9th INTERNATIONAL WORKSHOP on PRECIPITATION IN URBAN AREAS
Urban Challenges in Rainfall Analysis
6-9 December 2012, St. Moritz, Switzerland

BIAS CORRECTION OF PRECIPITATION DATA FROM THE CLM REGIONAL CLIMATE MODEL FOR WATER MANAGEMENT MODEL APPLICATIONS

by

M. Quirmbach[(1)], E. Freistühler[(1)], I. Papadakis[(1)] and A. Pfister[(2)]

[(1)] *dr. papadakis GmbH, Werksstraße 15, D-45527 Hattingen, Germany* (m.quirmbach@drpapadakis.de)

[(2)] *Emschergenossenschaft/Lippeverband, Kronprinzenstraße 24, D-45128 Essen, Germany* (pfister.angela@eglv.de)

ABSTRACT

The precipitation data of the Regional Climate Model CLM are used for the water management impact models within the *dynaklim* networking and research project. For this purpose, it is necessary to apply a bias correction to the CLM precipitation data. First, the bias assessed for varying temporal resolutions and precipitation characteristics is described. Subsequently, a method for the bias correction is introduced. The developed methodology is a modified form of the so-called quantile mapping. The focus lies on the corrections of the dry days and the heavy rainfall events. They are considered separately, deviating from other quantile mapping procedures.

Keywords: bias correction, regional climate model, CLM, *dynaklim*, rainfall, quantile mapping

1 INTRODUCTION

The networking and research project *dynaklim* – Dynamic Adaptation of Regional Planning and Development Processes to the Effects of Climate Change in the Emscher-Lippe-Region (North Rhine-Westphalia, Germany) – carries out multi-disciplinary research on 'dynamic' adaption to the effects of climate change. The project focuses on the potential impacts on the regional water balance and the possible options to adapt for population, economy and environment.

Results from different water management and economic impact models are used for these investigations. These impact models themselves, however, require climate data from measurements and regional climate models as input – especially on precipitation.

The Global Climate Model ECHAM5 provides humidity data that are too high to the Regional Climate Models with the amount of overestimation varying strongly from region to region all over Germany. As a consequence, precipitation amounts that are too high are featured in the Dynamic Regional Climate Models (Hollweg et al., 2008). This systematic error (bias) can also be seen in the Regional Climate Model CLM which is used in the research project *dynaklim*.

A correction of the precipitation model data is imperatively necessary for water management model applications which integrate the absolute values of these data, since the errors in the precipitation model data directly affect the resultant runoff to a significant degree. For the Regional Climate Model CLM in the Emscher-Lippe Region (ELR), the amount of bias in the model data in comparison to the measurement data concerning the precipitation parameter was examined. Taking the investigation results as a basis, a methodology was developed to correct the precipitation model data so that they meet the requirements of water management applications.

2 DATA ANALYSIS AND QUANTIFICATION OF PRECIPITATION BIAS

2.1 Data basis and scope of investigation

The reference period selected for examining the precipitation bias was the period from 1961 - 1990. This is regarded as the WMO standard normal period. Within this time period the modeled precipitation from both runs, CLM_C20_1_D3 (CLM1) and CLM_C20_2_D3 (CLM2) of the Regional Climate Model CLM

(Lautenschlager et al., 2009) was compared with the precipitation measured in the Emscher-Lippe Region (ELR). The investigations are carried out for the 18 grid areas displayed in *Figure 1* and the rain gauges in the ELR marked in red.

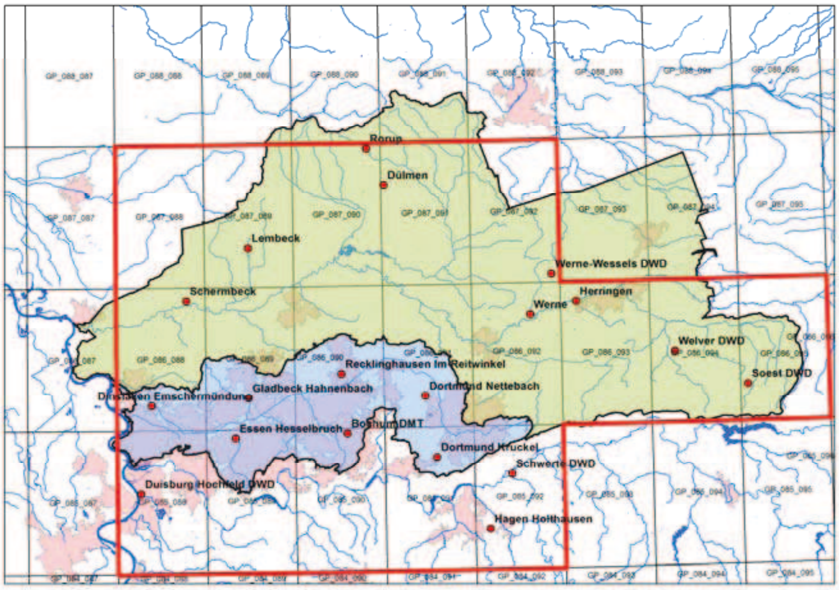


Figure 1 – Position of the CLM grid areas (red frame) including the Emscher (blue) and the Lippe (green) catchment areas; the rain gauges are marked in red

The investigations on bias were carried out for different temporal resolutions and characteristics of precipitation. For this, the mean yearly rainfall totals of the measurement and model data were compared to each other, as were the mean monthly rainfall totals of the individual calendar months. The focus of this contribution lies on the comparison of daily rainfall. This encompasses daily rainfall totals, the number of dry days, dry periods as well as the frequency of heavy rainfall events which is especially relevant for urban systems. Both the daily as well as the hourly rainfall totals were checked with regard to the extent to which they realistically reflect the properties of the measurement data within the reference period. It became clear that only the daily totals realistically reflect the precipitation. Therefore, a bias correction based on the daily totals is carried out. The correction factors established are then applied to the hourly totals.

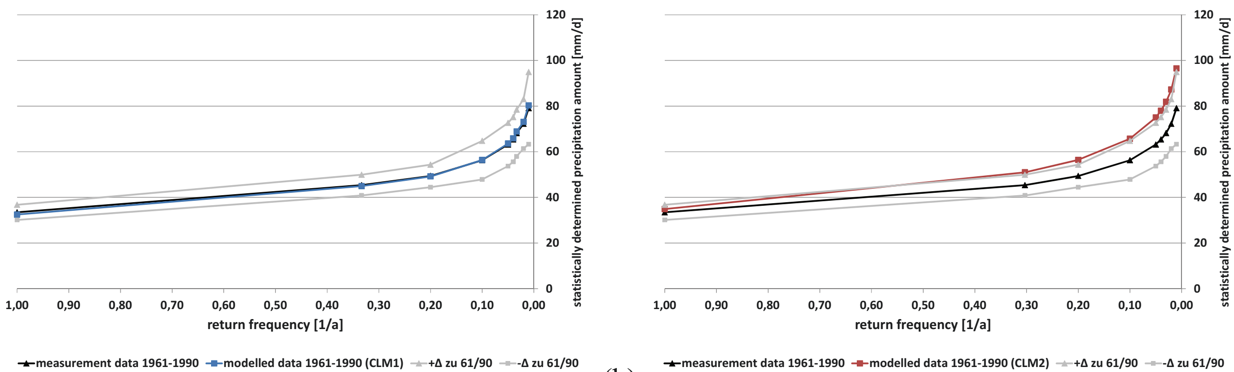
In addition to the bias correction, an additional downscaling of the precipitation model data to smaller time steps (Δt = 5 min) and smaller grid areas (1 km²) is necessary for the use of the precipitation model data as input series in impact models in small scales, e.g. in urban systems. This procedural step is described in a separate paper (Tessendorf et al., 2012).

2.2 Bias assessed in different precipitation parameters

The number of rainfall days created from the global model in the CLM is too high due to the high rate of humidity. The model data show a number of dry days ($R \leq 0.1$ mm/d) which is 40% lower (109 days mean) than assessed by the measured data (184 days mean). The model data also show a significant overestimation of the daily totals between 0.5 mm/d $< R \leq$ 20 mm/d, especially in the range from 5 mm/d $< R \leq$ 10 mm/d. The model data therefore imply an increase in frequency and amount of rainfall which also affects the mean monthly and yearly rainfall totals. The yearly rainfall totals in the model are overestimated by a mean rate of +34%. Depending on the grid area, the overestimation ranges from +17% to +65%. Thus the mean monthly rainfall totals are also overestimated accordingly. The bias in the mean monthly rainfall totals varies depending on the month and the region. Despite the bias, the mean monthly rainfall totals still reflect the seasonal distribution pattern of the measured data quite well.

Comparing the statistic evaluation of heavy rainfall events, the CLM1 time series shows similar daily rainfall totals for the different return periods T_n as they are provided by the measured data. The CLM2 time series, however, shows significantly increased daily totals. Unlike all the previously mentioned investigations the characteristics of the measured data do not deviate from those of the model data here; however, the two CLM time series do reflect different heavy rainfall patterns.

Figure 2a displays the statistically determined daily rainfall totals of the CLM1 time series and the measured time series for different return frequencies (n = 1/Tn). The values provided in both time series are nearly identical. The statistically determined daily rainfall totals of the CLM2 time series in *Figure 2b*, however, are approximately 10% - 20% higher than the statistically determined daily rainfall totals of the measured time series. What needs to be taken into consideration here though is the strong random component inherent in rare heavy rainfall events regarding their spatial and temporal occurrence. This large natural variability is also reflected in the statistical evaluations at the 18 rain gauges used in the ELR. The 10% - 20% deviations assessed in the CLM2 time series approximately match the range of tolerance as given in the statistic evaluations of the KOSTRA-DWD-2000 (DWD, 2005). The range of tolerance is marked in grey in *Figure 2*. The evaluation of heavy rainfall events from the CLM2 time series thus also provides possible results that lie within the range of tolerance - even before a bias correction is carried out. The differences in the two CLM time series should remain intact after a bias correction since they don't represent a systematic error but rather two different yet possible realizations within the given range of tolerance.



(a) (b)

Figure 2 – Comparison of the statistically determined precipitation amounts during the reference period for various return frequencies, (a) CLM1 vs. measured data; (b) CLM2 vs. measured data

3 DESCRIPTION OF METHOD

The analysis of the measured and the model data has shown that the precipitation bias varies both spatially as well as seasonally. Correction of the precipitation amounts is therefore carried out separately for regions that show similar characteristics. Apart from that specific correction factors are determined per calendar month.

The developed methodology is based on so-called quantile mapping which also serves as the basis for other bias correction methods in precipitation data of regional climate models (Piani et al., 2010; Muddelsee et al., 2010; Thermeßl et al., 2011). Different methods can be applied in order to transform model data into corrected data (Gudmundsson et al., 2012). Here, the mean values and the frequency distribution of the climatic modeling data are adapted to the distributions of the measured values. The empirical distribution functions of the measured and modeled data are compared to each other. The relation measured value / model value of every quantile represents the correction factor of the respective quantile by which all model values of the same size are then multiplied by during the actual correction step. This non-parameterised transformation is used to the 97% quantile for the method developed (*Figure 3b*). Values beyond the 97% quantile are calculated by means of a parameterised transformation using an area-wise linear regression (*Figure 3c*). Since linear regression is open for extrapolation it will also be possible to correct future rainfall events that are stronger than those in the reference period. *Table I* provide an overview of individual procedural steps for the data pool up to the 97% quantile and beyond the 97% quantile.

Table I – General Approach for Bias Correction

Procedural Step	**Correction up to the 97% quantile**	**Correction beyond the 97% quantile**
Methodology	Comparison of individual quantiles of measured and model data (non-parameterized transformation)	Area-wise linear regression between measured and model data (parameterized transformation)
Space	Initially, individual grid areas / gauges, later on aggregation into regions	Regional correction
Season	Monthly	Hydrologic half year period
"Random model noise"	Estimation of a monthly "dry value" → Correction of daily totals by the dry value	Correction of daily totals by the dry value
CLM model run	CLM1 / CLM2 combined	CLM1 / CLM2 combined
Time step	Correction factor derived from daily totals → apply to hourly totals	Correction factor derived from the daily totals → apply to hourly totals

With the developed modified form of quantile mapping, the focus lies on the corrections of the dry days and the heavy rainfall events. They are considered separately, deviating from other quantile mapping procedures.

The model showed a significant underestimation regarding the dry days. The CLM data feature multiple very small rainfall intensities (R < 0.1 mm/h) over longer periods of time. These intensities are not considered plausible and are interpreted as random model noise. However, the correct reflection of dry days is mandatory, for example with regard to the evaluation of dry periods, which leads to individual handling of the dry days during the bias correction. For the correction of the dry days, those quantiles per calendar month are determined which feature measured data close to zero. These quantiles are then used in the comparison between the CLM data and the data provided by the rain gauges. The precipitation amount of the corresponding quantile in the CLM data is equivalent to the threshold value which is then deducted from all CLM daily rainfall totals (*Figure 3a*).

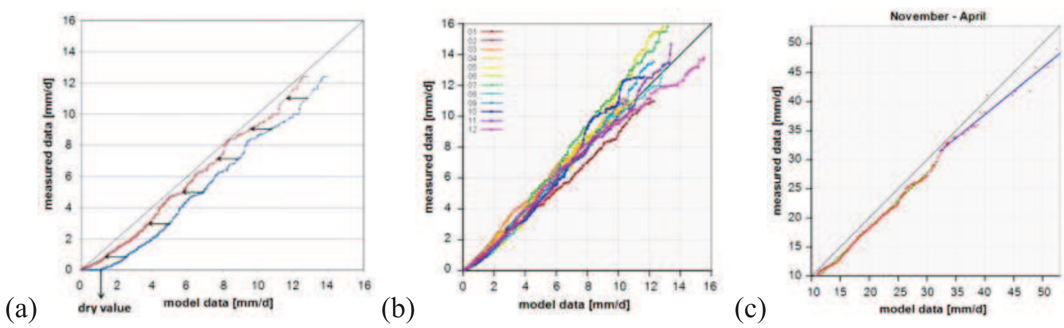


Figure 3 – General approach for bias correction, (a) Assessment of the dry value, (b) Quantile mapping (monthly) up to the 97% quantile, (c) Area-wise linear regression beyond the 97% quantile

Particularly in the case of extreme rainfall events, the random effects of spatial and temporal occurrence strongly influence the correction of the daily rainfall totals. Reasonable data aggregation is therefore carried out for heavy rainfall events, so that – in spite of the small number of extreme rainfall days – a data pool as large as possible is created for the estimation of the correction factors. The 97% quantile has proven suitable for the delimitation of extreme rainfall events from the remaining data pool. The regional observation of grid areas that show similar characteristics and which border with one another already facilitates aggregation for all data. In addition, both CLM runs are treated as one data pool. The model runs do feature significant differences for individual years but their features remain comparable when using a long-term mean value. Further useful aggregation of data that helps decrease the random component in the correction of rare heavy rainfall events is achieved by combined consideration of both CLM time series. An additional combination of the months of the hydrological winter season (November to April) and of the hydrological summer season (May to October) is carried out regarding the data that lie beyond the 97% quantile. The two seasons are nevertheless kept separate in order to be able to take into consideration the varying characteristics of short, convective heavy rainfall events during the summer season and long, advective heavy rainfall events during the winter season. The data up to the 97% quantile, however, are assessed per month.

4 CONCLUSIONS

The developed method has been applied successfully to the CLM precipitation data used in *dynaklim* so that now corrected daily rainfall totals are available covering the whole Emscher-Lippe Region. Apart from the water management related model applications, the corrected precipitation data are used for trend investigations in the Emscher-Lippe Region (Freistühler et al., 2012). The plausibility and quality of the data sets created was checked and proven by means of various criteria and parameters.

The method was not only tested in the Emscher-Lippe Region but also in the Ruhr catchment area. This catchment area features the high altitudes of the Sauerland Region – a characteristic not present in the Emscher-Lippe Region. The bias correction was successfully applied in the Ruhr catchment area as well which shows that applying the method to other regions proves to be unproblematic.

Application of the method to other Regional Climate Models cannot be examined within the framework of *dynaklim*. However, the methodology developed is a method that can generally be applied for bias correction. It essentially uses statistical adjustments of mean values and distribution functions. These adjustments are independent of limiting conditions such as grid size or temporal resolution. It is therefore plausible to assume that application of this method to other regional climate models is possible.

5 ACKNOWLEDGEMENTS

The *dynaklim* – project is part of the development programme KLIMZUG (climate change in regions) and is supported by the German Federal Ministry of Education and Research (BMBF).

6 REFERENCES

DWD, (2005) *KOSTRA-DWD-2000 Starkniederschlagshöhen für Deutschland (1951-2000) - Grundlagenbericht*, Selbstverlag des Deutschen Wetterdienstes, Offenbach am Main, Germany.

Freistühler, E., M. Quirmbach, I. Papadakis and A. Pfister, (2012) Trends in Precipitation Measurement and Climate Model Data and their Influence on the Assessment of Urban Systems. *Proceedings of the 9th International Workshop on Precipitation in Urban Areas*, 6.-9. December 2012, St. Moritz, Switzerland.

Gudmundsson, L., J.B. Bremnes, J.E. Haugen and T. Engen Skaugen, (2012) Technical Note: Downscaling RCM precipitation to the station scale using quantile mapping – a comparison of methods. *Hydr. Earth Syst. Sci. Discuss.*, 9, 6185-6201.

Hollweg, H.-D., U. Böhm, I. Fast, B. Hennemuth, K. Keuler, E. Keup-Thiel, M. Lautenschlager, S. Legutke, K. Radtke, B. Rockel, M. Schubert, A. Will, M. Woldt and C. Wunram, (2008) *Ensemble Simulations over Europe with the Regional Climate Model CLM forced with IPCC AR4 Global Scenarios*, Technical Report No. 3, Gruppe Modelle & Daten (M & D), Max Planck Institute for Meteorology, Hamburg.

Lautenschlager, M., K. Keuler, C. Wunram, E. Keup-Thiel, M. Schubert, A. Will, B. Rockel and U. Boehm, (2009) Climate Simulation with CLM, Climate of the 20th Century run no.1 / no.2, Scenario A1B run no.1 / no.2, Data Stream 3: European region MPI-M/MaD. World Data Center for Climate.

Muddelsee, M., D. Chirilia, T. Deutschländer, C. Döring, J. Haerter, S. Hagemann, H. Hoffmann, D. Jacob, P. Krahe, G. Lohmann, C. Moseley, E. Nilson, O. Panferov, T. Rath and B. Tinz, (2010) Climate Model Bias Correction und die Deutsche Anpassungsstrategie. *DMG Mitteilungen,* 03/2010, 2-7.

Piani, C., J.O. Haerterand E. Coppala, (2010) Statistical bias correction for daily precipitation in regional climate models over Europe. *Theor. Appl. Climatol.,* 99, 187-192.

Tessendorf, A., M. Quirmbach, T. Einfalt, E. Freistühler and M. Jessen, (2012) Statistical downscaling of CLM precipitation using adjusted radar data and objective weather types. *Proceedings of the 9th International Workshop on Precipitation in Urban Areas*, 6.-9. December 2012, St. Moritz, Switzerland.

Themeßl, M.J., A. Gobiet and A. Leuprecht, (2011) Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *Int. J. Climatol.*, 31, 1530-1544.