



Bias Adjustment of EURO-CORDEX High-resolution Simulations by Means of Quantile Delta Mapping: Evaluation on the Climate from the Near Past over Southeast Europe

Hristo Chervenkov* and Kiril Slavov

*hristo.tchervenkov@meteo.bg

National Institute of Meteorology and Hydrology, Bulgaria
1784 Sofia, “Tsarigradsko Shose” 66

Outline of the presentation

1. Introduction - Bias Correction Essence and Basic Assumptions
2. Main Aim, Methodology and Data
3. Data Flow and Performed Calculations
4. Obtained Results
 - Results for the individual EURO-CORDEX models
 - Results for the Multimodel Ensemble
5. Summary, Conclusions and Outlook
6. Acknowledgement

1. Introduction - Bias Correction Essence and Basic Assumptions

Regional climate models (RCMs) have been widely used in recent decades to downscale coarser general circulation model (GCM) simulations and provide fine-scale regional climate information, supporting detailed impact and adaptation assessment and planning. The RCMs' output, however, may exhibit large biases, inherited from the driving GCM, in addition to those introduced by the RCMs. Persistent biases can undermine the reliability of climate projections. Various methods have been created to minimize and correct these biases as sources of error in subsequent modelling chains. The evaluation of the effectiveness of different correction procedures under systematically varying climate conditions, conducted by many researchers, has shown that statistical distribution-based methods outperform all other investigated techniques. Regardless of their variety and underlying premises, their fundamental assumption are:

- The considered climate model produces skillful input for a bias correction (BC), including a plausible representation of climate change
- Stationarity, or time invariance, of the model biases

2. Main Aim, Methodology and Data (1 from 3)

Our main aim is to assess the added value of the bias correction, and thus we have applied quantile delta mapping (QDM), which is distribution-based method, trend-preserving variant of the widely used quantile mapping, on the primary output for the daily minimum, mean and maximum temperature as well as precipitation sum (further: t_n , t_g , t_x , and rr) from 14 combinations of driving GCMs and RCMs from EURO-CORDEX project for the period **1976–2005** over Southeast Europe.

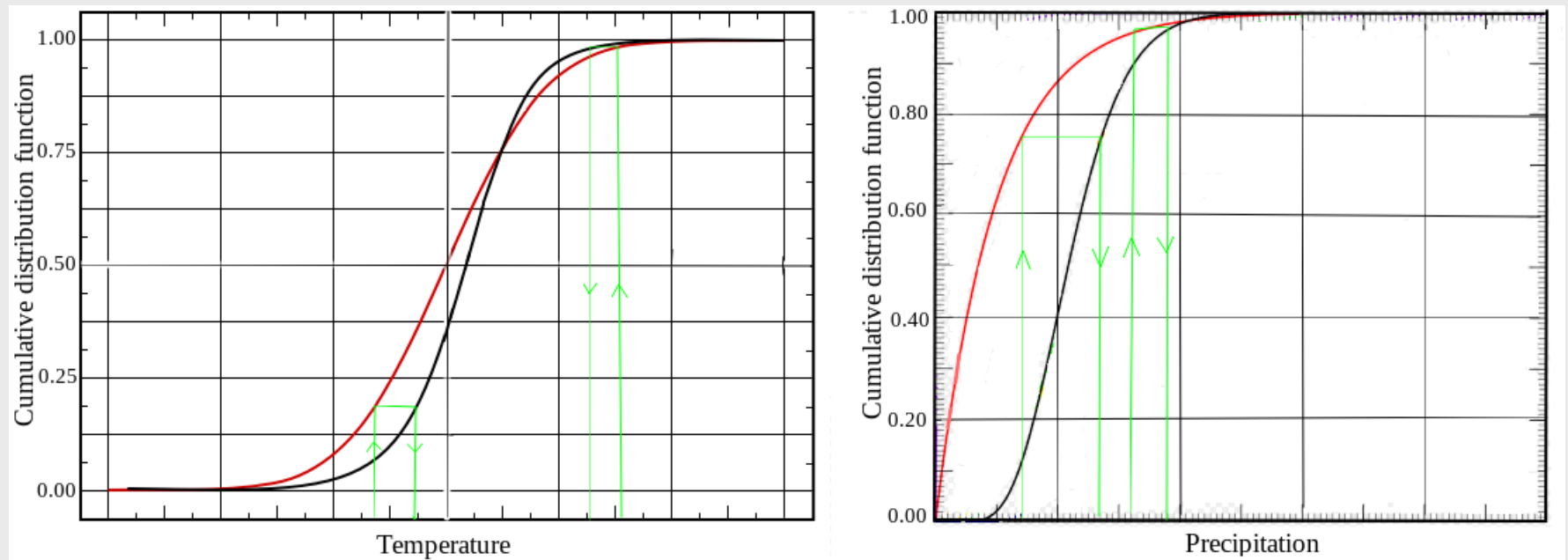
The main idea of most statistical BC methods is to correct the cumulative distribution function (CDF) of the raw RCM-simulated climate values to match that of the reference. The correction of the CDF is usually done by first creating a transfer function using data from the control (or calibration) period in the simulations and the reference, and subsequently applying this function to the raw output of the future simulations. Only data from the control period are used to construct the transfer function; data from future model projections are disregarded.

2. Main Aim, Methodology and Data (2 from 3)

The QDM can be classified as distribution-based (calibrated on CDFs rather than on paired data), direct (predictor and predictand are the same parameters), and parameter-free (using empirical quantiles, rather than theoretical CDFs). QDM adjusts for the combined GCM/RCM bias based on the climate simulation setup. Distribution-wise BC techniques that do not depend on temporal correlation between the model and the reference are the only ones that allow for such an application.

Remarks:

- Data for the near past/present climate could be also BC using control data from the (far) past.
- Any 'reliable enough' (*DanubeADAPT final product?*) dataset can be used as reference, not only (assimilated) observations.



Quantile mapping from the CDF of the simulated (in red) to the CDF of the reference (in black) for the temperature (on left) and precipitation (on right). The green arrows depict the replacement of the quantile of the simulated value by the quantile of the reference distribution corresponding to the same probability.

2. Main Aim, Methodology and Data (3 from 3)

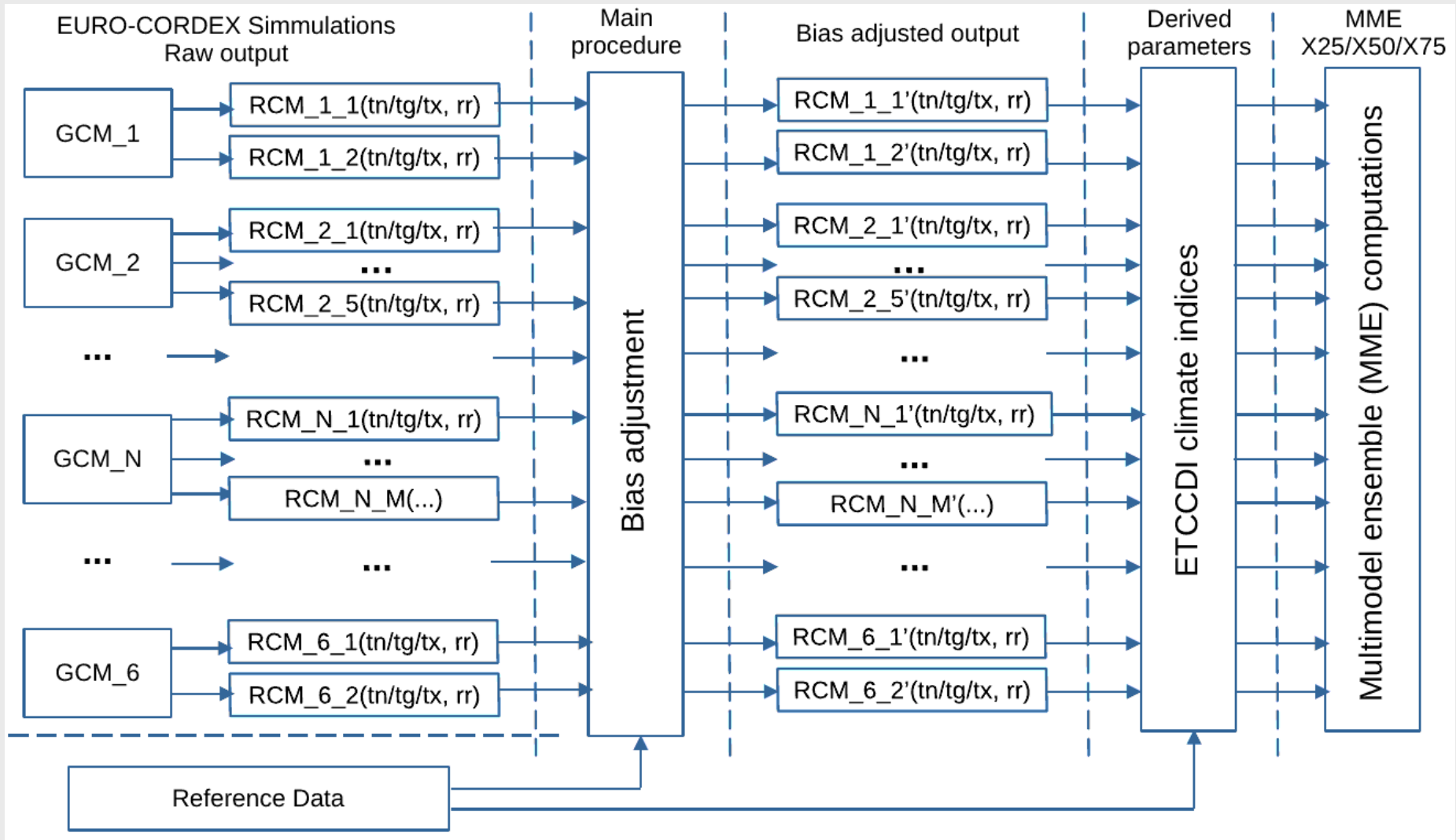
Driving GCM/Institution, Country	RCM/Institution, Country	No.
CNRM-CERFACS-CNRM-CM5/Meteo France, France	CNRM-ALADIN63/Meteo France, France	1
	KNMI-RACMO22E/KNMI, Netherlands	2
ICHEC-EC-EARTH/EC-EARTH Consortium	CLMcom-CCLM4-8-17/CLM Community & contr.	3
	DMI-HIRHAM5/DMI, Denmark	4
	GERICS-REMO2015/GERICS, Germany	5
	KNMI-RACMO22E/KNMI, Netherlands	6
IPSL-IPSL-CM5A-MR/IPSL, France	IPSL-WRF381P/IPSL, France	7
MOHC-HadGEM2-ES/MOHC, UK	CLMcom-CCLM4-8-17/CLM Community & contr.	8
	DMI-HIRHAM5/DMI, Denmark	9
	GERICS-REMO2015/GERICS, Germany	10
	KNMI-RACMO22E/KNMI, Netherlands	11
MPI-M-MPI-ESM-LR/MPI-M, Germany	MPI-CSC-REMO2009/MPI-M, Germany	12
NCC-NorESM1-M/NCC, Norway	DMI-HIRHAM5/DMI, Denmark	13
	GERICS-REMO2015/GERICS, Germany	14

Descriptive list of the used 14 EURO-CORDEX GCM/RCM combinations

3. Data Flow and Performed Calculations (1 from 2)

- The study utilizes the Climate Data Store (CDS) of the Copernicus Climate Change Service (C3S) as information source for all primary data.
- The calibration period is selected **1951-1975**, which represents the longest continuous span available in the CDS prior to the historical period of **1976-2005**
- The data for tn, tg, tx, and rr from ERA5-Land used as reference and from all 14 GCM/RCM EURO-CORDEX combinations are downloaded from the CDS (Python API)
- Selection of the domain, EURO-CORDEX data reprojection (from the native rotated grid to the standard ERA5-Land grid), time aggregation of the ERA5-Land data.
- Bias adjustment for each model one-by-one by means of CLI-based BiasAdjustCXX (parallelized via OpenMP, 4 threads)
- Batch processing (MPI-M cdo for all netCDF file operations; Linux bash shell scripts)

3. Data Flow and Performed Calculations (2 from 2)

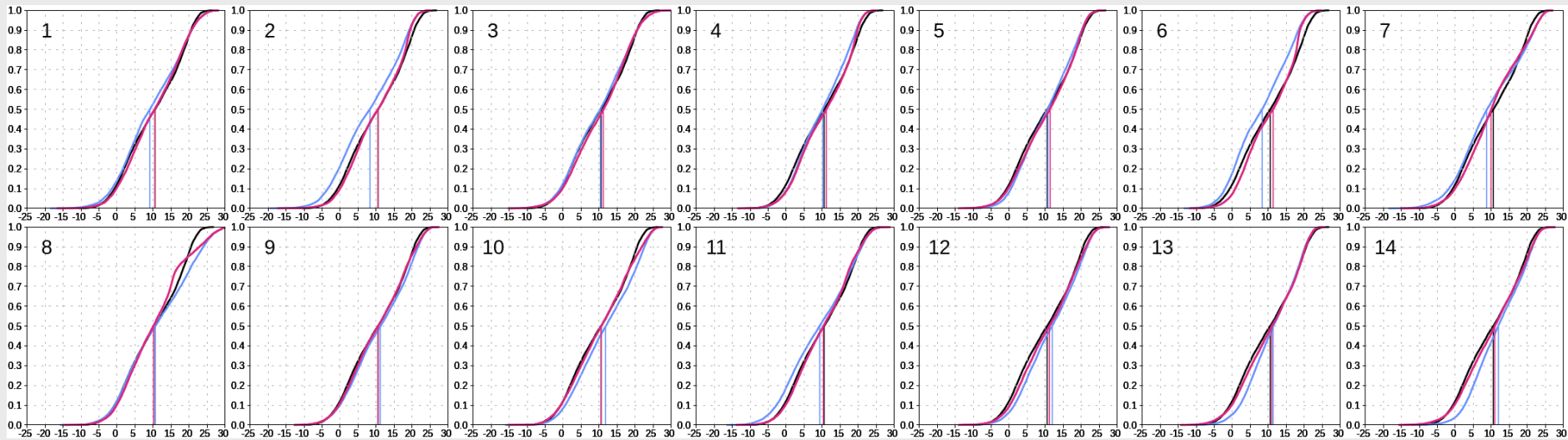


Generalized workflow chart of the performed investigation

4. Obtained results

- Results for the individual EURO-CORDEX models

Cumulative Probability Functions (1 from 2)

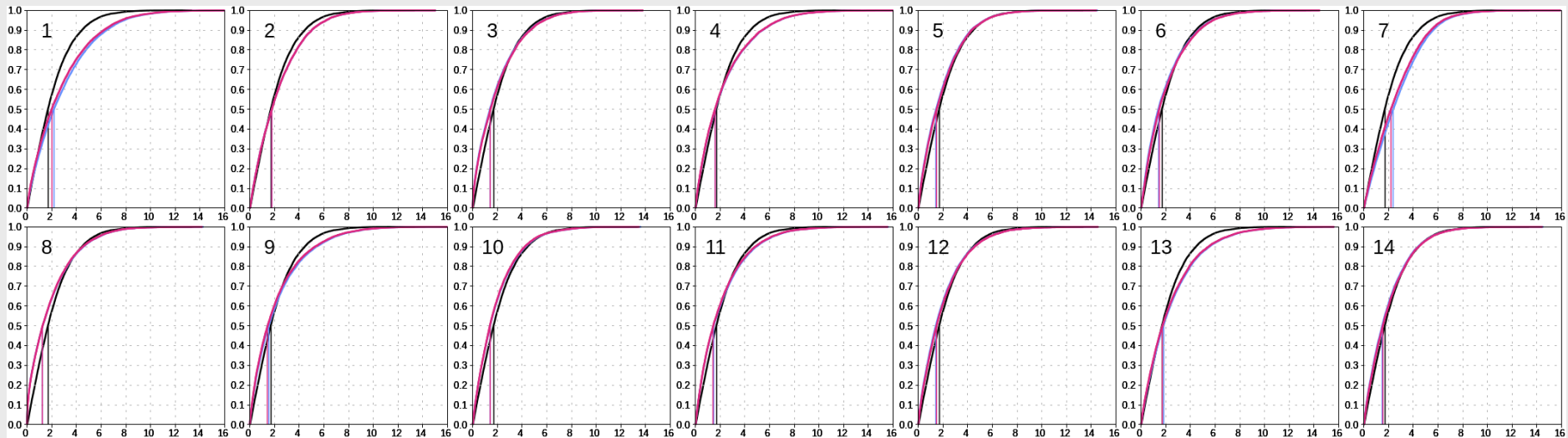


Empirical CDFs of the mean temperature (unit: °C) of the reference, raw, and corrected output in black, blue, and red correspondingly. The vertical bars show the medians' values. The models' numeration is according to the table.

4. Obtained results

- Results for the individual EURO-CORDEX models

Cumulative Probability Functions (2 from 2)

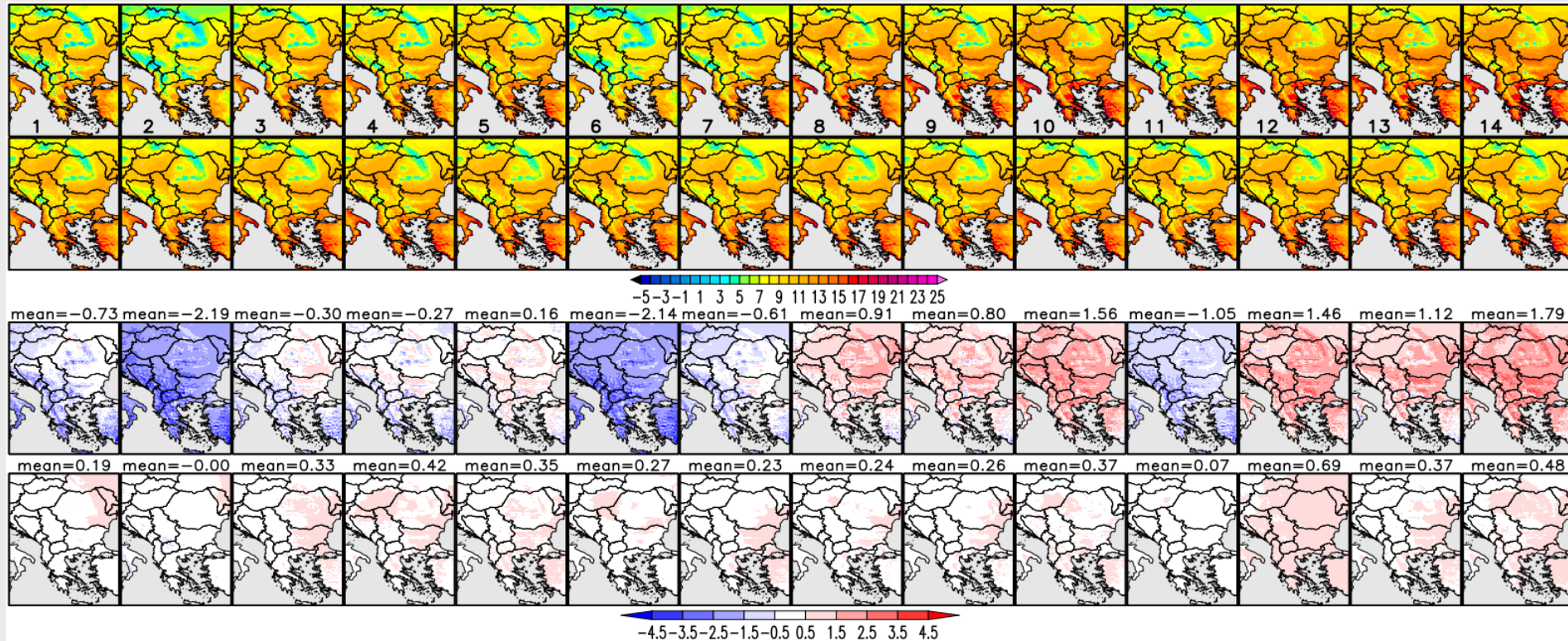


Empirical CDFs of the daily precipitation sum (unit: mm) of the reference, raw, and corrected output in black, blue, and red correspondingly. The vertical bars show the medians' values. The models' numeration is according to the table.

4. Obtained results

- Results for the individual EURO-CORDEX models

Multiyear Annual Statistics (1 from 2)

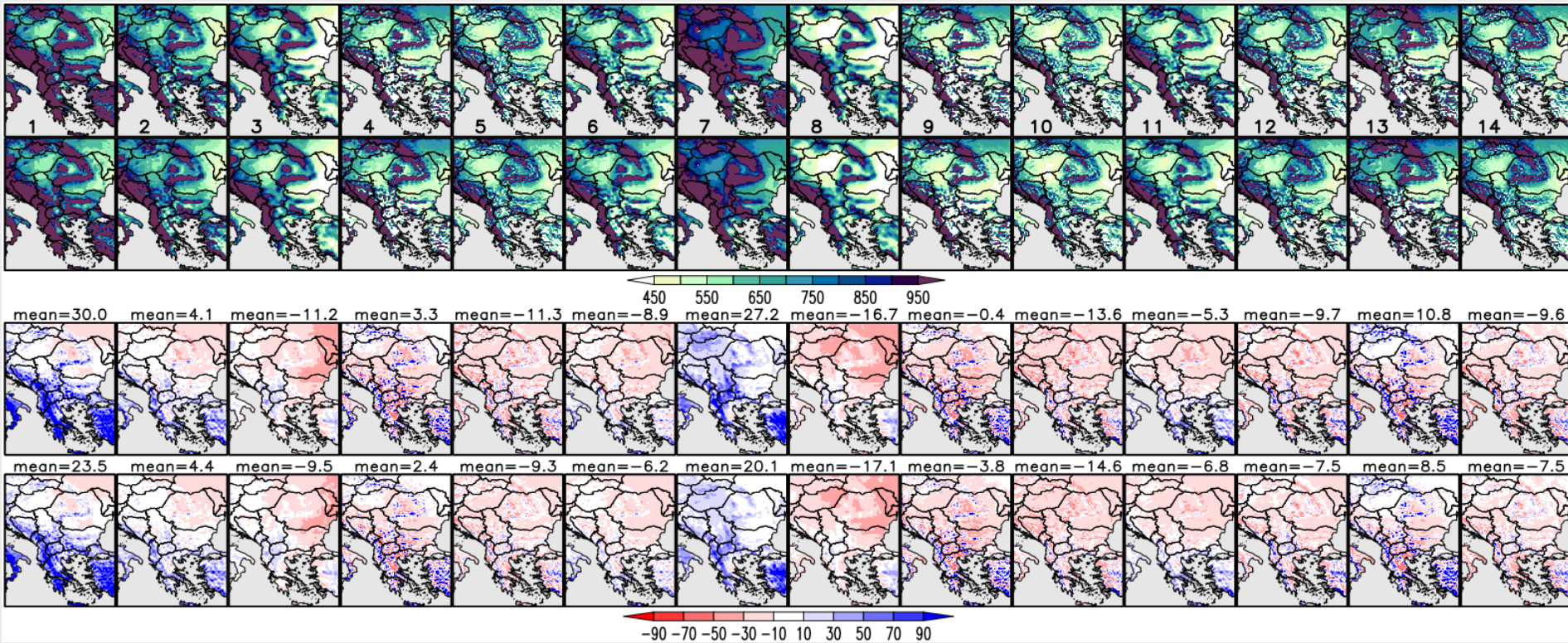


Raw output (first row), bias corrected output (second row) of the multiyear mean temperature (unit: °C) for 1976–2005 as well as their differences in respect to the reference in the third and fourth row correspondingly. The field mean values for the differences are reported over each subplot. The models' numeration is according to the Table.

4. Obtained results

- Results for the individual EURO-CORDEX models

Multiyear Annual Statistics (2 from 2)

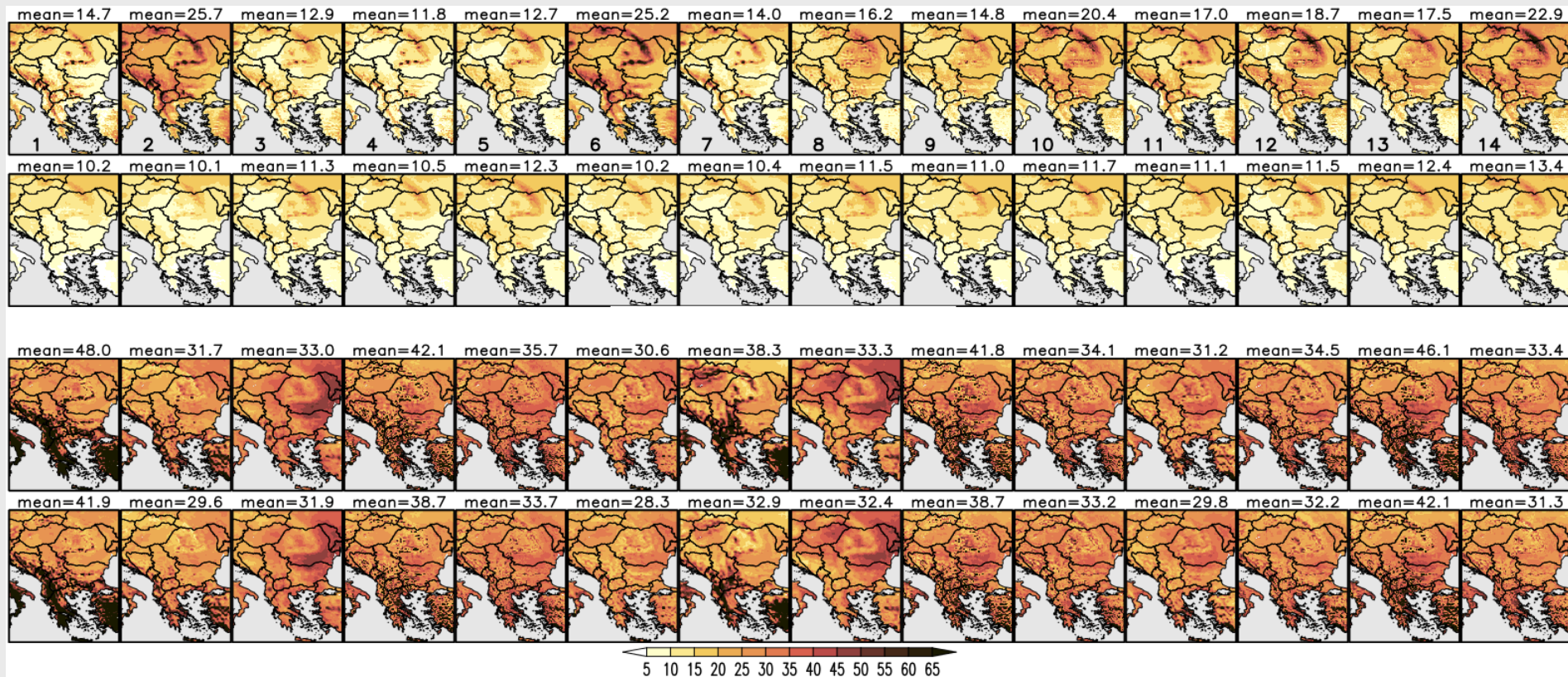


Same as previous but for the multiyear annual precipitation sum (unit: mm). Relative (in %) instead of absolute differences are used.

4. Obtained results

- Results for the individual EURO-CORDEX models

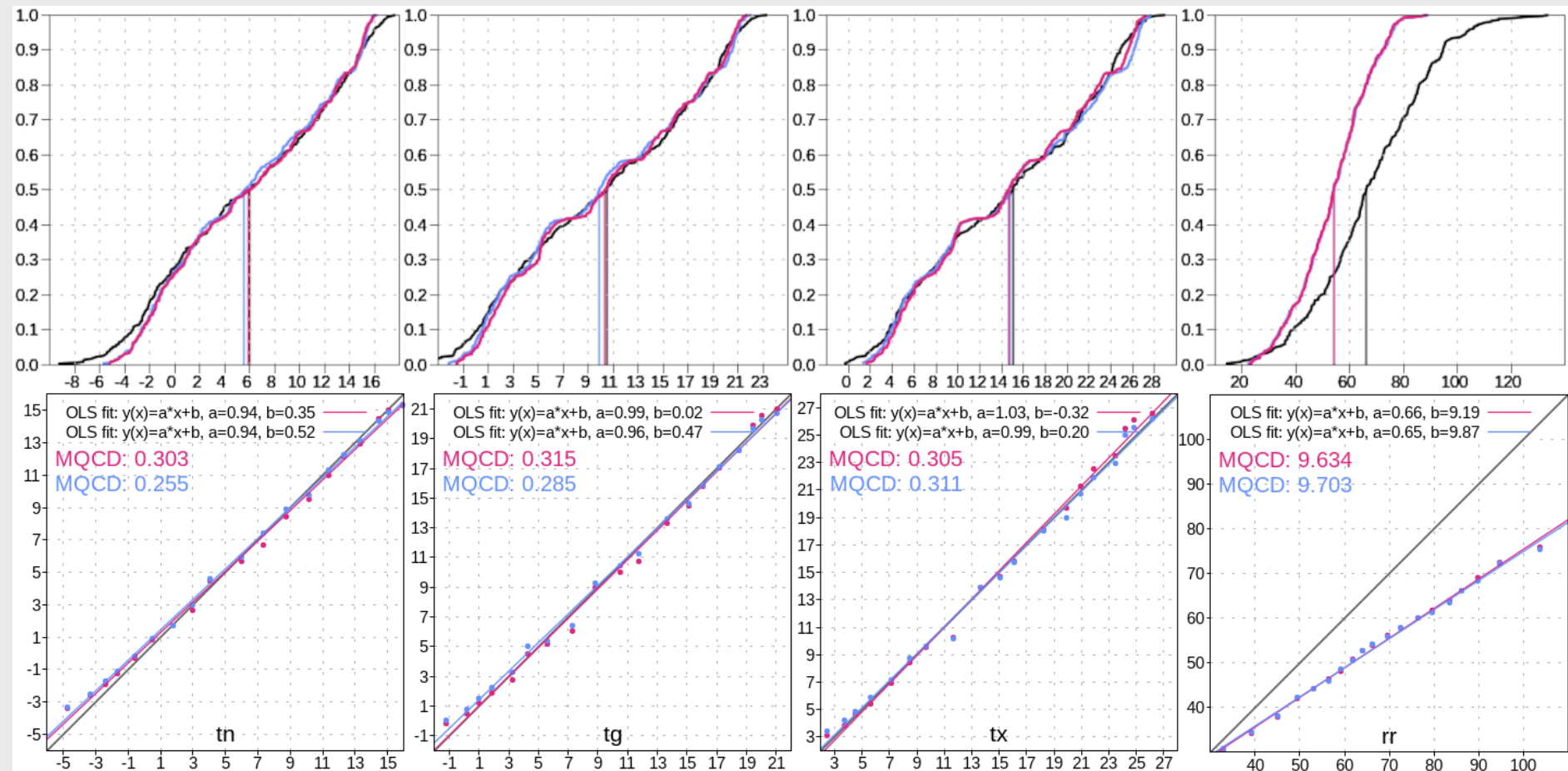
Multiyear Normalized Root Mean Square Error



Multiyear NRMSE (in %) of the raw output (first and third row), and the bias corrected output (second and fourth row) of the annual mean temperature (upper group) and annual precipitation sum (lower group). The models' numeration is according to the table.

4. Obtained results

- Results for the Multimodel Ensemble - Cumulative Probability Functions and QQ plots



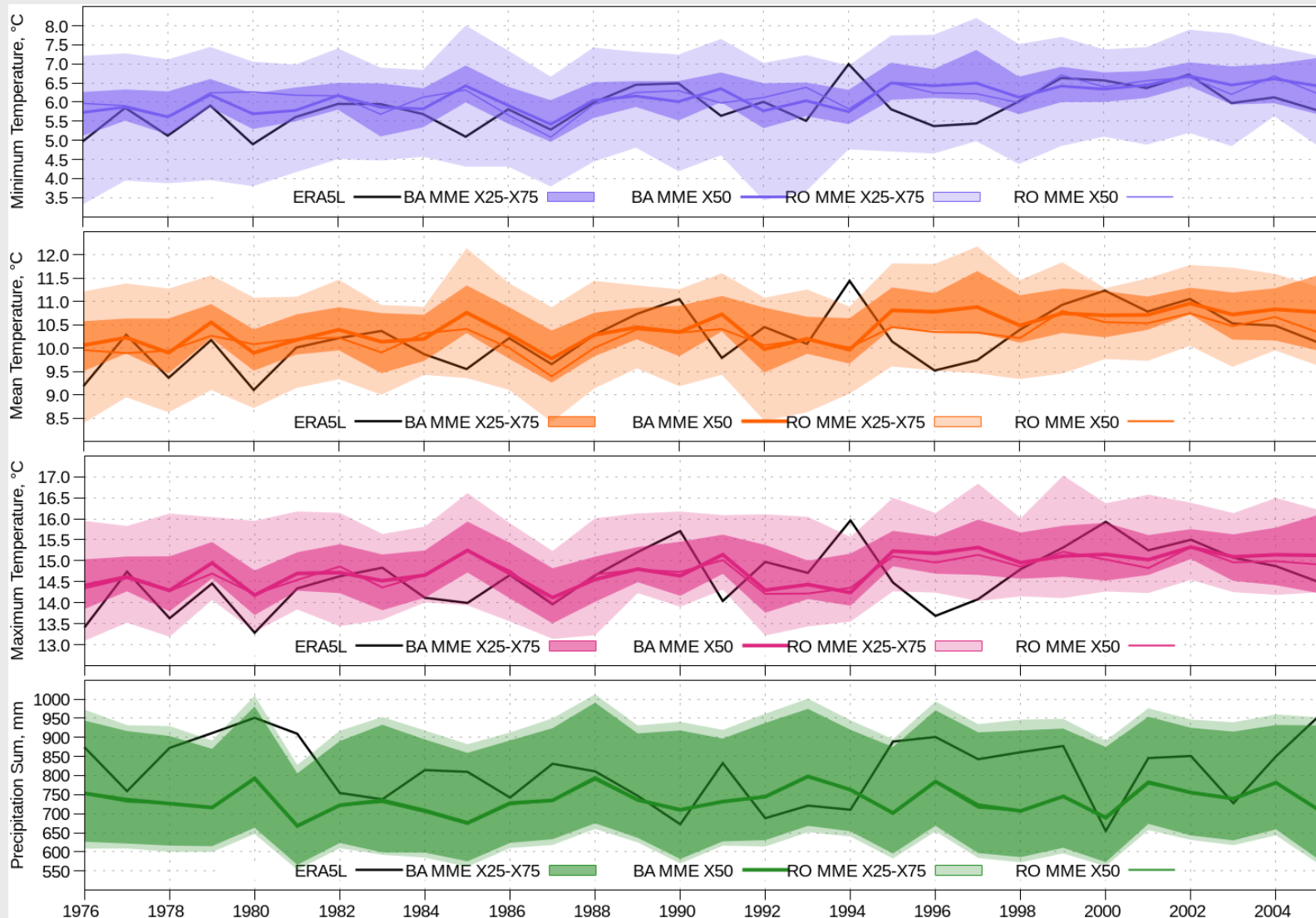
First row: empirical CDFs of the monthly mean values of the minimum, mean, and maximum temperature (unit: °C) as well as precipitation sum (unit: mm) of the reference, multimodel median of the raw and corrected output in black, blue, and red correspondingly.

The vertical bars show the medians' values. The models' numeration is according to the table.

Second row: QQ plots (reference on X-axis, raw/bias corrected output on Y axis), reversed color!

4. Obtained results

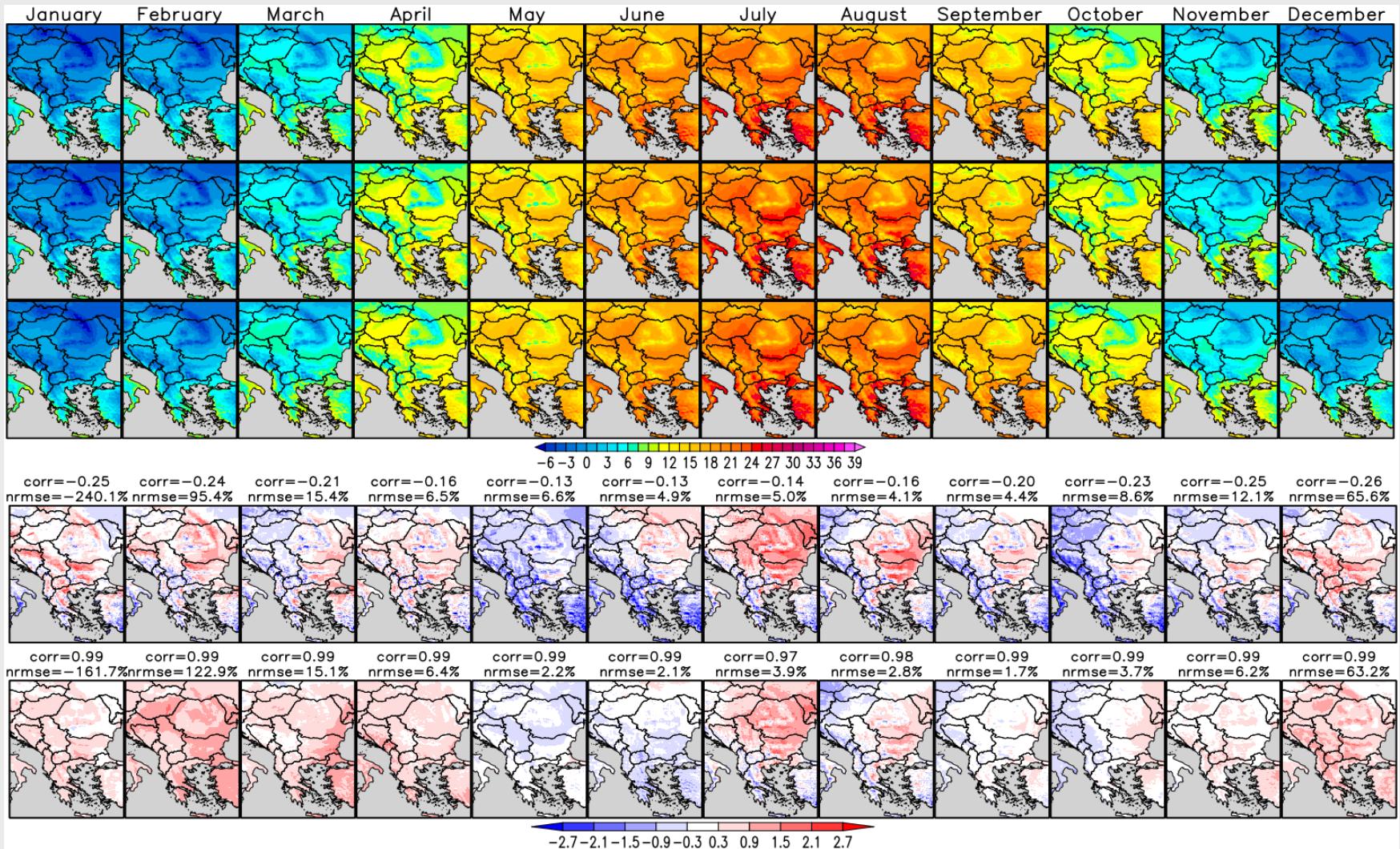
- Results for the Multimodel Ensemble - Field mean values of the MME X25, X50, and X75



Evolution of the field mean values of the daily minimum, mean and maximum temperature as well as annual precipitation sum of the MME from the raw model (RO), bias corrected output and ERA5-Land reference according to figure labels

4. Obtained results

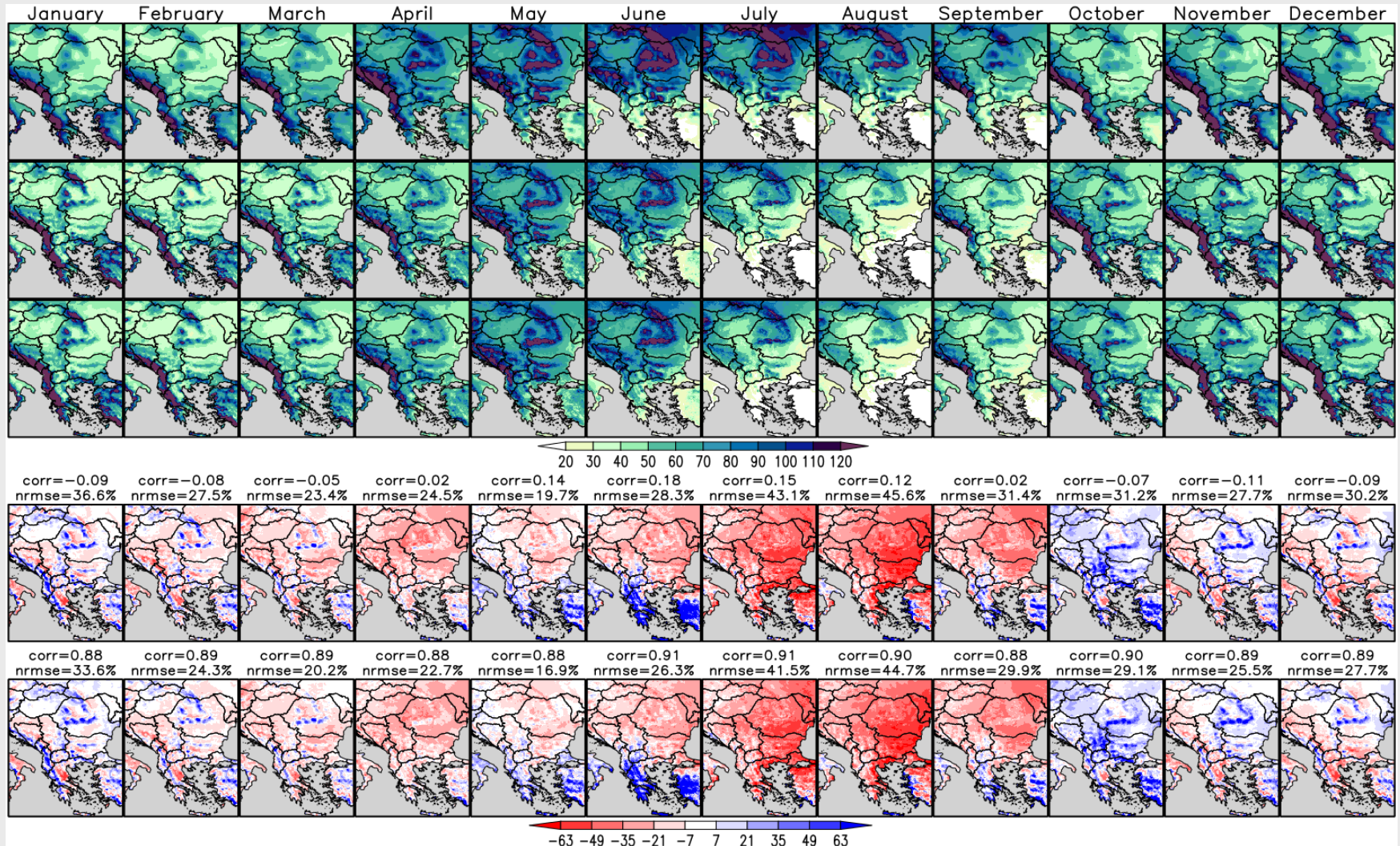
- Results for the Multimodel Ensemble – Multiyear monthly statistics of the MME X50 (1 from 2)



Multiyear mean values for 1976–2005 of the mean temperature (unit: °C) of the reference, MME X50 of the raw and MME X50 of the corrected output on the first, second and third row correspondingly. The absolute bias of the uncorrected and the corrected output are shown on the fourth and fifth row correspondingly. The field means of the corr and NRMSE are reported over each bias subplot.

4. Obtained results

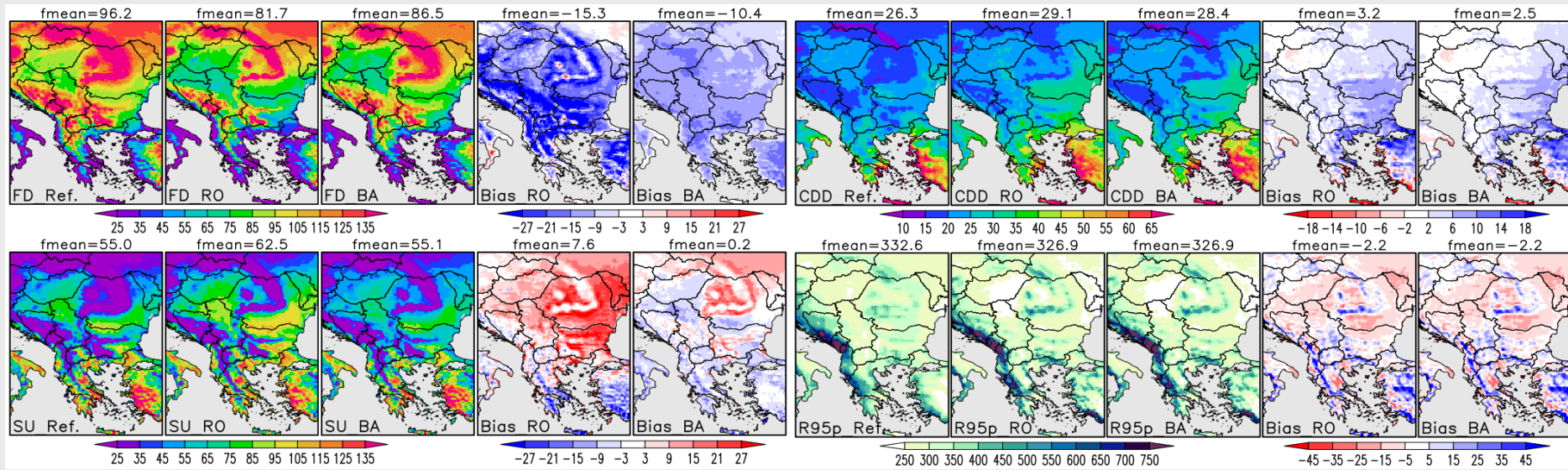
- Results for the Multimodel Ensemble – Multiyear monthly statistics of the MME X50 (2 from 2)



Same as previous figure but for the precipitation sum (units: mm). Relative (in %) instead absolute biases are presented.

4. Obtained results

- Results for the Multimodel Ensemble – MME X50 of key ETCCDI climate indices



Multiyear mean values for 1976–2005 of the reference, raw and bias corrected output as well as absolute biases according to the subplot titles of the temperature- (left group) and precipitation-based (right group) indices. The units of the frost days (FD), summer units (SU), and consecutive dry days (CDD) are days; for the wet days R95p — mm. Relative (in %) instead of absolute bias of the R95p is presented.

5. Summary, Conclusions and Outlook (1 from 2)

We present our preliminary results from applying the selected QDM BC method to temperature and precipitation data from a set of 14 EURO-CORDEX GCM/RCM combinations for the period 1976–2005 over Southeast Europe. Thus far four main conclusions can be outlined:

- The selected BC method QDM demonstrates very high performance in the reduction of the combined GCM/RCM bias of the minimum, mean, and maximum temperature as well as precipitation sum from all considered EURO-CORDEX GCM/RCM combinations.
- The impact of the QDM is more clearly expressed in single GCM/RCM simulations than in MME X50, indirectly supporting the methodological reasonability of climatological assessments based on (equally weighted) MME.
- A more comprehensive analysis would be required to thoroughly assess the performance of the QDM, especially for derived parameters that are nonlinear functions of temperature and precipitation, such as the percentile-based ETCCDI climate indices.
- The designed data processing chain, whose core is the command-line tool BiasAdjustCXX, demonstrates its ability to process large amounts of data efficiently and quickly.



5. Summary, Conclusions and Outlook (2 from 2)

Finally, we should express some words of caution:

- The reference data, sourced from ERA5-Land, may not be completely error-free; they result from a complex modelling chain.
- Care must be taken when performing bias correction of model data, regardless of the method used; hence this adds a further layer of uncertainty to regional climate simulations.

Nevertheless, the present study is a significant step toward our goal of constructing a bias corrected dataset for the near-past climate of Southeast Europe based on the series of EURO-CORDEX simulations for impact studies.

The logical next one is to extend this dataset with the scenario-driven projections for the future climate, and this is the subject of our forthcoming work.



6. Acknowledgment

Hence, this study is based entirely on freely accessible data and software, we express our gratitude to all institutions that provide them (ECMWF, primary GCM/RCM contributors, Copernicus CDS, MPI-M). Special thanks to B.T. Schwertfeger for the cooperation. The Scientific colour map lajolla is used in this study to prevent visual distortion of the data and exclusion of readers with colour-vision deficiencies.



THANK YOU FOR YOUR ATTENTION!