On the relation of CMIP6 GCMs errors at RCM driving boundary condition zones and inner domain for Central Europe region

Eva Holtanová

eva.holtanova@matfyz.cuni.cz

Charles University Faculty of Mathematics and Physics: Univerzita Karlova Matematicko-fyzikalni fakulta
https://orcid.org/0000-0002-8393-7119

Michal Belda

Charles University Faculty of Mathematics and Physics: Univerzita Karlova Matematicko-fyzikalni fakulta

Natália Machado Crespo

Charles University Faculty of Mathematics and Physics: Univerzita Karlova Matematicko-fyzikalni fakulta

Tomáš Halenka

Charles University Faculty of Mathematics and Physics: Univerzita Karlova Matematicko-fyzikalni fakulta

Research Article

Keywords: boundary conditions, global climate models, CMIP6, model assessment, CNRM-ESM2-1, radiosondes

Posted Date: December 28th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3779508/v1

Version of Record: A version of this preprint was published at Climate Dynamics on April 10th, 2024. See the published version at https://doi.org/10.1007/s00382-024-07216-z.

Abstract

Global climate models (GCMs) are important tools for studying the climate system and climate change projections. Due to their coarse spatial resolution, downscaling is necessary on regional scale. Regional climate models (RCMs) represent a common solution for this issue. Nevertheless, outputs of RCMs are influenced by the boundary conditions provided by GCMs. This study evaluates CMIP6 GCMs regarding the variables relevant as RCM boundary conditions. Special focus is on the simulation of CNRM-ESM2-1, which is being used as a driving model for convection-permitting Aladin-Climate/CZ RCM, used as one source feeding new Czech climate change scenarios. The analysis is conducted over the boundaries and inside the RCM integration domain. Firstly, an evaluation of CFSR and ERA5 reanalyses against radiosondes is performed in order to choose an appropriate reference dataset for upper air variables. Revealed high correlation between both reanalysis and radiosondes slightly decreases at the most upper tropospheric levels. ERA5 is then chosen as the reference for the boundary analysis. Over the inner domain, simulated mean annual cycle of impact-relevant variables is validated against E-OBS. The CNRM-ESM2-1 performs well in terms of near-surface variables over the Czech Republic, but it exhibits larger errors along the boundaries, especially for air temperature and specific humidity. Weak statistical relationship between the GCM performance over the boundaries in the upper levels and over the inner domain suggests that the nested RCM simulation does not necessarily have to be influenced by the biases in the driving data.

1 Introduction

Global climate models (GCMs) represent invaluable instruments for various purposes, in particular, analysis of climate system dynamics (e.g. Yang et al. 2022; Dai and Deng 2022), evolution of past climates (e.g. Askjær et al. 2022; Wang et al. 2023) and climate change projections (e.g. Coppola et al. 2021; Belda et al. 2016). The newest set of GCM simulations has been produced under the Coupled Model Intercomparison Project phase 6 (CMIP6) initiative coordinated by the World Climate Research Programme's (WCRP) Working Group on Coupled Modelling (Eyring et al. 2016). This multi-model ensemble enables the evaluation of uncertainty related to various sources (a comprehensive discussion of the uncertainties has recently been presented in Abramowitz et al. 2019).

Some studies compared the GCMs contributing to both CMIP5 and CMIP6 in order to evaluate the progress between the model generations in terms of model performance. For example, Cannon et al. (2020) and Fernandez-Granja et al. (2021) evaluated pairs of CMIP5-CMIP6 GCMs according to their skill in simulating observed patterns of atmospheric circulation; they concluded that the simulated atmospheric circulation is substantially improved in the new generation of models. Overall differences in model performance and ensemble spread of CMIP6 simulations in comparison to CMIP5 are hard to evaluate and explain due to some disparities between the ensembles. These include a higher number of contributing models in CMIP6, differences in model complexity (e.g. inclusion of wider scale of bio-geo-chemical processes in some earth-system models), etc.

The coarse spatial resolution of GCMs limits their use for applications on regional and local scales and consequently, some kind of downscaling is necessary (see e.g. Takayabu et al. 2016; Giorgi 2019; Maraun et al. 2015). One widely used approach is dynamical downscaling using regional climate models (RCMs). Clearly, the outputs of RCMs are influenced by the boundary conditions provided at the lateral boundaries of the integration domain (Prein et al. 2019; Takayabu et al. 2016; Holtanová et al. 2019, 2014; Crhová and Holtanová 2018; Christensen and Kjellstrom 2020). The magnitude of this influence is a subject of ongoing research and depends on various aspects including the geographical region, temporal scale, climatic variable, etc.

To overcome the problem of biases in driving data it is possible to apply some correction to the GCM outputs during pre-processing or to combine the multi-model GCM data with reanalysis products to provide more credible driving data for RCM simulations. A review of published methods is provided by Adachi and Tomita (2020). The complexity of various approaches evolved in time from very simple to quite complex methods modifying not only mean values but the whole simulated distribution (Adachi and Tomita 2020). Improvements in various aspects of RCM simulations after bias correction

of driving data have been shown by recent studies including Lim et al. (2019) and Kim et al. (2020). On the other hand, some studies pointed out that even complex approaches to bias correction do not lead to major improvements in RCM simulations (e.g. Rocheta et al. 2017). Kim et al. (2021) found out that the RCM-simulated inter-variable dependencies are not improved by univariate bias correction of boundary conditions and propose the use of more sophisticated bias correction methods that could help preserve the physical consistency of boundary conditions in future studies. Rocheta et al. (2020) highlight the dampening effect of the relaxation zone and interpolations of lateral boundary conditions, which reduce the influence of bias correction applied to the driving GCM data on the interior of the nested domain. Kim et al. (2021) also mention that the relaxation zone and the RCM dynamics improve possible inconsistencies in bias-corrected boundary conditions.

Besides using bias-corrected or combined GCM-reanalysis driving fields, many experiments all over the world with conventional approaches using raw GCM data are still being undertaken (e.g. EURO-CORDEX initiative). Therefore, a deeper understanding of the influence of errors in boundary conditions on RCM simulation is needed. Jury et al. (2015) attempted to answer the question of whether the evaluation of GCM-simulated near-surface meteorological variables inside the integration domain of a regional climate model driven by a GCM is relevant for the evaluation of the quality of the driving data and possibly as a hint for the choice of driving model. Analyzing CMIP5 GCMs over the domain used for EURO-CORDEX RCM experiments, they concluded that there is generally low correlation and no clear relationship between near-surface parameters and upper-air variables simulated by the GCMs that serve as boundary conditions for nested simulations. Therefore they recommend a multi-variate evaluation of GCM model performance including the driving data on mid-tropospheric pressure levels. Recently, studies evaluating CMIP5/6 GCMs with the aim of selecting the most suitable driving models for dynamical downscaling (e.g. Di Virgilio et al. 2022; Merrifield et al. 2023; Palmer et al. 2023; Sobolowski et al. 2023) concentrated on large scale climatic features, large scale circulation or regional means of selected variables.

In the present study, we take an alternative approach of evaluating the GCMs over the boundaries of the RCM integration domain concentrating on variables that are used as the RCM boundary conditions. A similar approach was also used in Jury et al. (2015), Xu et al. (2017), and Zhang et al. (2022). Our analysis is a part of the Czech national project Prediction, Evaluation and Research for Understanding National sensitivity and impacts of drought and climate change for Czechia (PERUN). One of the sources feeding the updated Czech climate change scenarios will be simulations of the Aladin-Climate/CZ regional model (Termonia et al. 2018; Mašek et al. 2023) in convection-permitting mode driven by the CNRM-ESM2-1 GCM (Séférian et al. 2016, 2019). Therefore we concentrate on CNRM-ESM2-1 (concretely the run denoted as r1i1p1f2) and put its results in the context of the suite of other available CMIP6 GCMs. We also evaluate the spread of perturbed initial conditions ensemble of CNRM-ESM2-1 (9 members) as an estimate of uncertainty arising from natural climate variability. Further, an important goal of this study is to investigate the potential relation between the performance of the boundary conditions and the inside-domain performance in simulating the near-surface variables over the Czech Republic. This assessment focuses on air temperature, precipitation, relative humidity and global radiation, which represent climatic conditions relevant to sectors of agriculture and hydrology. The driving GCM for the PERUN Aladin-Climate/CZ simulation has already been selected, so the focus is not on model selection, but rather analysis of biases in boundary conditions and their propagation into the inner domain.

The paper is structured as follows: section 2 describes the datasets (comprising observation, reanalyses and simulation data) and statistical methods for GCMs evaluation; section 3 presents the results, with a section discussing the reanalyses uncertainties and the GCMs performance in different parts of the RCM domain (i.e. boundaries and interior); and section 4 presents discussions and conclusions from the study.

2 Data and methods

2.1 Reanalysis and radiosondes

As the GCM evaluation over the boundaries of Aladin-Climate/CZ integration domain presented here concentrates on upper air parameters, reanalysis data are used as a reference. ERA5 reanalysis provides the most up-to-date state-of-the-art reanalysis dataset (Hersbach et al. 2020) and proved useful and trustworthy in many applications, however, most of the studies concentrate on near-surface parameters (e.g. Urban et al. 2021; Bandhauer et al. 2022) or convective parameters (e.g. Taszarek et al. 2021; Walawender et al. 2017; Varga et al. 2022). A comprehensive evaluation of upper air parameters is, as far as our knowledge, missing. Therefore, to investigate possible errors in the reanalysis data, in the first step of our analysis, we evaluate the ERA5 dataset and compare it to the NCEP Climate Forecast System. Moreover, we analyze differences between the two reanalyses and selected radiosonde data.

We compare time series of selected meteorological variables (air temperature, specific humidity and horizontal components of the wind) derived from the ERA5 reanalysis (0.25º of horizontal resolution; Hersbach et al. 2020, 2023) and the NCEP Climate Forecast System (CFSR) reanalysis (0.5º of horizontal resolution; Saha et al. 2010a,b, 2012, 2014) with selected radiosondes from the University of Wyoming (https://weather.uwyo.edu/upperair/sounding.html). Data at 0000 UTC and 1200 UTC in three tropospheric levels (850, 500 and 300 hPa) were used for the evaluation in four different locations, as seen in Fig. 1. We selected sounding locations as close as possible to each of the four boundaries of the Aladin-Climate/CZ domain studied here. Moreover, only soundings with approximately 30% (specifically 29% for 03354 location, 54% for 10035, 36% for 26477 and 29% for 16622) of data available for the period 1990–2014 were selected and then reanalysis data were intersected for the same timesteps for the nearest latitude and longitude grid point from the radiosonde location. Mean absolute error (MAE) and Spearman (s) correlation coefficients were computed for the assessment.

2.2 CMIP6 ensemble

An ensemble of CMIP6 GCM simulations is used (Table 1). Further information and references can be found in Tebaldi et al. (2021). The time series of monthly mean air temperature (further abbreviated as TA), mean specific humidity (HUS) and the horizontal components of wind speed vector (UA and VA) were retrieved via Deutsches Klimarechenzentrum (DKRZ) for the vertical levels of 850 hPa, 500 hPa and 300 hPa. The outputs from the experiment denoted as ‘historical’ are used for the reference period 1990–2014.

The choice of GCMs is based on the availability of data. We use preferably the simulations denoted as r1i1p1f1, but “f2” or “f3” is used in case “f1” was missing. For a detailed explanation of the “ripf” notation, we refer to the CMIP6 protocol and associated documentation. Further, nine perturbed initial conditions ensemble members are used for CNRM-ESM2-1 (hereafter “CNRM-INI”), and we focus on CNRM-ESM2-1 r1i1p1f2 simulation (further as CNRM-r1), i.e. the driving simulation for the Aladin-Climate/CZ as described above.

We conduct our analysis over the boundaries of the Aladin-Climate/CZ integration domain (Fig. 1). To be able to select corresponding GCM grid points, we first re-gridded the GCM data to the common grid in the Lambert tangent projection with horizontal step 5-times larger than the RCM integration; a simple bilinear interpolation was used for this purpose. We did not interpolate to the original nested integration resolution (approximately 2.8 km) because we considered it to be too fine resolution for the GCMs. From these re-gridded data, we masked the integration domain and chose the grid points at the four boundaries: western, eastern, southern, and northern.

For the evaluation of CMIP6 performance in simulating the near-surface climatic variables in the inner part of the Aladin-Climate/CZ integration domain, that covers the Czech Republic (Fig. 1), we use the E-OBS data (version 27, Cornes et al. 2018) with 0.25° horizontal resolution as the reference. The variables analyzed are near-surface air temperature (TAS), precipitation (PR), relative humidity (HUR) and global radiation (RSDS). The CMIP6 data for the inner domain were retrieved from the ETH Zurich CMIP6 Next Generation archive (Brunner et al., 2020), and a similar interpolation process was also applied for the same period (1990–2014) before the areal averages (delimited by 11.85°-19.15° of longitude and 48.25°-51.25° of latitude; Fig. 1) of monthly mean values were calculated. This area covers the Czech Republic and has been previously used for GCM and RCM evaluation over this country (e.g. Holtanová et al. 2022, 2012). The GCM subset for the

inner domain includes 14 GCMs less than the sample available for the upper air variables (for the lists of the models see Table 1).

Table 1
The list of CMIP6 GCMs used in present study. The GCMs denoted in italics are missing for the analysis over the inner domain.

GCM Acronym	Modelling center	horizontal resolution (lat x lon or grid spacing in kilometres)
ACCESS-CM2	Australian Community Climate and Earth System Simulator	1.25°x 1.875°
ACCESS-ESM1-5		1.25°x 1.875°
AWI-CM-1-1-MR	Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research (AWI)	0.9° x 0.9°
AWI-ESM-1-1-LR		~ 200 km
BCC-CSM2-MR	Beijing Climate Center	0.45° × 0.45°
BCC-ESM1		~ 250 km
CAMS-CSM1-0	Chinese Academy of Meteorological Sciences, China	1.1° x 1.1°
CanESM5	Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change Canada, Victoria, Canada	2.8° x 2.8°
CAS-ESM2-0	Chinese Academy of Sciences, Beijing, China	100 km
CESM2-FV2	National Center for Atmospheric Research (NCAR), Climate and Global Dynamics Laboratory, Boulder, USA	250 km
CESM2-WACCM-FV2		250 km
CESM2-WACCM		1.25° x 0.9°
CESM2		100 km
CIESM	Department of Earth System Science, Tsinghua University, Beijing, China	100 km
CMCC-CM2-HR4	Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC), Lecce, Italy	100 km
CMCC-CM2-SR5		100 km
CMCC-ESM2		100 km
CNRM-CM6-1-HR	Centre National de Recherches Meteorologiques (CNRM) and Centre Europeen de Recherche et de Formation Avancee en Calcul Scientifique (CERFACS), Toulouse, France	100 km
CNRM-CM6-1		250 km
CNRM-ESM2-1		1.4° x 1.4°

GCM Acronym	Modelling center	horizontal resolution (lat x lon or grid spacing in kilometres)
E3SM-1-0	Lawrence Livermore National Laboratory (LLNL), Livermore, CA, USA; Argonne National Laboratory (ANL), Argonne, IL, USA; Brookhaven National Laboratory (BNL), Upton, NY, USA; Los Alamos National Laboratory (LANL), Los Alamos, NM, USA; Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA, USA; Oak Ridge National Laboratory (ORNL), Oak Ridge, TN, USA; Pacific Northwest National Laboratory (PNNL), Richland, WA, USA; Sandia National Laboratories (SNL), Albuquerque, NM, USA	1° x 1°
E3SM-1-1-ECA		1° x 1°
E3SM-1-1		1° x 1°
EC-Earth3-AerChem	EC-Earth consortium, Rossby Center, Swedish Meteorological and Hydrological Institute/SMHI, Norrkoping, Sweden	100 km
EC-Earth3-CC		100 km
EC-Earth3-Veg-LR		250 km
EC-Earth3-Veg		100 km
EC-Earth3		0.7° x 0.7°
FGOALS-f3-L	Chinese Academy of Sciences, Beijing, China	100 km
FGOALS-g3		2° x 2.3°
FIO-ESM-2-0	First Institute of Oceanography (FIO) and Qingdao National Laboratory for Marine Science and Technology (QNLM), Qingdao, China	1.25° x 0.9°
GFDL-CM4	National Oceanic and Atmospheric Administration, Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA	100 km
GFDL-ESM4	National Oceanic and Atmospheric Administration, Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA	1.25° x 1°
GISS-E2-1-G-CC	Goddard Institute for Space Studies (GISS), New York, NY, USA	250 km
GISS-E2-1-G		2.5° x 2°
GISS-E2-1-H		200 km
HadGEM3-GC31-LL	Met Office Hadley Centre, UK	250 km
HadGEM3-GC31-MM		0.6° × 0.8°
IITM-ESM	Centre for Climate Change Research (CCCR) at the Indian Institute of Tropical Meteorology	0.9° x 0.9°
INM-CM4-8	Institute for Numerical Mathematics (INM), Russian Academy of Science, Moscow, Russia	100 km
INM-CM5-0		2° x 1.5°
IPSL-CM5A2-INCA	Institut Pierre Simon Laplace (IPSL), Paris, France	500 km

GCM Acronym	Modelling center	horizontal resolution (lat x lon or grid spacing in kilometres)
IPSL-CM6A-LR-INCA		250 km
IPSL-CM6A-LR		2.5° x 1.25°
KACE-1-0-G	National Institute of Meteorological Sciences/Korea Meteorological Administration, Republic of Korea	250 km
KIOST-ESM	Korea Institute of Ocean Science and Technology (KIOST), Busan, Republic of Korea	250 km
MCM-UA-1-0	Department of Geosciences, University of Arizona, Tucson, USA	3.75° x 2.2°
MIROC-ES2L	Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Kanagawa, Japan, Atmosphere and Ocean Research Institute (AORI), The University of Tokyo, Chiba, Japan, National Institute for Environmental Studies (NIES), Ibaraki, Japan, and RIKEN Center for Computational Science, Hyogo, Japan (MIROC)	500 km
MIROC6		250 km
MPI-ESM-1-2-HAM	ETH Zurich, Switzerland; Max Planck Institut fur Meteorologie, Germany; Forschungszentrum Julich, Germany; University of Oxford, UK; Finnish Meteorological Institute, Finland; Leibniz Institute for Tropospheric Research, Germany; Center for Climate Systems Modeling (C2SM) at ETH Zurich, Switzerland (HAMMOZ-Consortium)	250 km
MPI-ESM1-2-HR	Max Planck Institute for Meteorology, Germany	0.94° × 0.94°
MPI-ESM1-2-LR		~ 200 km
MRI-ESM2-0	Meteorological Research Institute, Japan	1.1° x 1.1°
NESM3	Nanjing University of Information Science and Technology (NUIST), Nanjing, China	1.9° x 1.9°
NorCPM1	NorESM Climate modeling Consortium, Oslo, Norway	250 km
NorESM2-LM		250 km
NorESM2-MM		1.25° x 0.9°
SAM0-UNICON	Seoul National University (SNU), Seoul, Republic of Korea	100 km
TaiESM1	Research Center for Environmental Changes, Academia Sinica, Nankang, Taipei, Taiwan	1.25° x 0.9°
UKESM1-0-LL	Met Office Hadley Centre, Exeter, UK	1.9° x 1.25°

2.3 GCMs evaluation metrics

For the GCM evaluation over the boundaries of the Aladin-Climate/CZ integration domain a metric based on the index introduced in Reichler and Kim (2008) is used. The calculation is done for each boundary (eastern, northern, western,

southern) separately. First, for each season *s*, variable *v*, model *m* and grid point *i* the differences of GCM simulated (s_{svmi}) and ERA5-given mean seasonal values (o_{svi}) normalized by the ERA5 standard deviation (σ^2_{svi}) are calculated. These differences are then summed up over all grid points of the boundary (Eq. 1):

$$e^2_{svm} = \sum_{i=1}^{n} \frac{\left(s_{svmi} - o_{svi}\right)^2}{\sigma^2_{svi}}$$

1

The resulting "error" (e^2_{svm}) is then divided by the multi-model mean value, i.e. the resulting metric gives the information about the magnitude of error of each of the models in comparison to the multi-model mean error. This quantity is further denoted as RK Index (or RKI). It is important to note that the higher RKI, the more distant the model quantity is from the reference data.

For the GCM evaluation over the inner part of the domain (defined above in Section 2.2) we use root mean square error (RMSE) of the mean annual cycle calculated from monthly mean values, i.e. twelve monthly mean values are taken into account for each of the studied variables. We use the long-term mean values to avoid the internal climate variability influencing the model evaluation. The relative error (RE, defined by Gleckler et al. 2008), calculated as the difference between individual GCM RMSE and multi-model median RMSE normalized by the multi-model median RMSE is then used to compare the performance among different variables. The interpretation of RE values is that more negative values imply relatively better performance, and more positive values imply relatively worse performance, in comparison to the multi-model median RMSE.

3 Results

3.1 ERA5 and CFSR reanalysis: observational uncertainty

In this section an evaluation of both studied reanalysis (ERA5 and CFSR) is presented. The aim is to verify how well these two commonly used gridded datasets perform for the four tropospheric parameters studied here (TA, UA, VA, HUS). Therefore, the reanalyses are compared to observations represented by soundings. Table 1 shows the Spearman correlation and mean absolute error (MAE) for the selected variables in different vertical levels and locations.

The correlation is mostly above 0.9, with only a few exceptions, mainly in the case of specific humidity at 500 hPa and 300 hPa levels (Table 2). Lower correlation is also found at the 850 hPa level for the wind components in the southernmost location (16622 Thessaloniki). In general, both reanalyses have similar correlations taking into account the same variable/level; in most cases, the values are basically the same (e.g. TA, see Table 2).

Regarding MAE, lower values imply better reanalysis performance. The MAE for specific humidity is generally lower at the 300 hPa level, which might be attributed to generally lower humidity in the upper troposphere. The MAE for wind speed components is increasing with height. This might be because we are comparing a single grid point from the reanalysis with the radiosonde data, which might drift away from the grid point location as it gets to the upper troposphere. Another issue is that the reanalysis data represent the whole gridbox whereas the radiosonde data have the character of point-value. For air temperature, the reanalyses are more accurate at mid-upper levels, except for the northeastern location (26477 Velikie Luki). Based on the results in Table 1 we can conclude that both assessed reanalyses perform well. The performance is very similar for both reanalyses, so we do not inflate large observational uncertainty in the GCM assessment by choosing only ERA5 as the reference data. Moreover, a positive agreement between ERA5 with radiosondes was found by Varga et al. (2022), who studied convective environmental parameters over Central Europe; they also found correlation coefficients higher than 0.9 for multiple variables and low error characteristics for a 26-year period analysis.

Table 2
Spearman's correlation coefficient (s) and mean absolute error (MAE) for temperature (ta, ºC), specific humidity (HUS, g/kg) and horizontal wind velocity components (UA and VA, m/s) at different levels, between reanalyses [i.e. ERA5 and CFSR (in parenthesis)] and four selected radiosonde locations. Single value indicates the same result for both ERA5 and CFSR.

Variable	TA			HUS			UA			VA		
Level	850	500	300	850	500	300	850	500	300	850	500	300
16622 LGTS Thessaloniki (Airport)												
s	0.99	0.99	0.99	0.94 (0.96)	0.88 (0.90)	0.84 (0.73)	0.88 (0.91)	0.96	0.98 (0.97)	0.91 (0.93)	0.97	0.97
MAE	0.72 (0.66)	0.52	0.48 (0.55)	0.67 (0.61)	0.17 (0.15)	0.04	1.52 (1.30)	1.60	1.75 (2.04)	1.41 (1.21)	1.33 (1.39)	1.80 (1.94)
03354 Nottingham												
s	0.99 (0.98)	0.99	0.99	0.94 (0.93)	0.91 (0.90)	0.91 (0.87)	0.98	0.99 (0.98)	0.99	0.97	0.99 (0.98)	0.99 (0.98)
MAE	0.53 (0.59)	0.38 (0.42)	0.41 (0.47)	0.47 (0.50)	0.14 (0.15)	0.02	1.32 (1.35)	1.37 (1.51)	1.70 (1.97)	1.31 (1.37)	1.47 (1.61)	1.86 (2.34)
10035 Schleswig												
s	0.99	0.99	0.99	0.94	0.89 (0.90)	0.89 (0.87)	0.98	0.98	0.99 (0.98)	0.96	0.97 (0.98)	0.98
MAE	0.50 (0.53)	0.46 (0.49)	0.45 (0.48)	0.49	0.16 (0.15)	0.02	1.34 (1.37)	1.60	1.88 (2.05)	1.32	1.73 (1.77)	2.20 (2.33)
26477 ULOL Velikie Luki												
s	0.99	0.98	0.95	0.96 (0.97)	0.92	0.80 (0.81)	0.96	0.98	0.98	0.96 (0.97)	0.97	0.98
MAE	0.86 (0.76)	1.01 (1.02)	1.26 (1.18)	0.48 (0.43)	0.17 (0.16)	0.04	1.40 (1.26)	1.41 (1.46)	1.73 (1.98)	1.35 (1.27)	1.49 (1.56)	1.67 (1.92)

3.2 GCM evaluation

3.2.2 Model performance in simulating upper air parameters along the boundaries

Figure S.1-S.4 and Fig. 2 show the values of the RK Index for each of the boundaries and the summed up values over all the boundaries. They enable comparison of the whole multi-model ensemble (large boxplots), CNRM-INI ensemble (small boxplots) and CNRM-r1 simulation (i.e. one of the CNRM-INI ensemble, black crosses). Different colors of boxplots correspond to different atmospheric pressure levels. The range of the vertical axis is set to 0–20, even though there are several outliers with RKI higher than this value. Up to 16 outliers have RKI higher than 20 for each variable-boundary combination. However, these are not shown for the sake of better readability of the figures.

There is no clear seasonal course in RKI values for any of the four studied variables and boundaries. Further, the performance of the full CMIP6 ensemble is not remarkably better or worse in any of the three studied pressure levels. However, regarding the CNRM-INI ensemble, for TA and HUS at 300 and 500 hPa levels the RKI values are outlying very far from the rest of the multi-model ensemble (Fig. 2a,b, Figs. S.1-S.4). For VA and UA this feature is seen only in summer and autumn with less extreme RKI values, the performance of CNRM-INI ensemble is more similar to the rest of CMIP6 GCMs (Fig. 2c,d, Figs. S.1-S.4). At the 850 hPa level, the RKI of CNRM-INI ensemble decreases in comparison with the two upper levels, in some cases it is even lower than the multi-model median (especially for VA and UA). The range of internal climate

variability, estimated here by the spread of the CNRM-INI ensemble, is generally smaller than the spread of the whole CMIP6 ensemble. However, in some cases, it is comparable or even larger (e.g. HUS and VA in the autumn, TA at 500 hPa and 300 hPa in the autumn and summer; Fig. 2).

Generally, the CNRM-r1 is rather exceptional, since it rarely behaves like a “mean” model. In most cases, its RKI is either lower than for most other GCMs, or much higher. In this sense, especially for the western and northern boundaries (Fig. S.1 and S.4), the whole CNRM-INI ensemble, including CNRM-r1, has much larger errors of TA and HUS at the 300 hPa level than the rest of GCMs. On the other hand, smaller errors are seen for UA and VA, especially in the southern boundary (Fig. S.2).

The heatmaps in Fig. 3a show RKI values for individual GCMs summed up over the four boundaries, the three vertical levels and the four seasons. Figure 3b then shows the values from Fig. 3a summed up over the four studied variables. The lower the values, the better overall performance of individual GCMs. The best performing GCMs include CMCC-ESM2, all variants of EC-Earth3 (except for the “Veg-LR”), FIO-ESM-2 and HadGEM3-GC31-MM. On the other hand, the largest RKI values are seen for CNRM-CM6-1-HR, IITM-ESM, KIOST-ESM, MCM-UA-1-0, FGOALS-g3 and NorESM2-LM (Fig. 3).

For some GCMs the performance varies considerably among variables. For instance, CIESM has an excellent representation of TA and UA compared to other models, but really poor representation for HUS. Another similar example is GISS-E2-1-H (Fig. 3a). Focusing on CNRM-r1, the worst represented variable is TA (Fig. 3a), mainly at upper levels, as seen in Fig. 2. The range of RKI for the CNRM-INI ensemble is relatively small compared to the overall spread of the multi-model ensemble (for TA values are between 80–110, HUS 40–55, VA 48–58 and UA 30–38, not shown in Fig. 3).

We also investigated mutual correlations between RKI at the three studied pressure levels in individual seasons (not shown). For TA, VA and UA the Spearman’s correlation coefficient is statistically significant on the 0.05 significance level for all pairs of pressure levels in all seasons, with the correlations between 500 hPa and the other two levels higher than 0.7. The correlation between 300 hPa and 850 hPa is mostly lower than 0.5, in several cases even only 0.3, however, even these relatively low values are assessed as statistically significant. Regarding HUS, the correlations of RKI between pressure levels are lower than for the other variables and not statistically significant; in eight out of twelve cases it is lower than 0.4, in four of these eight cases it is even slightly negative. These results are in accordance with Jury et al. (2015), who found strong positive correlations between 500 hPa level and upper/lower ones among the same atmospheric variables.

3.2.3 Model performance over the Czech Republic territory

Similarly to the boundaries at upper atmospheric levels analysis, the performance of GCMs in simulating the near-surface variables in the inner domain covering the Czech Republic varies considerably between individual variables (see for example IPSL-CM5A-INCA or the variants of CEMS2 in Fig. 3c). Overall, the best performing GCM is again the EC-Earth3 (all analyzed variants). Further, CanESM5, CMCC-ESM2 and both variants of GISS-E-2-1 belong also to the better half of the models in terms of all four studied variables (Fig. 3c).

CNRM-INI ensemble, including the CNRM-r1 performs relatively well, with the RE for precipitation around − 0.3, and for TA and HURS very close to zero (that is, the RMSE very close to the median of the multi-model ensemble). RSDS has the largest RE of CNRM-r1, with values ~ 0.6. It is important to mention that the RE values for each variable among the members of CNRMI-INI ensemble are similar. Figure 3d shows the RE summed up for all variables for each GCM. This figure highlights the good performance of the other three models (along with the previously mentioned ones): CESM2-FV2, CESM2-WACCM-FV2 and SAM0-UNICON, with RE of -0.4 (Fig. 3d).

Further, we investigate possible statistical relationships between the performance along the boundaries of the integration domain and the Czech territory. For this, the values of RKI are summed up over all seasons, boundaries and levels, and the Spearman correlation is computed between this summed-up RKI (Fig. 3a) and the RE for individual variables (Fig. 3c). The correlations are generally low, ranging from 0.4 to less (Fig. 4), even though statistically significant in some cases. The

scatterplots in Fig. 4 enhance the large variability in data. Hence, the graphical evaluation (Fig. 4) suggests a rather weak relationship between the indices.

In terms of correlation between pairs of RKI (RE) along the boundaries (inner domain), the highest correlation (0.8) is found between the RKI for VA and UA, and between RE for PR and HURS (0.6) (Fig. 4), which is in agreement with expectations. However, other pairs exhibit low correlations, similarly to the pairs between the “boundary” and “inner” variables. It is important to note that this analysis illustrated in Fig. 4 focuses on the errors compared to the reference data and not on the values of the variables.

4 Discussion and Conclusions

This study aimed to evaluate the GCMs from CMIP6 over the boundaries of the Aladin-Climate/CZ integration domain as well as their potential relation with the inside-domain performance in simulating the near-surface variables over the Czech Republic. A first assessment between ERA5 and CSFR reanalyses and radiosondes was performed; it was found that their correlation with radiosondes is very high in general, even though there are some differences depending on the variable. For instance, specific humidity is better correlated at lower levels, while wind speed components are better correlated at upper levels (Table 2). Further, in the north-eastern and south-eastern locations (Velikie Luki and Thessaloniki), the correlations are generally lower. The mean absolute error showed discrepancies between radiosonde and reanalysis that are much lower than the typical values of the parameters themselves, so we conclude that the performance of both reanalyses is satisfactory. It is worth mentioning that the amount of available data from radiosonde decreases with height, especially for HUS. Therefore large uncertainties should be taken into account for the GCM evaluation. The ERA5 dataset was chosen for further evaluation of GCMs errors over the boundaries and E-OBS for the region covering the Czech Republic territory.

The CNRM-r1 simulation, that is the Aladin-Climate/CZ driving simulation, has been previously found to perform satisfactorily in terms of large-scale climatological features over Europe (Palmer et al. 2023). Here we concentrate on the evaluation of biases directly in the coupling zone of the nested regional simulation that could potentially affect the RCM simulation. Our results showed that the CNRM-INI ensemble, including the CNRM-r1, is rather exceptional, rarely behaving as a “mean” model. In most cases, CNRM-INI members have RKI either smaller or much larger than the rest of the CMIP6 GCMs. Most profoundly this is seen for TA and HUS at 300 hPa, for which the CNRM-INI ensemble has much larger errors than the rest of GCMs. On the other hand, it has smaller errors for UA and VA, especially in winter and spring.

Besides the evaluation of the simulation of CNRM-r1, which is of special interest for the Czech Republic’s climate change scenarios, this study could contribute to the discussion of possible strategies of selecting driving GCMs for the EURO-CORDEX downscaling experiments (Katragkou et al., 2023), since we concentrate on variables relevant for RCM boundary conditions. From our results, the best-performing GCMs along the boundaries include CMCC-ESM2, all variants of EC-Earth3 (except for the “Veg-LR”), FIO-ESM-2 and HadGEM3-GC31-MM. For some GCMs, the performance varies considerably among variables. Similarly to the boundaries, the performance of GCMs in simulating the near-surface variables in the inner domain covering the Czech Republic varies considerably between individual variables. Overall, the best performing GCM is again the EC-Earth3 (all variants), followed by CanESM5, CMCC-ESM2, GISS-E-2-1, CESM2-FV2, CESM2-WACCM-FV2 and SAM0-UNICON.

The topic of GCM subselection for regional impact studies and for downscaling purposes has been widely discussed also in connection to the “hot model problem” (Hausfather et al., 2022), i.e. the fact that the new generation of CMIP6 GCMs includes models with higher equilibrium climate sensitivities (ECS) than the previous CMIP5 GCMs. Palmer et al. (2023) point to the potential tension between selection based on plausible values of ECS and regional-performance-based selection of models. Our results indicate a rather weak relationship between GCM errors (both along the boundaries in upper levels and in the CZ domain near the surface) and ECS. Confirming the concern of Palmers et al. (2023), the overall best-performing EC-Earth model (with several different configurations) has ECS outside the range of 2–4.5°C, which has been evaluated as likely realistic by the last IPCC report (Forster et al. 2021).

The spread of the whole multi-model ensemble shows the magnitude of uncertainty in GCM outputs arising from structural uncertainty (Abramowitz et al. 2019). The CNRM-INI represents an estimate of internal variability, however, it consists of only nine members from one GCM, so the results might unfold differently if larger ensembles of different GCMs were used, most probably increasing the spread. Nevertheless, our results show that RKI for the CNRM-INI ensemble (summed up along all boundaries, variables and levels) is quite high (around three times higher than EC-Earth), whereas over the Czech Republic territory, it performs relatively well, belonging to better half of the models according to RE, except for RSDS (Fig. 3) .

On one hand, a study over Eastern Africa found better results when forcing an RCM with the best score GCM (Pickler and Mölg 2021), on the other hand over Europe, generally, a good GCM performance over the boundaries or in large-scale phenomena does not guarantee a good result within the RCM integration domain. The influence of GCM on nested RCM simulation is known to be variable and season-dependent. For example, Bartók et al. (2017) have shown that the patterns of RSDS simulated by RCM are weakly connected to values of RSDS simulated by the driving GCM. This is due to complex interactions of cloudiness, albedo, evapotranspiration, and other parameters and processes. Moreover, Vautard et al. (2021) investigated the EURO-CORDEX ensemble of RCM simulations driven by several CMIP5 GCMs and showed that for RSDS, the influence of an RCM is a more important source of bias than the driving GCM, but for temperature and precipitation, they found the dominance of the boundary conditions. Furthermore, some studies mentioned in Section 1 claim that the relaxation zone and the coupling itself can handle potential inconsistencies in bias-corrected boundary conditions.

In general, this study has shown that the GCM performance in simulating the upper air atmospheric variables that are used as RCM boundary conditions relates rather weakly to the GCM performance in simulating the near-surface parameters in the inner domain (in terms of parameters relevant for impact studies). This implies that good performance of boundary conditions cannot automatically guarantee better performance of the downscaling system and, even more importantly, the other way around. This is justified by the fact that boundary conditions are subject to further dynamical modifications, i.e. since RCMs operate with finer spatial resolution, the errors might be processed and dynamically improved. Hence, the near-surface values in the inner domain will have less connection to the boundary conditions, and potential errors entering the domain on the boundaries are suppressed. But the result cannot be anticipated, individual combinations of GCM-RCM will behave differently.

Declarations

Acknowledgments

This work used resources of the Deutsches Klimarechenzentrum (DKRZ) granted by its Scientific Steering Committee (WLA) under project MCS LOVE CCS in the framework of IS-ENES3 Analysis Platform activity. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model outputs. For CMIP the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We thank the Copernicus Climate Change Service, ECMWF, NCEP, the University of Wyoming, the ETH Zurich CMIP6 next generation archive (Brunner et al., 2020) and E-OBS dataset from the EU-FP6 project UERRA (http://www.uerra.eu) and the data providers in the ECA&D project (https://www.ecad.eu) for the data availability. The work has been done in the frame of the project Prediction, Evaluation and Research for Understanding National sensitivity and impacts of drought and climate change for Czechia (PERUN) funded by the Technology Agency of the Czech Republic (Grant No. SS02030040).

Funding

This research was funded by the Technology Agency of the Czech Republic (Grant No. SS02030040, PERUN).

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Author contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Eva Holtanová, Natália Machado Crespo and Michal Belda. The first draft of the manuscript was written by Eva Holtanová and Natália Machado Crespo and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data Availability

The data used for this research are publicly available at: a) CFSR: https://rda.ucar.edu/datasets/ds094.2/; b) ERA5: 10.24381/cds.6860a573; c) E-OBS: https://www.ecad.eu/; d) Radiosondes: https://weather.uwyo.edu/upperair/sounding.html; e) CMIP6 GCM: https://esgf-node.llnl.gov/projects/cmip6/. Further post-processed datasets generated during the current study are available under request.

References

1. Abramowitz G, Herger N, Gutmann E, Hammerling D, Knutti R et al (2019) ESD Reviews: Model dependence in multi-model climate ensembles: weighting, sub-selection and out-of-sample testing. Earth Syst Dynam 10:91–105. https://doi.org/10.5194/esd-10-91-2019
2. Adachi SA, Tomita H (2020) Methodology of the constraint condition in dynamical downscaling for regional climate evaluation: A review. J Geophys Res Atmos 125. https://doi.org/10.1029/2019JD032166. e2019JD032166
3. Askjær TG, Zhang Q, Schenk F, Ljungqvist FC, Lu Z et al (2022) Multi-centennial Holocene climate variability in proxy records and transient model simulations. Q Sci Rev 296:107801. https://doi.org/10.1016/j.quascirev.2022.107801
4. Bandhauer M, Isotta F, Lakatos M, Lussana C, Båserud L et al (2022) Evaluation of daily precipitation analyses in E-OBS (v19. 0e) and ERA5 by comparison to regional high-resolution datasets in European regions. Int J Climatol 42:727–747. https://doi.org/10.1002/joc.7269
5. Bartók B, Wild M, Folini D, Lüthi D, Kotlarski S et al (2017) Projected changes in surface solar radiation in CMIP5 global climate models and in EURO-CORDEX regional climate models for Europe. Clim dyn 49:2665–2683. https://doi.org/10.1007/s00382-016-3471-2
6. Belda M, Holtanová E, Kalvová J, Halenka T (2016) Global warming-induced changes in climate zones based on CMIP5 projections. Clim Res 71:17–31. https://doi.org/10.3354/cr01418
7. Brunner L, Hauser M, Lorenz R, Beyerle U (2020) The ETH Zurich CMIP6 next generation archive: technical documentation. 10.5281/zenodo.3734128
8. Cannon AJ (2020) Reductions in daily continental-scale atmospheric circulation biases between generations of global climate models: CMIP5 to CMIP6. https://dx.doi.org/10.1088/1748-9326/ab7e4f. Environmental Research Letters 15
9. Christensen OB, Kjellström E (2020) Partitioning uncertainty components of mean climate and climate change in a large ensemble of European regional climate model projections. Clim Dyn 54:4293–4308. https://doi.org/10.1007/s00382-020-05229-y
10. Coppola E, Raffaele F, Giorgi F, Giuliani G, Xuejie G, Ciarlo JM et al (2021) Climate hazard indices projections based on CORDEX-CORE, CMIP5 and CMIP6 ensemble. Clim Dyn 57:1293–1383. https://doi.org/10.1007/s00382-021-05640-z.)
11. Cornes R, van der Schrier G, van den Besselaar EJM, Jones PD (2018) An Ensemble Version of the E-OBS Temperature and Precipitation Datasets. J Geophys Res Atmos 123:9391–9409. 10.1029/2017JD028200
12. Crhová L, Holtanová E (2018) Simulated relationship between air temperature and precipitation over Europe: sensitivity to the choice of RCM and GCM. Int J Climatol 38:1595–1604. 10.1002/joc.5256

13. Dai and Deng (2022) Recent Eurasian winter cooling partly caused by internal multidecadal variability amplified by Arctic sea ice–air interactions. Clim dyn 58:3261–3277. https://doi.org/10.1007/s00382-021-06095-y
14. Di Virgilio G, Ji F, Tam E, Nishant N, Evans JP et al (2022) Selecting CMIP6 GCMs for CORDEX dynamical downscaling: Model performance, independence, and climate change signals. *Earth's Future 10*: e2021EF002625. https://doi.org/10.1029/2021EF002625
15. Eyring V, Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ, Taylor KE (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci Model Dev 9:1937–1958. https://doi.org/10.5194/gmd-9-1937-2016
16. Fernandez-Granja JA, Casanueva A, Bedia J, Fernández J (2021) Improved atmospheric circulation over Europe by the new generation of CMIP6 Earth System Models. Clim Dyn 56:3527–3540. https://doi.org/10.1007/s00382-021-05652-9
17. Forster P, Storelvmo T, Armour M, Collins W, Dufresne J-L et al (2021) The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte V, Zhai P, Pirani A, Connors SL, Péan C, Berger S, Caud N, Chen Y, Goldfarb L, Gomis MI, Huang M, Leitzell K, Lonnoy E, Matthews JBR, Maycock TK, Waterfield T, Yelekçi O, Yu R, Zhou B (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 923–1054, 10.1017/9781009157896.009
18. Giorgi F (2019) Thirty years of regional climate modeling: Where are we and where are we going next? J Geophys Res Atmos 124:5696–5723. https://doi.org/10.1029/2018JD030094
19. Gleckler PJ, Taylor KE, Doutriaux C (2008) Performance metrics for climate models. J Geophys Res 113:D06104. 10.1029/2007JD008972
20. Hausfather Z, Marvel K, Schmidt GA, Nielsen-Gammon JW, Zelinka M (2022) Climate simulations: Recognize the 'hot model' problem. Nature 605:26–29. 10.1038/d41586-022-01192-2
21. Hersbach H, Bell B, Berrisford P, Biavati G, Horányi A et al (2023) ERA5 monthly averaged data on pressure levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). 10.24381/cds.6860a573 (Accessed on 16-May-2023)
22. Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A et al (2020) The ERA5 global reanalysis. Q J R Meteorol Soc 146(730):1999–2049. https://doi.org/10.1002/qj.3803
23. Hollander M, Wolfe DA (1973) Nonparametric Statistical Methods. John Wiley & Sons, New York
24. Holtanová E, Belda M, Halenka T (2022) Projected changes in mean annual cycle of temperature and precipitation over the Czech Republic: Comparison of CMIP5 and CMIP6. Front Earth Sci 10:1018661. 10.3389/feart.2022.1018661
25. Holtanová E, Kalvová J, Pišoft P, Mikšovský J (2014) Uncertainty in regional climate model outputs over the Czech Republic: the role of nested and driving models. Int J Climatol 34:27–35. 10.1002/joc.3663
26. Holtanová E, Mendlik T, Koláček J, Horová I, Mikšovský J (2019) Similarities within a multi-model ensemble: functional data analysis framework. Geosci Model Dev 12:735–747. 10.5194/gmd-12-735-2019
27. Holtanová E, Mikšovský J, Kalvová J, Pišoft P, Motl M (2012) Performance of ENSEMBLES regional climate models over Central Europe using various metrics. Theoret Appl Climatol 108:463–470. 10.1007/s00704-011-0542-5
28. Jury MW, Prein AF, Truhetz H, Gobiet A (2015) Evaluation of CMIP5 models in the context of dynamical downscaling over Europe. J Clim 28:5575–5582. https://doi.org/10.1175/JCLI-D-14-00430.1
29. Katragkou E, Sobolowski SP, Teichmann C, Solmon F, Pavlidis V et al (2023) Delivering an improved framework for the new generation of CMIP6-driven EURO-CORDEX regional climate simulations. Bulletin of American Mereological Society, under review
30. Kim Y, Evans JP, Sharma A, Rocheta E (2021) Spatial, temporal, and multivariate bias in regional climate model simulations. *Geophysical Research Letters 48*: e2020GL092058. https://doi.org/10.1029/2020GL092058

31. Kim Y, Rocheta E, Evans JP, Sharma A (2020) Impact of bias correction of regional climate model boundary conditions on the simulation of precipitation extremes. Clim Dyn 55:3507–3526. https://doi.org/10.1007/s00382-020-05462-5
32. Lim C-M, Yhang Y-B, Ham S (2019) Application of GCM Bias Correction to RCM Simulations of East Asian Winter Climate. Atmosphere 10:382. https://doi.org/10.3390/atmos10070382
33. Maraun D, Widmann M, Gutiérrez JM, Kotlarski S, Chandler RE et al (2015) VALUE: A framework to validate downscaling approaches for climate change studies. Earth's Future 3:1–14. 10.1002/2014EF000259
34. Mašek J, Brožková R et al (2023) Using ALARO configuration of the ALADIN system for the convection-permitting climate simulations over Central Europe. Journal of Applied Meteorology and Climatology, in preparation
35. Merrifield AL, Brunner L, Lorenz R, Humphrey V, Knutti R (2023) Climate model Selection by Independence, Performance, and Spread (ClimSIPS v1.0.1) for regional applications. Geosci Model Dev 16:4715–4747. https://doi.org/10.5194/gmd-16-4715-2023
36. Palmer TE, McSweeney CF, Booth BBB, Priestley MDK, Davini P et al (2023) Performance-based sub-selection of CMIP6 models for impact assessments in Europe. Earth Syst Dynam 14:457–483. https://doi.org/10.5194/esd-14-457-2023
37. Pickler C, Mölg T (2021) General circulation model selection technique for downscaling: Exemplary application to East Africa. J Geophys Res Atmos 126. https://doi.org/10.1029/2020JD033033. e2020JD033033
38. Prein AF, Bukovsky MS, Mearns LO, Bruyère CL, Done JM (2019) Simulating North American Weather Types With Regional Climate Models. Front Environ Sci 7:36. 10.3389/fenvs.2019.00036
39. Reichler T, Kim J (2008) How Well Do Coupled Models Simulate Today's Climate? Bull Am Meteorol Soc 89:303–312. https://doi.org/10.1175/BAMS-89-3-303
40. Rocheta E, Evans JP, Sharma A (2017) Can Bias Correction of Regional Climate Model Lateral Boundary Conditions Improve Low-Frequency Rainfall Variability? J Clim 30:9785–9806. https://doi.org/10.1175/JCLI-D-16-0654.1
41. Rocheta E, Evans JP, Sharma A (2020) Correcting lateral boundary biases in regional climate modeling – The effect of the relaxation zone. Clim Dyn 55:2511–2521. https://doi.org/10.1007/s00382-020-05393-1
42. Saha S, Moorthi S, Pan HL, Wu X, Wang J et al (2010a) The NCEP climate forecast system reanalysis. Bull Am Meteorol Soc 91:1015–1058. https://doi.org/10.1175/2010BAMS3001.1
43. Saha S et al (2010b) NCEP Climate Forecast System Reanalysis (CFSR) Monthly Products, January 1979 to December 2010 [Dataset]. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. https://doi.org/10.5065/D6DN438J. Accessed 16 May 2023
44. Saha S et al (2012) NCEP Climate Forecast System Version 2 (CFSv2) Monthly Products [Dataset]. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. https://doi.org/10.5065/D69021ZF. Accessed 16 May 2023
45. Saha S et al (2014) The NCEP Climate Forecast System Version 2. J Clim 27:2185–2208. 10.1175/JCLI-D-12-00823.1
46. Séférian R, Nabat P, Michou M, Saint-Martin D, Voldoire A et al (2019) Evaluation of CNRM Earth-System model, CNRM-ESM2-1: role of Earth system processes in present-day and future climate. J Adv Model Earth Syst 11:4182–4227. https://doi.org/10.1029/2019MS001791
47. Séférian R, Delire C, Decharme B, Voldoire A, Salas y Melia D et al (2016) Development and evaluation of CNRM Earth system model – CNRM-ESM1. Geosci Model Dev 9:1423–1453. https://doi.org/10.5194/gmd-9-1423-2016
48. Sobolowski S, Somot S, Fernandez J, Evin G, Maraun D, Kotlarski S, Jury M, Benestad RE, Teichmann C, Christensen OB, Bülow K, Buonomo E, Katragkou E, Steger C, Sørland S, Nikulin G, McSweeney C, Dobler A, Palmer T, Wilke R, Boe J, Brunner L, Ribes A, Qasmi S, Nabat P, Sevault F, Oudar T, Brands S 2023, EUROCORDEX CMIP6 GCM Selection and Ensemble Design: Best Practices and Recommendations. https://doi.org/10.5281/zenodo.7673400
49. Takayabu I, Kanamaru H, Dairaku K, Benestad R, von Storch H, Christensen JH (2016) Reconsidering the quality and utility of downscaling. J Meteor Soc Japan 94:31–45. https://doi.org/10.2151/jmsj.2015-042

50. Taszarek M, Pilguj N, Allen JT, Gensini V, Brooks HE, Szuster P (2021) Comparison of convective parameters derived from ERA5 and MERRA-2 with rawinsonde data over Europe and North America. J Clim 34:3211–3237. https://doi.org/10.1175/JCLI-D-20-0484.1
51. Tebaldi C, Debeire K, Eyring V, Fischer E, Fyfe J et al (2021) Climate model projections from the Scenario Model Intercomparison Project (ScenarioMIP) of CMIP6. Earth Syst Dynam 12:253–293. https://doi.org/10.5194/esd-12-253-2021
52. Termonia P, Fischer C, Bazile E, Bouyssel F, Brožková R et al (2018) The ALADIN System and its canonical model configurations AROME CY41T1 and ALARO CY40T1. Geosci Model Dev 11:257–281. https://doi.org/10.5194/gmd-11-257-2018
53. Urban A, Di Napoli C, Cloke HL, Kyselý J, Pappenberger F et al (2021) Evaluation of the ERA5 reanalysis-based Universal Thermal Climate Index on mortality data in Europe. Environ Res 198:111227. https://doi.org/10.1016/j.envres.2021.111227
54. Varga ÁJ, Breuer H (2022) Evaluation of convective parameters derived from pressure level and native ERA5 data and different resolution WRF climate simulations over Central Europe. Clim Dyn 58:1569–1585. https://doi.org/10.1007/s00382-021-05979-3
55. Wang J, Lu H, Cheng J, Zhao C (2023) Global terrestrial monsoon area variations since Last Glacial Maximum based on TraCE21ka and PMIP4-CMIP6 simulations. Glob Planet Change 104308. https://doi.org/10.1016/j.gloplacha.2023.104308
56. Walawender E, Kielar R, Ustrnul Z (2017) Use of RegCM gridded dataset for thunderstorm favorable conditions analysis over Poland—climatological approach. Theor Appl Climatol 127:229–240. https://doi.org/10.1007/s00704-015-1620-x
57. Xu J, Gao Y, Chen D, Xiao L, Ou T (2017) Evaluation of global climate models for downscaling applications centred over the Tibetan Plateau. Int J Climatol 37:657–671. https://doi.org/10.1002/joc.4731
58. Yang H, Lu J, Wang Q, Shi X, Lohmann G et al (2022) Decoding the dynamics of poleward shifting climate zones using aqua-planet model simulations. Clim Dyn 58:3513–3526. https://doi.org/10.1007/s00382-021-06112-0. *)*
59. Zhang MZ, Xu Z, Han Y, Guo W (2022) Evaluation of CMIP6 models toward dynamical downscaling over 14 CORDEX domains. Clim Dyn. https://doi.org/10.1007/s00382-022-06355-5

Figures

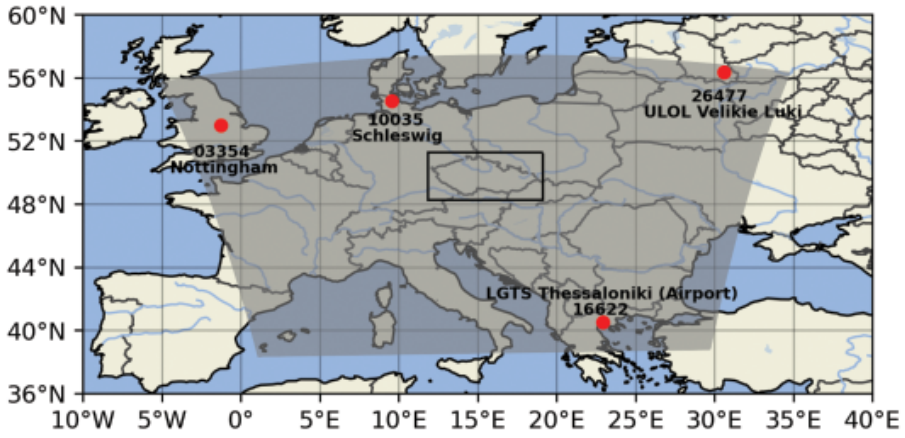

Figure 1

Integration domain of Aladin-Climate/CZ (shaded in grey), the selected locations with available soundings between the period 1990-2014 (red dots, and their respective names), and the inner domain covering the Czech Republic territory (black

rectangle).

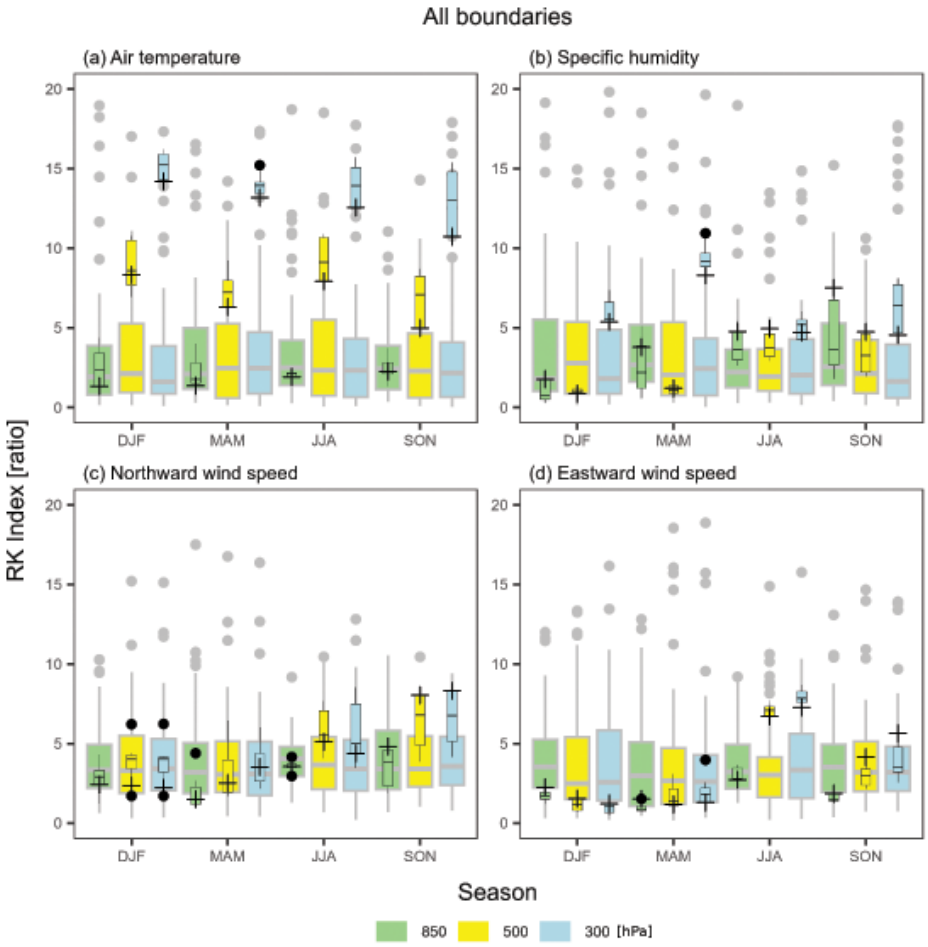

Figure 2

RK Index of seasonal mean of (a) air temperature, (b) specific humidity, (c) northward (v) wind speed and (d) eastward (u) wind speed in the period 1990–2014 summed up over all the boundaries of the studied domain (see Figure 1). Boxplots show intra-ensemble statistical distribution (median, lower and upper quartiles, whiskers representing 1.5 * IQR, and outliers as dots) of CMIP6 (larger boxplots with grey lines and outliers as grey dots) and perturbed initial conditions mini-ensemble CNRM-INI (smaller boxplots with black lines and outliers as black dots). The CNRM-r1 simulation is shown using black crosses. Green, yellow and blue colors denote, respectively, 850, 500 and 300 hPa levels.

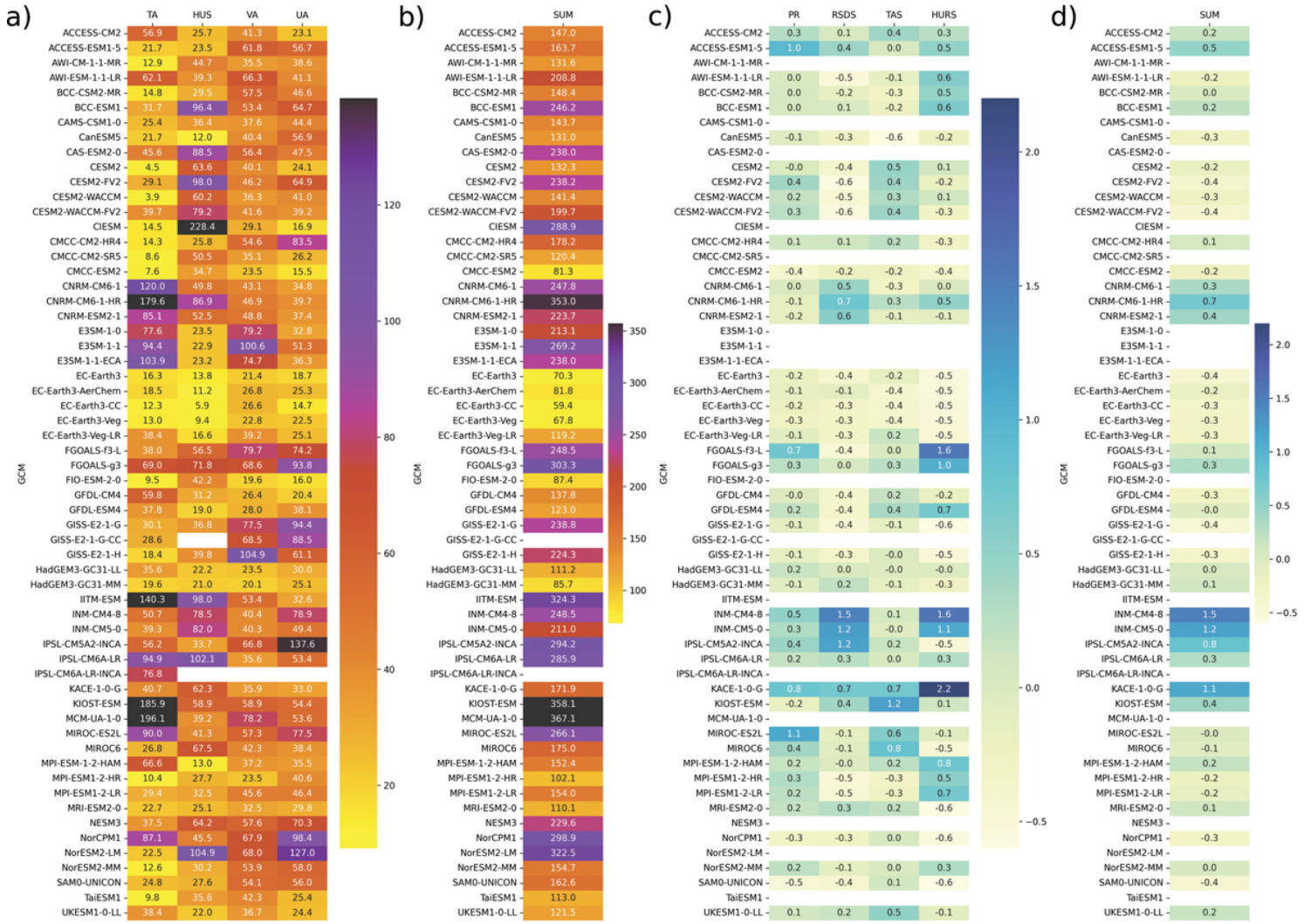

Figure 3

Heatmap for the RKI summed up for all boundaries, seasons and levels for the individual GCMs: a) for the individual variables and b) the sum of the four variables in a). Heatmap for the relative error RE of GCMs c) based on RMSE of mean annual cycle for individual surface variables and d) based on RMSE summed up for the four variables over the Czech Republic territory.

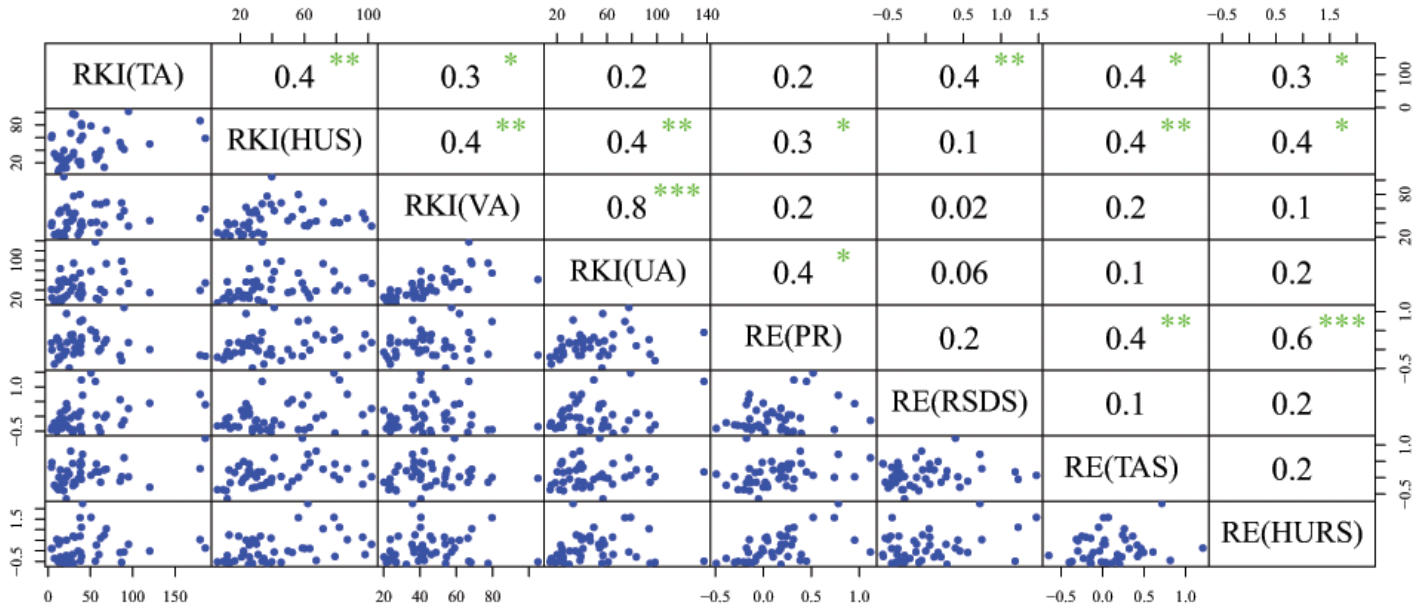

Figure 4

Correlation matrix between the RKI summed up over all seasons, boundaries and levels and the RE for individual variables. Above diagonal, the Spearman correlation coefficients among respective metrics are shown. The stars correspond to the intervals of p-values resulting from testing the null hypothesis of zero correlation, i.e. ***: 0 - 0.001, **: 0.001 - 0.01, *: 0.01-0.05. No star implies p-value higher than 0.05. The method of the statistical test can be found in Hollander and Wolfe (1973), pages 185-194.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Holtanovaetalsubmittedsupplement.pdf