Construction of a daily meteorological database for climate change related impact studies

Laura DOBOR[1], Zoltán BARCZA[1,2], Ágnes HAVASI[3], Tomas HLÁSNY[4,5]

[1] Department of Meteorology, Eötvös Loránd University, H-1117 Budapest, Pázmány P. sétány 1/A, Hungary. e-mail: doborl@nimbus.elte.hu, bzoli@.elte.hu
[2]Institute of Ecology and Botany, Centre for Ecological Research, Hungarian Academy of Sciences, H-2163 Vácrátót, Alkotmány u. 2-4.
[3]Department of Applied Analysis and Computational Mathematics, Eötvös Loránd University, H-1117 Budapest, Pázmány P. sétány 1/C, Hungary. e-mail: hagi@nimbus.elte.hu
[4]National Forest Centre – Forest Research Institute, Slovakia, T. G. Masaryka 22, 960 92 Zvolen, e-mail: hlasny@nlcsk.org
[5]Czech University of Life Sciences, Faculty of Forestry and Environment, Czech Republic, Kamýcká 129 165 21 Praha 6 - Suchdol

Abstract — Adaptation to climate change, as well as mitigation of the strength of climate change, needs robust impact studies based on reliable input data. Our aim was to create climate scenarios that contain essential daily meteorological data (precipitation, maximum and minimum temperature) for the 1951-2100 period, and to establish an open access database for the scientific community. We utilized the results of ten RCM experiments that were produced and made accessible within the framework of the ENSEMBLES FP6 project. Every climate model suffers from systematic errors (e.g., under- or overestimation of temperature and the amount and frequency of precipitation), which may prevent the impact studies from direct application of the RCM results. There are several bias correction strategies to correct those errors in the RCM datasets. To quantify the systematic errors in the RCM results the modified E-OBS datasets were used, which were also created within the framework of the ENSEMBLES FP6 project.

In order to create the best possible observation based dataset for the bias correction and for the reconstruction of past meteorological conditions, we corrected the E-OBS database with the CRU TS 1.2 (disseminated by the Climatic Research Unit, University of East Anglia, UK) high resolution gridded dataset. CRU TS 1.2 contains monthly averages for selected meteorological parameters at a spatial resolution of 1/6 × 1/6 degree. We used a bias correction method based on the fitting of cumulative distribution functions, which allowed for correcting both the intensity and the frequency of the precipitation.

The resulting database – the so called FORESEE database – contains daily meteorological data based on ten RCM results for 2010-2100, and the E-OBS database corrected by using CRU data for the 1951-2009 period. The target area of the FORESEE database is Central Europe.

In the future, we plan to create a website making the FORESEE database broadly accessible.

Key words*: climate change, impact studies, bias correction, ENSEMBLES, precipitation frequency correction*

1. Introduction

Climate change is one of the most important political, scientific and social challenges of the 21st century. Despite significant efforts and international agreements it seems that anthropogenic greenhouse gas emission can not be reduced to the desired level. It means that mankind is probably unable to reduce the future impacts of climate change and to avoid significant environmental changes.

The effects of the changing climate could impact any living being all over the Earth (IPCC, 2007a). To reduce the damage at different temporal and spatial scales we have to estimate the impacts of the changing climate and develop potential adaptation strategies.

As many natural processes depend on the meteorological conditions, we heavily need a meteorological database that covers both the past and possible future conditions.

The changes in the meteorological conditions in the past could be calculated from direct measurements or observation based interpolated datasets. For the estimation of future conditions climate model results could be used. Global climate models (GCMs) are run at coarse spatial resolution so they are unable to capture sub-grid cell features and fine scale, topography-induced local characteristics. Because of this, GCM results are not feasible to demonstrate and scrutinize local-scale changes.

In order to address the above described problem several downscaling methods are developed. In case of dynamical downscaling, higher spatial resolution models (so called regional climate models, RCMs) are run using GCM outputs as driver (Giorgi, 1990). The RCMs are able to simulate smaller-scale atmospheric processes, due to their higher resolution topography and physics (IPCC, 2007b).

In the past years several international research institutes developed and ran climate models, and most of the results are available through the Internet. Virtually anyone could download RCM results for any target area, but not every user is aware about the availability and characteristics of the model outputs. For a climate researcher it is clear that every climate model result contains systematic errors (Christensen et al, 2008). These systematic errors are not causing problems if only the expected changes are sought, assuming that the systematic errors in the past and in the future are equal (which is a realistic assumption according to Maraun, 2012). However, problems arise when the user requires realistic daily data to perform a climate change related impact study.

In order to develop adaptation strategies the potential impacts of climate changes have to be estimated. For this reason many researcher from different scientific areas (biologists, agronomists, hydrologists etc.) would like to execute climate change related impact studies. Usually for these studies scientific area-specific models are used (for example a crop model or a hydrological model) but high temporal resolution meteorological data for the investigated terms are always required for the models. For climate change related impact studies reliable meteorological data are heavily needed. Because the investigated time period often covers periods in the future, RCM results provide plausible solutions (Dosio and Paruolo, 2011). To eliminate the existing systematic errors in climate model outputs several bias correction methods have been developed (Ines and Hansen, 2007; Déqué, 2007, Piani et al., 2009). Our aim is to create a database for researchers who need daily meteorological data to execute an impact study in their research.

2. Data

In order to create the essential meteorological data needed for a typical impact study (daily precipitation, maximum and minimum temperature) we used RCM results and observation based data for the 1951-2100 period.

We selected 10 RCM-GCM couplings to perform a state-of-the-art bias correction on the daily maximum-, minimum temperature and precipitation time series for the1951-2100 period (Table 1). The RCM results were produced and made accessible within the framework of the ENSEMBLES FP6 project (van der Linden et al., 2009).

For the 1951-2009 period we used the daily E-OBS database (also created within the framework of the ENSEMBLES FP6 project, Haylock et al., 2008) and the monthly CRU TS 1.2 (disseminated by the Climatic Research Unit, University of East Anglia, UK; Mitchell et al., 2004) high resolution gridded dataset.

Availability of an observation-based dataset is a prerequisite for any bias correction. For practical reasons we used version 3 of E-OBS for 1951-1961 and a newer version (precipitation: version 5; temperature: version 7) for the 1961-1990 period. In order to create the best possible observation based dataset for the bias correction and for the reconstruction of past meteorological conditions first we corrected the E-OBS database using the monthly CRU TS 1.2 dataset. We used the CRU dataset to fit the monthly averages and totals because it seems more realistic than the E-OBS climatology for the target area.

The climate model outputs were created within the same project, but there were differences between the datasets (i.e., regarding the used calendar and coordinate system). At the beginning of the work the RCM results were converted to use a 365 days calendar for the study area, Central Europe (Fig.1). Then all of the data were interpolated to a common 1/6 × 1/6 degree horizontal resolution grid with an inverse distance technique.

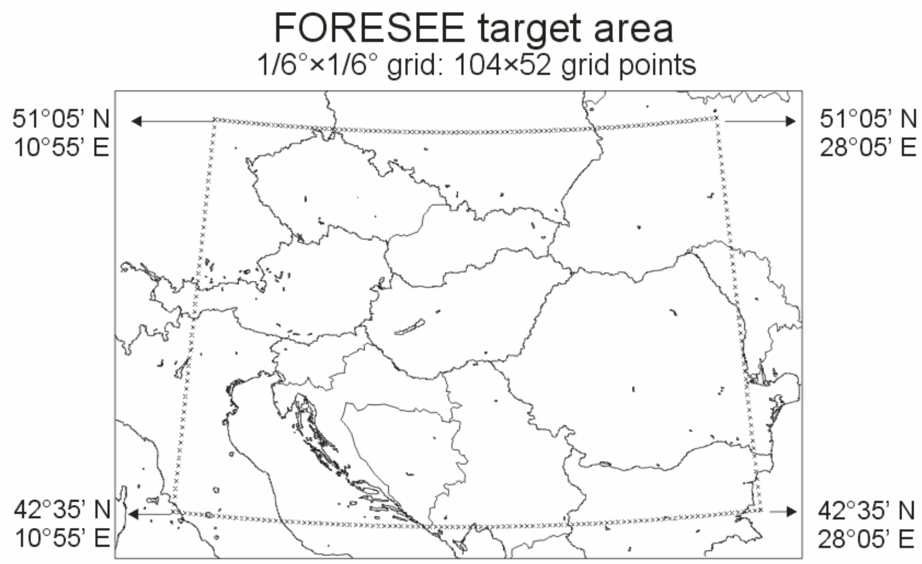

Fig.1. The selected target area of the FORESEE database.

3. Bias correction

For the correction the 1951-2009 period was selected as a reference period – that is the longest period for which observation based data and RCM results are also available. First the climate model results and the observation based datasets were compared for the reference period. Based on the monthly comparison correction factors were defined. These correction factors are then applied to the daily climate model results for the past and also for the future. In case of the temperature the correction means shifting, while in case of rate of precipitation it means multiplication.

The correction of precipitation is a more complex procedure because in a given month precipitation is not just characterized by the sum but also the frequency of precipitation (number of wet days). Apparently, the systematic errors do not just affect the rate of the precipitation. It is evident that for an impact study the frequency of the precipitation is also essential. As an example, is it clearly not irrelevant that the given precipitation amount fall down in 2-3 specific days or is evenly distributed through 15-20 days.

In order to adopt a state-of-the-art bias correction method we corrected not just the rate but also the frequency of precipitation (Ines and Hansen, 2007; Déqué, 2007).

The applied method is based on the cumulative density function (cdf) fitting technique. Here the correction method is demonstrated through precipitation, because the temperature correction is almost the same as the second part of the precipitation correction.

The first part of the correction is the fitting of monthly number of wet days (when the precipitation is not less than 0.1 mm/day). Monthly ratios were determined between the observed and the modelled monthly wet days based on the 1951-2009 period pixel by pixel. At this point two cases have to be discriminated: the model could underestimate and also overestimate the number of wet days in a given month. In case of overestimation wet days were eliminated, while in the other case wet days were created artificially. When wet days were created, the precipitation was set to 0.1 mm for the required number of days. The days were selected in random fashion. When dry days were created, the wet days were neglected starting from days with the least precipitation rate. Fig.2. shows cumulative density functions (cdf) in different stages of the precipitation correction for one month and one selected pixel.

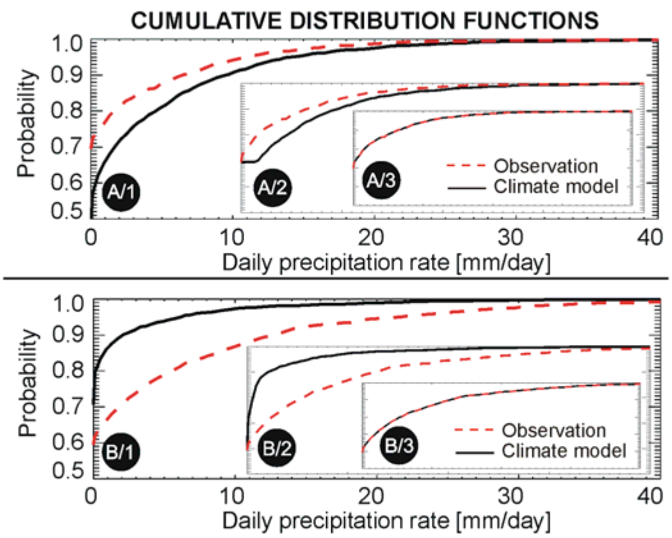

Fig.2. Cumulative distribution function changes during the two steps of the precipitation bias correction. A1-A3 shows the case where the climate model overestimates the monthly number of wet days. B1-B3 shows the case where the climate model underestimates the monthly number of wet days.

In Fig.2 the red dashed line represents the corrected E-OBS database; obviously this line is not changing between the stages (A-B/1-3). The black solid line represents the climate model result. The diagram at the bottom shows when the model underestimated the wet days (B/1-3). The upper one shows an overestimation case (A/1-3).

The first stage is the starting point when the climate model is not yet corrected. The differences between the observed and the modelled number of wet days can be clearly seen on the diagrams A/1 and B/1 of Fig.2. The diagrams A/2 and B/2 show the cdf-s after the correction of precipitation frequency, which was discussed previously.

The second step of the bias correction is the correction of the amount which is executed by cdf fitting. Quantile functions were determined also month by month using 1000 partitions for the corrected E-OBS database and also for the climate model results pixel by pixel. The differences between the observed and the modelled values of the 1000 quantiles can be shown in a quantile-quantile (Q-Q) plot (Fig.3. A/1 and B/1). If the 1000 points are close to the 1:1 line in the Q-Q plot, the cdf-s are also fitted (Fig.3. A/2 and B/2). Thus the correction factors are defined as the ratio of the observed and the modelled quantile functions.

In a given day the precipitation rate is corrected based on the number of the closest quantile (1-1000). The amount of the precipitation is multiplied with the predefined correction factor. In Fig.2. the differences between the A/2-A/3 and B/2-B/3 diagrams show the result of the fitting.

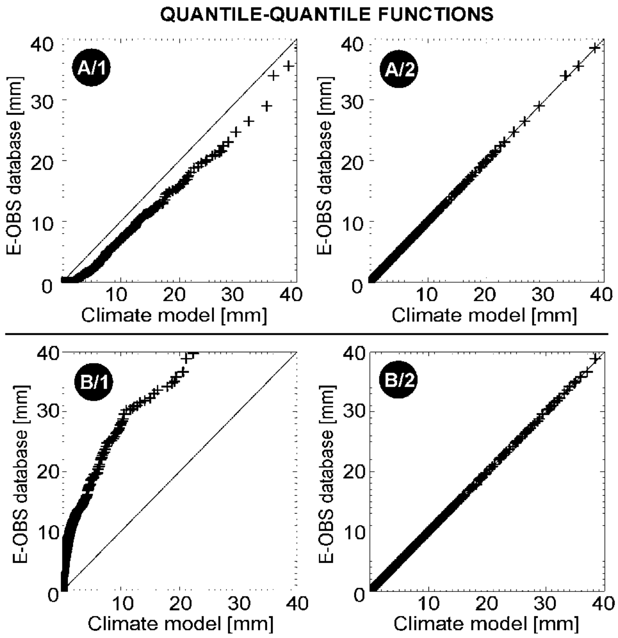

Fig.3. Q-Q plots before and after the correction of the precipitation amount.
A1-A2 shows the case where the climate model overestimates the monthly number of wet days.
B1-B2 shows the case where the climate model underestimates the monthly number of wet days.

4. Results

In the present study bias correction was performed on 10 RCM results retrieved from the database of the ENSEMBLES project (Table 1).

The final time series cover the 1951-2100 period. As the corrected RCM results are fully consistent with the observations used for the correction, in the 1951-2009 period the RCM results were dropped and the CRU corrected E-OBS database was used. This is a necessary requirement of scientists performing impact studies as they generally validate their models using observations. Validation would be impossible if the RCM results were used, as they differ from the realization of the past weather. The results of the bias correction are demonstrated in Fig.4. for all of the three variables (precipitation, maximum and minimum temperature).

In Fig.4 the upper graphs show the annual means/totals for the whole target area based on the non-corrected RCM results. The central graphs show the CRU-corrected E-OBS database before 2009 and the 10 bias corrected RCM results after 2009. The bottom graphs show the averages of the 10 models in case of the corrected (green solid line) and also the not corrected RCM results (black solid line) and finally the CRU-corrected E-OBS database (blue solid line).

Table 1. The list of the RCM-GCM combinations used in the present study.

RCM-GCM	**Developing institute**
ALADIN-ARPEGE	National Centre for Meteorological Research (CNRM)
CLM-HadCM3Q0	Swiss Federal Institute of Technology Zürich (ETHZ)
HadRM3Q0- HadCM3Q0	Hadley Centre for Climate Prediction and Research (HC)
HIRHAM5-ARPEGE	Danish Meteorological Institute (DMI)
HIRHAM5-ECHAM5	Danish Meteorological Institute (DMI)
RACMO-ECHAM5	Royal Netherlands Meteorological Institute (KNMI)
RCA-ECHAM5	Sweden's Meteorological and Hydrological Institute (SMHI)
RCA-HadCM3Q0	Sweden's Meteorological and Hydrological Institute (SMHI)
REGCM-ECHAM5	The Abdus Salam International Centre for Theoretical Physics (ICTP)
REMO-ECHAM5	Max Planck Institute for Meteorology (MPI)

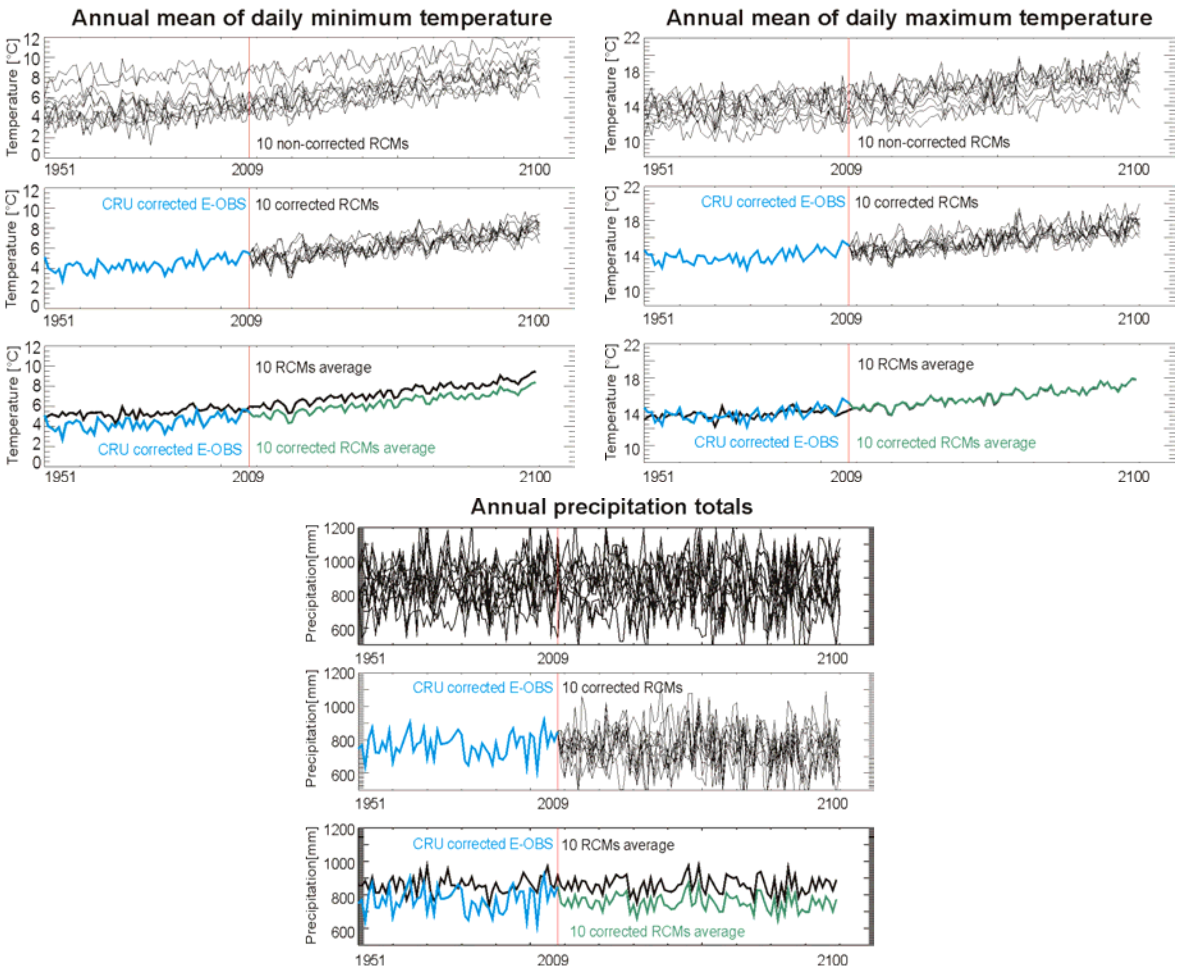

Fig.4. The annual average of daily minimum/maximum temperature and the annual precipitation sums for the whole target area before and after the bias correction.

Based on the results (Fig. 4) it is clear that the variability of the 10 models is appreciable. If the non-corrected RCMs results are compared to the observation based time series the need of the correction is clearly demonstrated. It can be seen in the figure that the bias corrected results are consistent with the observations. The presented charts refer to spatial averages, so the local differences are probably much higher.

The size of the created daily database is about 22 GB, which contains 33 NetCDF files. This means 10 files for every variable for the future (2009-2100) and one for the past (1951-2009). Monthly means are also created for the 1951-2100 time period. The size of these 30 NetCDF files is about 1 GB.

5. Conclusion

For climate change related impact studies RCM results may provide essential base data. However, the climate model result may not be used without further processing steps because every climate model output has intrinsic systematic errors. The bias correction is a process which eliminates the systematic error by using an observation based dataset.

In the present study we described the steps of the construction of a database that is aimed to support impact studies in Central Europe. Based on the presented bias corrected RCMs a new database was built, which is called the FORESEE database (Open Database **For** Climat**e** Change Related Impact **S**tudies in C**e**ntral **E**urope). We are planning to create a homepage for FORESEE, and set the database accessible for the scientific community.

Acknowledgements **—** The Project is supported by the European Union and co-financed by the European Social Fund (grant agreement no. TAMOP 4.2.1/B-09/1/KMR-2010-0003). We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://eca.knmi.nl). The ENSEMBLES data used in this work was funded by the EU FP6 Integrated Project ENSEMBLES (Contract number 505539) whose support is gratefully acknowledged. This research was funded by the Hungarian Scientific Research Foundation (OTKA K104816 and K67672), by the Biodiversity Virtual e-Laboratory (BioVeL) and by the CarpathCC. The authors gratefully acknowledge the monthly high resolution dataset for the Climatic Research Unit.

References

Christensen, J.H., Boberg, F., Christensen, O.B., Lucas-Picher, P., 2008: On the need for bias correction of regional climate change projections of temperature and precipitation. Geophys. Res. Lett. 35., L20709.

Déqué, M., 2007: Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values. *Glob. Plan. Change*, 57: 16-26.

Dosio, A. and *Paruolo, P.,* 2011: Bias correction of the ENSEMBLES high-resolution climate change projections for use by impact models: Evaluation on the present climate. J. Geophys. Res., 116, 1–22.

Giorgi, F. 1990: Simulation of regional climate using a limited area model nested in general circulation model, J. Clim., 3, 941–963.

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

Ines, A. and *Hansen, J.,* 2006: Bias correction of daily GCM rainfall for crop simulation studies. Agric. For. Meteorol., 138, 44-53.

IPCC, 2007a: Climate change 2007: Impacts, adaptation and vulnerability. In: M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden and C.E. Hanson (eds) Contribution of Working

Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, 2007 Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. 976 pp

IPCC, 2007b: Climate change 2007: The physical science basis. In: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L. (eds), Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996p.

Maraun, D., 2012: Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums, Geophys. Res. Lett., 39, L06706.

Mitchell T., Carter T.R., Jones P.D., Hulme M., New M., 2004: A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901-2000) and 16 scenarios (2001-2100). Tyndall Centre. Working Paper 55.

Piani, C., Haerter, J.O., Coppola, E., 2009: Statistical bias correction for daily precipitation in regional climate models over Europe. Theor. Appl. Climatol., 99, 187–192.

van der Linden, P., Mitchell, J.F.B., 2009: ENSEMBLES: Climate Change and Its Impacts: Summary of research and results from the ENSEMBLES project. Office, p. 160.