SCIENCE AT THE ENVIRONMENTAL RESEARCH STATION SCHNEEFERNERHAUS/ ZUGSPITZE

Prof. Dr. Michael Bittner; Ed.
(Coordinator Science Team UFS)

Umwelt
Forschungsstation
Schneefernerhaus

Table of contents

11 Statistical downscaling of future global climate change scenarios for Alpine high mountain regions

Andreas Philipp, Christoph Beck, Severin Kaspar, Stefanie Seubert and Jucundus Jacobeit

Institute for Geography, University of Augsburg

Keywords: Climate change, Statistical Downscaling, High mountain climate

Abstract

Global climate change is expected to show considerable impacts on the European Alpine high mountain region. However, global climate models show limitations concerning regional and local scales. Therefore, dynamical downscaling and statistical downscaling techniques have to be applied. Exemplarily for an especially simple method of statistical downscaling the Reference Class Forecast (RCF) method is explained and applied to the climatic time series at Zugspitze and Sonnblick. As a very different method, the technique of Artificial Neural Network (ANN) for downscaling is introduced. An important step is to find out an optimal set of predictors. However, if predictor screening is done appropriately, remarkably high skill scores can be achieved, which allow for confidence on the projected future assessments. The results of the downscaling approach applied for future scenarios are discussed concerning the degree of warming at the presented example stations and the changes in precipitation which not only show reduced rainfall in future summers, how it is assumed widely in the literature, but partly suggest the possibility of future increases in summer rainfall for the Zugspitze at the northern edge of the Alps in contrast to the Sonnblick in the center of the Alps. Reasons are probably higher transport rates for humidity in warmer air masses which are more relevant at the edge of the Alpine ridge, while the interior might be affected more by increased anticyclonicity.

11.1 Intro: Alpine regions exposed to climate change

Climate change in the Alps has wide-ranged implications due to strong interrelations between the different spheres (atmosphere, hydrosphere, cryosphere, lithosphere, pedosphere and biosphere) additionally intensified by distinct spatial complexity. The following short overview aims to give an idea about the broad potential effects of possible future temperature und precipitation changes and thus outlines the significance of the subsequently presented downscaling efforts and results.

With an observed increase of +2 °C (total annual mean temperature) since the late 19th century, air temperature in the Alps rises twice as much as the Northern hemisphere average (Auer et al. 2007). Since 1980 the recent warming has further accelerated and a faster increase of annual mean temperature (+0.5 °C per decade) is observed (European Environment Agency 2009). Alpine glacier retreat follows that trend and also increased in speed from 1980 onward (European Environment Agency 2017a). Since the beginning of the 20th century Alpine glaciers have lost nearly half of their ice masses (European Environment Agency 2017a, Huss 2012), an almost complete loss of their current volume is estimated until 2100 (84 and 90 % under RCP 4.5 and 8.5, Radić et al. 2013).

While increasing summer temperatures are regarded most important for this development, further future warming is expected for all seasons and the whole Alpine region. Until the end of the 21th century air temperature is projected to rise to +3,3 °C on annual average (Gobiet et al. 2014). Models agree on the sign of the expected change and emphasize the robustness of the warming signal (Heinrich et al. 2013).

Precipitation changes at present vary stronger concerning seasonal and regional distributions as well as the observed period (Gobiet et al. 2014). Regarding the spatial distribution of annual means a north-west to south-east gradient from slight increases to significant decreases during the 20th century is shown (Brunetti et al. 2006). This north-south oriented distribution is expected to sharpen until the end of the 20th century with more precipitation in the northern Alps in winter, according to the Northern Europe pattern of change, and reduced precipitation in the southern Alps in summer, following the climate change signal in Southern Europe (Gobiet et al. 2014).

However, not only the enhanced warming or mean precipitation changes alone are critical, but also their implications for other meteorological and hydrological variables, mainly air humidity, precipitation variability and above all the resulting impacts on the hydrological storage terms, i.e duration and depth of snowcover and glacier mass balance. Slight warming of just a few centidegrees can produce strong changes in the water budget e.g. if the melting point is reached for certain areas. The above mentioned accelerated glacier melting since 1980 for example has increased the glacier contribution to late summer runoff of four main European rivers originating in the Alps, namely Danube, Rhine, Rhone and Po (by around 13%, Huss 2011). Thus also the lowland parts, especially of catchments with high portions of glacial melt water in late summer, will be affected by future glacier retreat and lacking runoff contributions on the long-term. Lowered ground water levels, restricted water availability for agriculture or limitations of ship traffic along the main European streams are mentioned as potential risks of the future warming-induced glacier retreat in the Alps (Huss 2011).

With the consequences for the local economic sectors in mind, strong implications are expected from future changes of snow cover, entailing limitations for hydrological power generation (Kobierska et al. 2012) or especially the winter tourism (Steiger et al. 2010). For the entire Alps a dramatic decrease of snow cover duration and amount is projected until the end of 21th century, mainly for altitudes below 1500–2000 m (Gobiet et al. 2014, Steiger et al. 2013). Above, gains are expected (due to potentially increasing heavy precipitation), which leads amongst others to an increased avalanche activity, e.g in the Western Alps in winter (Castebrunet et al. 2014).

Higher temperatures as well as less solid but more liquid precipitation during the winter half year have particular strong effects in high altitude regions with high relief energy: wet winters e.g can reinforce the landslide activity in spring (Stoffel 2014). Being identified as one of the European "susceptibility hotspots for weather-induced landslides" (European Environment Agency 2017a), shallow landslides like rock falls, debris flows/avalanches but also ice falls and snow avalanches are expected to rise with future temperature and precipitation changes (Stoffel 2014). This gives rise not only to remarkable risks for summer tourism or transhumance but possibly endangers settling in the Alps in general.

In conclusion mountain landscapes are characterized by an exceptional complexity of geofactors, enhancing the vulnerability to climate change. The above mentioned complex structure of interrelations is responsible for amplifying even small irregularities in the input variables, in this context temperature and precipitation. To know their future changes and associated temporal and regional variations as exactly as possible is an essential base for accurate assessments of adaptation strategies.

All the consequences for the inanimate parts of nature are more dramatic due to the high relief. Thus, retreating glaciers and permafrost regions change the hydrological cycle and cause land and rock slides, to mention only a few aspects which can affect also human society directly and harder than in low lands. The Alps besides the Pyrenees are identified as hotspots.

11.2 Global Climate Change and Modelling

Global climate models allow to estimate the reaction of the climate system of the earth to climate forcing factors in a quantitative way. However, in order to be able to use the information provided by climate models, it is absolute essential to understand in principle how they work and to know about the strength and weakness of this scientific tool. The core of each climate model is a general circulation scheme including physical laws of i) conservation of energy, ii) conservation of momentum and iii) conservation of mass. These three laws can be connected by the equation of state and transformed into a system of prognostic equations, which allows to calculate the state of the climate system based on its state at a time step before. In a strong-

11

ly simplified way it could be said, that e.g. the distribution of airmasses (measured as air pressure) is causing winds which itself redistribute air and change the pressure field in a second time step, which itself causes – now modified – winds causing a newly modified pressure field in a third time step and so on and so on. However, beside wind and pressure, it is of course neccesary to include density, heat, moisture, heat capacity etc., in order to represent all dependencies of the system. The resulting set of non-linear partial differential prognostic equations can be resolved not analytically but only by extensive numerical methods (which is the real reason for calling them numerical models). However, even if all processes of the atmosphere are included, there remains a fundamental problem for all models: the spatial, and depending on this, the temporal resolution of the calculations. In order to get a realistic description of the distribution of a certain variable, let's say air pressure, over the globe, a regular network of grid points, e.g. on 32 latitudes and 64 longitudes what was state of the art in 1990ies, is defined.

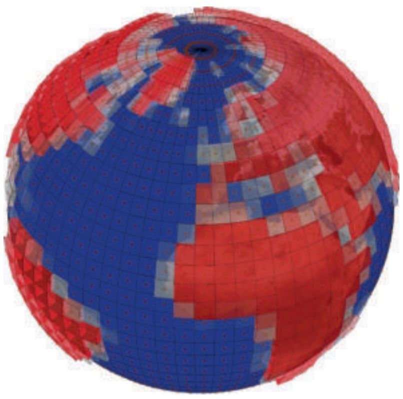

Fig. 1: Grid cells in T21 horizontal resolution used to describe the global land-sea-mask for GCMs. Values vary between 0.0 (blue) for grid cells completely covering ocean areas and 1.0 (red) for cells completely on land. Note that pixels covering coast lines as for western Africa are mixture pixels, half sea half land, indicating the low level of detail for such a low resolution.

However, with a limited number of discrete grid points for describing e.g. a low pressure system, only a limited degree of detail or only systems of a certain minimum size can be described. Although there is the possibility to represent the pressure changes over space by a set of overlaid continuous sinus functions, following the scheme of Fourier series, which is called the spectral representation and is used for some effective calculations in numerical models, only a limited number of functions can be used which also allows only for a limited degree of detail. In order to increase the detail level, the number of grid points or, equivalently, the number of wave functions, most often described by the so called triangular truncation number (e.g. T21), must be increased.

Tab. 1: model resolutions

Truncation	lat × lon	km at Equator	deg at Equator
T21	32x64	625	5.625
T42	64x128	310	2.8125
T62	94x192	210	1.875
T63	96x192	210	1.875
T85	128x256	155	1.4
T106	160x320	125	1.125
T255	256x512	60	0.703125
T382	576x1152	38	0.313
T799	800x1600	25	0.225

Table 1 shows a set of most commonly used horizontal spatial resolutions for numerical circulation models. It is apparent, that decreasing the distance between two grid points by a half, at the same time the number of grid points quadruplicates. If an increase of the number of levels in the vertical direction is considered additionally, it is clear that the computational effort increases dramatically, when the spatial resolution and thus the number of calculations for all grid points or spectral functions is increased. (National Center for Atmospheric Research Staff 2017).

However, the situation is even more difficult, because of the so called Courant-Friedrich-Levy (CFL) criteria.

$$|u \cdot dt/dx| \leq 0$$

This criterion is saying that the time step dt for the prognostic computations must be smaller than the distance between two grid points dx divided by the speed of flow in the model, e.g. wind u. Since the speed of flow in the model is something which is given by the physical circumstances, e.g. the subtropical jet stream in the upper troposphere, only the spatial (dx) and temporal resolution (*dt*) can be changed by the modeller. If a certain spatial resolution is chosen, it subsequently demands for a certain maximum time step length with which the set of prognostic equations is solved repeatedly. The principle of the CFL criterion is saying nothing else than that a flow of mass, impulse or energy should not be faster, than that it reaches a distance not more than one grid point far within one time step, or in other words, no transport process within the model should skip a grid point along its way. What happens, when this criterion is violated can be seen in Fig. 2: the variables show a totally unrealistic pattern and extreme values outside the physically plausible range. The model „crashes".

Essentially it is the CFL criterion which prevents the model resolution becoming much more increased, while a limited capacity of compute power is available, even though, it is numerical modelling for which the most extensive compute clusters are build. However, when the spatial (and accordingly the temporal) resolution is limited, small and short term phenomena in the atmosphere, like thunderstorm cells or showers of rain, may be missed because they happen on the sub grid scale.

Another problem which is limiting the detail level of general circulation models, is just the complexity of certain phenomena in the atmosphere, which cannot be expressed by a reasonably small set of equations. One prominent example is precipitation. The process of generating precipitation is such complex that, apart from its high spatial variability, it is virtually impracticable to simulate it with the necessary precision directly. Especially the initiation and growth of cloud and later rain droplets interacting with condensation nuclei or its changes between ice and liquid phase is a highly complex subject of microphysics which has to consider effects down to the molecular scale. Calculation of all these effects, if they are understood at all, for the whole globe is simply impossible. In order to still include those processes, rough, empirical estimations of their quantitative dependence on its most important influences, described by certain parameters are used. Without these so called parametrisations, a climate model which should include most of the relevant processes is unimaginable and impracticable.

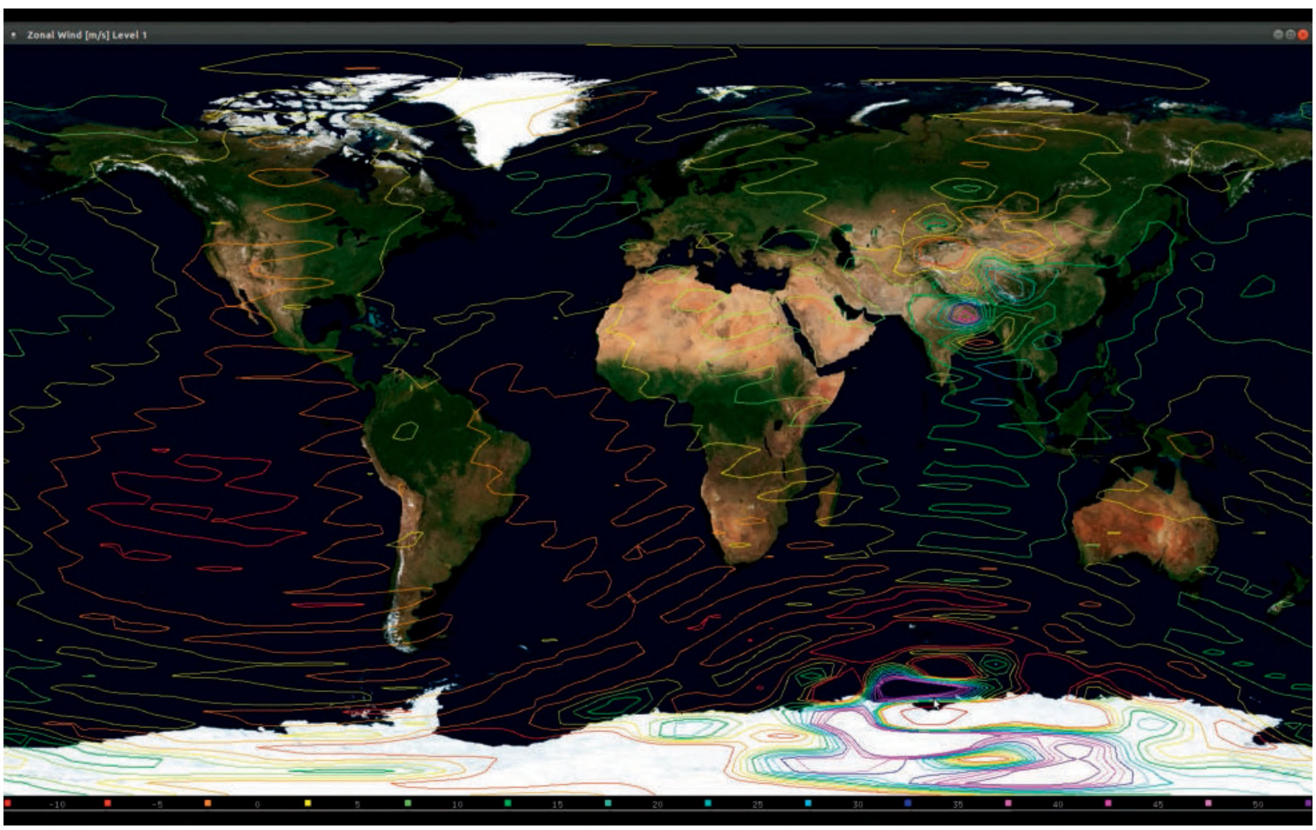

Fig. 2: Zonal wind component at surface level as displayed by the PLASIM interactive model environment (Fraedrich et al. 2005) immediately before a model crash due to violation of the Courant-Friedrichs-Levy criterium (see text). Negative values (easterly wind) is colored orange to red (–15 m/s), positive values (westerly wind) yellow to violet (55 m/s). Model resolution was T21, time step intervall was 60 minutes. Note the unrealistic wave like pattern of alternating easterly and westerly winds extending from a center at the Antarctic coast with common high wind speeds, where the CFL-criterium was violated first.

However, both the spatial and the processual limited resolution obstructs a high level of detail of the model output, concerning the spatial details but also the processual details. This problem is not the same for all variables, thus, the global pressure field is much less affected than the temperature or the precipitation field (Raisanen 2007), since the complexity of the relevant processes is lower. Additionally the vertical level is important, thus, temperature in the mid and high troposphere can be modelled much more precisely than directly above the surface in the planetary boundary layer because there the influences are much more divers.

To conclude this, it can be said that general circulation models allow to simulate the global atmospheric state realistically on a rough level of detail. However the high demand for compute power, which is caused by the spatio-temporal and processual resolution, when simulating processes directly, prevents from a sufficient detail level of some variables especially near the surface, where applications of model outpout data mostly take place. In order to solve this problem several downscaling methods have been developed.

11.3 Downscaling methodology

Downscaling techniques try to close the gap between the low level of detail provided by general circulation models and the needed high level of detail for applications, which is caused mainly by the difference in spatial scale between global models and local applications.

In principle two main groups of downscaling techniques may be discerned: dynamical downscaling using regional climate models (RCM) on the one hand and statistical downscaling on the other. The latter using transfer functions to apply empirically determined dependencies between local variables of interest and global or large scale circulation data generated by GCMs for distinct scenarios. Other ways to categorize downscaling techniques are suggested by Maraun et al. (2010), however for the sake of clarity the two-fold distinction between dynamical and statistical methods is preferred here. Both ways have their advantages and disadvantages. For dynamical downscaling a highly resolved numerical simulation model, the RCM, is nested into a GCM. This means that only a certain area of the earth is simulated on a high resolution level, while the boundary conditions at the borders of the RCM are determined by the GCM output data. Regional climate models for dynamical downscaling need extensive compute power, usually not quite as much as GCMs do, since only a region of interest is simulated and not the full globe. However, the saved amount of computations is considerably reduced because of the increase in spatial resolution. Thus dynamical downscaling without parallelized high performance compute environments is not feasible. Another disadvantage for dynamical downscaling is the problem of discrepancies between spatial, temporal and processual detail at the boundary between highly resolved regional models and low resolution GCMs. Since at the border of the spatial domain of a RCM the highly resolved RCM data and the less resolved GCM data may show unrealistic steep gradients of the climate variables, this border has to be excluded from interpretation, leading to the necessity to configure a much larger RCM domain than actually needed. In order to achieve a smooth transition between GCM and RCM, usually a whole series of RCMs of successively increasing resolution are nested into a GCM and into itself by increasing the resolution slowly from step to step in order to reduce the inhomogeneities at the borders. This of course increases the demand of compute time additionally. RCMs also cannot avoid parametrisations, like GCMs do. However, RCM parametrisations can be tuned to a much higher level of detail than GCM ones. Thus e.g. precipitation can be much more realistically simulated concerning dynamical (e.g. orographic) or thermal convection. However, in areas of extreme relief energy, as it is the case for high mountain regions, simplifying assumptions may lead to unrealistic results. Even though the parametrisations of a RCM reflect empirical dependencies observed in the past, RCMs are able to simulate meteorological conditions that have not been observed before, because the driving circulation and most of the exchange processes are still simulated on basis of physical laws which are universally valid. However, remaining errors of RCM estimations are usually subject to so called model output statistics (MOS) for statistical correction of the results e.g. concerning the mean and distributions of the target variables, by shifting or scaling of the simulated values.

The full dependence on observations of the past is the most important disadvantage of statistical downscaling methods. Statistical downscaling is using significant relationships between the local target variable of interest, the so called predictand, and the large scale or synoptic atmospheric conditions, described by so called predictor variables. If a certain manifestation of the variables of such a relationship was not observed in the past, it may lead to erroneous predictions of the state of the target variable for a scenario where new conditions of the predictors appear. If the relationship is a continously linear function, the problem might be neglectible. However, unfortunately most relationships in the atmospheric sciences are non-linear. Using statistical methods which account for non-linear functions, which is shown below, therefore is of large importance. They also cannot extrapolate relationships beyond known parts of the spectra, but much better capture the links within the observed boundaries.

Sometimes only atmospheric circulation is used as predictor, usually described by the pressure field of the troposphere, since this is the variable which can be best simulated by GCMs and shows the highest spatial autocorrelation among all climate elements. However, other GCM variables are also used, and it is the main task of the so called calibration step in the statistical downscaling scheme to find optimal predictor combinations of variables in order to maximize the statistically explained fraction of variance of the predictand. However, the performance of statistical models not only depends on the kind of variables used (pressure, temperature, humidity etc.) but also on the size and position of the area from which the predictor is taken, the predictor domain. Additionally it is important which atmospheric level is used. Also building seasonal subsamples is of great relevance, since some processes are realized differently depending on the annual cycle. Many additional factors may be subject to the optimization of the statistical downscaling models.

The core of statistical downscaling is the method chosen as transfer function. As a basic method often multiple linear regression (LM) is used to model the predictand. In order to estimate a whole grid point field of predictants at once by linear regression, canonical correlation models may be used. However, LM relie on a normal distribution of the residuals, i.e. the unexplained variation of the target variable, which is often not fulfilled. Therefore it is often a better choice to use generalized linear models (GLM). Apart from these linear methods, several approaches exist to model non-linear functions of dependence between predictors and predictands. A much more simple method than regression techniques which actually is able to use non-linear relationships is the analog method (AM) presented e.g. by Zorita and von Storch (1990). The idea of this method is rather intuitive: for any situation of the circulation in the GCM where the state of a local target variable is of interest, chose the most similar situation from observation data in the past and use the referring value of the target variable from this analog situation as the downscaling result. It turns out that the performance of this method can be as good as that of regression models, while it can be applied without statistical prerequisites.

Exemplarily for the possible large bandwidth of the method spectrum used for downscaling, two selected methods, will be described more in detail below, since they have been applied recently in intensive downscaling studies in high mountain areas, i.e. for the Zugspitze and the Sonnblick. The first is the more simplistic reference class forecast method (RCF) based on circulation type classification (CTC), the second the more sophisticated technique of artificial neural networks (ANN).

11.4 Circulation Type Classification

A derivate of the analog method is the so called reference class forecast method (RCF) and its functional principle is described straightforwardly. This method is based on weather or circulation type classifications (CTC) of observed predictor fields (often pressure maps) from the past. Most prominent examples for often used weather type classifications are e.g. the Lamb classification for Great Britain (Lamb 1972) or the Hess-Brezowsky classification for central Europe (Hess and Brezowsky 1977) produced manually by assigning daily weather maps to subjectively defined weather types or the automatically produced classification of the german meteorological service DWD (Bissoli and Dittmann 2003). However many more classifications are available (see e.g. Huth et al. 2008).

In order to model a certain state, e.g. in the future, the expected predictor field for this situation, e.g. as simulated by a GCM, is assigned to its most similar class. The predictand value is then chosen as the mean of the target variable for this reference class in the past.

Even though this method is rather simple and easy to apply, it can show considerable good skill (see below). Its main limitations result from the fact, that this method assumes, that all relevant climatic changes in the (future) scenario are caused by changes in the frequencies of weather and circulation types. However, such actualistic principles are used for most of the downscaling methods based on known states of the large scale predictors. And even dynamical downscaling is affected by this problem, since many assumptions for setting up RCMs are based on observations of the past. Besides, the skill of RCF depends on several methodological technical factors, mainly concerning the underlying circulation type classification.

There are plenty of methods for weather and circulation type classifications that can be used for downscaling and the question arises which method is suited to achieve the best classification in order to maximize the downscaling skill, i. e. to minimize the model errors. A comprehensive comparison of classification methods was done within the COST Action 733 "Harmonisation and Applications of Weather Type Classifications for European regions" and the main results of these studies are also relevant for downscaling applications. First it has been found that the classifications resulting from 33 examined classification methods are surprisingly dissimilar among each other, i. e. they do not have significantly more in common than classifications based on purely randomly defined types, except for some methods based on non-hierarchical cluster analysis (Philipp et al. 2016). Second it has been found that there is no single method which can always provide the best skill in discriminating different states of a predictand variable, however methods based on cluster analysis (more generally spoken those using optimization algorithms) show a tendency to higher skills compared to others (Beck and Philipp 2010). This does not mean that cluster analysis is always the best, but it suggests that it is not the worst choice to consider them for downscaling.

Besides the classification method, the skill depends on the selection of the predictors, which is generally true for all statistical models. This includes the climatological variables used as predictors (e. g. pressure, wind components, large scale temperature, humidity etc.) as well as the location where they were measured. This includes the atmospheric level (near the surface or in the middle or upper troposphere) as well as the topographic region, i. e. the location, shape and size of the section of the grid point field used, what is usually called the model domain. The optimal model domain has been examined by Beck et al. (2013). It has been found that there are actually systematic preferences, e. g. for temperature the domain should be larger than for precipitation and west-east-elongated domains are often superior, however in order to achieve the optimal model for a certain predictand, it is necessary to empirically test potential configurations by a systematic predictor screening. An additional way to optimize a classification for a certain target is to include the target variable already during the classification process in the calibration step. Then the classification produces types with members that are not only similar concerning the pressure field e. g., but also concerning the target variable, like temperature e. g., which is called a conditional classification scheme. It could be shown that this method can improve classification based downscaling schemes in general if an optimal weighting between the predictor and the predicant variable is found (Lutz et al. 2011).

11.5 Artificial Neural Networks

Compared to RCF artificial neural networks (ANN) represent the other end of the scale of complexity of transfer functions for downscaling. As the name suggests this method tries to imitate the neural network of a brain consisting of neurons which receive a signal generated originally by a sensory organ at the one side and – depending on the result of an activation function – further transmit the signal to the next neurons on the other side. The strength of the forwarded signal depends on the transmitting neuron. Thus, after moving through the network, a signal can be filtered out competely or amplified leading to a corresponding reaction at the output of the system.

In statistics this principle is realized by defining an array of nodes connected by weigthing factors. The array consists i) of an input layer, where the input neurons imitate the sensory organs and receive the values of the predictors, ii) the neurons responsible for the transport of the signals, which are called hidden neurons since they are not directly connected to external data and iii) the output layer where the output neurons (often only one) represent the predictands. In order to automatically get the desired result in the output layer in dependence from a certain state of the predictors in the input layer, the weighting coefficients of the links between the layers have to be optimized by a learning algorithm. The most often used learning algorithm is the so called backward progagation (backprob) scheme, where the estimation error of the output neuron is propagated backwards through the net in order to optimize the weights of the links to the hidden neurons and then of the links to the input neurons. In order to reach optimal weights for all pairs of observed predictor and predictand data, this is done repeatedly for all elements (days) of the training data subset. Each time the network was shown all training data once (a so called training epoch) the weigths are adjusted a bit better to learn the dependencies of the target variable. However the training has to be stopped before the absolut optimum is reached in order to keep a minimum level of abstraction in the net and avoid overfitting. Overfitting is given, when the network has memorized all single situations it was trained for but is not able to perform well for a new situation it hasn't seen before. Therefore, in order to decide for the stopping, an independent data subset (the validation subset) has to be kept aside the training data subset, which is only used to check the current network for its skill with unknown data. Otherwise the skill would not be representative for the application to the scenario GCM data.

11.6 Combination of ANN and CTC

Since CTC based downscaling is extremely fast compared to ANN training, it has been tested whether it is possible to combine both methods and still reach or even outperform the performance of ANNs. Therefore a non-hierarchical classification scheme is applied to the variables determined by predictor screening and the target variable in order to train an ANN for each class separately. The idea is, that it might be easier for a network to discern different factors for dry, normal and wet conditions and therefore reach a higher level of detail and model performance.

11.7 Skill

Before applying downscaling methods to GCM data in order to derive estimates of possible future regional or local scale climate change it is necessary to assess the confidence that can be attributed to the simulated future climate. This is typically done by determining the skill of the downscaling approaches via model validation experiments.

Such a model validation comprises firstly the derivation of estimates of the predictand variable (e. g. daily mean air temperature at a certain location) by applying the downscaling model to large-scale atmospheric predictor fields from available observational or reanalysis data sets and – secondly – the comparison of the downscaling model output to observed predictand data.

These comparisons utilize suitable performance measures which quantify the accordance or the mismatch between observed and modeled data. For continuous predictands commonly used measures are for instance the mean error, the mean absolute error, the mean squared error or the correlation coefficient. Furthermore, based on the comparison of these measures estimated for the downscaling model and for a reference model (e. g. utilizing the climatological mean of the predictand) respectively, skill scores can be calculated (e. g. the mean squared skill score based on the mean squared errors of the downscaling model and the reference) indicating in how far the downscaling model outperforms the reference model and thus is suitable to provide valuable future climate simulations.

Exemplary, the calculation of the mean squared skill score (MSSS) based on the mean squared error estimated for the downscaling model (MSEmod) and for the climatological reference (MSEclim) is illustrated in the following three equations (see also Wilks 2006).

$$MSE_{\mathrm{mod}} = \frac{1}{n} \sum_{k=1}^{n} (y_{\mathrm{k}} - o_{\mathrm{k}})^2$$

$$MSE_{\mathrm{clim}} = \frac{1}{n} \sum_{k=1}^{n} (\overline{o} - o_{\mathrm{k}})^2$$

$$MSSS = 1 - \frac{MSE_{\mathrm{mod}}}{MSE_{\mathrm{clim}}}$$

With *n* being the number of observations, *o* being the observed values of the predictand, $\overline{o}$ being the mean of the observed values and *y* being the simulated predictand values. Thus, an *MSSS* of 1 indicates a perfect model while an *MSSS* less or equal zero indicates a model performing equal or even worse than the climatological reference.

Those skill scores essentially describe the absolute errors of the model, i. e. the magnitude of the difference between o and y. However, it can happen, that there is a systematic error and the model values are always too high or too low compared to the observations by a certain amount, called bias. Then the model output can be corrected by subtracting the bias afterwards. Repeating the skill score calculation with bias corrected model results then increases the performance. The same might be true for the scaling, and dividing by a correction scale coefficient might improve the performance of the model.

In order to estimate the performance of the downscaling model directly, besides of any systematic bias or scale, i. e. just evaluating the coincidence of positive and negative anomalies relatively but not concerning the absolute values, correlation coefficients may be used. For that, the output values of the downscaling model for the historical reference period are correlated with the actually observed values of the target variable. Correlation coefficients can give information on how many percent of variance in the target variable is captured by the model if the squared correlation coefficient is considered : $r^2 = d$, where d is called the coefficient of determination.

However, as the skill estimate may be artificially high when calculated on the basis of the data that has been used to fit the downscaling model, it is necessary to validate the model on data that has not been used for model calibration. This is done by so called cross-validation. Here, the time period for which predictor data and observed predictand data are available is divided into two or more non-overlapping sub-intervals and the downscaling models are then in turn calibrated using all but one of these sub-intervals and validated in the remaining independent sub-interval. Variants of this cross validation approach include the use of varying lengths of the sub-intervals and varying methods for defining the sub-samples used for calibration and validation, including random sampling techniques.

For the overall performance of the regional to local scale future climate simulations beside the skill of the statistical downscaling approach it is in addition of crucial importance how well the GCMs simulate the large-scale input data for the statistical downscaling models. For instance, it is well known that many GCMs feature warm and as well cold biases in sea surface temperatures over different parts of the North Atlantic leading to an incorrect representation of the large-scale atmospheric circulation over Europe (Keeley et al. 2012). Such biases – model errors relative to observations – need to be considered utilizing varying approaches for bias correction (see for example Teutschbein and Seibert 2021 for a review of common approaches). For instance, one rather simple method – linear scaling (Lenderink et al. 2007) – uses the differences in the mean between GCM and observations (or reanalyses) for bias correction of the GCM output.

In addition, climate processes are partly differently represented in GCMs from different climate modelling groups leading to accordingly diverging future climate projections. To account for this source of uncertainty so called multi-model ensembles comprising projections from several GCMs are used to derive quantitative estimates of the range of uncertainty in future climate projections.

Furthermore, uncertainties may also arise from differences between projections of variants of the same GCM run with varying values of certain model parameters or run from varying start dates. So called perturbed physics ensembles are used to determine the range of uncertainty in future projections related to variants in model parameters whereas initial condition ensembles consider the effect of differing start dates.

Finally, uncertainties are also due to varying properties of the statistical downscaling models. For instance, using different techniques for cross validation or using different time periods for calibrating the final model which is then applied to GCM data may lead to differing statistical models and accordingly to varying future regional climate simulations. Thus, in addition to the above mentioned numerical ensembles as well statistical ensembles have been introduced to quantify uncertainty in climate projections.

11.8 Case study Zugspitze and Sonnblick

The success of developing a useful downscaling scheme for locations in high mountain areas is especially affected by two factors concerning the mountain relief. On the one hand high mountain places can be exposed to the free atmosphere, compared to the low land. This means that the large scale circulation and associated atmospheric parameters which can be simulated well by GCMs on the synoptic scale play a larger role for the variability of the local target variable than atmospheric processes working near the surface on the meso- and microscale, i.e. the subgrid scale for GCMs. This means that the downscaling models based on large scale GCM output as discussed above should perform especially well. On the other hand, the high relief energy of the mountains also includes areas affected strongly by GCM-subgrid processes, like luv/lee effects, channeling of wind, gradients in friction or radiation energy uptake, slope winds, mountain plane winds or increased turbulence which is mixing air from the boundary layer into the area up-stream of the measurement site (footprint area) etc. Thus, locations in high mountain areas which are less exposed to the free atmosphere but significantly affected by meso- and microscale processes, may be even less suitable for downscaling than low land stations.

A high mountain station especially suited for downscaling is the meteorological observatory at the Zugspitze (2962 m) maintained by the German weather service DWD. It offers a long time series of observation data (see Tab. 2) even though not all data can be used as explained below. Moreover, the location of the Zugspitze as a comparatively high peak, exposed at the northern edge of the high mountain range makes it a promising object for downscaling. Another high mountain observatory maintained with long records is the Sonnblick observatory at the Hoher Sonnblick (3106 m) operated by the Austrian meteorological service (ZAMG) and located in the Alpine main ridge. Compared to the Zugspitze the Hoher Sonnblick is surrounded by mountain ridges in all directions.

Tab. 2: Start time of daily records for the target variables precipitation and temperature at Zugspitze and Hoher Sonnblick

Station	Variable	Start of available records
Zugspitze	Precipitation	01.01.1901
	Temperature	01.08.1900
Hoher Sonnblick	Precipitation	01.08.1890
	Temperature	01.10.1886

In order to calibrate and validate the downscaling models, large scale circulation data have been obtained from the 20th century reanalysis dataset version 2 (Compo et al. 2011). They have been generated by a weather forecast model initiated by historical and recent station and radiosonde observations and are used as equivalent to large scale GCM output, however not for certain scenarios but for the past where also the target variables temperature and precipitation from Zugspitze and Hoher Sonnblick are known. The variables which have been examined as potential predictors include air pressure at sea level (slp), thickness of the layer between 850 hPa and 500 hPa (thi), geopotential height of several pressure levels (hgt), as a common way to describe air pressure distribution in upper levels, zonal wind speed (uwnd), meridional wind speed (vwnd), total horizontal wind speed (swnd), vertical wind speed (omega), air temperature (air), spezific humidity (shum), relative humidity (rhum), zonal and meridional moisture flux (umf and vmf) as well as vorticity (vor) and divergence (div) of the wind field. In order to cover the whole vertical extent of the troposphere but at the same time restrict the number of variants, the variables have been extracted for the levels 850 hPa, 700 hPa, 500 hPa and 250 hPa.

In order to obtain useful and robust models, it is necessary to reduce the set of potential predictors to the most important ones. Even though the influence of less important variables is reduced by weighting in the classification scheme as well as during training of the neural network, it is much more effective and avoids to end up in less stable solutions of the model optimisation process if they are excluded and if the set of predictors is kept as small as possible. However, to obtain still a well-performing model it is very important to find out which variables

Tab 3: Optimal predictor combinations determined by screening through all possible configurations of variables (air pressure at sea level (slp), thickness of the layer between 850 hPa and 500 hPa (thi), geopotential height of several pressure levels (hgt), as a common way to describe air pressure distribution in upper levels, zonal wind speed (uwnd), meridional wind speed (vwnd), total horizontal wind speed (swnd), vertical wind speed (omega), air temperature (air), spezific humidity (shum), relative humidity (rhum), zonal and meridional moisture flux (umf and vmf) as well as vorticity (vor) and divergence (div)), atmospheric pressure level (level), start and end of the domain given in degrees east (lon) and degrees north (lat) for each season (December, January, February (DJF), March, April, May (MAM), June, July, August (JJA) and September, October, November (SON)) and each target variable (precipitation (prc) and temperature (tmp) at Zugspitze (Zug) and Hoher Sonnblick (Son)).

			slp	thi	hgt	uwnd	vwnd	swnd	omega	air	shum	rhum	umf	vmf	vor	div
Zug prc	DJF	level					0850		0700		0850	0700				
		lon					0:14		0:16		-2:14	4:14				
		lat					42:52		42:52		38:52	40:56				
	MAM	level				0850	0850				0700	0700				
		lon				-2:16	-4:16				-4:16	6:12				
		lat				40:50	38:50				40:52	44:52				
	JJA	level							0500		0850	0500			0850	0850
		lon							-12:16		0:14	0:14			0:20	0:14
		lat							44:52		42:52	42:52			42:52	40:54
	SON	level				0850			0700		0850	0700			0850	0700
		lon				2:12			6:20		2:12	6:16			-4:24	6:16
		lat				44:48			44:52		42:52	42:52			38:52	40:50
Zug tmp	DJF	level					0850			0700						0850
		lon		6:16			2:20			6:16						2:14
		lat		46:52			40:52			44:52						42:50
	MAM	level								0850					0500	
		lon		4:14						0:14					4:18	
		lat		44:52						42:52					42:52	
	JJA	level								0850					0850	0850
		lon		4:16						0:14					0:22	0:16
		lat		46:52						44:52					40:52	42:52
	SON	level					0850			0850					0500	
		lon					2:16			-4:14					0:18	
		lat					42:50			38:56					44:56	
Son prc	DJF	level							0700		0850	0700			0850	0700
		lon							8:20		4:16	8:16			4:18	4:22
		lat							46:52		44:52	44:52			36:50	40:52
	MAM	level							0500		0700	0850			0850	0850
		lon							2:18		6:14	4:16			-2:14	2:18
		lat							42:50		44:52	44:50			38:56	40:50
	JJA	level				0850			0700		0700	0700				0850
		lon				-4:16			-4:20		2:16	6:16				-4:16
		lat				36:50			40:54		44:54	44:52				40:52
	SON	level									0850	0700			0850	0700
		lon									0:16	4:18			-8:20	-4:22
		lat									38:52	40:52			38:56	42:52
Son tmp	DJF	level								0850					0850	0850
		lon		8:18						6:14					-4:20	8:20
		lat		44:52						42:50					38:50	42:50
	MAM	level								0850					0500	0850
		lon		4:18						4:16					4:20	8:16
		lat		42:54						44:50					42:54	44:52
	JJA	level								0850					0500	
		lon		6:18						4:18					8:24	
		lat		44:54						42:52					40:50	
	SON	level								0850					0850	
		lon								4:18					6:22	
		lat								42:50					36:50	

are the most important ones by a systematic predictor screening. This is done by applying a fast training algorithm (resilient backward propagation) iteratively to all possible predictor combinations including the different tropospheric levels and varying the domains defined by the sector latitudes and longitudes. Tab. 3 shows the results of the predictor screening. As is clearly apparent, precipitation can be predicted best by variables describing large scale humidity (shum, rhum), horizontal wind (uwnd, vwnd) and variables associated with vertical circulation (omega, div and vor) for both stations. Station temperature is best explained by the large scale temperature itself (thi, air) as well as vorticity and divergence, which are also associated with cyclonic activity.

The model calibration has been done not only once for the whole available period starting around the beginning of the 20th century but for 30-year subperiods, which are shifted by 1 year steps through the overall period in order to examine the eventually varying link between target variable and predictors (see Tab. 3). In order to get robust results for each subperiod, the ANN for calibration was initialized 15 times. For each of these 15 ensemble members the spearman corrleation coefficient between estimated and observed values for the target variable has been calculated and aggregated to an ensemble mean for the referring subperiod. The results in Fig. 3 reveal that in all cases there are so called non-stationarities (Hertig et a. 2015), i. e. the model performance is not constant. As expected, temperature estimation works generally better than the precipitation models. However, may be most striking is a drop in skill in the middle of the 20th century for precipitation at the Hoher Sonnblick. Such pronounced slumps are usually indicators for inhomogeneities, i. e. abrupt jumps in measurement time series, often caused by a change in the instrumentation or location of the instruments. However, another prominent feature, beside that, is a generally increase in skill over time for both variables in both places for all four seasons. A possible explanation is an increase in measurement and recording quality over time. In order to take this into account, the final calibration has been done for the period 1970 to 2000.

The skill of the final models, including circulation type classification, neural networks and their combination, is shown in Fig. 4. In order to compare the skill for the two main target variables precipitation and temperature to subordinated ones, the performance for specific humidity and wind at the station Zugspitze is included here. In some cases the RTF method based on circulation type classifications (triangles) is somewhat better than the ANN (circles). However, this is true only for the calibration skill (violet) and not for the real skill estimated with the validation data (green) which is the relevant one. Regarding the latter, i. e. looking only on the green symbols, the CTC based method is always the worst (except for the wind) and the ANN always the best. Further on it is striking that the skill for precipitation and temperature at the Zugspitze is always better than the skill for the Hoher Sonnblick throughout all seasons. A possible reason is assumed to be the topographical position of the Hoher Sonnblick within the Alpine main ridge and associated larger influence from the surface, modifying the direct forcing of the large scale predictors to some degree, although the skill for the Sonnblick is still remarkably high. Regarding the skill of the method combination ANN+CTC it turns out, that it is not as performant as

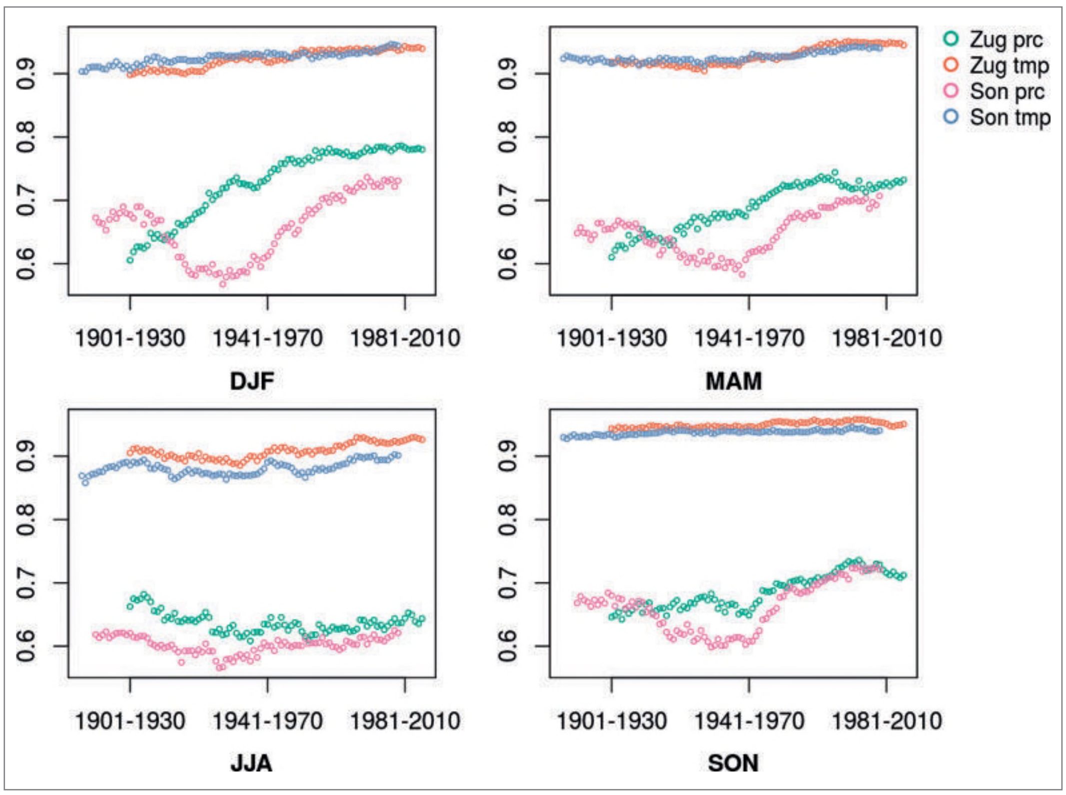

Fig. 3: Stationarity of the relation between predictors (see Tab. 3) and target variables: ensemble mean spearman correlation coefficients (ordinate) between observed values of the target variable and those estimated by artificial neural networks in shifting 30-year subperiods between 1901 to 2010 (abscissa) for winter (DJF), spring (MAM), summer (JJA) and autumn (SON).

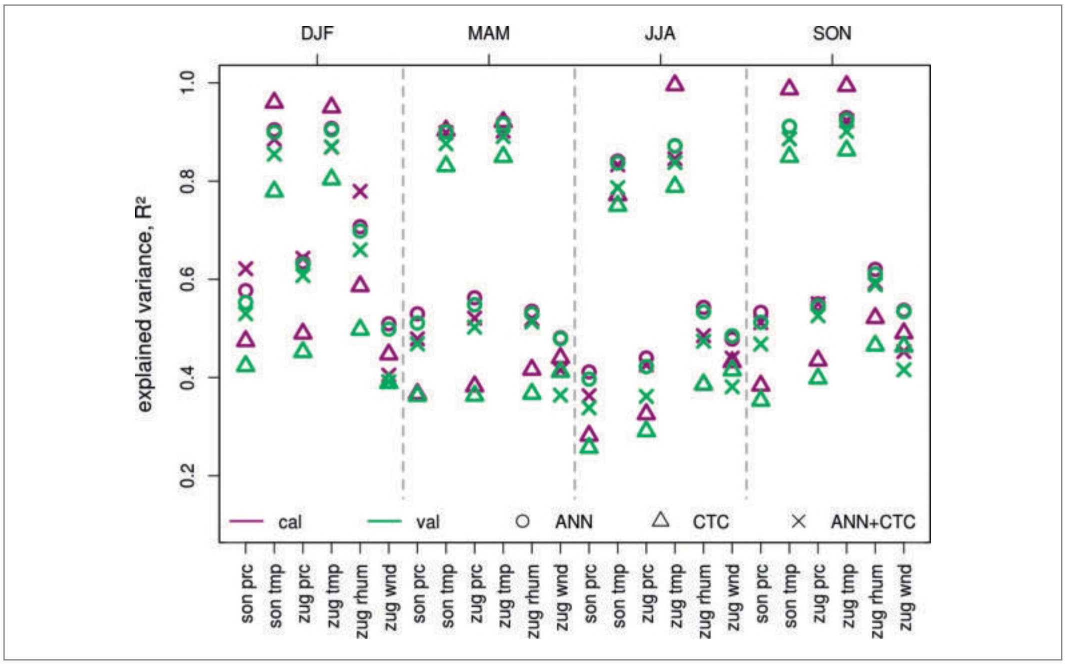

Fig. 4: Comparison of skill of ANN-, CTC- and combined downscaling models for various target variables (prc: precipitation, tmp: temperature, rhum: relative humidity and wnd: wind speed), station (zug: Zugspitze, son: Hoher Sonnblick) and season (DJF: winter, MAM: spring, JJA: summer, SON: autumn). Circles show the result of artificial neural networks (ANN), triangles those for circulation types (CTC) and crosses for the combined method (see text). Violet symbols denote the mean coefficients for 15 calibration subsamples, green symbols those for 15 validation subsamples in the period 1970 to 2000.

ANN alone. The networks in fact do learn faster within the classes, however finally they do not reach the skill level of single neural networks trained for a long time (several days up to weeks of compute time).

Even though the variance fraction of the target variables that can be explained by the downscaling models, as apparent from Fig. 4, is extraordinary high for temperature (ca. 90 %) it is rather limited for precipitation, where only roughly 50 % of the variablity can be simulated, which is still high compared to other studies where 30 % or 40 % are reached (e. g. Cavazos and Hewitson 2005). However, it is still possible to estimate the precipitation changes that are caused by changes of the large scale predictors only, even if there might be other tendencies due to other factors. Thus the precipitation estimation must not be interpreted as the expected real change but only as one impulse for changes among others.

Keeping this in mind, the models are applied by feeding them with GCM output data for several scenarios. A common collection of GCM simulations is available by the Coupled Model Intercomparison Project Phase 5 (CMIP5) framework. Tab. 4 contains the models and their realisations for the scenarios. The first scenario is called "historical" (hist) and includes the boundary conditions for the global climate system since 1850 up to now. Since GCMs, even if they have reached high performance in the last few decades, can not simulate the observed climate absolutely perfect but sometimes show systematic discrepancies, the historical run of a GCM can serve as a reference run in order to determine the relative changes of any other scenario. Thus it is assumed, that the difference between the reference and the scenario runs can be transferred to the real world climate system, even if there is a general bias which is found in all model runs, but not relevant if the changes are of interest. This method is sometimes called the delta approach, and should be considered for all model interpretations. The future scenarios used in this study are the so called RCP4.5 and the RCP8.5 scenarios, the former assuming a global increase of radiative forcing by 4.5 W/m^2 and the latter by 8.5 W/m^2 in the year 2100 relative to the pre-industrial year 1850. RCP means representative concentration pathway and points out that it reflects not the emissions of radiation-relevant trace gases, but their actual concentration and its effect on the radiation budget. For the MPI model three realisations were available which helps to increase the robustness of the results by running the downscaling models with data from more than one GCM run, thus building an ensemble of runs, and calculating the ensemble mean result.

Tab. 4: Overview of the general circulation models and their scenarios used for driving the downscaling models. All model datasets offer a historical run (hist) and those for representative concentration pathways (RCP) for radiation-relevant trace gases leading to an increase of 4.5 respectively 8.5 W/m^2 in the radiative forcing of the climate system. For the German model three ensemble members are available for the RCP scenarios.

GCM model	Responsible Institute	Realisations/Scenarios
MPI-ESM-LR	Max-Planck-Institute, Germany	Hist, RCP4.5 (3 ens. members), RCP8.5 (3 ens. members)
HadGEM2-CC	Met Office, United Kingdom	Hist, RCP4.5, RCP8.5
ACCESS1-0	CSIRO (Commonwealth Scientific and Industrial Research Organisation) und BOM (Bureau of Meteorology), Australia	Hist, RCP4.5, RCP8.5
CMCC-CMS	CMCC (Centro Euro-Mediterraneo per i Cambiamenti Climatici), Italy	Hist, RCP4.5, RCP8.5
IPSL-CM5A-LR	IPSL (Institut Pierre Simon Laplace), France	Hist, RCP4.5, RCP8.5

The time series of annual mean temperatures and annual precipitation sums produced by driving the downscaling models with the GCM scenario output data for the respective predictor variables is shown in Fig. 6. While the thin lines, respesenting the single ensemble members result, allow to estimate the spread of the simulated target variables, the ensemble means point out their long term evolution. The temperature time series (Fig. 6a and b) show a strong trend as expected, which has been positively tested for significance by the trend-noise-ratio using the 5 % uncertainty level. According to these results the station annual mean temperatures at the Zugspitze summit will approach the freezing point in the year 2100 for the RCP8.5 scenario which will have dramatic consequences for the whole environment in this region. At the Sonnblick the absolute temperature is generally a bit lower, but the changes are also significant.

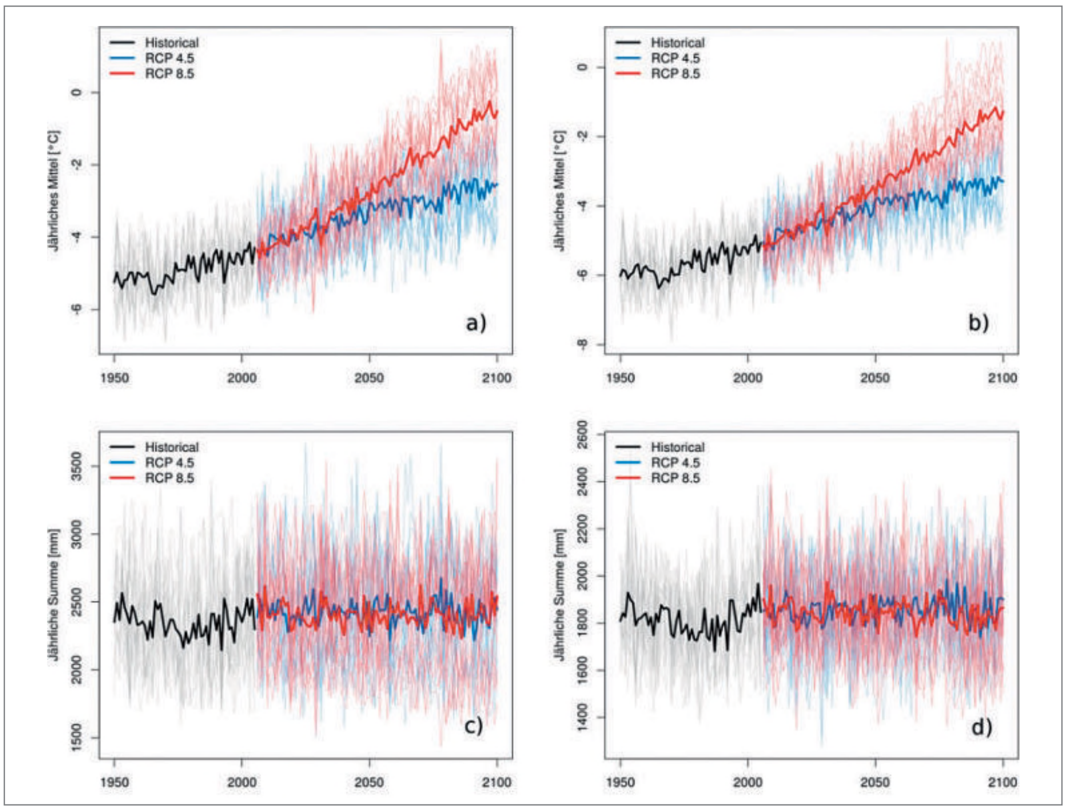

Fig. 5: Evolution of target variables temperature and precipitation simulated by downscaling models using circulation type classifications, artificial neural networks and their combination in the period 1970 to 2100 aggregated to annual values (means of temperature, sums for precipitation). Thin lines denote single time series of 15 cross validated model runs for each of the seven GCM runs, while the thick line represents the overall ensemble mean. a) Zugspitze temperature, b) Sonnblick temperature, c) Zugspitze precipitation, d) Sonnblick precipitation.

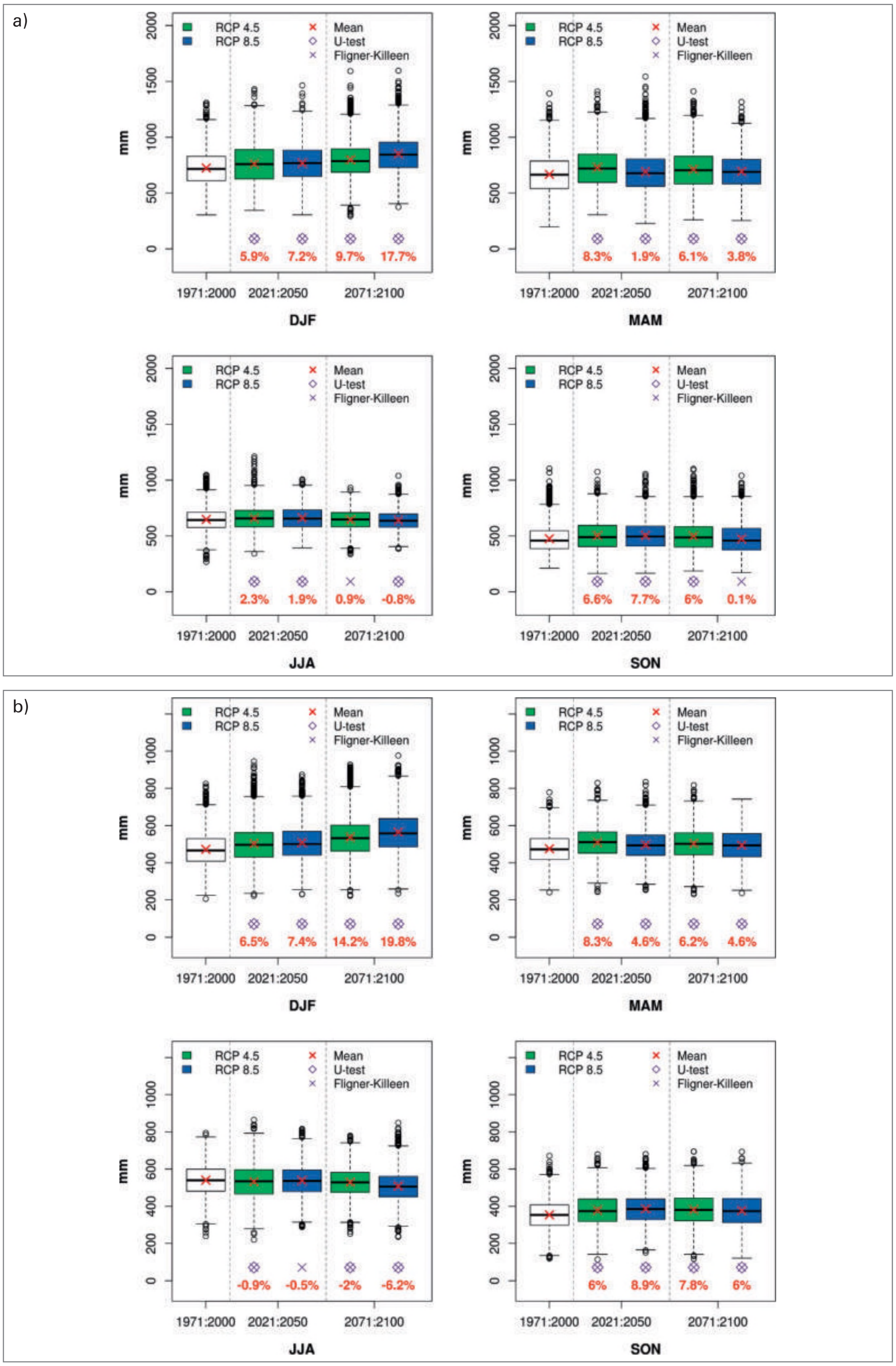

Fig. 6: Precipitation changes for a) Zugspitze and b) Hoher Sonnblick discerned by seasons: Red cross shows the mean and boxplots the distribution of seasonal precipitation sums estimated by 15 cross validation runs of neural network downscaling models driven by 7 GCM runs for the historical reference period 1971 to 2000 and the RCP4.5 (green) and RCP8.5 (blue) scenarios in the time range 2021 to 2050 and 2071 to 2100. Red numbers denote the changes in percent compared to the historical reference period. A violet circle indicates a significant difference in the central tendency according to the U-test, a violet cross a significant change in variablity according to the Fligner-Killeen-test on the 5 % uncertainty significance level.

Regarding precipitation, no significant trend can be observed for both RCP scenarios if the annual sum is considered at both stations (Fig. 6c and d).

However, if the precipitation is analysed separately for each season, different results are apparent. Fig. 6 shows the respective changes for seasonal precipitation sums for the 30-year periods 2021 to 2050 and 2071 to 2100 compared to the reference period 1971 to 2000 as obtained by the ANN models, which give the most reliable results compared to the other methods. According to the U-test there are significant changes for both stations in all seasons. In winter (DJF), spring (MAM) and autmn (SON) both, Zugspitze as well as Hoher Sonnblick show a similar tendency. In winter precipitation is constantly increasing, stronger for the RCP8.5, somewhat slower for the RCP4.5 scenario and somewhat stronger for the Hoher Sonnblick than for the Zugspitze. In spring (MAM) and autumn (SON) an initial strong increase in precipitation is followed by either constant or even decreasing rates of growth, except for RCP8.5 at the Zugspitze and RCP4.5 at Hoher Sonnblick with an accelerating increase. Regarding the summer (JJA) Hoher Sonnblick shows a growing decrease in precipitation for both scenarios. The Zugspitze, however, is characterized by an initial increase followed by conditions comparable to the historical period or even below.

Two antagonistic main processes forcing the moisture budget and precipitation are supposed to be responsible for such diverse climate change signals: i) increasing storage capacity for water vapor in a warmer troposphere and therefore increased moisture transport into the atmosphere with increased precipitation sums and ii) strengthening of the Mediterranean subtropical anticyclonic high pressure cell (Jacobeit et al. 2017) with increased subsidence and stable conditions suppressing precipitation. Thus it is fairly likely that at the beginning of a global warming period, the moisture storage effect dominates at the Zugspitze in summer until the effects of changing circulation dynamics prevail. The difference to Hoher Sonnblick, where increasing precipitation reduction prevails in summer, might be explained by its position in the interior of the Alpine ridge where air masses already have lost considerable amounts of their moisture content due to luv effects at the edge of the mountain ridge. Maybe the effect of increased moisture transport is therefore less important in the center of the Alps, a topic for further investigations. In any case, high mountain climate is indicated to be changed considerably according to the presented results with strong implications for the whole ecosystem of the Alps to be expected.

References

Auer I., Böhm R., Jurkovic A., Lipa W., Orlik A., Potzmann R., Schöner W., Ungersböck M., Matulla C. Briffa K., Jones P. D., Efthymiadis D., Brunetti M., Nanni T., Maugeri M., Mercalli L., Mestre O., Moisselin J.-M., Begert M., Müller-Westermeier G., Kveton V., Bochnicek O., Stastny P., Lapin M., Szalai S., Szentimrey T., Cegnar T., Dolinar M., Gajic-Capka M., Zaninovic K., Majstorovic Z., Nieplova E. (2007): HISTALP – Historical instrumental climatological surface time series of the Greater Alpine Region 1760–2003. International Journal of Climatology 27, 17–46.

Bavay M., Lehning M., Jonas T., Löwe H. (2009) Simulations of future snow cover and discharge in Alpine headwater catchments. Hydrol Process 23: 95–108. doi:10.1002/hyp.7195

Beck C. and A. Philipp (2010): Evaluation and comparison of circulation type classifications for the European domain. Physics and Chemistry of the Earth, 35, 374–387. DOI: 10.1016/j.pce.2010.01.001

Beck C., A. Philipp and F. Streicher (2016): The effect of domain size on the relationship between circulation type classifications and surface climate. Int. J. Climatol. 36: 2692–270.

Bissolli, P., Dittmann, E., 2003. Objektive Wetterlagenklassen (Objective weather types). In: Klimastatusbericht 2003. DWD (Hrsg.). Offenbach 2004, Germany (in German).

Brunetti, M., M. Maugeri, T. Nanni, I. Auer, R. Böhm & W. Schöner (2006): Precipitation variability and changes in the greater Alpine region over the 1800–2003 period. Journal of Geophysical Research, Atmospheres, D11107,1–29.

Castebrunet, H., N. Eckert, G. Giraud, Y. Durand & S. Morin (2014): Projected changes of snow conditions and avalanche activity in a warming climate: the French Alps over the 2020–2050 and 2070–2100 periods, The Cryosphere, 8(5), 1673–1697.

Cavazos T. and B.C. Hewitson (2005): Performance of NCEP – NCAR reanalysis variables in statistical downscaling of daily precipitation. Clim. Res., 28, 95–107.

European Environment Agency (2017a): Climate change, impacts and vulnerability in Europe 2016. An indicator-based report. EEA Report 1/2017, Luxembourg.

European Environment Agency (2017b): Climate change adaption and disaster risk reduction in Europe. Enhancing coherence of the knowledge base, policies and practices. EEA Report 15/2017, Luxembourg.

Fraedrich K., H. Jansen, E. Kirk, U. Luksch and F. Lunkeit (2005): The Planet Simulator: Towards a user friendly model. Meteorologische Zeitschrift, 14/3, 299–304.

Gobiet, A., S. Kotlarski, M. Beniston, G. Heinrich, J. Rajczak, M. Stoffel (2014): Science of the Total Environment, 493,1138–1151.

Heinrich, G., A. Gobiet, H. Truhetz, T. Mendlik (2013): Expected climate change and its uncertainty in the Alpine region: extended uncertainty assessment of the reclip: century and ENSEMBLES multi-model dataset. Wegener Center Scientific Reoprt 50.

Hertig, E., C. Beck, H. Wanner and J. Jacobeit (2015): A review of non-stationarities in climate variability of the last century with focus on the North Atlantic-European sector. Earth-Science Reviews, doi:10.1016/j.earscirev.2015.04.009.

Hess, P., and H. Brezowsky, 1977: Katalog der Großwetterlagen Europas 1881–1976, 3e verbesserte und ergänzte Auflage. Berichte des Deutschen Wetterdienstes 113, Offenbach, Germany, 70 pp.

Huss, M. (2011): Present and future contribution of glacier storage change to runoff from macroscale drainage basins in Europe. Water resources research, 47, W07511.

Huss, M. (2012): Extrapolating glacier mass balance to the mountain-range scale: the European Alps 1900–2100, The Cryosphere, 6(4), 713–727.

Huth R., C. Beck, A. Philipp, M. Demuzere, Z. Ustrnul, M. Cahynová, J. Kyselý, and O.-E. Tveito (2008): Classifications of atmospheric circulation patterns: recent advances and applications. Annals of the New York Academy of Sciences. 1146, 105–152.

Keeley, S., Sutton, R., Shaffrey, L., (2012), The impact of North Atlantic sea surface temperature errors on the simulation of North Atlantic European region climate. Q. J. R. Meteorol. Soc., 138, 1774–783.

Kobierska F, Jonas T, Zappa M, Bavay M, Magnusson J, Bernasconi SM (2012) Future runoff from a partly glacierized watershed in Central Switzerland: a two-model approach. Adv Water Resour, 55, 204–214.

Lamb, H., 1972: British Isles Weather Types and a Register of Daily Sequence of Circulation Patterns, 1861–1971. Geophysical Memoirs, Vol. 116, Her Majesty's Stationery Office, 85 pp.

Lenderink, G., Buishand, A., Van Deursen, W., (2007), Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach. Hydrol. Earth Syst. Sci. 11 (3), 1145– 1159.

Lutz K., J. Jacobeit, A. Philipp, S. Seubert, H. Kunstmann & P. Laux (2011): Comparison and evaluation of statistical downscaling techniques for station-based precipitation in the Middle East (2011), International Journal of Climatology, 32, 1579–1595. DOI: 10.1002/joc.2381

National Center for Atmospheric Research Staff (Eds), 2017: The Climate Data Guide: Common Spectral Model Grid Resolutions. Retrieved from https://climatedataguide.ucar.edu/climate-model-evaluation/common-spectral-model-grid-resolutions.

Philipp A., C. Beck, R. Huth and J. Jacobeit (2016): Development and comparison of circulation type classifications using the COST 733 dataset and software. Int. J. Climatol. 36: 2673–2691.

Räisanen J. (2007): How reliable are climate models? Tellus, 59A, 2–29.

Rigling, A., C. Bigler, B. Eilmann, E. Feldmeyer-Christe, U. Gimmi, C. Ginzler, U. Graf, P. Mayer, G. Vacchiano, P. Weber, T. Wohlgemuth, R. Zweifel, and M. Dobbertin, 2013: Driving factors of a vegetation shift from Scots pine to pubescent oak in dry Alpine forests. Global Change Biology, 19, 229–240

Steiger R (2010) The impact of climate change on ski season length and snowmaking requirements in Tyrol, Austria. Clim Res 43: 251–262. doi:10.3354/cr00941

Stoffel, M., Tiranti, D. and Huggel, C., 2014, 'Climate change impacts on mass movements – Case studies from the European Alps', Science of The Total Environment 493, 1255–1266 (DOI: 10.1016/j.scitotenv.2014.02.102).

Strasser U., T. Marke, L. Braun, H.Escher-Vetter, I. Juen, M. Kuhn, F. Maussion, C. Mayer, L. Nicholson, K. Niedertscheider, R. Sailer, J. Stötter, M. Weber and G. Kaser (2018): The Rofental: a high Alpine research basin (1890–3770 m a.s.l.) in the Ötztal Alps (Austria) with over 150 years of hydrometeorological and glaciological observations. Earth Syst. Sci. Data, 10, 151–171.

Teutschbein, C., Seibert, J., (2012), Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. Journal of Hydrology, 456–457, 12–29.

Wilks, D.S., (2006), Statistical Methods in the Atmospheric Sciences, 2nd edn. Elsevier: Academic Press, 627 pp.

Zorita E., H. von Storch (1990): The Analog Method as a Simple Statistical Downscaling Technique: Comparison with More Complicated Methods. Journal of Climate, 12, 2474–2489.