Processing and analysing an ensemble of climate projections for the joint research project KLIWAS

F. Imbery, S. Plagemann, and J. Namyslo
German Weather Service, Offenbach, Germany

Correspondence to: F. Imbery (florian.imbery@dwd.de)

Received: 15 January 2013 – Revised: 9 May 2013 – Accepted: 15 May 2013 – Published: 12 June 2013

Abstract. The research programme KLIWAS, funded by the German Federal Ministry of Transport, Building and Urban affairs is focussed on climate change and its impacts on waterways and navigation for Germany in the 21th century. In order to derive sound statements about the range of possible future climate changes, KLIWAS use hydro-meteorological information derived from a wide variety of global and regional climate models.

In the framework of KLIWAS emphasis is taken on the quantification of uncertainties in climate model output. Therefore, a 19-member ensemble of climate model runs was used. On the basis of the SRES-scenario A1B the probabilities of changes in air temperature, precipitation amount, global radiation and several climate indices were computed for near (2021 to 2050) and distant (2071–2100) scenario horizons.

Furthermore, statistical downscaling techniques, including approved bias correction methods, were used to provide a spatial high-resolution sub-ensemble of eight climate model simulations for climate change impact investigations.

1 Introduction

The assessment of hydrological impacts of climate change demands accurate and high resolution climate model simulation results. For hydrological applications, the key input variables are typically precipitation, air temperature and radiation, at high spatial resolution or at station scale, and with daily or sub-daily temporal resolution (Kotlarski et al., 2012; Nilson et al., 2010).

Today it is common that only one climate model run can’t give the full range of information about uncertainties of possible future climate conditions. Therefore, the use of climate model ensembles is recommended to investigate possible impacts of climate change (Stott and Forest, 2007; Christensen et al., 2010; Deque and Somot, 2010; Giorgi et al., 2009).

Typically, the spatial resolution of global climate models (GCM) is not suitable for hydrological demands on catchment scale (Fowler et al., 2007). Therefore regional climate models (RCM) are used to downscale the simulation results of global climate models. Currently, their typical spatial resolution lies between 10 and 50 km. By using regional climate models, the influence of mountains and also the impact of different land-uses on the local climate, for example, are represented in considerably better detail. Likewise the temporal variability and also the intensity of precipitation amounts are much better captured.

Climate models produce a reduced representation of natural processes such that a systematic error (bias) arises. When using climate model simulation results for impact models, these internal errors must be quantified and corrected. The correction can be realised using the comparison between the control run of the climate model and observed data (e.g. 1961–1990). Depending on the variables to be corrected different methods of bias-correction can be used (Piani et al., 2010a, b; Fowler et al., 2007; Themeßl et al., 2012).

Bias correction procedures assume that model biases do not change over time. Although it is not possible to test whether this assumption is fulfilled for future climate conditions, bias stationarity is a main concern when using bias correction procedures (Maraun, 2012; Teutschbein and Seibert, 2012). On the other side it is important to be aware that the consistency between different variables may get lost after the
application of statistical correction terms. Hence, for studies based on climate model output it is important to be aware of possible disadvantages using bias-corrected model data. While corrected RCM datasets are required as input for impact investigations in most cases, studies on climate change signals can also be carried out with uncorrected RCM data.

In this study, we introduce an ensemble of high resolution climate projection runs covering Middle Europe for the time period 1961–2100. Meteorological variables of eight climate projections from different GCM-RCM combinations and three Re-Analysis data driven model runs are down-scaled and corrected on a consistent high resolution grid (5 km × 5 km) using different methods based on statistical and physical equations. In this context, existing methods developed by the DWD (German Weather Service) were modified and new methods developed. The results were evaluated and compared to the original model output. This ensemble can be used as input for various climate change impact investigations.

Secondly, a statistical method was developed for analysing climate model ensembles. With this technique it is possible to detect climate change signals as well as uncertainties in the used ensembles. For the investigations on climate change signals in this work we used original model output of 19 different GCM-RCM Combinations.

2 Considered area and methods

2.1 Domain

The region of interest in this study includes Germany and the hydrological catchment areas of the rivers Rhine, Elbe, Oder and partly Danube (approximately from 45 to 56° N and 3 to 20° E). A Lambert Conformal Conic map projection on a spatial resolution of 5 km with the European Terrestrial Reference System (ETRS-LCC) (Annoni et al., 2003) is the spatial basis for this study.

2.2 Data

An ensemble of regional climate models (RCM) is used, which were driven by several GCM on the basis of climate simulations of the years 1961–2000 (control run) and of the years 2001–2100 with SRES A1B climate change scenario (projection run). This multi-model ensemble is composed of different RCM developed in the European framework project ENSEMBLES (Van der Linden and Mitchell, 2009), two so-called consortial runs of CLM (Keufer et al., 2009a, b, c; Lautenschlager et al., 2009) and two additional REMO runs (Jacob et al., 2005a, b, 2009a, b). In addition 3 RCM runs, forced by lateral boundary conditions of the Re-Analysis ERA40 data set (Uppula et al., 2005) are used in this study to assess uncertainties of the models. An overview of the used RCM runs can be found in Table 1. Currently, we regionalised and bias-corrected eight climate projections of this ensemble (so-called KLIWAS ensemble, marked with [x] in Table 1).

Gridded observational data for air temperature and precipitation amount were obtained from the German Weather Service DWD. They were developed in context of KLIWAS and the HYRAS pilot project, funded by the Federal Institute of Hydrology (Bundesanstalt für Gewässerkunde – BfG). These high-resolution (1 km × 1 km) gridded daily data sets are referred to as DWD/BfG-HYRAS v2.0 precipitation data set (HYRAS-PRE) for precipitation amount (Rauthe et al., 2013) and DWD/BfG-HYRAS v0.21 temperature data set (HYRAS-TAS) for the air temperature (Rauthe et al., 2012). Both variables have the same geodetic reference system as the targeted map projection in this study, therefore it is only necessary to aggregate the grid information in 5 km resolution.

Furthermore, for gridded global radiation information the CM-SAF Surface Radiation MVIRI Data Set 1.0 (Posselt et al., 2011) is used and interpolated with distance weighted average.

2.3 Regionalisation

2.3.1 Precipitation

The regionalisation of precipitation amount is partly adopted from the observational interpolation method REGNIE (Rauthe et al., 2013). In a first step background fields are constructed by a multiple linear regression and for optimising the explained variance by the Wards- and k-means-clustering. Once the clustered areas in the domain are certain, a multiple Regression with the predictor variables, geographical longitude, latitude, height above sea level and exposition of hillside, could be calculated by the use of the Cholesky decomposition (e.g. Kanzow, 2005). The resulting regression coefficients are applied for the whole cluster in spatial original (e.g. 25 km) and high resolution scale (5 km). The residuals are assessed for each month as a mean of long-term (30 yr) period with a moving average of 10 yr. Afterwards all residuals and predictors are interpolated conservative on the high resolution grid. Consequently the high resolution values are computed related to the 3 × 3 adjacent RCM grid cells. The required relations are the monthly background fields and are used for the daily precipitation amount by inverse distance weighting.

2.3.2 Air temperature

For air temperature a multiple linear regression is used the spatial high resolution regionalisation. Thereby the potential predictor height above sea level is predefined. Due to a screening method (e.g. Urban and Mayerl, 2008) and cross-validation (e.g. Michaelson, 1987) the other potential predictors (geographical longitude and latitude, sea level pressure, humidity in 700 hPa, air temperature in 850 hPa, thickness 500–925 hPa (geopotential height) and the products of air
Table 1. Overview of climate simulations of the years 1961–2000 for the control run (C20), projection runs for the years 2001–2100 based on the scenario A1B and Re-Analysis driven runs (ERA40) used in this study. The symbol X in the last column indicates the projections that are regionalised and bias-corrected.

<table>
<thead>
<tr>
<th>Control run/ SRES scenario/ reanalysis driven run</th>
<th>GCM</th>
<th>RCM</th>
<th>KLIWAS ensemble</th>
</tr>
</thead>
<tbody>
<tr>
<td>HadCM3Q0 (HC)</td>
<td>CLM2.4.6 (ETHZ)</td>
<td>HadRM3Q0 (HC)</td>
<td>X</td>
</tr>
<tr>
<td>HadCM3Q16 (HC)</td>
<td>HadRM3Q16 (HC)</td>
<td>RCA3 (C4I)</td>
<td></td>
</tr>
<tr>
<td>HadCM3Q3 (HC)</td>
<td>HadRM3Q3 (HC)</td>
<td>RCA3 (SMHI)</td>
<td></td>
</tr>
<tr>
<td>BCM2 (NERSC)</td>
<td>RCA3 (SMHI)</td>
<td>HIRHAM5 (DMI)</td>
<td>X</td>
</tr>
<tr>
<td>ECHAM5-r3 (MPI-M)</td>
<td>RCA3 (SMHI)</td>
<td>RegCM3 (ICTP)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>HIRHAM5 (DMI)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RACMO2 (KNMI)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>REMO5.7 (MPI-M)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>ECHAM5-r2 (MPI-M)</td>
<td>REMO5.8 (MPI-M)</td>
<td>CLM2.4.11 (GKSS)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CLM2.4.11 (GKSS)</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>REMO5.7 (MPI-M)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>ARPEGE (CNRM)</td>
<td>HIRHAM5 (DMI)</td>
<td>RM5.1 (CNRM)</td>
<td></td>
</tr>
<tr>
<td>ERA40</td>
<td>CLM2.4.6 (ETHZ)</td>
<td>REMO5.7 (MPI-M)</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>REMO5.7 (MPI-M)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RM4.5 (CNRM)</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Temperature in 850 hPa multiply with thickness 500–925 hPa and sea level pressure multiply with thickness 500–925 hPa are chosen for regression to minimise the mean square error. The resulting regression coefficients apply for the whole cluster in spatial original (e.g. 25 km) and high resolution scale (5 km), which are used for calculation of residuals. Last, all chosen predictors and residuals are interpolated conservative to find the air temperature in high resolution grid.

2.3.3 Global radiation

The chosen method for regionalisation of global radiation is not a statistical approach but is based on an empirical procedure. First, for each day of the year and for different Linke turbidity factors, which represent the full spectrum radiation attenuations (Rigollier et al., 2000), the global radiation for clear sky at each grid is calculated. Considering geographical longitude, latitude and height above sea level, solar altitude angle and altitude correction are included (Remund et al., 2003; Whiteman, 2000). The results for the RCM original grid and the targeted high resolution grid are stored in so-called background fields. Then the RCM global radiation dataset is randomized in global radiation for clear sky taking into account cloudiness fraction. These values are compared to the background fields and evaluated. If these computed results are between two global radiation values for clear sky and defined Linke turbidity factors, a simple numeric approximation calculates the correct turbidity. After this, the Linke turbidity factors are interpolated using inverse distance weighting. At last, the procedure is executed in the opposite direction to define the global radiation in the case of the total cloud cover in spatial resolution of 5 km × 5 km.

2.4 Bias correction

2.4.1 Precipitation

For daily precipitation amounts in climate simulations ($P$) two different bias correction methods were carried out, using model data and observational data (HYRAS) for the time period 1961–1990 for calculating correction terms for future scenarios 2000–2100.
On the one side, a linear scaling approach (Hashino et al., 2007) is used. For every grid $i$ and every month $j$ a correction factor $f$ is computed by calculating the mean ratio between observed reference data ($R$) and simulated data ($C$) of the control period.

\[
f = R^j_i \cdot C^{-1}_j \\
P^* = f \cdot P
\]

New daily values $P^*$ are calculated with the monthly correction factor $f$.

Also, an empirical quantile mapping approach was carried out. In this case, empirical distributions for observed ($F_{oi}$) and simulated ($F_{sj}$) precipitation amounts were calculated for each grid and month (Hashino et al., 2007; Piani et al., 2010b). With $F_{oi}$ and $F_{sj}$ a transfer function can be obtained, with which a distribution depending correction

\[
P^*_j = [F_{oi}^{-1} F_{sj}](P^*_j)
\]

can be carried out.

### 2.4.2 Air temperature

For simulated daily mean air temperature ($T_a$), a linear scaling approach was carried out. In contrast to precipitation amount, the bias of temperature was calculated as the difference of monthly mean values of simulated control data ($C$) and observed reference data ($R$) for each grid.

\[
\text{bias} = R^j_i - C^j_i \\
T^*_a = T_a - \text{bias}
\]

### 2.4.3 Global radiation

Values of global radiation were corrected similar to precipitation by computing a correction factor with the ratio of simulated control data and observed reference data (see Eq. 1).

### 2.5 Uncertainties in climate projections

For the consideration of relative changes (trends, climate change signals) of mean values of meteorological variables, a bias-correction is not generally necessary. The use of original simulation data gives the possibility to access more ensemble members which lead to a better assumption of uncertainties for the investigated variables.

To represent the bandwidth of possible future climate conditions, an ensemble of 19 climate projections, based on the SRES emission scenario A1B (IPCC, 2007) was analysed (see Table 1). This ensemble consists of RCM runs in a spatial resolution of 25 km x 25 km that were computed over the time period 1961–2100.

Percentiles of this ensemble were evaluated for different meteorological variables and indices for the time periods 2021–2050 and 2071–2100 relative to the control period 1961–1990, respectively. A conservative lower threshold (15%) is defined that is almost certainly exceeded, and an upper limit (85%) is set that expresses unexpectedly drastic changes. The span between these two thresholds covers 70% of the investigated ensemble and can be understood as a range in which occurrence can be expected in this scenario.

### 3 Results

#### 3.1 Post-processing of climate model data

In this section we present some examples of the verification results for the post-processed RCM data. As described in Sect. 2.4.1, for precipitation amounts two different bias-correction methods were used. In Table 1, the processed model runs are indicated with a [x]. In Fig. 1, the results for the linear scaling approach and the quantile mapping method are compared on four different grid points for the ECHAM5-r3 REMO5.7 projection. The selected grid points represent different landscapes in the KLIWAS domain: coastal regions (North Sea, Schleswig-Holstein west coast, 1), hilly landscape, dominated by croplands (Thuringian Forest, 2), mid-latitude mountain regions (Black Forest, 3) and alpine regions (Berchtesgaden National Park, 4). For the time period 1961–1990, all precipitation events are represented in the qq-plots. Dots near the 1 : 1 line represent a perfect accordance to the reference data, dots under this line can be interpreted as an underestimation of precipitation amounts.

A slightly improvement can be seen for the simulation data corrected by the linear scaling method. Improved results can be achieved for all regions with the quantile mapping correction. For both methods, limits of improvement can be seen in the highest parts of the distribution (about upmost three percents). This is in agreement with other investigations (e.g. Hashino et al., 2007; Theimerl et al., 2012).

Different metrics are used for the verification of the processed data. For example, the Perkins score (Perkins et al., 2007) is used for testing the quality of the spatial representation of the bias corrected projections. In a first step, for the whole investigated area the probability density functions (PDF) of simulated data and reference data were calculated for each month. Afterwards, the cumulative minimum value of both distributions was compared. This can be expressed as

\[
\text{Skill}_{PDF} = \sum^n_i \min(Z_m, Z_o)
\]

where $n$ is the number of bins used to calculate the PDF for a given region, $Z_m$ is the frequency of values in a given bin from the model data, and $Z_o$ is the frequency of values in a given bin from the observed data. If the simulation data reproduce the reference data perfectly, the skill equals one. In Fig. 2 the simulated air temperature of two ERA40-driven model runs (CLM 2.4.6 and REMO 5.7, both with and without a linear scaling correction) is tested with the Perkins score for the time period 1961–1990.
Figure 1. qq-Plots for different bias-correction methods for precipitation for the A1B-ECHAM5r3-REMO5.7 run. Uncorrected (blue), linear-scaling method (orange) and quantile mapping approach (red) in comparison with HYRAS-PRE for 1961–1990 on four different regions in the KLIWAS area. (1) Coastal regions (north sea, Schleswig-Holstein west coast), (2) hilly landscape, dominated by croplands (Thuringian Forest), (3) mid-latitude mountain regions (Black Forest) and (4) alpine regions (Berchtesgaden National Park).

Figure 2. Monthly Perkins scores for regionalised and bias-corrected (linear-scaling) mean air temperature for two ERA40-driven model runs (CLM 2.4.6 and REMO 5.7) for the time period 1961–1990. Dashed lines indicate uncorrected regionalised model runs (ds), solid lines represent regionalised and bias corrected (bc) runs.
Figure 3. Projected relative change [%] of mean summer precipitation (JJA, upper figure), and mean winter precipitation (DJF, below); periods 2021–2050 (left), 2071–2100 (right), with respect to the control run 1961–1990 based on a 19 member ensemble of regional climate projections (see Table 1). Blue colours indicate increasing precipitation, red colours decreasing precipitation.

While the two uncorrected runs show a strong deviation to reference data with values between 0.4 and 0.8, the corrected runs are in good agreement to the observations. The REMO Data have best conformance during summer with a score over 0.95, whereas CLM have highest values during the months October to December.

3.2 Climate change signals

As pointed out in Sect. 2.5, we use an ensemble of 19 climate models to analyse projected changes for several variables and indices. Furthermore, with this technique it is possible to quantify the uncertainties within the given ensemble. As an example, Fig. 3 gives an overview of projected relative changes of mean precipitation for a near (2021–2050) and a far (2071–2100) future in relation to the reference period (1961–1990). With respect to precipitation, a differentiation between the seasons is necessary. By the end of the century, winters throughout Germany are likely to become wetter, whereas summers could become drier. Following the investigated climate projections, mean precipitation in the hot season should decrease only weakly until 2050, but in the last quarter of the century it could decrease by nearly 20 percent. However, changes during the first half of the century display regional variation; the amounts of precipitation in summer may even remain nearly unchanged in some regions. In winter, increases in precipitation up to 10 percent in the near future (2021–2050) and 15 percent in the more distant future (2071–2100) are expected.

The DWD has established a “German Climate Atlas” (http://www.deutscher-klimaatlas.de), which allows to analyse more variables and a wide range of climate indices with the presented technique.

4 Conclusions

We established a 19 member ensemble of climate projections. With this ensemble it is possible to investigate signals of changes due to possible future climate conditions and the underlying uncertainties of climate projections.

Furthermore we regionalised and corrected air temperature, precipitation and global radiation of a subgroup of eight climate projections of this ensemble. These post-processed climate model runs are an important basis for further investigations on the impacts of climate change, e.g. hydrological modelling. Next steps of our work will be the regionalisation and correction of further variables such as air humidity and
wind speed and to expand the ensemble with new projections based on the RCP scenarios (Giorgi et al., 2009).

Acknowledgements. This research has been carried out in the project 1.02 “Provision of Application-oriented and Evaluated Climate Projection Data” within the departmental research programme “KLIWAS—Impacts of climate change on waterways and navigation” that is financed by the German Federal Ministry of Transport, Building and Urban Development.

The ENSEMBLES data used in this work was funded by the EU FP6 Integrated Project ENSEMBLES (contract 505539) whose support is gratefully acknowledged.

Edited by: B. Ahrens
Reviewed by: M. Zimmer and one anonymous referee

References


Jacob, D.: REMO Climate of the 20th Century Run, UBA Project, 0.088 Degree Resolution, Run no. 006210, 1H Data, World Data Center for Climate, CERA-DB “REMO_UBA,C20,1,R006210_1H”, http://cera-www.dkrz.de/WDCC/Compact.jsp?acronym=REMO_UBA,C20,1,R006210_1H, 2005a.

Jacob, D.: REMO A1B Scenario Run, UBA Project, 0.088 Degree Resolution, Run no. 006211, 1H Data, World Data Center for Climate, CERA-DB “REMO_UBA,A1B,1,R006211_1H”, http://cera-www.dkrz.de/WDCC/Compact.jsp?acronym=REMO_UBA,A1B,1,R006211_1H, 2005b.


Van der Linden, P. and Mitchell, J. F. B.: ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project, Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK, 2009.