

IDŐJÁRÁS

*Quarterly Journal of the Hungarian Meteorological Service
Vol. 125, No. 1, January – March, 2021, pp. 123–135*

Geostatistical comparison of UERRA MESCAN-SURFEX daily temperatures against independent data sets

Hristo Chervenkov* and Kiril Slavov

*National Institute of Meteorology and Hydrology – Bulgarian Academy of Sciences
66, Tsarigradsko Shose Blvd Sofia 1784, Bulgaria*

**Corresponding author E-mail: hristo.tchervenkov@meteo.bg*

(Manuscript received in final form April 23, 2020)

Abstract— Regional reanalyses are attractive new sources of meteorological data for the growing society of the end users, due to their physical consistency, dynamical coherency, and multivariate products at higher than the global reanalyses spatio-temporal resolution. The assessment and quantification of uncertainties of the products of the regional reanalyses and their added value are crucial for the interpretation. Hence these products could be also incorporated in the regional climatology, consistent comparisons of their long-term timeseries against independent and representative data sets have to be performed. The present study could be considered as step ahead in this direction - the MESCAN-SURFEX, which is the product with the most detailed spatial structure among all others in the UERRA (Uncertainties of Ensembles in Regional Reanalysis) project, is compared against two gridded observational data sets in South-east Europe: the well-known regional CARPATCLIM and the product of the Bulgarian National Institute of Meteorology and Hydrology ProData. The comparison aims to assess the skill of MESCAN-SURFEX to reproduce the climatological field of the mean temperature. Additionally, the daily extreme temperatures are estimated using the MESCAN-SURFEX output on sub daily basis and the results are compared against their CARPATCLIM- and ProData-counterparts. The computation of the mean and extreme temperatures with the MESCAN/SURFEX data are performed for the whole time span of this product and the comparison against the references for the whole time span of each of them on daily basis. The main conclusion of the study, which agrees with the outcomes of more detailed recent evaluations, is that MESCAN-SURFEX reproduces realistically the regional temperature field over Southeast Europe. According to the mean temperature, the differences remains under certain limits (RMSE generally below 2 °C) without, at least not apparent, systematic and spatial pattern. The estimation of the extreme temperatures produces results with biases comparable to the biases of the mean temperature, which makes the proposed method applicable for certain cases.

Key-words: UERRA, MESCAN-SURFEX, CARPATCLIM, ProData, uncertainties estimation, mean and extreme temperatures

1. Introduction

Daily minimum, mean, and maximum temperatures, noted subsequently hereafter for the sake of brevity t_n , t_d , and t_x , as well as the daily precipitation sum, are essential climate variables (ECVs) particularly involved in determining climate change impacts on society and ecosystems (*Birsan et al.*, 2014). They are also widely used, both in regional and global scales, as input parameters for computation of climate indices as the ETCCDI (Climate Extreme Indices) dataset (*Chervenkov and Slavov*, 2019; *Cheval et al.*, 2014; *Lakatos et al.*, 2013b; *Sillmann and Röckner*, 2008). The growing demand of the user community for high resolution meteorological data, both in space and time, could be addressed by various downscaling methods, with variable success depending on meteorological parameter, method applied, and spatial scale (*Kaiser-Weiss, et al.*, 2019). Atmospheric reanalysis (RA) provides a physically consistent and dynamically coherent description of the state of the atmosphere. The main merit of RA are that they provide a multivariate, spatially complete, and coherent record of the atmospheric state – far more complete than any observational dataset is able to achieve (*Ridal et al.*, 2018). This is the basic reason for the remarkable success of the widely used global RA, serving tens of thousands of users in need of climatological information (*Kaiser-Weiss, et al.*, 2019). Recently, due to the progress of the scientific knowledge and computation power, regional reanalyses (RRA) which can ingest additional observations and achieve higher resolution by means of regional numerical weather prediction (NWP) models, become possible. Hence wind speed and solar irradiation are strongly fluctuating parameters requiring the evaluation of their fields at highly-resolved spatial (a few kilometers) and temporal (hours) scales, the RRA products could address also the needs of the growing user community in the energy sector (*Niermann et al.*, 2017).

The objective of the project-driven collaborative initiative UERRA (Uncertainties in Ensembles of RRA; www.uerra.eu) is to produce ensembles of European RRA of ECVs for several decades and estimate the associated uncertainties in the data sets (*Ridal et al.*, 2018; *Unden*, 2018). It also includes recovery of historical (last century) data and creation of user friendly data services. Within UERRA, three different NWP models have been employed to generate European regional reanalyses and subsequent surface reanalysis products. All models share the same boundary conditions (provided by RA ERA-Interim and ERA40) and domain (CORDEX-EUR11, European domain of the Coordinated Regional Climate Downscaling Experiment, <http://www.euro-cordex.net>; *Kaiser-Weiss, et al.*, 2019; *Niermann et al.*, 2017).

The assessment and quantification of uncertainties of the RRA products and their added value is crucial for the interpretation. The proper use in applications and downstream services depends on the knowledge of the quality of the RRA and the representation of uncertainties. The information content of

the RRA UERRA and their uncertainties are statistically assessed by comparison against observation-based independent data sets, with several different methods, at user-relevant scales (*Niermann et al.*, 2017). One of these methods ('Method C') consist of comparison against gridded station observations. Advantage of the method, as stated in *Niermann et al.* (2017) is that the aggregation of data over selected regions in the European domain and over time provides one simple measure which can be used as an easy-to-interpret metric for the overall quality of the RRA product. For the pan-European approach, gridded observational data E-OBS (*Cornes et al.*, 2018) based on a dense network of stations covering Europe is used to assess reanalysis results. Beside the fact that the spatial aggregation from the finer (UERRA) to coarser (E-OBS) grid obscures possible local problems, E-OBS itself is prone to quality and reliability issues over regions with sparse station coverage as Southeast (SE) Europe (*Cornes et al.*, 2018) This problem is addressed in *Niermann et al.* (2017) performing additional comparisons against regional gridded observational datasets (APGD, NGCD, ROCADA). Hence these sources, being products of National Meteorological Services (NMS) and/or regional initiatives, they generally incorporate many more station series compared to E-OBS. Therefore, it would be expected to produce gridded datasets that are closer to the 'true' climate field (*Cornes*, 2016). Only one of the aforementioned products, however, the Romanian ROCADA (*Birsan and Dumitrescu*, 2014), is situated in SE Europe. Although ROCADA contains many variables, only evaluations for drought-related quantities and not for the temperature are presented in the UERRA-related documentation.

The present work, which is not comprehensive evaluation study, is dedicated to the comparison of the long-term temperature field, reproduced with UERRA MESCAN-SURFEX with the regional gridded dataset CARPATCLIM and the gridded dataset of the Bulgarian NMS ProData. It is worth emphasizing, however, that generally, all observations come with uncertainties, which will impact more or less the output from any such comparison (*Kaiser-Weiss et al.*, 2019). The domains of CARPATCLIM and ProData are in SE Europe, the full time spans with daily temporal resolution of both datasets are considered. The work also try to answer the question: How realistic are the daily minimum and maximum temperature obtained from the MESCAN-SURFEX output in 6-hour temporal resolution?

The article is structured as follows. The considered datasets are described in Section 2 — MESCAN-SURFEX in Subsection 2.1 and CARPATCLIM and ProData in Subsection 2.2. The applied numerical techniques for estimation of t_n , t_d , and t_x , which are rather simple indeed, are explained in Section 3. The core of the article is in Section 4, where the performed comparisons and the obtained results are presented. The concise concluding remarks are in Section 5.

2. MESCOAN-SURFEX and reference datasets

2.1. Short Description of MESCOAN-SURFEX

The MESCOAN-SURFEX system analysis uses the 2D-analysis system MESCOAN (Soci *et al.*, 2016) and the land surface platform SURFEX (Bazile *et al.*, 2017) to generate a coherent surface and soil analysis. The UERRA-NWP HARMONIE-ALADIN at 11 km grid spacing is used as a starting point to further downscaling. For the forcing for the SURFEX surface and soil model, downscaling was only done through interpolation by Météo-France. These downscaled fields are refined with additional data in a surface analysis with observations that are not used in the 3D reanalysis, resulting in the MESCOAN-SURFEX at 5.5 km grid spacing (Bazile *et al.*, 2017). It is run offline, i.e., without feedback to the atmospheric analysis. Beside the other parameters, MESCOAN-SURFEX produces temperature at 2 m above the surface in 6-hour temporal resolution, i.e., at 00, 06, 12, and 18 UTC for the period 1961–2018.

2.2. Reference datasets - CARPATCLIM and ProData

We use two independent data sets as reference in this study; both of them are based on surface measurements and are in form of gridded digital maps. Climate of the Carpathian Region (CARPATCLIM; <http://www.carpatclim-eu.org>) is a collaborative international project over the area, developed with the joint effort of NMSs from all the Carpathian countries. Main aim of the project was to enhance the climatic information in the region by providing comprehensive, temporally and spatially homogeneous data sets of the main meteorological variables, and the corresponding metadata (Cheval *et al.*, 2014; Lakatos *et al.*, 2013a). Besides the common software, the harmonization of the results across country borders was promoted also by near border data exchange. The database of CARPATCLIM is produced at daily temporal resolution, covers the period 1961–2010 for the Carpathian Region (44°N–50°N and 17°E–27°E) with 0.1° grid spacing. It provides relevant outcomes, suitable for various applications in the regional climatology (Birsan *et al.*, 2014; Lakatos *et al.*, 2013b). Comprehensive studies as Spinoni *et al.* (2015) prove the suitability of CARPATCLIM for elaboration of objective climatologies and estimation of the trends of key variables. It is worth emphasizing, that CARPATCLIM is used also as reference in the UERRA project (Cornes, 2016), although not for evaluation of MESCOAN-SURFEX.

ProData is a product of the Bulgarian NMS and runs operationally since the second half of 2013. ProData assimilates ground-based in situ measurements from more than 140 automatic weather and hydrological stations, satellite-derived products, and auxiliary data and produces estimates of 10 near-surface meteorological parameters. The ProData domain covers Bulgaria entirely with

horizontal resolution of 0.045° , which is very close to the native resolution of MESCAN-SURFEX. The main purpose of the system is to serve as a reliable source of consistent meteorological information with high spatial and temporal (1 hour) resolution with minimal latency from the input data acquisition time (Chervenkov *et al.*, 2017). The everyday practical experience as well as validation against independent data (synoptic measurements and satellite-derived products, see Chervenkov and Slavov (2021) for details) reveals the high performance skill of the system. Beyond its operative applications, the system is used recently in climatological studies (Chervenkov and Slavov, 2020; Ivanov and Chervenkov, 2019) as source of valuable information.

The main characteristics of the considered datasets are listed in *Table 1*, and the orography of the model domain as well as the spatial extent of CARPATCLIM and ProData is shown in *Fig. 1*.

Table 1. Overview of the main characteristics of the considered datasets

	spatial/temporal coverage	spatial/temporal resolution
MESCAN-SURFEX	CORDEX-EUR11/1961–2018	5.5 km×5.5 km/6-hourly
CARPATCLIM	Carpathian Basin/1961–2010	0.1°×0.1°/daily
ProData	Bulgaria/2014–2018	0.045°×0.045°/1-hourly

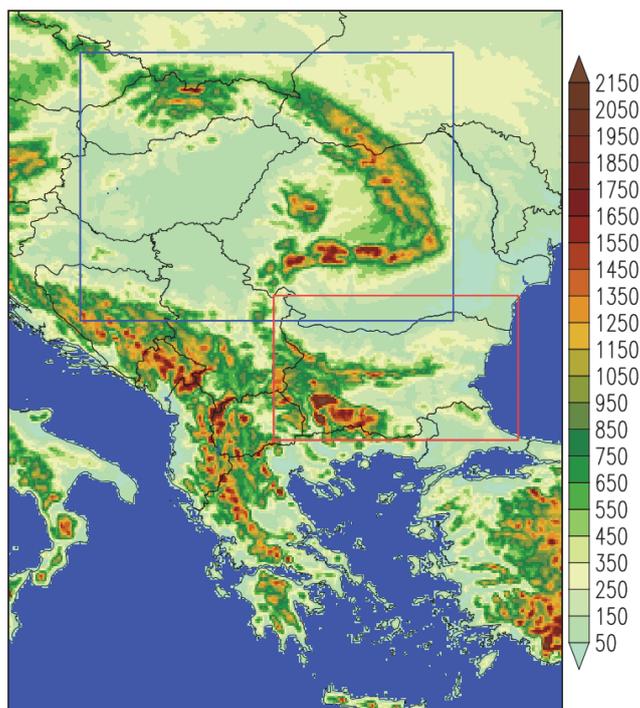


Fig. 1. Orography of the model domain in 0.05° resolution. The borders of CARPATCLIM and ProData are shown in blue and red, respectively.

3. Estimation of the temperatures

The complex nature of missing daily t_n , t_d , and t_x is a problem in many climatological applications. There are different ways to calculate daily mean temperature from data collected at different times of the day. In many countries, including the US, the approach is to average the minimum and maximum temperatures observed. In other countries, a linear combination of measurements taken at different times of the day is used, sometimes including the minimum and maximum as well. For example, the Nordic countries each have a different linear combination of data, depending on the frequency of recorded observations (*Ma and Guttorp, 2013*), while Germany employs yet another linear combination of data (see *Dall'Amico and Hornsteiner (2006)* and citations therein). The World Meteorological Organization defines the mean daily temperature as the 'mean of the temperature observed at 24 equidistant times in the course of a continuous interval of 24 hours, or a combination of temperatures observed at less numerous times, so arranged as to depart as little as possible from the mean defined above' (*Dall'Amico and Hornsteiner, 2006*). In the present work, the mean daily temperature is computed as average from five values: the MESCAN-SURFEX-output at 06, 12, 18 UTC and both values which bound the current day – at 00 UTC of the same day and 00 UTC of the next one. This method is most natural, physically consistent, and, not at last, close to definition above in the considered case. The methods used for estimates of daily temperature minima and maxima vary also greatly in complexity and sophistication. Traditionally, tabulated values have been published for a given region or country. At the other end of the scale, some groups have analyzed detailed daily time series of temperatures to prepare synthetic data (*Huld et al., 2006*). Researchers have relied on a variety of techniques to estimate missing data depending on the region, time of year, spatial distribution of neighboring stations (*Allen and DeGaetano, 2001*), or more recently, remote sensing products. Other methods rely on temporal interpolation by means of polynomial fitting or implementation of piecewise cosine functions (*Huld et al., 2006*). Such methods attempt to reproduce the diurnal temperature cycle using additionally data for the sunrise/sunset times. Hence our task is limited to the estimation of the extreme temperatures only, the most simple solution, based on the polynomial fitting, is to find a local polynomial of 2nd order around the time horizon with the minimum/maximum (hereafter minimum for the sake of brevity) temperatures:

$$T(t) = At^2 + Bt + C, \quad (1)$$

where T is the temperature, t is the time, and A , B , and C are the coefficients of the polynomial, which passes through the point of the minimum temperature and the two neighboring points. This popular and simple procedure, which is ought

to find the minimum in a single iteration, is technically called parabolic interpolation (*Press et al.*, 1986; *Stoer and Bulirsch*, 2002). If we apply the quadratic Lagrange interpolating polynomials, then:

$$T(t) = T_1 \frac{(t-t_2)(t-t_3)}{(t_1-t_2)(t_1-t_3)} + T_2 \frac{(t-t_1)(t-t_3)}{(t_2-t_1)(t_2-t_3)} + T_3 \frac{(t-t_1)(t-t_2)}{(t_3-t_1)(t_3-t_2)}, \quad (2)$$

where $T_1 = T(t=t_1)$, $T_2 = T(t=t_2)$, and $T_3 = T(t=t_3)$. The time of minimum temperature t_{\min} can be found by simple differentiation of the polynomial in Eq. (2):

$$t_{\min} = \frac{1}{2}(t_1 + t_2) + \frac{1}{2} \frac{(T_1 - T_2)(T_2 - T_3)(T_3 - T_1)}{(t_2 - t_3)T_1 + (t_3 - t_1)T_2 + (t_1 - t_2)T_3} \quad (3)$$

Finally, we get the minimum temperature T_{\min} , substituting Eq.(3) in Eq.(2): $T_{\min} = T(t=t_{\min})$.

This is the generic case for sufficiently smooth functions, in particular the near-surface temperature in relatively calm weather conditions. The method is also applicable if the minimum temperature is in internal point (i.e., not at 00 UTC of the current or 00 UTC of the next day), and its accuracy depends on the distance between the interpolation knots. From meteorological point of view, however, its main weakness is obvious: every (sudden enough) change of the diurnal thermal cycle, caused by events with characteristic time scale smaller than the distance between the knots (in our case: 6h), leads to deviation which the method can not encounter. Typical examples are the various convection-related events, which are frequent for the considered domain. Thus, the method could be treated only as pragmatic workaround in cases of absence of data for the extreme temperatures, like the present situation with MESCAN-SURFEX. According to our main idea, if the biases of the extreme temperatures are comparable with the bias of the mean temperature, which is computed much more consistently, the method, at least in certain extent, could be accepted as justified. Thus, we will try to estimate empirically the validity of this approach for long-term data series, comparing its outcomes with reference data, in order to assess its suitability for climatological applications.

4. Comparison and discussion

The UERRA reanalysis datasets are freely available. The analyses are downloadable from the Copernicus Climate Data Store (CDS; <https://cds.climate.copernicus.eu>) and, applying the functionality of the python package cdsapi, we get directly the necessary data.

Due to pragmatic reasons, we selected a sub-domain over Southeast Europe, shown in *Fig. 1*, constraining the computations over the land only. The MESCAN-SURFEX dataset has a native resolution 5500×5500 m (see *Table 1* again) and Lambert Conic Conformal projection, and it is distributed in GRIB2 format, which can be handled with specific tools, at best the *ecCodes* library of the European Centre for Medium Range Weather Forecasts (ECMWF). In order to prepare testbed for further computations, we regridded the temperature data to the regular $0.05^\circ \times 0.05^\circ$ grid with 301×321 gridcells using a first-order conservative remapping procedure. The computation of the daily mean, as well as the extreme temperatures, is performed according to the methodology described in Section 2 for the full available time span of the MESCAN-SURFEX data set 1961–2018. For the present analysis, we mapped the MESCAN-SURFEX intercept over CARPATCLIM and ProData in the native grid of both products. All GRIB and netCDF file manipulations are performed with the powerful and convenient collection of the Climate Data Operators (<https://code.zmaw.de/projects/cdo>), compiled with *ecCodes*.

Comparison of the cumulative and density probability functions of the reference and the model is frequently used within climate modeling (*Harding et al.*, 2015). The quantile-quantile (Q-Q) plot, which is a plot of the quantiles of the first data set against the quantiles of the second one, is a commonly used technique for checking whether two data sets are distributed differently (*Gadzhev et al.*, 2020). The method could be generalized for two dimensional samples, limiting the comparison for some key quantiles, for example, the 10th, 25th (lower quartile), 50th (median), 75th (upper quartile), and 90th percentiles, traditionally noted as X10, X25, X50, X75, and X90, respectively. Similar approach, based on 95% quantiles and interquantiles (between the 10% and the 90% quantiles) is applied in the comprehensive UERRA-evaluation study (*Niermann et al.*, 2017) for comparison against gridded station observations. The quantiles of the minimum, mean, and maximum temperatures from MESCAN-SURFEX for the time span of CARPATCLIM (1961–2010) and ProData (2014–2018) are superimposed to their corresponding counterparts from CARPATCLIM and ProData, as shown in *Figs. 2* and *3*.

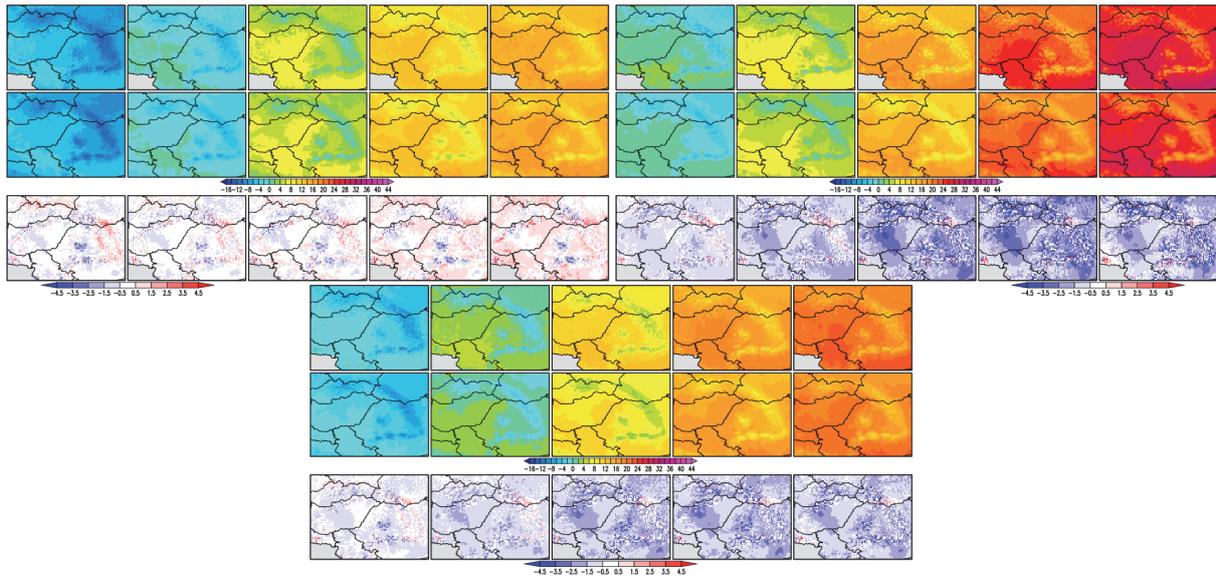


Fig. 2. Upper left corner: tnX10, tnX25, tnX50, tnX75, and tnX90 from CARPATCLIM (first row, reference) and MESCAN-SURFEX (second row), as well as the bias between them (third row); upper right corner: same as in the upper left corner, but for tx; bottom middle: same as in the upper left corner, but for td. The units are °C.

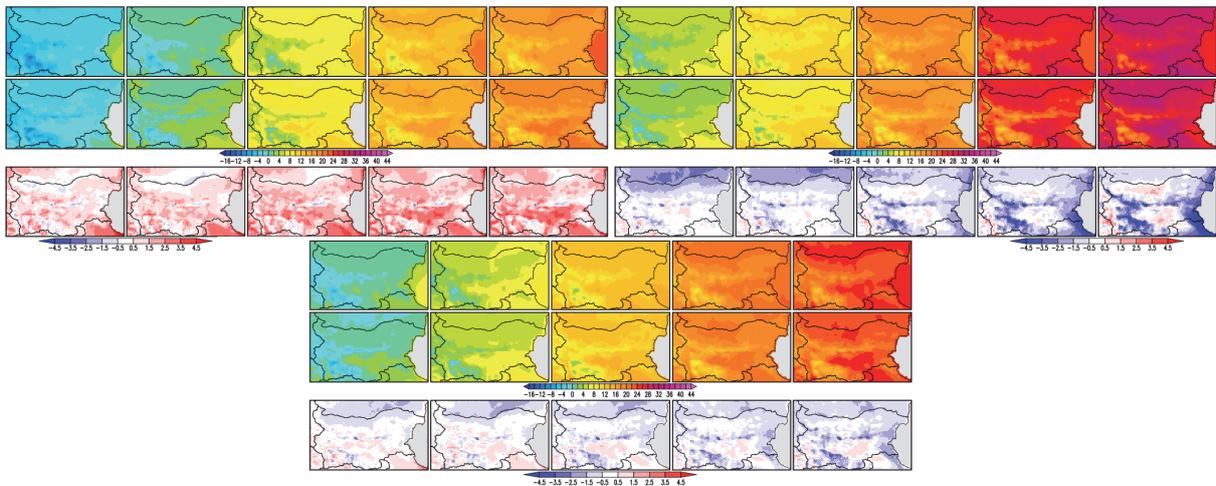


Fig. 3. Same as *Fig. 2*, but for ProData used as reference.

Fig. 2 shows that the minimum temperature is relatively well reproduced: the spatially prevailing bias is between $-0.5\text{ }^{\circ}\text{C}$ and $0.5\text{ }^{\circ}\text{C}$. The maximum temperature is, however, negatively biased with values about roughly $-3\text{ }^{\circ}\text{C}$ – $-2\text{ }^{\circ}\text{C}$. Mean temperatures are also underestimated with average biases of about $-1.5\text{ }^{\circ}\text{C}$. There is neither apparent geographical structure of the bias, nor clear dependence of the magnitude of the deviation from the value of the percentile. Thus, the bias for the upper tail of the distribution of the maximum temperature (txX90) is not generally bigger than the median txX50. The spatially dominating bias of the minimum temperature in *Fig. 3* is positive, but, surprisingly, the extreme low values, i.e., tnX10, are with the smallest bias of

about 0.5 °C–1.5 °C. The spatial distributions of the tx-percentiles in *Fig. 3* are the single ones with clear altitudinal dependence — hence the bias over the flat regions is near to zero, its value over the mountains is negative, with increasing tendency from the lower to the higher percentiles. Over the Rila Mountain, in the southwestern corner of the domain, the model underestimates the reference with more than 4 °C. It is worth emphasizing that this issue could be explained with the poor station coverage there, which affects the ProData product. Same could be the reason for the strong disagreement over the southeastern corner of the domain; hence ProData utilizes only Bulgarian (and not Turkish) observations. The bias of the mean temperature is without clear spatial structure; its values are generally between -1.5 °C and 1.5 °C.

Popular statistical scores as the root mean square error (RMSE) and mean bias (MB) are frequently used in the climatology as estimators of the model score (*Chervenkov and Slavov, 2019*) and uncertainty measures in the UERRA project (*Cornes, 2016; Kaiser-Weiss et al., 2019; Niermann et al., 2019*). The spatial distributions of the MB and RMSE of the comparison of tn, td, and tx between UERRA and reference datasets are shown in *Fig. 4*. *Fig. 4* reveals that the mean and maximum MESCAN-SURFEX-temperatures are, compared with CARPATCLIM, negatively biased with average values of roughly -1.5 °C. The spatially dominating MB of td is near zero. The biggest bias is (according its absolute value and among all performed comparisons) the positive bias of the comparison MESCAN-SURFEX-tn – ProData-tn, shown in the fifth column of *Fig. 4*. Its values are, over limited area indeed, up to 3.5 °C. The peak values of the MB for td and tx are negative and clearly situated over the mountain heights.

The RMSE of both panels in *Fig. 4* are generally in the interval 0.5–2.5 °C with slightly smaller values for td in the evaluation against ProData. A vertical gradient, not well expressed indeed, could be distinguished, mainly over the main Carpathian ridge for tn and the Balkan and Rila Mountain in Bulgaria.

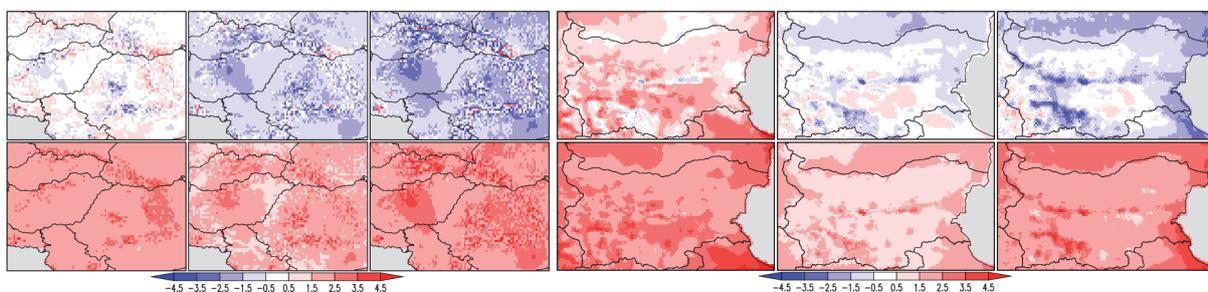


Fig. 4. MB (first row) and RMSE (second row) of the performed comparisons between MESCAN-SURFEX and CARPATCLIM (left panel) and ProData (right panel). tn, td, and tx are in the first, second, and third column of each panel, respectively. The units are °C.

The comparison results of any RRA-validation study depend on the location and the time scale and spatial scale considered, and thus, any generalization could be accepted with caution (*Kaiser-Weiss et al.*, 2019). The applied method, namely comparison against gridded observational datasets, has also limitations, rooted mainly in the specific weaknesses of the reference data (*Niermann et al.*, 2017). As it was demonstrated, however, confidence in both products can be gained when they agree in their general spatial patterns, or in their magnitude, or when their differences can be explained (e.g., when differences get larger for areas of poor data coverage used for the gridded fields).

5. Conclusion

The RRA within UERRA and more especially MESCAN-SURFEX, which has the highest horizontal spacing, constitute an attractive new source of highly-resolved spatial and temporal meteorological data. In many cases, for example for the vertical wind profiles, this source has practically no alternative. The performed comparisons of the mean temperature with the regional data set CARPATCLIM and the national ProData are novel. They reveals that the differences remains under certain limits without, at least not apparent, systematic and spatial pattern. Our opinion is that a part of these biases could be attributed to the principally different nature of the considered data sets, which are practically unavoidable. Obviously, the RRA product MESCAN-SURFEX is able to capture the regional temperature distribution in SE Europe. This conclusion agrees generally with the outcomes of the comprehensive evaluation studies (*Cornes*, 2016; *Niermann et al.*, 2017) adding new results for the region of SE Europe.

The applied method for estimation of the minimum and maximum temperatures, rather simple and schematic, as emphasized in Section 3, is intended as pragmatic workaround in case of absence of data for these parameters. Nevertheless, the obtained results show biases generally comparable to the biases of the mean temperature. Such approach seems justified for specific applications when the highest available resolution is crucial (and the use of *tn* and *tx* from coarser resolution RRA, as for example HARMONIE, is not an option) and/or in case of 'moderate extremes', causing potential pitfalls for threshold applications such as climate indices (*Kaiser-Weiss et al.*, 2019).

Conceptually, the UERRA project is the necessary prerequisite before the start of exploitation of unified pan-European RRA, similarly to the global RA. Subsequently, the positive consequences to the end user community will be manifold.

With the global RA running at higher resolution (i.e., ERA-5 with grid resolution of about 30 km), the RRA will benefit from higher quality boundary conditions, and at the same time, the question of added value and estimation of the uncertainties will arise anew (*Kaiser-Weiss et al.*, 2019).

The output parameters computed in this study, more specifically the mean daily temperature, could be used as a surrogate for computation for some climate indicators, as for example, agrometeorological indices (Chervenkov and Slavov 2019; Harding et al., 2015). The data sets with the daily mean, as well as the extreme temperatures, are in standard form (GrADS binary/descriptor files and netCDF files) could be supplied from the corresponding author upon request.

Acknowledgments: This work is based on the UERRA EU FP7 Collaborative Project, Grant agreement 607193. Hence this study is entirely based on free available data and software, the authors would like to express their deep gratitude to the organizations and institutions (UERRA project, CARPATCLIM, MPI-M, UNI-DATA, Copernicus Data Store), which provides free of charge software and data. Without their innovative data services and tools this work would be not possible. Personal thanks to I. Tsonevsky from ECMWF for the cooperation. Not at least we thank to the anonymous reviewers for their comments and suggestions which led to an overall improvement of the original manuscript.

References

- Allen, R.J. and DeGaetano, A.T., 2001: Estimating Missing Daily Temperature Extremes Using an Optimized Regression Approach. *Int. J. Climatol.* 21, 1305–1319.
<https://doi.org/10.1002/joc.679>
- Bazile, E., Abida, R., Verelle, A., Le Moigne, P., and Szczypta, C., 2017: MESCAN-SURFEX surface analysis, deliverable D2.8 of the UERRA project.
<http://www.uerra.eu/publications/deliverable-reports.html>
- Birsan, M.-V. and Dumitrescu, A., 2014: ROCADA: Romanian daily gridded climatic dataset (1961–2013) V1.0. Administratia Nationala de Meteorologie, Bucuresti, Romania.
<https://doi.org/10.1594/PANGAEA.833627>
- Birsan, M.-V., Dumitrescu, A., Micu, D. M., and Cheval, S., 2014: Changes in annual temperature extremes in the Carpathians since AD 1961. *Nat. Hazards*, 74, 1899–1910.
<https://doi.org/10.1007/s11069-014-1290-5>
- Chervenkov, H. and Slavov, K., 2019: STARDEX and ETCCDI Climate Indices Based on E-OBS and CARPATCLIM. Part Two: ClimData in Use, in G. Nikolov et al. (Eds.): NMA 2018, LNCS 11189, 368–374. <https://doi.org/10.1007/978-3-030-10692-8>
- Chervenkov, H. and Slavov, K., 2021: Modelled versus Satellite Retrieved Estimation of the Direct Normal Irradiance and the Sunshine Duration over Bulgaria, *Advanced Computing in Industrial Mathematics, Studies in Computational Intelligence*, vol. 961, I. Georgiev, K. Georgiev (eds), Springer (in press)
- Chervenkov, H. and Slavov, K., 2021: Solar Radiation Modelling for Bulgaria Based on Assimilated Surface Data, *Advanced Computing in Industrial Mathematics, Studies in Computational Intelligence*, vol. XXX, I. Georgiev, K. Georgiev (eds), Springer (in press)
- Chervenkov, H., Spiridonov, V., Artinyan, E., Neytchev, P., Slavov, K., and Stoyanova, M., 2017: The Operative System ProData — Part One: Current Stage and Recent Improvements. *Bul. J. Meteo Hydro* 22, 73–86.
- Cheval, S., Birsan, M.-V., and Dumitrescu, A., 2014: Climate variability in the Carpathian Mountains Region over 1961–2010. *Glob. Planet. Change* 118, 85–96.
<https://doi.org/10.1016/j.gloplacha.2014.04.005>
- Cornes, R., 2016: Assessment of the potential for enhancing the gridding resolution in parts of Europe, together with more comprehensive comparisons with NMHS derived gridded products, deliverable D1.11 of the UERRA project. <http://www.uerra.eu/publications/deliverable-reports.html>

- Cornes, R., van der Schrier, G., van den Besselaar, E.J.M. and Jones, P.D., 2018: An Ensemble Version of the E-OBS Temperature and Precipitation Datasets. *J. Geophys. Res. Atmos.*, 123. <https://doi.org/10.1029/2017JD028200>
- Dall'Amico, M. and Hornsteiner, M., 2006: Simple Method for Estimating Daily and Monthly Mean Temperatures from Daily Minima and Maxima. *Int. J. Climatol.* 26, 1929–1936 <https://doi.org/10.1002/joc.1363>
- Gadzhev G., Ivanov V., Valcheva R., Ganey K., and Chervenkov H. (2021) HPC Simulations of the Present and Projected Future Climate of the Balkan Region. In: Dimov I., Fidanova S. (eds) Advances in High Performance Computing. HPC 2019. Studies in Computational Intelligence, vol 902, 234-248 Springer, Cham. https://doi.org/10.1007/978-3-030-55347-0_20
- Harding, A. E., Rivington, M., Mineter, M. J., and Tett, S. F. B., 2015: Agrometeorological indices and climate model uncertainty over the UK. *Climat. Change* 128, 113–126. <https://doi.org/10.1007/s10584-014-1296-8>
- Huld, T. A., Šúri, M., Dunlop, E., and Fabio, M., 2006: Estimating average daytime and daily temperature profiles within Europe. *Environ. Modell. Software* 21, 1650–1661. <https://doi.org/10.1016/j.envsoft.2005.07.010>
- Ivanov, V. and Chervenkov, H., 2021: Modelling Human Biometeorological Conditions Using Meteorological Data from Reanalysis and Objective Analysis —Preliminary Results, *Advanced Computing in Industrial Mathematics, Studies in Computational Intelligence*, vol. 961, I. Georgiev, K. Georgiev (eds), Springer (in press)
- Kaiser-Weiss, A. K., Borsche, M., Niermann, D., Kaspar, F., Lussana, C., Isotta, F. A., van den Beselaar, E., van der Schrier, G., and Undén, P. Added value of regional reanalyses for climatological applications 2019: *Environ. Res. Commun.* 1, 071004 <https://doi.org/10.1088/2515-7620/ab2ec3>
- Lakatos, M., Bihari, Z., Szentimrey, T., Szalai, S. and the CARPATCLIM project Team, 2013a: Climate of the Carpathian Region - summary of the CarpatClim project. 3th EMS Annual Meeting & 11th European Conference on Applications of Meteorology (ECAM) — 09–13 September 2013, Reading, United Kingdom. <http://presentations.copernicus.org/EMS2013-501presentation.pdf>
- Lakatos, M., Szentimrey, T., Bihari, Z., and Szalai, S. 2013b: Investigation of climate extremes in the Carpathian region on harmonized data. *Int. Scient. Conf. on Environmental Changes and Adaptation Strategies*, Sep., 2013.
- Ma, Y. and Guttorp, P., 2013: Estimating daily mean temperature from synoptic climate observations. *Int. J. Climatol.* 33, 1264–1269. <https://doi.org/10.1002/joc.3510>
- Niermann D., Borsche M., Kaiser-Weiss A. K., and Kaspar, F., 2019: Evaluating renewable energy relevant parameters of COSMO-REA6 by comparing against station observations, satellites and other reanalyses. *Met. Zeit.*, Vol. 28, No. 4, 347–360 <https://doi.org/10.1127/metz/2019/0945>
- Niermann, D., Borsche, M., van den Besselaar, E., Lussana, C., Isotta, F., Frei, C., Kaiser-Weiss, C., Cantarello, L, Tveito, O. E. van der Schrier, G., Cornes, R., de Vreede E., and Davie, J., 2017: Scientific report on assessment of regional analysis against independent data sets. deliverable D3.6 of the UERRA project, <http://www.uerra.eu/publications/deliverable-reports.html>
- Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T., 1986: Numerical Recipes: the Art of Scientific Computing. Cambridge University Press, Cambridge, UK. [https://doi.org/10.1016/S0003-2670\(00\)82860-3](https://doi.org/10.1016/S0003-2670(00)82860-3)
- Ridal, M., Schimanke, S., and Hopsch, S., 2018: Documentation of the RRA system: UERRA. deliverable D322 Lot1.1.1.2 in the scope of the Copernicus service C3S 322 Lot1, available via Copernicus.
- Sillmann, J. and Röckner, E., 2008: Indices for extreme events in projections of anthropogenic climate change *Climatic Change* 86, <https://doi.org/10.1007/s10584-007-9308-6>
- Spinoni, J., Szalai, S., Szentimrey, T., Lakatos, M., Bihari, Z., Nagy, A., Németh, A., Kovács, T., Mihic, D., Dacic, M., Petrovic, P., Kržič, A., Hiebl, J., Auer, I., Milkovic, J., Štěpánek, P., Zahradníček, P., Kilar, P., Limanowka, D., Pyrc, R., Cheval, S., Birsan, M.-V., Dumitrescu, A., Deak, G., Matei, M., Antolovic, I., Nejedlík, P., Štastný, P., Kajaba, P., Bochníček, O., Galo, D., Mikulová, K., Nabyvanets, Y., Skrynyk, O., Krakovska, S., Gnatiuk, N., Tolasz, R., Antofie, T.,

- and *Vogt, J.* 2015: Climate of the Carpathian Region in the period 1961–2010: climatologies and trends of 10 variables. *Int. J. Climatol.* 35,1322–1341. <https://doi.org/10.1002/joc.4059>
- Soci, C., Bazile, E., Besson, F., and Landelius, T.*, 2016: High-resolution precipitation re-analysis system for climatological purposes. *Tellus A:* 68, <https://doi.org/10.3402/tellusa.v68.29879>
- Stoer, J. and Bulirsch, R.*, 2002: Introduction to Numerical Analysis, Texts in Applied Mathematics, 3rd ed., Vol. 12, Springer, New York. <https://doi.org/10.1007/978-0-387-21738-3>
- Unden, P.*, 2018: UERRA: Final Report, Project No.: 607193, Period number: 3rd Ref: 607193 UERRA Final Report-13 20180319 112103 CET.pdf