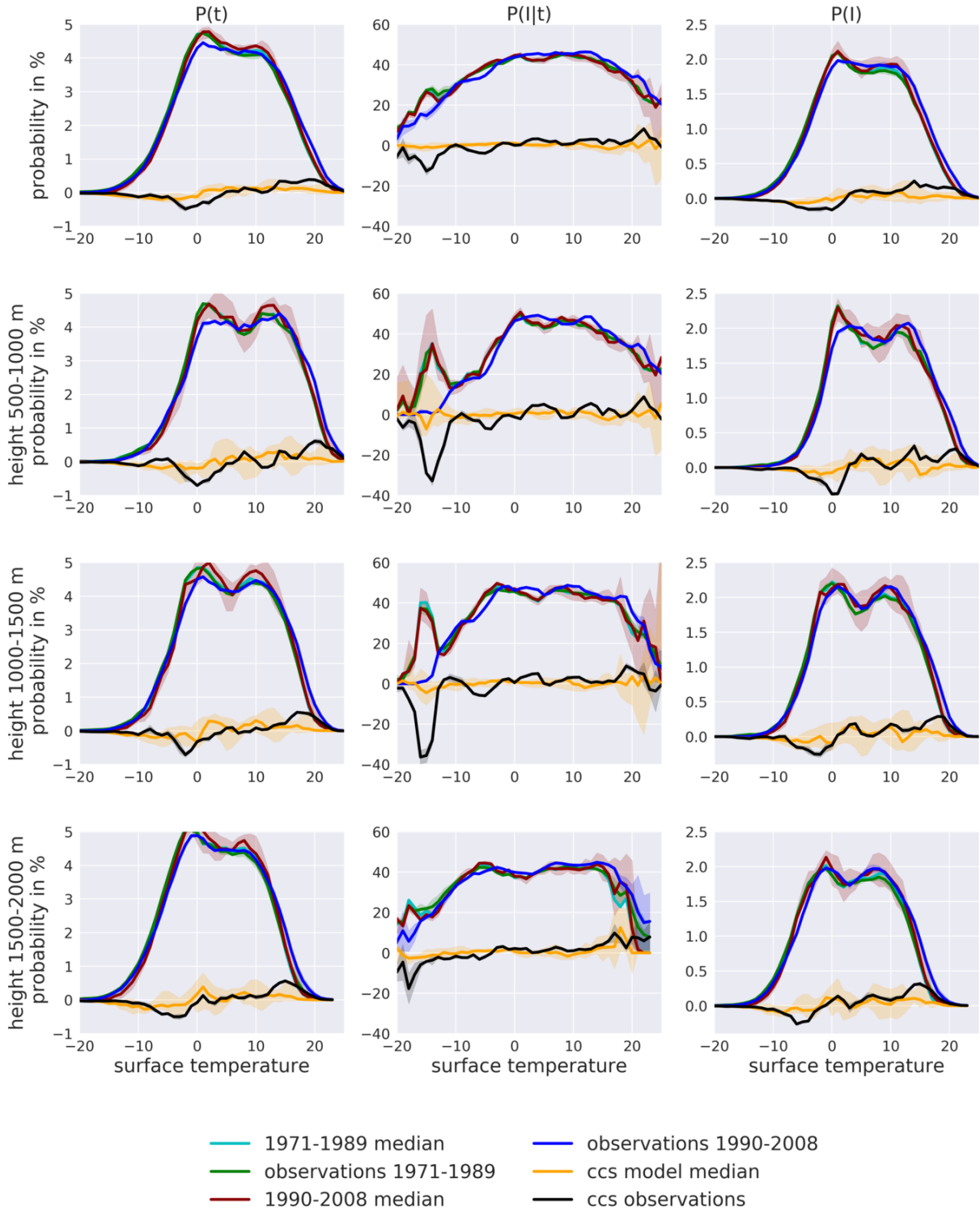
Figure S10: Bias-adjusted temperature dependent probability distributions, P(I), P(I|T) and P(I) for precipitation frequencies, at different altitudes. For the periods 1971-1989 and 1990-2008.

precipitation frequency > 25 mm/day

Figure S11: Temperature dependent probability distributions, P(I), P(I|T) and P(I) for heavy precipitation frequencies, at different altitudes. For the periods 1971-1989 and 1990-2008.
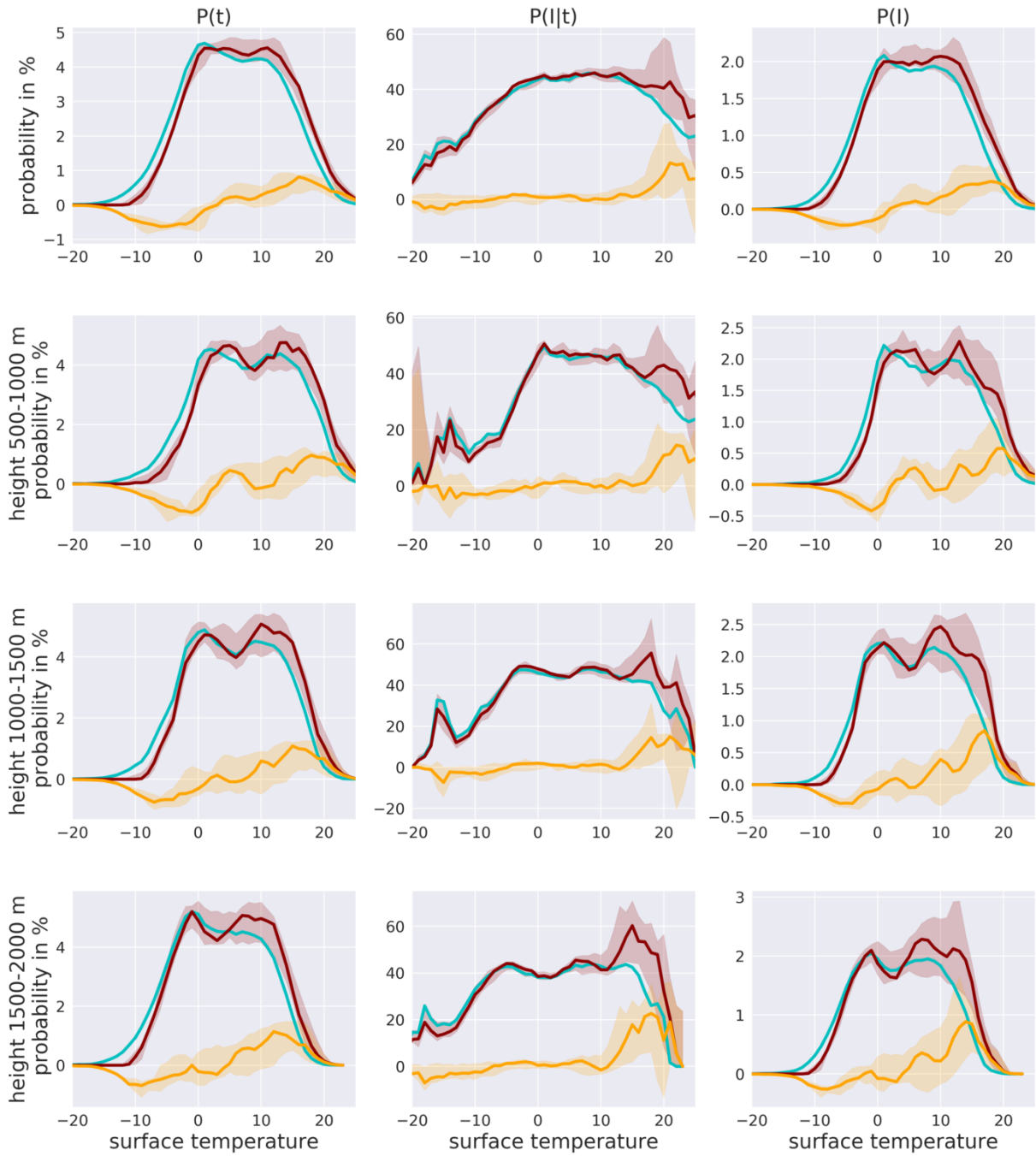
precipitation frequency > 25 mm/day

Figure S12: Bias-adjusted temperature dependent probability distributions, P(I), P(I|T) and P(I) for precipitation frequencies, at different altitudes. For the periods 1971-1989 and 1990-2008.

snow flux frequency > 1 mm/day

Figure S13: Temperature dependent probability distributions, P(T), P(I|T) and P(I) for snowfall frequencies, at different altitudes. For the periods 1971-1989 and 1990-2008.

snow flux frequency > 1 mm/day

Figure S14: Bias-adjusted temperature dependent probability distributions, P(T), P(I|T) and P(I) for snowfall frequencies, at different altitudes. For the periods 1971-1989 and 1990-2008.
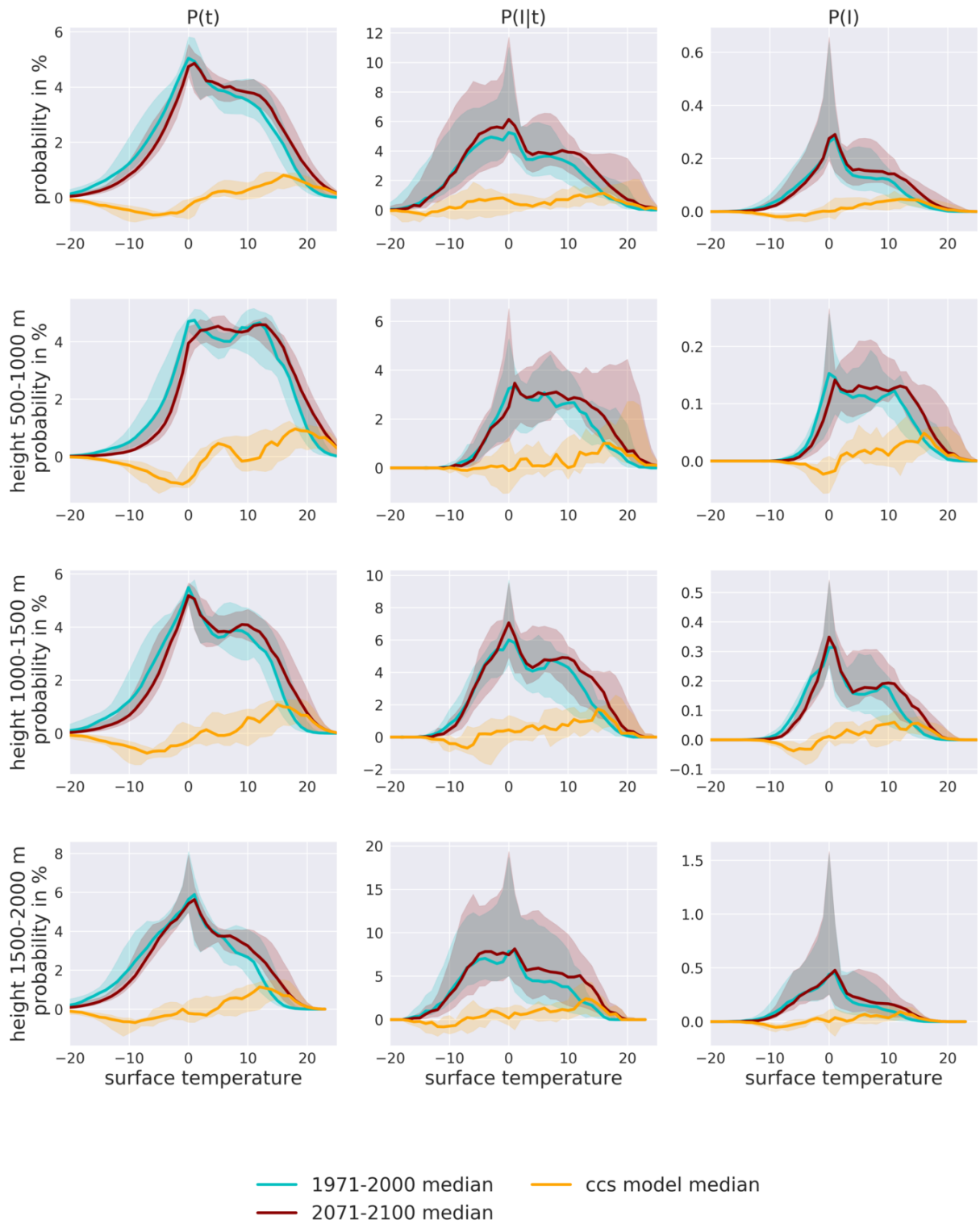
snow flux frequency > 25 mm/day

Figure S15: Temperature dependent probability distributions, P(T), P(I|T) and P(I) for heavy snowfall frequencies, at different altitudes. For the periods 1971-1989 and 1990-2008.

snow flux frequency > 25 mm/day

Figure S16: Bias-adjusted temperature dependent probability distributions, P(T), P(I|T) and P(I) for heavy snowfall frequencies, at different altitudes. For the periods 1971-1989 and 1990-2008.

precipitation frequency > 1 mm/day

Figure S17: Temperature dependent probability distributions, P(T), P(I|T) and P(I) for precipitation frequencies, at different altitudes. For the periods 1971-2000 and 2071-2100.
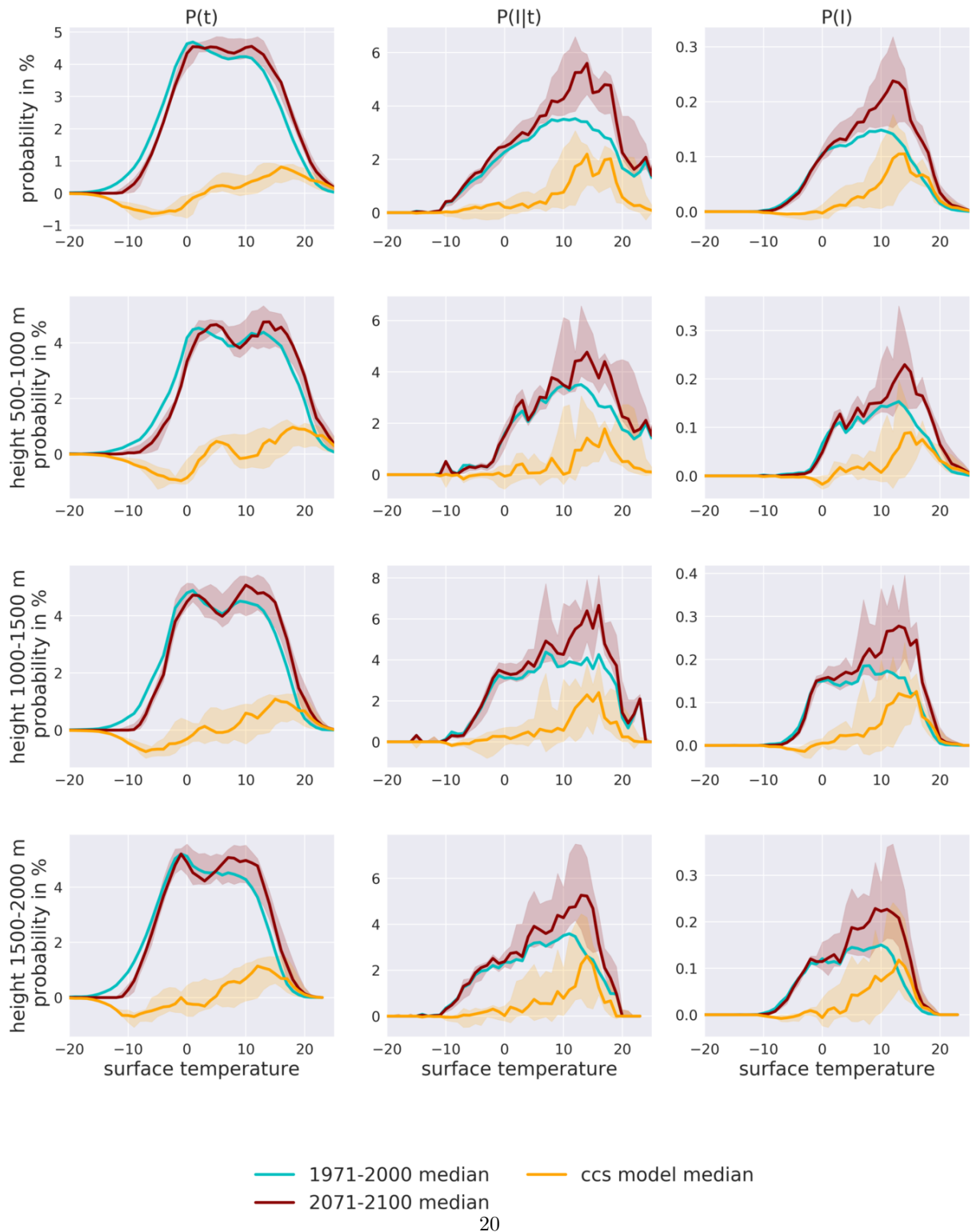
precipitation frequency > 1 mm/day

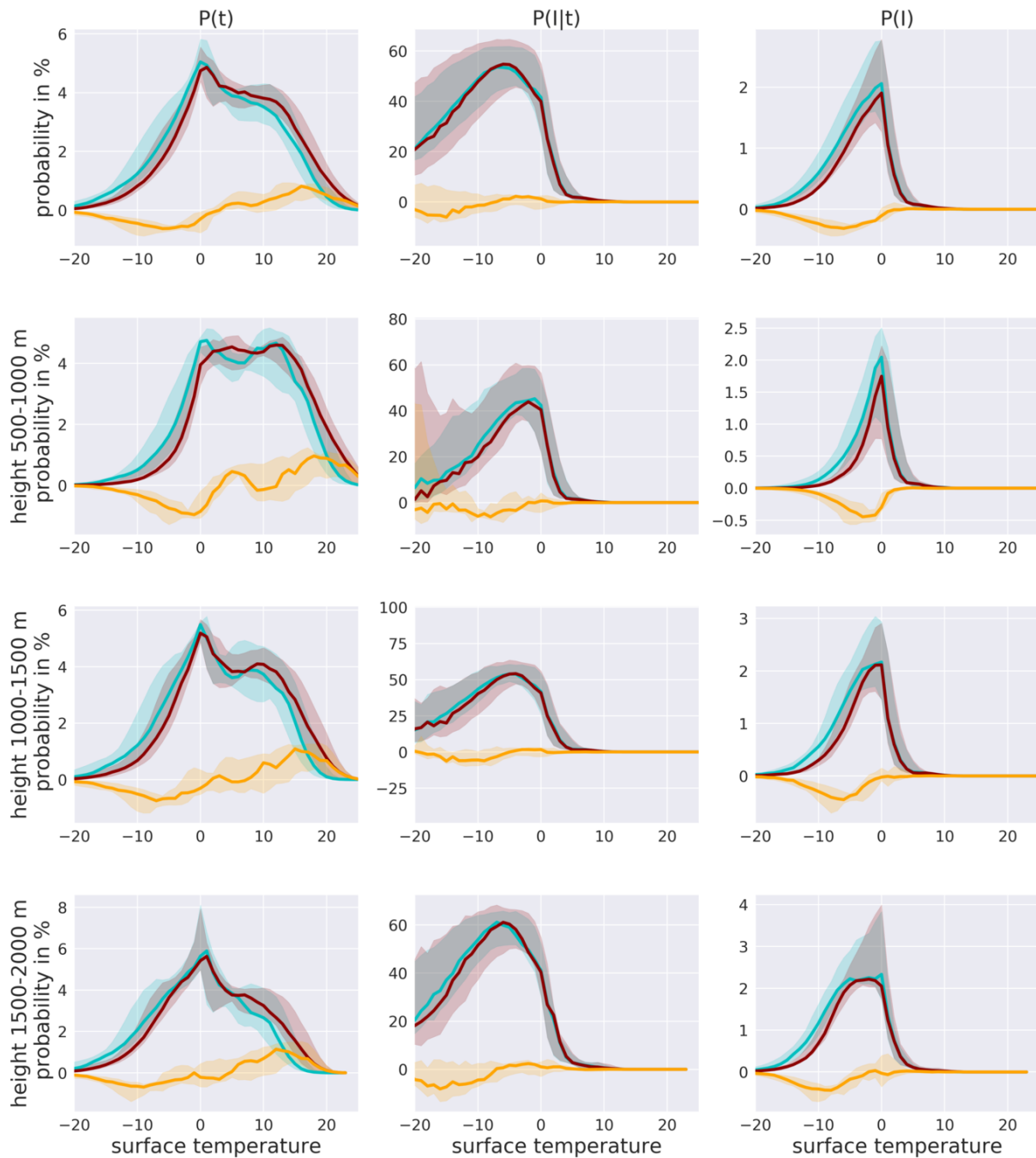Figure S18: Bias-adjusted temperature dependent probability distributions, P(T), P(I|T) and P(I) for precipitation frequencies, at different altitudes. For the periods 1971-2000 and 2071-2100.

precipitation frequency > 25 mm/day

Figure S19: Temperature dependent probability distributions, P(T), P(I|T) and P(I) for heavy precipitation frequencies, at different altitudes. For the periods 1971-2000 and 2071-2100.

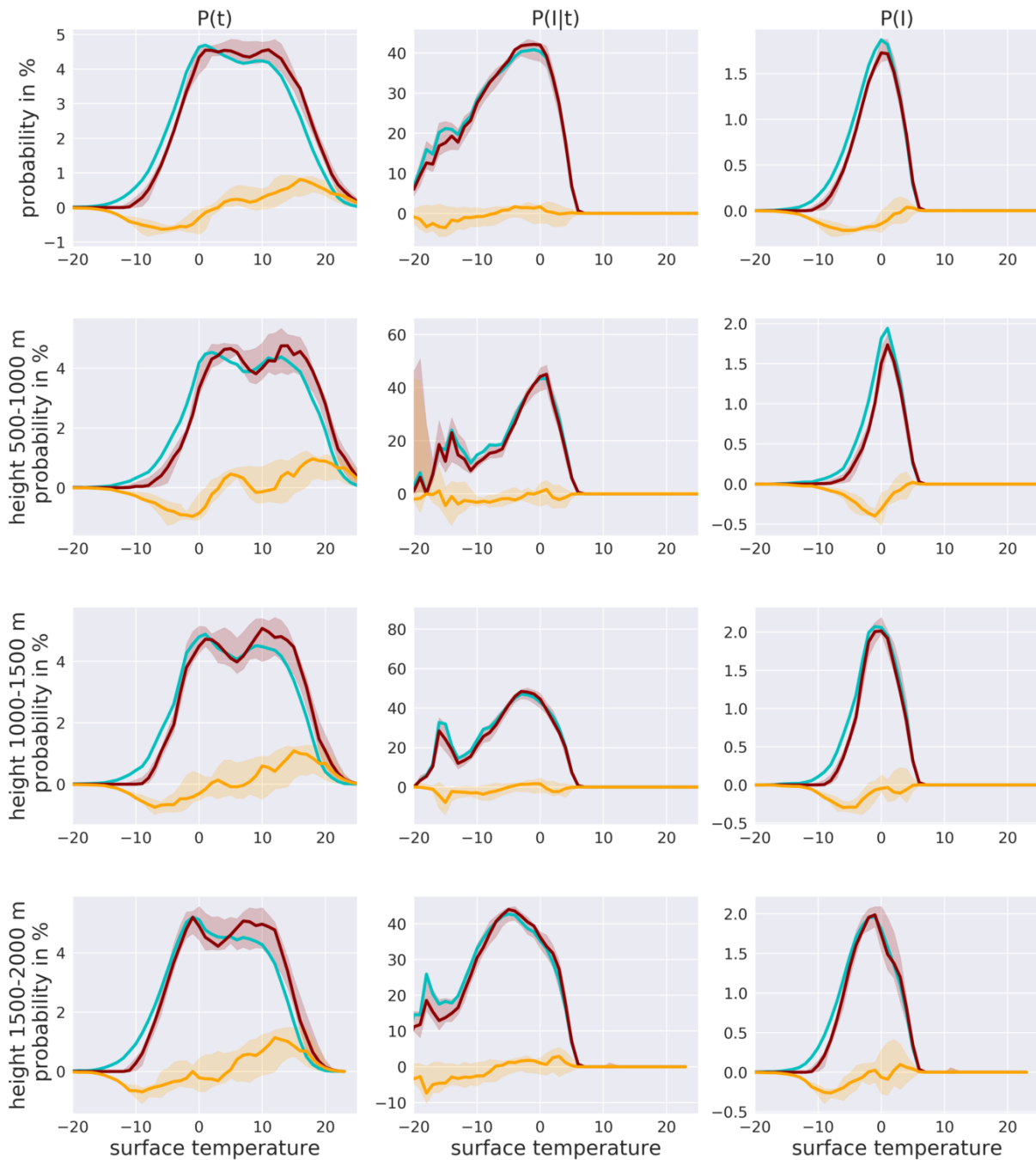# precipitation frequency > 25 mm/day

Figure S20: Bias-adjusted temperature dependent probability distributions, P(T), P(I|T) and P(I) for heavy precipitation frequencies, at different altitudes. For the periods 1971-2000 and 2071-2100.

snow flux frequency > 1 mm/day

Figure S21: Temperature dependent probability distributions, P(T), P(I|T) and P(I) for snowfall frequencies, at different altitudes. For the periods 1971-2000 and 2071-2100.
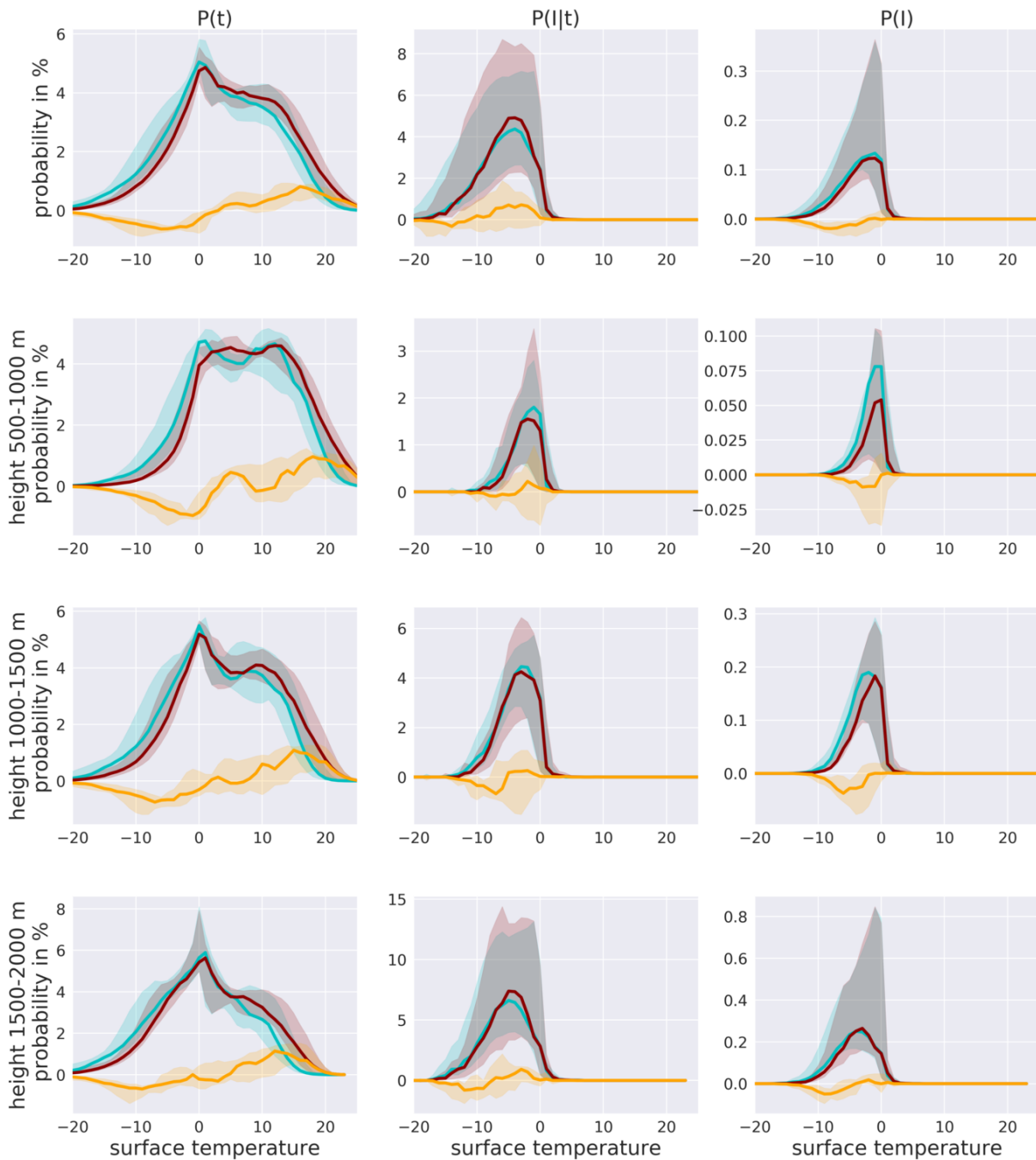
snow flux frequency > 1 mm/day

Figure S22: Bias-adjusted temperature dependent probability distributions, P(T), P(I|T) and P(I) for snowfall frequencies, at different altitudes. For the periods 1971-2000 and 2071-2100.

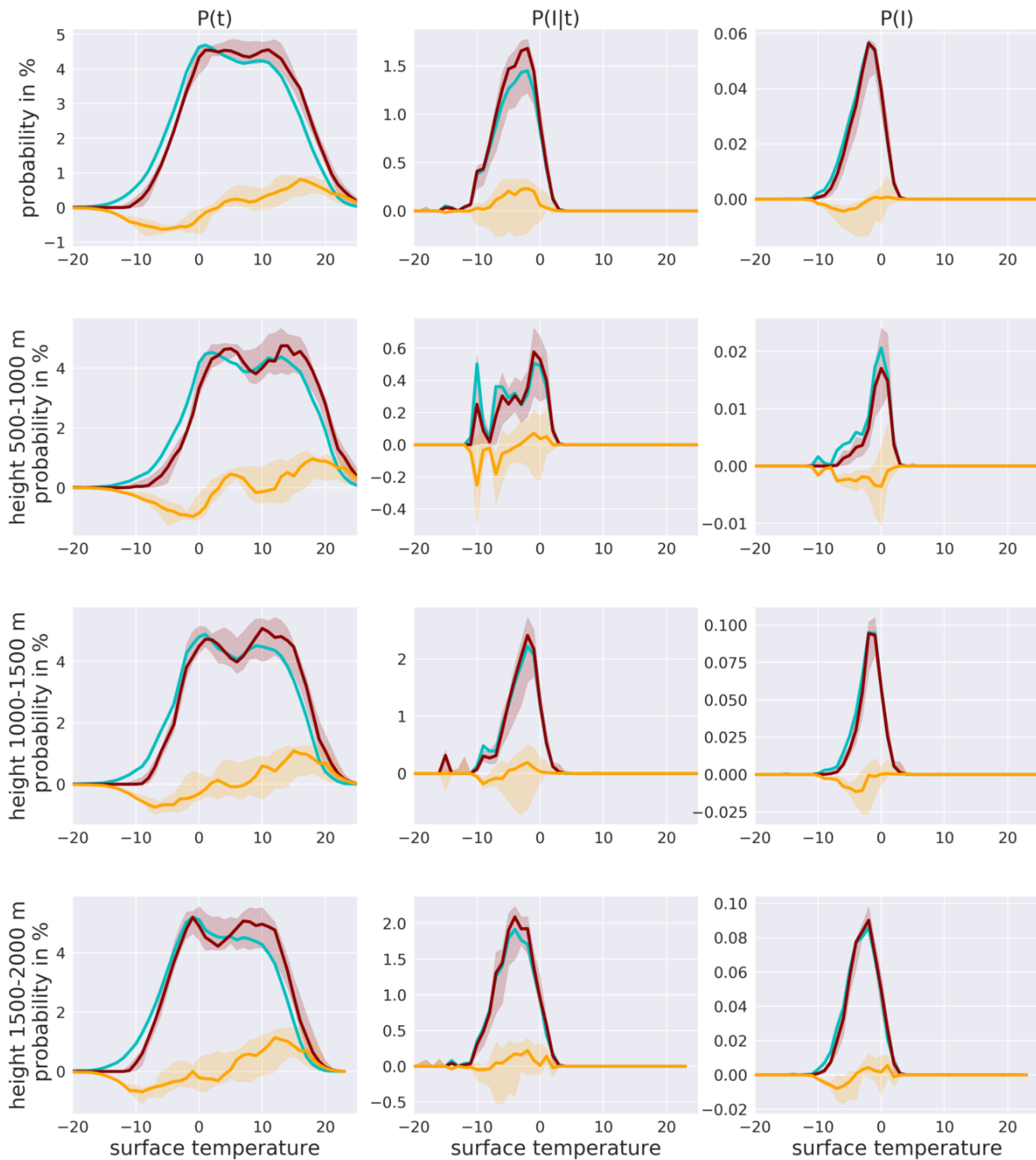Figure S23: Temperature dependent probability distributions, P(T), P(I|T) and P(I) for heavy snowfall frequencies, at different altitudes. For the periods 1971-2000 and 2071-2100.

snow flux frequency > 25 mm/day

Figure S24: Bias-adjusted temperature dependent probability distributions, P(T), P(I|T) and P(I) for heavy snowfall frequencies, at different altitudes. For the periods 1971-2000 and 2071-2100.

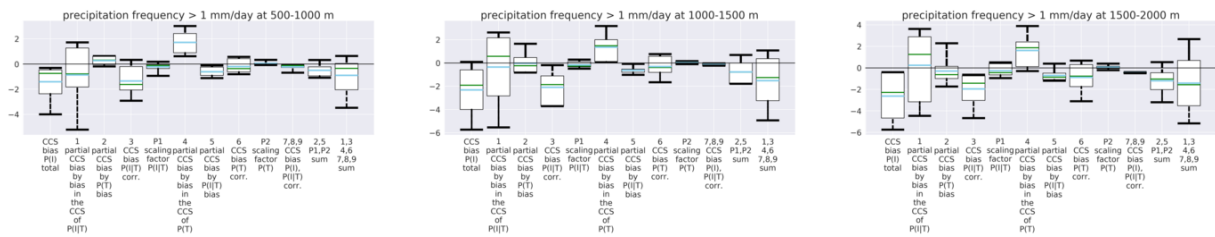Figure S25: Difference in the percentage climate change signal of precipitation frequencies between the model ensemble and the observations within the periods 1971-1989 and 1990-2008 at different altitudes. The Boxblots show the interquartile range of the model ensemble. The whiskers indicate the ensemble spread. The ensemble (mean) median is illustrated by the blue (dark green) line.
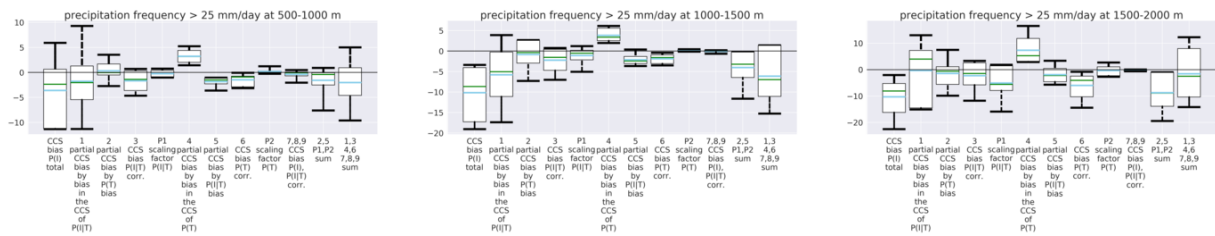

Figure S26: Difference in the percentage climate change signal of heavy precipitation frequencies between the model ensemble and the observations within the periods 1971-1989 and 1990-2008 at different altitudes.

Figure S27: Difference in the percentage climate change signal of snowfall frequencies between the model ensemble and the observations within the periods 1971-1989 and 1990-2008 at different altitudes.
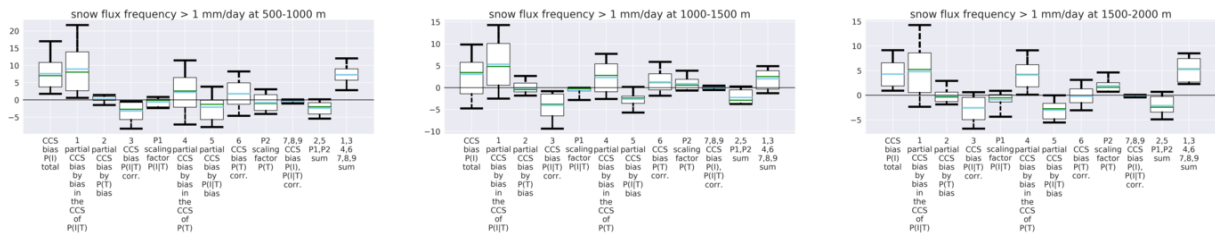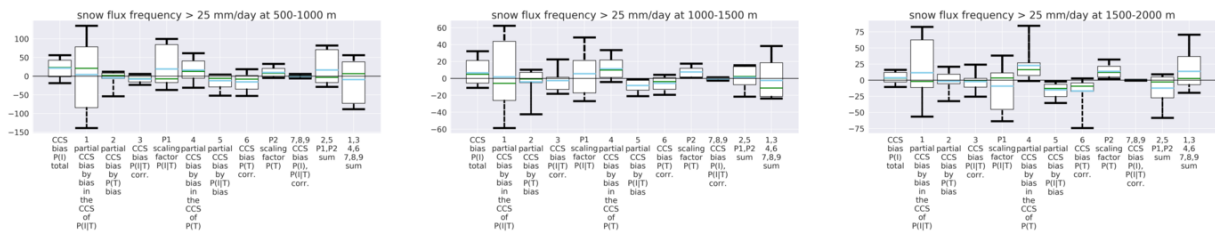

Figure S28: Difference in the percentage climate change signal of heavy snowfall frequencies between the model ensemble and the observations within the periods 1971-1989 and 1990-2008 at different altitudes.

		P(I) hist	P(I) future	P(I) CCS	P(T) dep. CCS	P(I|T) dep. CCS	corr. dep. CCS
SFE1-EOBS	all-all	0.17	0.16	-7.03	-6.98	-0.07	0.02
	500-1000	0.13	0.11	-12.06	-10.08	-2.12	0.14
	1000-1500	0.18	0.17	-4.14	-6.56	2.58	-0.15
	1500-2000	0.20	0.19	-5.46	-6.12	0.51	0.15
SFE1-HISTALP	all-all	0.17	0.16	-8.79	-8.84	-0.13	0.19
	500-1000	0.13	0.11	-14.37	-11.67	-3.17	0.47
	1000-1500	0.19	0.17	-7.21	-9.28	1.89	0.18
	1500-2000	0.21	0.19	-6.44	-7.63	1.01	0.19
SFE2-EOBS	all-all	0.12	0.11	-8.49	-7.06	-1.54	0.10
	500-1000	0.08	0.07	-14.06	-9.34	-5.06	0.35
	1000-1500	0.13	0.12	-7.11	-6.74	-0.33	-0.03
	1500-2000	0.16	0.15	-6.10	-7.04	0.75	0.20
SFE2-HISTALP	all-all	0.13	0.11	-11.58	-10.07	-1.82	0.32
	500-1000	0.08	0.07	-18.39	-12.56	-6.60	0.77
	1000-1500	0.14	0.12	-11.56	-10.81	-1.11	0.37
	1500-2000	0.17	0.15	-8.59	-9.20	0.35	0.25
ENSEMBLE mean	all-all	0.21	0.20	-3.86	-5.46	1.66	-0.06
	500-1000	0.13	0.12	-7.12	-7.81	0.71	-0.03
	1000-1500	0.22	0.21	-4.44	-5.50	1.04	0.03
	1500-2000	0.31	0.30	-2.35	-4.04	1.66	0.04

Table S7: Snowfall frequency changes within the historic period

		P(I) hist	P(I) future	P(I) CCS	P(T) dep. CCS	P(I|T) dep. CCS	corr. dep. CCS
SFE1-EOBS	all-all	0.00	0.00	-11.44	-7.13	-4.29	-0.02
	500-1000	0.00	0.00	-27.15	-8.67	-17.06	-1.42
	1000-1500	0.00	0.00	-10.78	-8.30	-1.97	-0.51
	1500-2000	0.01	0.00	-7.00	-7.96	0.90	0.05
SFE1-HISTALP	all-all	0.00	0.00	-18.26	-10.67	-8.62	1.03
	500-1000	0.00	0.00	-42.09	-13.25	-32.04	3.20
	1000-1500	0.00	0.00	-21.79	-12.40	-11.14	1.74
	1500-2000	0.01	0.01	-10.30	-9.99	-0.86	0.56
SFE2-EOBS	all-all	0.00	0.00	-14.64	-6.97	-7.83	0.17
	500-1000	0.00	0.00	-40.43	-6.20	-33.15	-1.08
	1000-1500	0.00	0.00	-15.63	-8.33	-6.93	-0.38
	1500-2000	0.00	0.00	-8.23	-8.82	0.49	0.09
SFE2-HISTALP	all-all	0.00	0.00	-20.29	-11.21	-10.32	1.24
	500-1000	0.00	0.00	-54.23	-12.78	-45.83	4.38
	1000-1500	0.00	0.00	-26.27	-13.35	-15.20	2.28
	1500-2000	0.01	0.00	-11.11	-10.91	-0.74	0.55
ENSEMBLE mean	all-all	0.02	0.02	-8.06	-5.63	-2.38	-0.05
	500-1000	0.00	0.00	-19.51	-7.25	-12.96	0.70
	1000-1500	0.01	0.01	-12.00	-5.02	-7.15	0.17
	1500-2000	0.03	0.03	-5.51	-4.09	-1.58	0.16

Table S8: Heavy snowfall frequency changes within the historic period
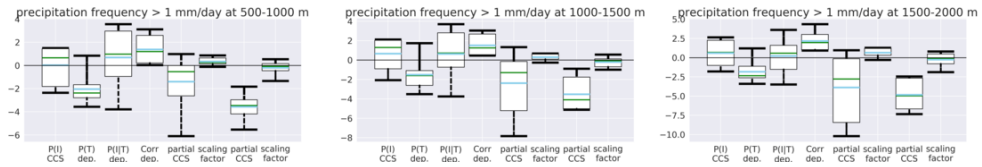

Figure S29: Precipitation frequency changes by the end of the century at different altitudes. Boxplots as in Fig. S25.

		P(I) hist	P(I) future	P(I) CCS	P(T) dep. CCS	P(I|T) dep. CCS	corr. dep. CCS
ENSEMBLE mean	all-all	0.50	0.51	0.37	-1.69	1.33	0.73
	500-1000	0.48	0.49	0.03	-2.04	0.69	1.38
	1000-1500	0.51	0.52	0.66	-1.50	0.64	1.52
	1500-2000	0.52	0.54	0.59	-1.82	0.23	2.18

Table S9: Precipitation frequency changes by the end of the century at different altitudes
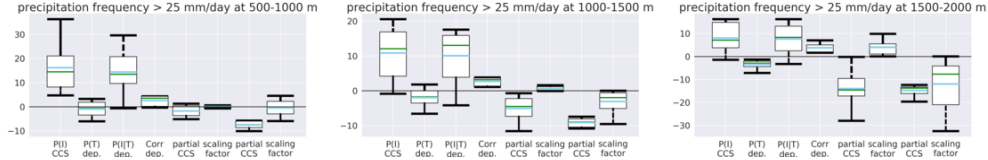

Figure S30: Heavy precipitation frequency changes by the end of the century at different altitudes. Boxplots as in Fig. 25.

		P(I) hist	P(I) future	P(I) CCS	P(T) dep. CCS	P(I|T) dep. CCS	corr. dep. CCS
ENSEMBLE mean	all-all	0.04	0.04	10.77	-4.52	13.65	1.63
	500-1000	0.02	0.03	16.20	-0.87	14.44	2.63
	1000-1500	0.04	0.04	10.86	-1.92	10.07	2.72
	1500-2000	0.05	0.06	8.06	-3.37	7.63	3.81

Table S10: Heavy precipitation frequency changes by the end of the century at different altitudes
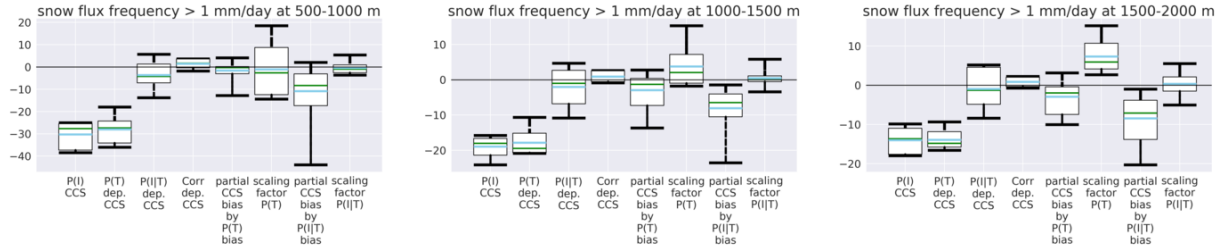

Figure S31: Snowfall frequency changes by the end of the century at different altitudes. Boxplots as in Fig. 25.

		P(I) hist	P(I) future	P(I) CCS	P(T) dep. CCS	P(I|T) dep. CCS	corr. dep. CCS
ENSEMBLE mean	all-all	0.20	0.17	-18.73	-19.09	0.14	0.22
	500-1000	0.13	0.09	-30.24	-28.00	-3.67	1.43
	1000-1500	0.21	0.18	-18.94	-17.84	-2.01	0.91
	1500-2000	0.29	0.26	-13.99	-13.91	-0.92	0.85

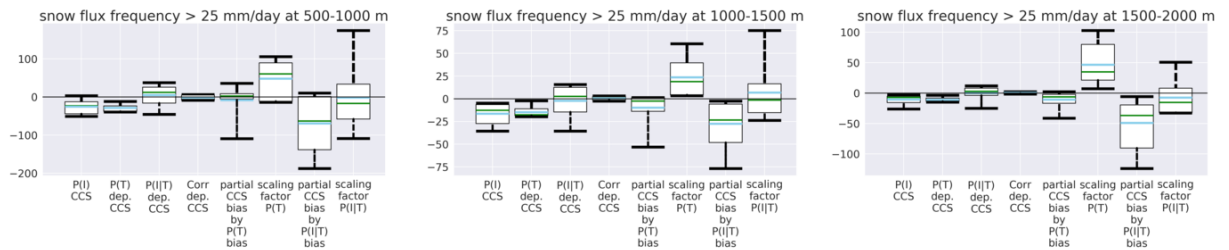Table S11: Snowfall frequency changes by the end of the century at different altitudes

Figure S32: Heavy snowfall frequency changes by the end of the century at different altitudes. Boxplots as in Fig. 25.

		P(I) hist	P(I) future	P(I) CCS	P(T) dep. CCS	P(I|T) dep. CCS	corr. dep. CCS
ENSEMBLE mean	all-all	0.01	0.01	-13.29	-19.18	7.22	-1.33
	500-1000	0.00	0.00	-25.39	-28.46	4.84	-1.77
	1000-1500	0.01	0.01	-16.41	-14.69	-2.26	0.54
	1500-2000	0.03	0.03	-10.47	-11.03	-0.02	0.59

Table S12: Heavy snowfall frequency changes by the end of the century at different altitudes