ORIGINAL PAPER

Improving the operational forecasts of outdoor Universal Thermal Climate Index with post-processing

Danijela Kuzmanović[1] · Jana Banko[2] · Gregor Skok[1]

Abstract

The Universal Thermal Climate Index (UTCI) is a thermal comfort index that describes how the human body experiences ambient conditions. It has units of temperature and considers physiological aspects of the human body. It takes into account the effect of air temperature, humidity, wind, radiation, and clothes. It is increasingly used in many countries as a measure of thermal comfort for outdoor conditions, and its value is calculated as part of the operational meteorological forecast. At the same time, forecasts of outdoor UTCI tend to have a relatively large error caused by the error of meteorological forecasts. In Slovenia, there is a relatively dense network of meteorological stations. Crucially, at these stations, global solar radiation measurements are performed continuously, which makes estimating the actual value of the UTCI more accurate compared to the situation where no radiation measurements are available. We used seven years of measurements in hourly resolution from 42 stations to first verify the operational UTCI forecast for the first forecast day and, secondly, to try to improve the forecast via post-processing. We used two machine-learning methods, linear regression, and neural networks. Both methods have successfully reduced the error in the operational UTCI forecasts. Both methods reduced the daily mean error from about 2.6°C to almost zero, while the daily mean absolute error decreased from 5°C to 3°C for the neural network and 3.5°C for linear regression. Both methods, especially the neural network, also substantially reduced the dependence of the error on the time of the day.

Keywords UTCI forecasting · Thermal comfort · Verification · Post-processing

Introduction

The Universal Thermal Climate Index (UTCI) is a thermal comfort index that quantitatively describes how the human body experiences ambient conditions. It is based on an advanced thermo-physiological model (Pappenberger et al. 2015). It has units of temperature, and it takes into account the effect of air temperature, humidity, wind, radiation, and clothes (Bröde et al. 2012). One of its favorable properties is that it can be used to express thermal comfort throughout the entire range of thermal conditions (i.e., for hot and cold conditions, Zare et al. 2018).

Following the concept of an equivalent temperature, UTCI for a given combination of wind speed, radiation, humidity, and air temperature was defined as the air temperature of the reference environment, which, according to the thermo-physiological model, produces an equivalent dynamic physiological response (Bröde et al. 2012). The reference conditions are 10 m wind of 0.5 m/s, relative humidity of 50%, mean radiant temperature equal to the air temperature, and a person walking at a speed of 4 km/h (Bröde et al. 2012). It is based on a relatively complex Fiala multi-node model of human heat transfer and temperature regulation that can reproduce the human thermal response to a wide range of external climatic conditions (Fiala et al. 2012).

Due to its strengths as a bioclimatic index, the UTCI has been widely used in many studies in bioclimatology and in many related scientific disciplines (Błażejczyk and Kuchcik 2021). For example, assessment of regional and local bioclimate characteristics (Wu et al. 2010; Błażejczyk and Matzarakis 2007; Kingma et al. 2021; Folkerts et al. 2021; Eggeling et al. 2022), urban bioclimate (Czarnecka

✉ Gregor Skok
  Gregor.Skok@fmf.uni-lj.si

1 Faculty of Mathematics and Physics, University of Ljubljana, Jadranska Cesta 19, Ljubljana 1000, Slovenia

2 Slovenian Environment Agency, Vojkova 1b, Ljubljana 1000, Slovenia

et al. 2011; Bröde et al. 2013; Nowosad and Wereski 2013; Błażejczyk et al. 2014; Lukić and Milovanović 2020), recreation, tourism and sports (Liu et al. 2022; Krüger 2017; Lindner-Cendrowska and Błażejczyk 2018), epidemiology and health research (Nastos and Matzarakis 2011; Morabito et al. 2014; Urban and Kyselý 2014; Krzyżewska et al. 2017; Romaszko et al. 2017; Błażejczyk et al. 2018; Di Napoli et al. 2018; Skutecki et al. 2019; Urban et al 2021), *UTCI* mapping (Vinogradova 2019; Kuchcik et al. 2021; Milewski 2013; Di Napoli et al. 2018), assessment/forecast of bioclimatic changes (Rozbicki and Rozbicka 2018; Kuchcik et al. 2021; Głogowski et al. 2020; Brecht et al. 2020; Błażejczyk et al 2013; Emerton et al. 2022).

The UTCI is also increasingly used in many countries as a measure of thermal comfort for outdoor conditions and is calculated as part of the operational meteorological forecast (Di Napoli et al. 2021a). Some examples of such use of the UTCI in Europe are: the Czech Republic (since 2019, Novak 2013; Novák 2021), Italy (since 2007, Maracchi 2017), Poland (since 2010, Bröde et al. 2012; IMGW-PIB 2015), Portugal (since 2010, IPMA 2022), and Slovenia (since 2019, ARSO 2021).

At the same time, forecasts of outdoor UTCI tend to have a relatively large error primarily caused by the error of forecasts of relevant meteorological parameters that are used as input for the UTCI (e.g., Novák 2021; Pappenberger et al. 2015). Our goal was to analyze the errors of the operational outdoor UTCI forecasts for the first day and then try to reduce these errors via post-processing by using two machine-learning approaches. We used a relatively simple method (linear regression) and a more complex non-linear method (neural network). We tried to improve the forecasts for Slovenia, which has a relatively dense network of meteorological stations where the forecasted and observed UTCI values could be compared. We used hourly data for seven years (2013–2018 and 2020) at 42 stations.

Data and methods

Universal thermal climate index

As already mentioned, the UTCI is defined as an equivalent air temperature that would produce the same physiological response under a set of reference conditions (Bröde et al. 2012). The UTCI is a function of several environmental parameters and can be defined as follows:

$$UTCI(T_a, T_{mrt}, e, v_{10m}) = T_a + \text{Offset}(T_a, T_{mrt}, e, v_{10m}),$$
(1)

where T_a is air temperature, e is water vapour pressure, and v_{10m} is wind speed at 10 m height (Bröde et al. 2012). The function Offset represents the deviation of the UTCI from the actual air temperature. Lastly, the mean radiant temperature (T_{mrt}) is a measure of the total radiation from the atmosphere and the ground incident on an object from all directions, but instead of expressing this measure as a flux density, it is converted into a temperature via the Stefan-Boltzmann equation (Di Napoli et al. 2020).

The UTCI is based on the Fiala multi-node model of human thermoregulation (Fiala et al. 2012). However, calculating the UTCI by running the complete Fiala model is computationally expensive, and simpler calculation procedures, such as look-up tables (such as the table presented by Bröde et al. 2012) or polynomial regression methods, are usually used. For example, an approximation of the Offset function by a 6th-degree polynomial that depends on T_a, v_{10m}, e, and $T_{mrt} - T_a$, can be used (such an example, with 210 polynomial coefficients, was presented by Bröde et al. 2012). Such approximation is used by the BioKlima model and the ALADIN numerical weather prediction model.

Data description

Measurements at meteorological stations

Slovenia is situated at the intersection of four major European geographic regions (the Alps, the Dinarides, the Pannonian Plain, and the Mediterranean). It has very complex terrain with altitudes ranging from 0 to about 3000 m as well as multiple climate types (Kozjek et al. 2017). The meteorological data used in this study was collected on 42 automatic stations, which provide data on an hourly resolution. These stations are part of the Slovenian Environment Agency (SEA) measurement network, and their spatial distribution can be seen in Fig. 1. A list of all stations with some information is shown in Table S6 in the Supplementary Materials. Seven years of data (2013–2018 and 2020 - this period was determined by the availability of the ALADIN model data) from these stations were used in the study.

The SEA employs an automated quality control algorithm that checks for probable errors or inconsistencies in the measurements. For example, it flags values where unusually large changes occurred in a short time or values that are substantially different from values at nearby stations. The values flagged by the automated algorithm are then manually checked by a quality control expert who decides if a particular measurement should be kept in the archive or discarded. The missing values were not used in the analysis (no gap filling or interpolation was performed if a measurement was missing).
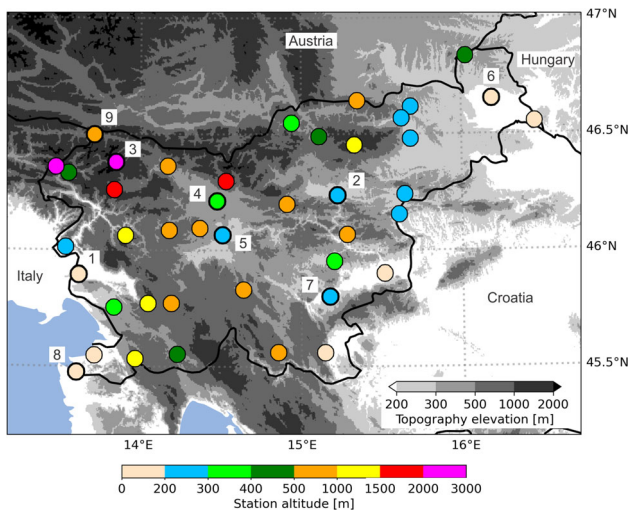
Fig. 1 The geographical distribution of meteorological stations used in the study. The circular markers are color-coded to reflect the altitude of each station, with altitudes indicated by the color bar below the figure. The nine stations that were individually analyzed are marked with bold circles and id numbers: 1. Bilje, 2. Celje-Medlog, 3. Kredarica, 4. Letalisce Jožeta Pučnika Ljubljana, 5. Ljubljana-Bežigrad, 6. Murska Sobota-Rakičan, 7. Novo Mesto, 8. Portorož-Letalisce, 9. Rateče. The gray shading represents the topography elevation from the ETOPO Global Relief Model (NOAA National Centers for Environmental Information 2022)

Since the stations span a large range of altitudes and multiple climate types, they represent a good dataset for evaluating the performance of the UTCI forecast as well as their post-processing. Crucially, measurements of global solar radiation are also performed at these stations. Having radiation measurements makes estimating the actual value of UTCI more reliable since T_{mrt} can be determined with more accuracy compared to the situation when no radiation measurements are available (in this case, other less accurate proxy parameters, need to be used).

For example, Novák (2021) compared the ALADIN-based UTCI forecasts in the Czech Republic to data from a single automatic weather station for a period of eleven months. Since no radiation measurements were available at the station, they used two proxies for T_{mrt} – they assumed that T_{mrt} is equal either to the air temperatures measured at 2 m or to the temperature measured at 5 cm above the ground. Another example is the study by Di Napoli et al. (2021b) where the authors tried to verify the T_{mrt} and UTCI values derived from ECMWF ERA5-HEAT reanalysis by comparing them to the values they obtained from 177 meteorological stations distributed around the world. Since measurements of radiation were not available at these stations, they had to rely on other parameters, such as air temperature and total cloud cover, to estimate the values of T_{mrt} and UTCI.

In addition to analyzing the combined data from all 42 stations, we also selected nine stations (marked with numbers

in Fig. 1), representing distinct geographical and climatic characteristics in Slovenia, to analyze how the UTCI forecast post-processing performs if done separately for each station (as opposed to all the stations together).

From the centrally located Ljubljana Basin, which exhibits a Subcontinental climate (Kozjek et al. 2017), we selected the Ljubljana-Bežigrad station in the capital city and the Jože Pučnik Airport Ljubljana station to represent both urban and countryside locations. From the NW part of Slovenia, we selected the Kredarica station, which has the highest altitude and an Alpine climate, and the Rateče station, located at the bottom of an alpine valley at a much lower altitude and exhibiting a Subalpine climate (Kozjek et al. 2017). From the SW part, which exhibits a Submediterranean climate (Kozjek et al. 2017), we selected the Portorož-Letališče station at the coast and the Bilje station located more inland. From the Celje Basin, we selected the Celje-Medlog station, while from the SE part of Slovenia, we selected the Novo Mesto station, with both stations exhibiting a Subcontinental climate (Kozjek et al. 2017). From the NE part of Slovenia, which also exhibits a Subcontinental climate (Kozjek et al. 2017), we selected the Murska Sobota-Rakičan station.

The software BioKlima 2.6 (Błażejczyk 2017) was used to calculate the UTCI values from the meteorological measurements. BioKlima is a user-friendly software for the Windows operating system that offers various methods for bioclimatic studies and can calculate around 60 different biometeorological and thermophysiological indices. It has been used successfully in previous studies for UTCI calculations (Liu et al. 2022; Rozbicki and Rozbicka 2018; Zare et al. 2018).

As T_{mrt} is not directly measured in Slovenia, BioKlima calculated it based on other parameters such as sun altitude and measurements of global solar radiation and temperature (Wu et al. 2010; Błażejczyk and Matzarakis 2007; Błażejczyk 2005). The Sun's altitude can be directly inputted or calculated by BioKlima from parameters such as month, day, hour, minutes, and latitude.

The input data for UTCI calculations in the BioKlima model consisted of month, day, hour, minutes, latitude, and measurements of air temperature, relative humidity, wind speed at 10 m height, and global solar radiation.

ALADIN model forecasts

The ALADIN model (Aire Limitée Adaptation Dynamique Développement International, Termonia et al. 2018) is a numerical weather prediction model that has been used operationally in Slovenia since 1997 by SEA. The model configuration includes a horizontal resolution of 4.4 km, 87 terrain-following sigma vertical levels, a domain with 421×421 grid points, and a time step of 180 s (Wang et al. 2018). The extent of the model domain is shown in Fig. S4 in the Supplementary Materials.

The UTCI calculation has recently been incorporated into the ALADIN using the 6th-degree polynomial approximation mentioned in "Universal thermal climate index". The forecasted value of UTCI is calculated in the ALADIN in a post-processing manner with the forecasted values of some of the model's basic variables used as input. The mean radiant temperature (T_{mrt}) is a key component in the calculation of UTCI, and in the ALADIN model, it is obtained by taking into account various shortwave and longwave fluxes (for details on how the T_{mrt} is calculated in ALADIN please refer to Weihs et al. 2011).

The SEA provided the data for the ALADIN-based forecasts. Seven years (2013–2018, and 2020) of archived operational model outputs were used for the analysis (the forecast data for 2019 was not readily available in the SEA archive, and providing it would demand a significant investment of both time and resources). The data was provided on a spatial resolution of 4.4 km and a temporal resolution of one hour. The data were provided for the first 24 hours of the forecasts, initialized at 00 UTC. The ALADIN forecast is integrated for 72 hours, but only the outputs for the first 24 hours are archived by SEA and could thus be used in the analysis. The model data from the grid point that was nearest to the location of the station was used for the analysis.

Methods

Verification metrics

The mean error (ME) and mean absolute error (MAE) were used as criteria to evaluate the model performance. ME is the difference between the average forecast and average observation and expresses the bias of the forecasts (Wilks 2006). It is defined as

$$ME = \frac{1}{N}\sum_{i=0}^{n}(f_i - o_i) \ , \tag{2}$$

where f_i and o_i and the forecasted and observed values, respectively, and N the number of data. ME closer to zero indicates a better forecast (von Storch and Zwiers 1999).

The MAE is the arithmetic average of the absolute values of the differences between the pairs of forecasted and observed values and can be interpreted as a typical magnitude for the forecast error in a given verification data set (Wilks 2006). It is defined as

$$MAE = \frac{1}{N}\sum_{i=1}^{n}|f_i - o_i| \ , \tag{3}$$

A lower MAE indicates a better forecast (von Storch and Zwiers 1999).

Neural network

Modern machine learning techniques, especially Neural Networks (NNs), are increasingly used to improve specific aspects of weather prediction (e.g. Reichstein et al. 2019; Palmer 2020; Schultz et al. 2021). Due to their strengths and successful use by researchers in many previous studies, we decided to use the NN to try to improve the accuracy of the ALADIN-based UTCI forecasts.

We used various parameters outputted by ALADIN (e.g., forecasted UTCI, air temperature, humidity, wind speed, cloudiness), as well as some other parameters (e.g., station altitude) as input variables for the NN. The output from the NN was a corrected value of UTCI forecast. To train the NN, we also used the UTCI values derived from the measurements at the stations.

To prepare the data for training the NN and using it for Linear Regression (LR, Goodfellow et al. 2016), we normalized the input and output variables to a range between 0 and 1 using the MinMaxScaler module from the Python scikit-learn library (Pedregosa et al. 2011). This normalization process ensures that all variables have a similar scale, which helps with the training of NN. Next, we randomly shuffled the data to remove any temporal sequences and divided it into training (90%) and test (10%) sets. The training set was used to train the NN (and LR), while the test set was used to evaluate its performance once the training had finished.

We used fully-connected feedforward NNs (Goodfellow et al. 2016) that consisted of several layers of neurons that are fully connected – namely, all the neurons in one layer are connected with all the neurons in the previous layer. The values of neurons in each layer were thus determined from all the values in the previous layer, while the input parameters determined the values of the neurons in the input layer. We implemented the NNs in Python with the TensorFlow library (Abadi et al. 2016)

The architecture of the NN consisted of an input layer, three hidden layers, and an output layer. The number of neurons in the input layer varied based on the number of input parameters, while the hidden layers had 5, 3, and 1 neurons, respectively. The output layer had only one neuron, whose value determined the forecasted UTCI value. We also experimented with more complex NN architectures consisting of many more neurons arranged in more layers but found that more complex NNs did not perform substantially better. We used the LeakyReLU activation function (Maas 2013) for the hidden layers and a linear activation function for the output layer. The Adam optimizer (Kingma and Ba 2014) and the MAE loss function were employed during training. The NN was trained for 100 epochs (testing showed that using more epochs did not improve the results).

We evaluated the performance of the NN and LR by calculating the ME and MAE for the test set. We trained the

NN and LR using data from all the stations together (in this case, we used a batch size of 1000 to speed up the training process). Additionally, we trained the NN and LR separately for the nine selected stations to assess their performance on a per-station basis with a batch size of 100.

To evaluate the performance of various approaches, we compared the forecasted UTCI values with the observed UTCI values. This comparison was made for the independent test set (which was not used for training the NN or LR) and was made separately for the NN-based forecasts, the LR-based forecasts, and the non-post-processed ALADIN-based forecasts of UTCI.

Linear regression

NNs represent a relatively advanced and complex machine-learning approach. We also wanted to assess whether employing such a complex approach yields substantial advantages over a simpler approach. This is why we also used Linear Regression (LR), which is one of the simplest methods that can be used to explain a relationship between a chosen dependent variable and one or more explanatory variables. In LR, the dependent variable can be expressed as a linear combination of explanatory variables as

$$Y = b_0 + b_1 X_1 + b_2 X_2 + ... + b_k X_k, \tag{4}$$

where Y is the dependent variable while $X_1, ..., X_k$ are the explanatory variables (Darlington and Hayes 2017). In our case, the explanatory variables are the same normalized input variables that were also used with the NNs. b_0 is the regression constant, and $b_1, ..., b_k$ are regression coefficients

(Darlington and Hayes 2017) that are determined using the least squares method that minimizes the sum of the squared differences between the expected and predicted outputs (Darlington and Hayes 2017).

The sign and size of a regression coefficient determine how a particular explanatory variable influences the dependent variable. For example, if the sign of the coefficient is positive, the dependent variable value will increase when the explanatory variable value increases, and vice versa. Moreover, if the coefficient's absolute value is larger, the influence on the dependent variable will also be larger, which can be used to determine which explanatory variables will tend to have a larger influence on the dependent variable.

We used the Python scikit-learn library (Pedregosa et al. 2011) to implement the LR, and we used the same input/output variables as with the NN.

Results

Verification of ALADIN-based UTCI forecasts

The first goal of this study was to verify operational ALADIN-based UTCI forecasts for Slovenia. Figure 2a shows the ME and MAE with respect to the time of the day. The solid line represents ME, while the dashed line represents MAE. The daily average ME of the UTCI forecasts is 2.56 °C, but its value changes a lot with respect to the time of the day. For example, at 7 am, the ME is almost 8 °C while it is close to zero in the evening. The average MAE is 5.02 °C. MAE also has a peak at 7 am, exceeding 8 °C, while it is about
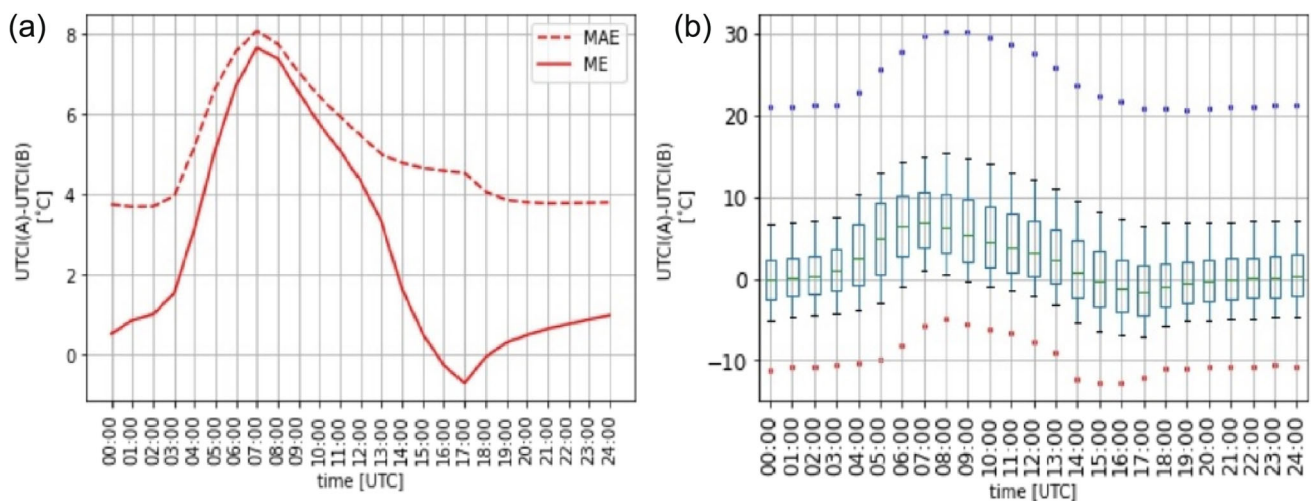

Fig. 2 Verification of ALADIN-based UTCI forecasts for all stations in the seven years that were analyzed (2013–2018 and 2020). (a) ME and MAE for each hour of the day. (b) The distribution of differences between the forecasted and observed UTCI values. The median is represented by the green line within the box, which represents the values between the 25th and 75th percentiles. The whiskers show the range from the 10th to 25th percentile and from the 75th to 90th percentile. The red and blue dots represent the 1st and 99th percentile, respectively

4 °C in the evening. Figure 2b shows the distribution of differences between the forecasted and observed UTCI values. We can note that the median follows the ME course, with maximal value 6.9 °C reached at 07:00 and minimal value -1.5 °C obtained at 17:00. In addition, we can observe that the interquartile range is larger during the daytime, with a maximum value of 8.4 °C reached at 5:00 and a minimum value of 4 °C obtained at midnight. The rather large distance between the 1st and 99th percentile (red and blue dots) of about 35 °C indicates issues with ALADIN forecasts. For example, in about 2% of cases, ALADIN overestimates the UTCI value for more than 20 °C, while in about 1% of cases, ALADIN underestimates the UTCI value for more than 10 °C.

The results in Fig. 2 can be compared to those from Novák (2021) who tried verifying the ALADIN-based UTCI forecasts in the Czech Republic. It is interesting to note that they got quite different results. For example, they found that the ALADIN model tended to underestimate the UTCI values

and that the largest error was during the night, while the error was smaller during the day and the smallest in the mornings and afternoons. It should also be stressed that the ALADIN model used in Slovenia (by SEA) is not identical to the ALADIN model used in the Czech Republic (by the Czech Hydrometeorological Institute) as the different branches of the original ALADIN model have been developed separately for quite a long time. Nevertheless, it is important to highlight the differences between the two studies. Novák (2021) used eleven months of data from a single station, while we used seven years of data from 42 stations. They also did not have any radiation measurements available at their station and used temperatures measured at heights of 2 m and 5 cm as proxies for T_{mrt}, while on our stations, measurements of global solar radiation were available.

Figure 3 shows the results for the nine selected stations. Kredarica stands out with the highest median with a maximal median of 18.57 °C at 7:00 and a wider spread, where the
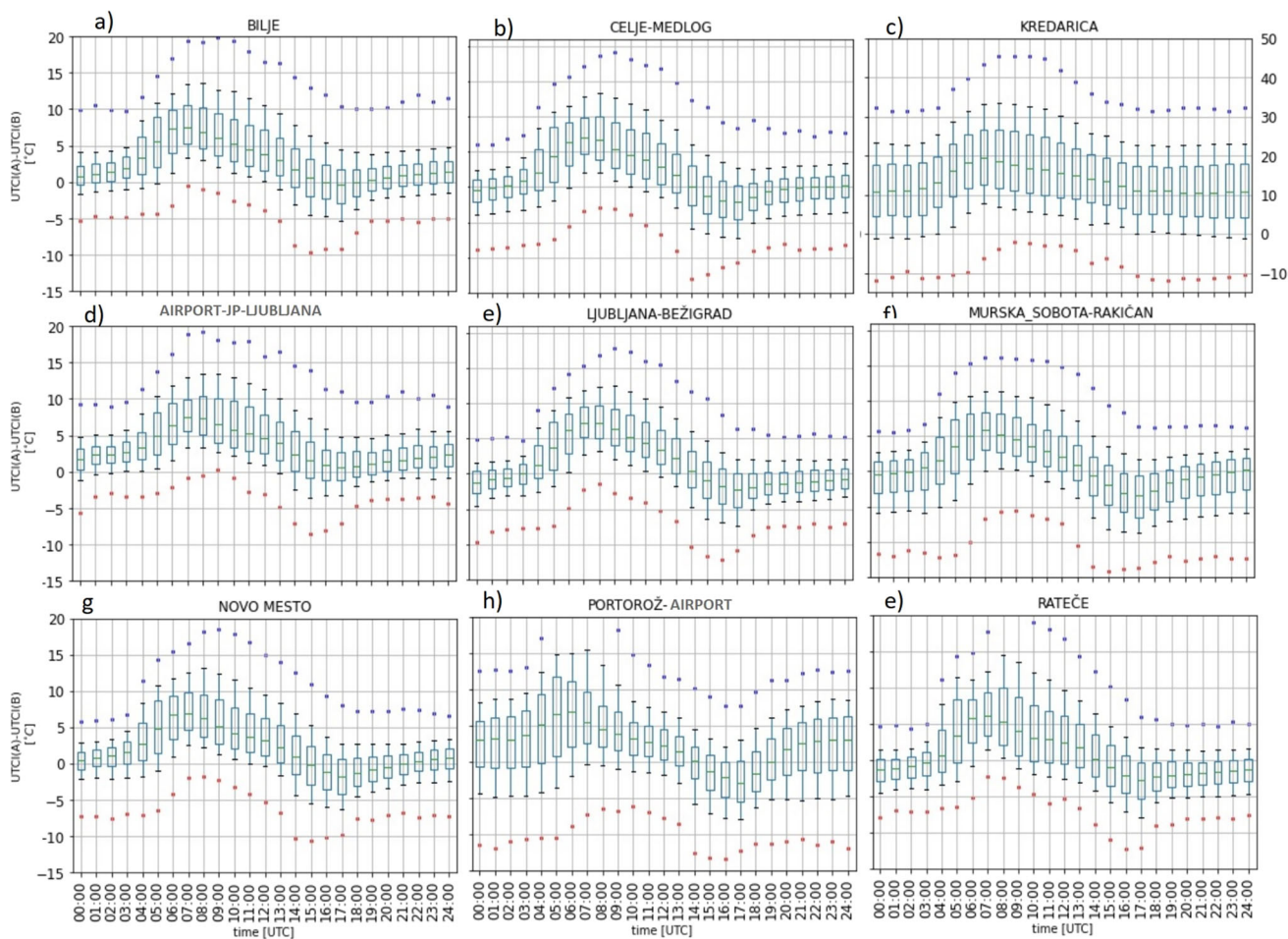

Fig. 3 Same as Fig. 2b, but for selected nine stations shown in bold font in Fig. 1. The y-axis of each subfigure has the same range for all stations from -15 °C to 20 °C, except for the Kredarica station, which has a range from -10 °C to 50 °C

maximal interquartile range was 15,85 °C at 10:00. In overall, we can note that its mean errors tend to be about 10 °C larger than at other stations. Kredarica is a high-altitude station located at an altitude of about 2500 m. The errors are likely caused by the difference between the station's altitude in the model and the actual altitude of the station. Portorož-Airport also shows a wider interquartile range in the first hours, with a maximal value of 10.86 °C reached at 5:00. This wide scatter of the errors is possibly influenced by wind forecasts or measurement errors near the sea.

The impact of station altitude on the error is evident in the box plots. To verify this, we categorized stations into two groups based on altitude: those below or equal to 509 m and those above. The results are shown in Fig. S1 in the Supplementary Materials. For both groups, the ME and MAE errors display a similar dependence on the time of the day (with the largest errors in the early morning), but the errors for the low altitude group tend to be 3 to 6 °C lower than the errors of the high altitude group.

An investigation of several cases with exceptionally large errors (two such examples are shown in the Supplementary Materials in Section S2) identified several possible causes for these differences:

- Incorrectly forecasted cloud cover, humidity, and surface radiation balance: These factors can have an impact on the calculation of T_{mrt}, particularly during the morning hours when the maximum error was observed. Based on the analysis we performed, it was observed that errors in T_{mrt}, and consequently in UTCI, tended to be the largest during periods when there were rapid changes in these values, such as during the transition from day to night.
- Incorrectly forecasted wind speed or temperature: Inaccurate forecasts of wind speed or temperature can also contribute to errors in the UTCI prediction.
- Large difference between the station's altitude in the model and the actual altitude of the station (the station altitudes are shown in Table S6 in the Supplementary Materials). This disparity in altitude can substantially impact the accuracy of radiation forecasts, as the actual altitude of the stations may deviate from the model's predictions.
- Specific micro-meteorological conditions at the locations of meteorological stations: Factors such as relief variability at a scale smaller than the spatial resolution can influence the accuracy of forecasts. For instance, stations located within urban heat islands may experience differences between measured meteorological parameters and the forecasted values, unlike stations in rural areas with more natural surroundings and no obstacles (although we did not observe a substantial urban heat islands effect since the station at Ljubljana-Bežigrad, which is situated in an urban environment, had one of the smallest errors).

Proximity to hills or the sea can also have a notable effect on the model's ability to predict the UTCI.
- Possible errors in measurements: It should be noted that the station data from 2014 onward has been quality controlled; however, there is still a possibility of errors being present in the measurements that could impact the verification of the forecasts.
- Approximations in determining solar elevation in the BioKlima model: The method used to determine solar elevation in BioKlima may involve approximations, which could influence the results.
- Lack of detailed explanations of calculations in BioKlima: The lack of comprehensive explanations regarding the calculations in BioKlima may pose challenges in understanding the accuracy and reliability of the verification.

We also investigated the seasonal dependence of errors. The results are shown in Fig. S2 in the Supplementary Materials. All four seasons exhibit the same hourly pattern of errors, with the error being the largest in the early morning hours. However, the exact timing, as well as the amplitude of the error maximum, varies across seasons.

In the summer, the largest ME of 8.34 °C was at 5:00, while in the autumn, the largest ME of 8.18 °C was at 7:00. For the winter, we observed the largest ME of 10.3 °C at 8:00, which was the largest ME of all seasons. During the spring, the ME peaked with 8.9 °C at 6:00. The timing of the error peak is probably linked to the time of sunrise, which comes later in the winter than it does in the summer – this is why the error peak in winter happens a few hours later than in the summer. A similar conclusion can also be reached for the MAE error. Generally, both errors tend to be the largest in winter and the smallest in summer. This result is somewhat surprising as we had anticipated larger errors during the summer when the model might encounter challenges in forecasting cloudiness due to convection. Previous analyses revealed the sensitivity of the UTCI to wind speed values, suggesting that the higher errors during winter may be attributed to differences in forecasted and measured wind speeds. The model might also have problems successfully forecasting a relatively thin radiation fog near the surface, which frequently develops in many regions in Slovenia during the night, especially in winter. Such fog affects how fast the temperature increases in the morning and also affects the amount of radiation that reaches the ground. Many regions in Slovenia are also affected by downslope mountain winds and cold-air pools, which frequently form during the night in complex terrain. Since these phenomena can be small and relatively shallow, they might not be forecasted correctly by the model, which could result in the temperature increasing too fast in the forecast in the morning.

Post-processing of UTCI forecasts

Post-processing the combined data from all the stations

The second goal of this study was to try to improve the ALADIN-based operational UTCI forecasts by using post-processing with LR and NN.

First, we wanted to try to obtain a single NN and LR model that could be used for post-processing on all the stations. Thus, we used the combined data from all 42 stations to train the NN and obtain the optimal LR coefficients.

We first tried to identify the relationship between different meteorological parameters and UTCI values by calculating correlation coefficients (not shown). Based on these findings, we prepared six different setups with different sets of input parameters that were used as predictors by the LR and NN (a detailed description of the setups is provided in Table S1 in the Supplementary Materials). Since the initial values of the weights in the NN are determined randomly, each training realization of the NN can produce a somewhat different NN. This is why we performed the NN training three times for each setup and selected the training realization with the smallest MAE error.

The performance of LR and NN for all six setups is shown in Fig. 4. This visual representation provides insights into the effectiveness of the post-processing methods for improving the accuracy of ALADIN-based UTCI forecasts. Numerical values of daily mean ME and MAE errors for each setup are given in the Supplementary Materials in Table S2.

From the graphs in Fig. 4, it is evident that both methods successfully improve the ALADIN forecast of UTCI in all six setups, resulting in smaller ME and MAE errors. For example, the ME and MAE errors of the uncorrected forecasts for the test set were 2.57 °C and 5.02 °C, respectively. On the other hand, the ME for the LR was zero for all six setups and the MAE was between 3.47 °C and 4.15 °C, while the ME for the NN was between 0.15 °C and 1.18 °C and the MAE between 3.03 °C and 3.77 °C.

Both methods, especially the NN, also substantially reduce the dependence of the error on the time of the day. For the NN, the forecasts of UTCI in the morning tend to be as accurate as those in the afternoon or night. The same could be said for the MAE error of LR, while its ME does exhibit some hourly dependence with overestimation in the first half of the day and underestimation in the second half.

According to the daily average values of ME and MAE (shown in the Supplementary Materials in Table S2), the NN method outperforms the LR method in terms of MAE, while the LR method tends to have a bit better ME. However, even though the daily average ME for the LR method is almost zero for all setups, LR exhibits a pronounced dependence of ME on the time of the day, which is not present in the case of NN.

The best results for both methods were obtained with Setup 5, which uses the following input parameters: ALADIN UTCI forecast, the hour of the day, forecasted air temperature, forecasted relative humidity, forecasted wind speed, forecasted T_{mrt}, station altitude, and the altitude of the station in the model. In this case, the daily average ME and MAE errors for LR were 0.00 °C and 3.47 °C, respectively, and for NN they were 0.15 °C and 3.03 °C, while the errors for the uncorrected ALADIN forecasts were 2.57 °C and 5.02 °C. This means that both post-processing methods reduced the ME from about 2.6 °C to almost zero, while the mean absolute error decreased from 5 °C to 3 °C for the NN and 3.5 °C for the LR. Interestingly, including station altitude and model altitude as input parameters yielded better results than using only the difference between these two quantities (this is the only difference between setups 1 and 5).

Analysis of the LR coefficients (shown in the Supplementary Materials in Table S3) revealed that the parameters that have the largest influence on the resulting UTCI value in the LR are the uncorrected ALADIN forecast of UTCI and forecasted air temperature. The coefficients for forecasted T_{mrt} are negative, with values around -0.15. The coefficients for forecasts of cloudiness, relative humidity, wind speed, and the hour of the day are all close to zero. The same is true for the difference between the model and station altitude, while the coefficients for the station and model altitudes tend to be around -0.25 and 0.15, respectively, indicating that the LR uses them.

Figure 5 shows the distribution of differences between the forecasted and observed UTCI values for Setup 5. In addition to moving the median value of the differences closer to zero, both LR and NN also manage to reduce the spread of differences. For example, the mean distance between the 25th and 75th percentile reduces from 6.1 °C to 5.0 °C (LR) and 4.2 °C (NN), the distance between the 10th and 90th percentile reduces from 13.4 °C to 10.5 °C (LR) and 9.2 °C (NN), while the distance between the 1st and 99th percentile reduces from 33.5 °C to 25.3 °C (LR) and 23.8 °C (NN).

Post-processing the data from individual stations

In "Post-processing the combined data from all the stations", we tried to obtain a single NN and LR model that would work well for all 42 stations. Here, we tried a different approach by training/optimizing the NN and LR models on data from each of the nine selected stations separately, thereby obtaining a different NN and LR model for each station. The goal was to see how well the post-processing methods perform when optimized and used on data from individual stations.

We used the same input parameters as in Setup 5 (which performed the best with combined data from all the stations) but excluded the station altitude and the station's altitude in the model since these are constant for a specific station. As
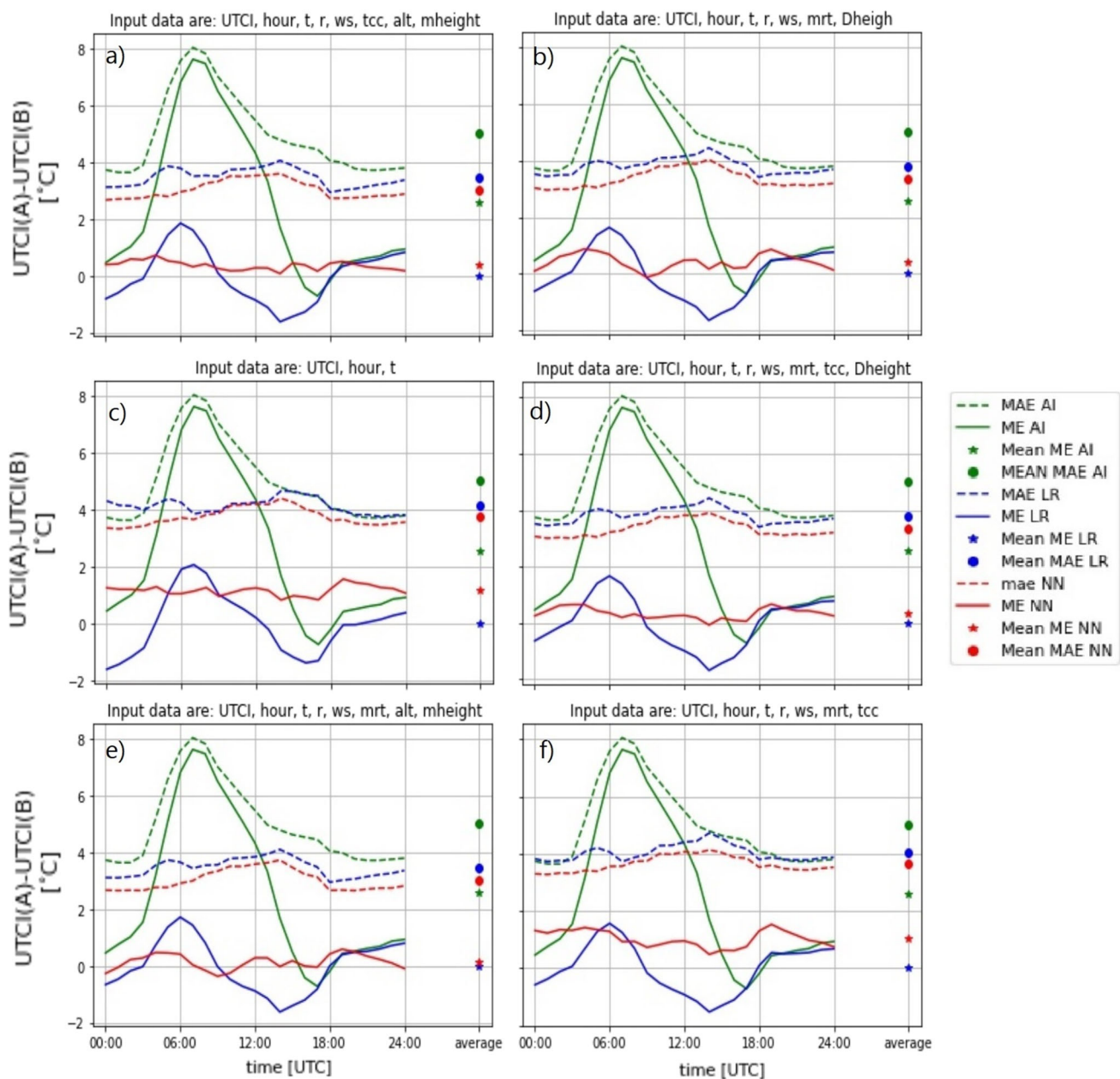
Fig. 4 The values of ME and MAE with respect to the hour of the day for all stations for the test set for the six different setups (see Table S1 in the Supplementary Materials for details). From (a) to (f), setups 1, 2, 3, 4, 5, and 6, respectively. The ME is depicted as a solid line and MAE as a dashed line. The results for LR are represented in blue, NN in red, while the results for the uncorrected ALADIN forecasts are shown in green color. The circles on the right side of each graph show the daily average ME and MAE values, which are also shown in Table S2 in the Supplementary Materials

before, we trained the NN for each station three times and selected the network with the lowest MAE. The results for all nine stations are depicted in Fig. 6, with the numerical values of daily average ME and MAE errors provided in the Supplementary Materials in Table S4.

From the graphs in Fig. 6, it is evident that both methods yield better results and smaller errors compared to uncorrected ALADIN forecasts for all nine selected stations. For example, for NN the ME reduction was between 0.65 °C and 12.68 °C, while the MAE reduction was between 1.57 °C and 7.27 °C. For LR the ME reduction was between 0.57 °C and 12.78 °C, while the MAE reduction was between 1.26 °C and 7.11 °C.

Similarly to the analysis done for all 42 stations, NN outperforms the LR in terms of MAE, while LR has a bit better daily average ME (0.00 °C vs. 0.16 °C, when averaged
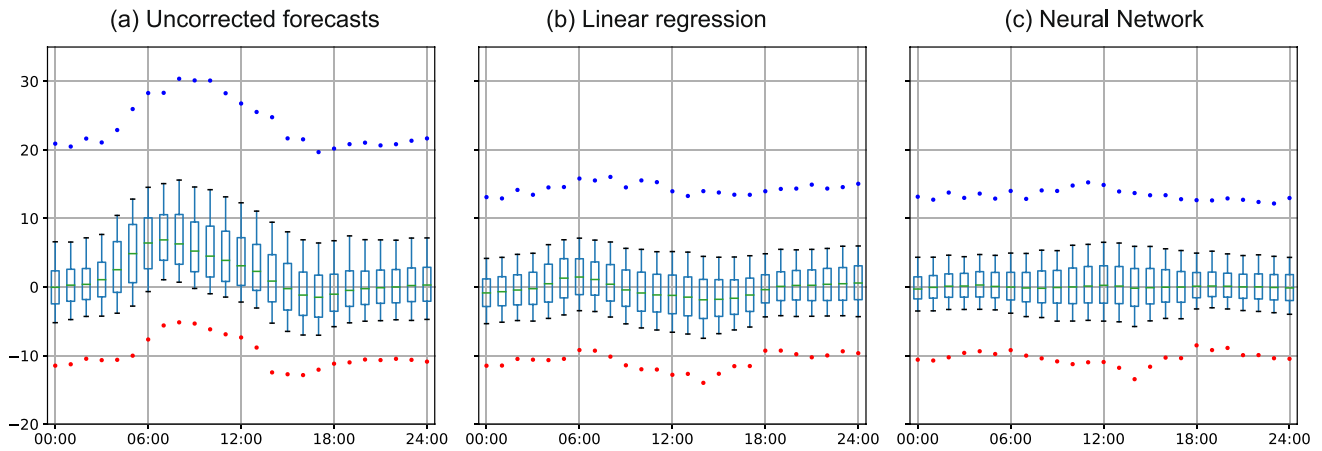
Fig. 5 The distribution of differences between the forecasted and observed UTCI values for the data from all the stations evaluated for the test set. Results for Setup 5 are shown. (a) the uncorrected forecasts. (b) the forecasts post-processed by LR. (c) the forecasts post-processed by NN

over the results for all six stations), although both methods have ME close to zero. Interestingly, in terms of MAE, both NN and LR substantially improve the forecasts in the morning hours (when the errors of the uncorrected forecasts are the largest), while during the afternoon and the night, the MAE improves only slightly. Consequently, after the post-processing, the largest MAE error for most stations is in the

afternoon and not in the morning, as is the case for uncorrected forecasts.

The station on Kredarica is a special case. As already mentioned, this is a high-altitude station located at an altitude of approximately 2500 m. The errors of the uncorrected forecasts are the largest at this station ($ME = 12.76$ °C, $MAE = 14.55$ °C), and while both NN and LR substan-
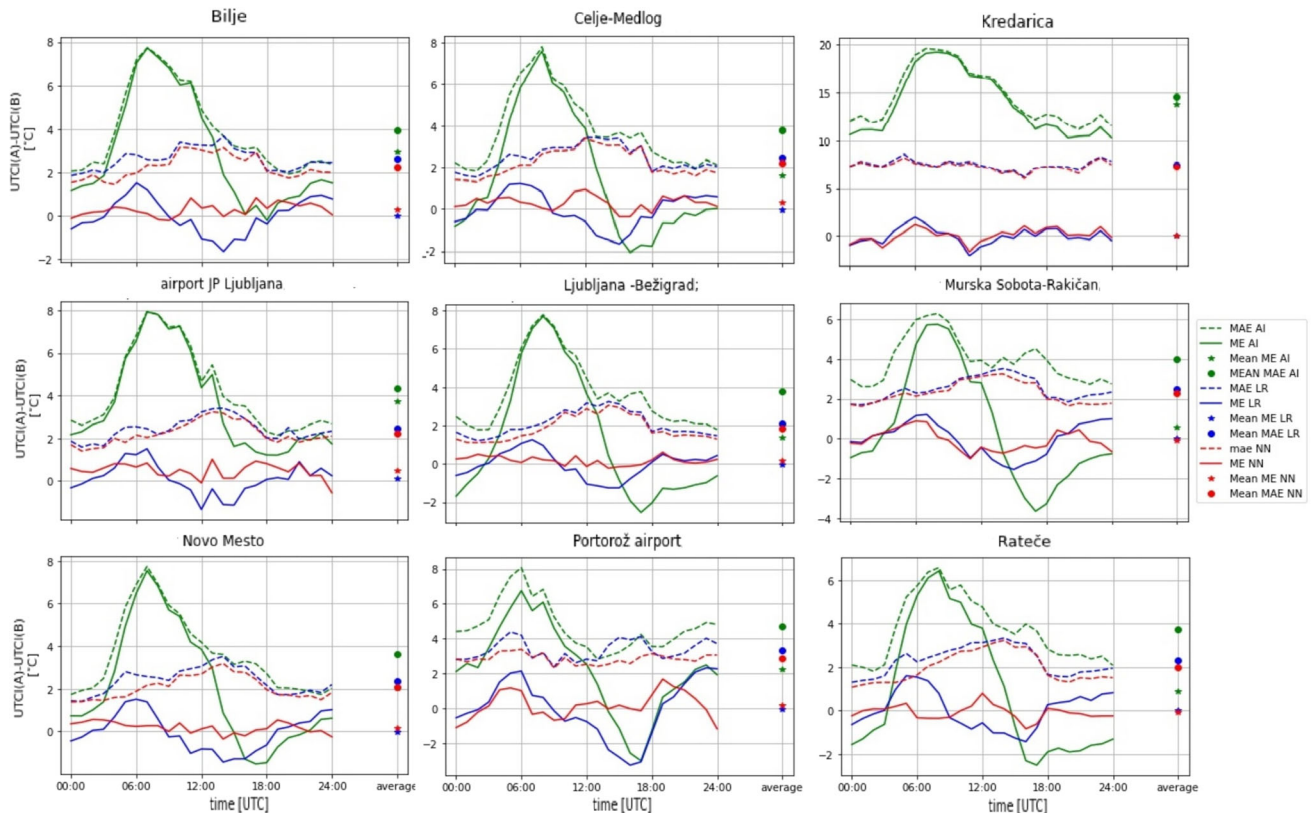

Fig. 6 Same as Fig. 4, but for the post-processing done separately on the data from each of the nine selected stations shown with bold font in Fig. 1. Setup 5 was used for all the stations

tially improve the forecasts, with ME being close to zero for both methods, the MAE error nevertheless remains large compared to the other stations (i.e., 7.28 °C and 7.44 °C, for the NN and LR, respectively). Contrary to other stations, the two methods successfully reduce the MAE errors for all hours of the day, and the error does not exhibit dependence on the time of the day.

On the other hand, out of the nine stations, the best results (lowest MAE) are observed for the Ljubljana-Bežigrad station, with MAE reduced from 3.75 °C to 1.86 °C for NN and 2.10 °C for LR.

Similarly to the analysis of the data from all stations, analysis of the size of LR coefficients (shown in the Supplementary Materials in Table S5) revealed that the parameters that have the largest influence on the resulting UTCI value in the LR, are the uncorrected ALADIN forecast of UTCI and the forecasted air temperature. The values of coefficients for other parameters tend to be close to zero.

We also analyzed the distribution of differences between the forecasted and observed UTCI values at three stations: Kredarica, Ljubljana-Bežigrad, and Portorož Airport. The results are shown in the Supplementary Materials in Fig. S3.

For the Kredarica station, which is a high-altitude station with the largest errors, the LR and NN managed to move the median value of the differences closer to zero but were unable to reduce the spread of differences. For example, in the uncorrected forecasts, the mean distances between the 25th-75th, 10th-90th, and 1st-99th percentiles were approximately 14 °C, 25 °C, and 44 °C, respectively, and LR and NN could not reduce these for more than 2 °C.

On the other hand, for the Ljubljana-Bežigrad and Portorož Airport stations, the post-processing methods managed to move the median value of the differences closer to zero and reduce the spread of differences. For example, at the Ljubljana-Bežigrad station, mean distances between the 25th-75th, 10th-90th, and 1st-99th percentiles were 4.1 °C, 8.2 °C, and 16.4 °C, respectively. The LR managed to reduce these to 3.2 °C, 6.2 °C, and 12.9 °C, and the NN to 2.7 °C, 5.8 °C, and 12.6 °C.

Discussion and conclusions

The goals of the study were to analyze errors in the operational forecasts of outdoor UTCI in Slovenia and to improve these forecasts by using post-processing.

The verification of the UTCI forecasts showed that the average ME and MAE errors were typically a few degrees Celsius (about 2.6 °C and 5 °C, respectively) but depended greatly on the time of the day and were the largest in the morning. There was also a prominent influence of the alti-tude – namely, the errors were markedly larger at the stations at higher altitudes. Moreover, the errors also exhibited a pronounced seasonal dependency, with the largest errors in winter and the smallest in summer. There were also some situations with very large errors (e.g., > 20 °C), which can be caused, for example, by incorrectly forecasted meteorological parameters that strongly influence the UTCI, like cloud cover, temperature, and wind.

Post-processing by NNs or LR substantially improved the accuracy of the UTCI forecasts. Both methods, especially the NNs, also substantially reduced the dependence of the error on the time of the day. Generally, NNs outperformed the LR method, but the difference between the two was not very large, which was surprising since we expected NNs to be substantially better than LR. The NNs we used were relatively simple, and we also tried to use more complex NNs, but they did not perform substantially better. It seems that the simple NNs already managed to correct the most obvious errors that frequently occur in the forecasts, while even the more complex NNs could not help much with the more complicated errors that might occur due to specific weather situations being substantially misdiagnosed in the forecasts.

Overall, the study provides important insights into the typical sizes and properties of errors of operational forecasts of outdoor UTCI and how these forecasts can be post-processed to increase accuracy. Although the study was based on data from a relatively small geographical region, the substantial number of stations used in the analysis, the relatively long analysis period of 7 years, the large climate variability and complex orography of Slovenia make the results more generally relevant. Moreover, the radiation measurements taken at the stations enhance the accuracy of estimating T_{mrt}, thereby making the estimation of the true UTCI and the forecast error more reliable. The two presented post-processing approaches show clear benefits by substantially increasing the accuracy of the UTCI forecasts and could also be used in other parts of the world. Their use could improve the accuracy of UTCI forecasts and thus help with early warning of extreme weather events related to thermal stress.

Declarations

Competing interests The authors declare no competing interests

References

Abadi M, Barham P, Chen J et al (2016) Tensorflow: A system for large-scale machine learning. In: 12th {USENIX} Symposium on operating systems design and implementation ({OSDI} 16), pp 265–283

ARSO (2021) Biovremenske vsebine - občutena temperatura. https://meteo.arso.gov.si/uploads/probase/www/sproduct/biomet/bulletin/sl/biovreme/. Accesed 17 Aug 2022

Błażejczyk K (2005) New indices to assess thermal risks outdoors. Environmental Ergonomics XI, Proc of the 11th International Conference. https://lucris.lub.lu.se/ws/portalfiles/portal/96090083/ProceedingsICEE2005.pdf#page=222

Błażejczyk K (2017) Bioklima - universal tool for bioclimatic and thermophysiological studies. https://www.igipz.pan.pl/Bioklima-zgik.html. Accesed 17 Aug 2022

Błażejczyk K, Kuchcik M (2021) UTCI applications in practice (methodological questions). Geogr Pol 94. https://doi.org/10.7163/GPol.0198

Błażejczyk A, Błażejczyk K, Baranowski J et al (2018) Heat stress mortality and desired adaptation responses of healthcare system in Poland. Int J Biometeorol 62. https://doi.org/10.1007/s00484-017-1423-0

Błażejczyk K, Idzikowska D, Błażejczyk A (2013) Forecast changes for heat and cold stress in Warsaw in the 21st century, and their possible influence on mortality risk. Papers on Global Change IGBP 20(1). https://doi.org/10.2478/igbp-2013-0002. http://journals.pan.pl/dlibra/publication/113805/edition/98863/content

Błażejczyk K, Kuchcik M, Błażejczyk A, et al (2014) Assessment of urban thermal stress by UTCI - experimental and modelling studies: an example from Poland. Erde 145. https://doi.org/10.12854/erde-145-3

Błażejczyk K, Matzarakis A (2007) Assessment of bioclimatic differentiation of poland based on the human heat balance. Geogr Pol 80

Brecht B, Schädler G, Schipper J (2020) Utci climatology and its future change in germany - an rcm ensemble approach. Meteorol Z 29. https://doi.org/10.1127/metz/2020/1010

Bröde P, Fiala D, Błażejczyk K et al (2012) Deriving the operational procedure for the Universal Thermal Climate Index (UTCI). Int J Biometeorol 56(3):481–494. https://doi.org/10.1007/s00484-011-0454-1. http://link.springer.com/10.1007/s00484-011-0454-1

Bröde P, Błażejczyk K, Fiala D et al (2013) The Universal Thermal Climate Index UTCI compared to ergonomics standards for assessing the thermal environment. Ind Health 51. https://doi.org/10.2486/indhealth.2012-0098

Czarnecka M, Mąkosza A, Nidzgorska-Lencewicz J, (2011) Variability of meteorological elements shaping biometeorological conditions in Szczecin, Poland. Theor Appl Climatol 104(1–2):101–110.

https://doi.org/10.1007/s00704-010-0326-3. http://link.springer.com/10.1007/s00704-010-0326-3

Darlington RB, Hayes AF (2017) Regression analysis and linear models: concepts, applications, and implementation. The Guilford Press, New York, USA

Di Napoli C, Pappenberger F, Cloke HL (2018) Assessing heat-related health risk in Europe via the Universal Thermal Climate Index (UTCI). Int J Biometeorol 62(7):1155–1165. https://doi.org/10.1007/s00484-018-1518-2. http://link.springer.com/10.1007/s00484-018-1518-2

Di Napoli C, Hogan RJ, Pappenberger F (2020) Mean radiant temperature from global-scale numerical weather prediction models. Int J Biometeorol 64(7):1233–1245. https://doi.org/10.1007/s00484-020-01900-5. http://link.springer.com/10.1007/s00484-020-01900-5

Di Napoli C, Messeri A, Novák M, et al (2021a) The universal thermal climate index as an operational forecasting tool of human biometeorological conditions in Europe. In: Applications of the universal thermal climate index UTCI in Biometeorology. Springer International Publishing, Cham, p 193–208. https://doi.org/10.1007/978-3-030-76716-7_10. https://link.springer.com/10.1007/978-3-030-76716-7_10

Di Napoli C, Barnard C, Prudhomme C et al (2021b) Era5-heat: a global gridded historical dataset of human thermal comfort indices from climate reanalysis. Geosci J 8. https://doi.org/10.1002/gdj3.102

Eggeling J, Rydenfält C, Kingma B et al (2022) The usability of ClimApp: a personalized thermal stress warning tool. Clim Serv 27(100):310. https://doi.org/10.1016/j.cliser.2022.100310. https://linkinghub.elsevier.com/retrieve/pii/S2405880722000280

Emerton R, Brimicombe C, Magnusson L et al (2022) Predicting the unprecedented: forecasting the June 2021 Pacific Northwest heatwave. Weather 77(8):272–279. https://doi.org/10.1002/wea.4257. https://rmets.onlinelibrary.wiley.com/doi/10.1002/wea.4257

Fiala D, Havenith G, Brode P et al (2012) UTCI- Fiala multi-node model of human heat transfer and temperature regulation. Int J Biometeorol 56. https://doi.org/10.1007/s00484-011-0424-7

Folkerts M, Boshuizen A, Gosselink G et al (2021) Predicted and user perceived heat strain using the ClimApp mobile tool for individualized alert and advice. Clim Risk Manag 34(100):381. https://doi.org/10.1016/j.crm.2021.100381. https://linkinghub.elsevier.com/retrieve/pii/S2212096321001108

Głogowski A, Brys K, Perona P (2020) Bioclimatic conditions of the lower silesia region (South-West Poland) from 1966 to 2017. Int J Biometeorol 65. https://doi.org/10.1007/s00484-020-01970-5

Goodfellow I, Bengio Y, Courville A (2016) Deep Learning. MIT Press. http://www.deeplearningbook.org

IMGW-PIB (2015) Prognoza biometeorologiczna dla polski. https://biometeo.imgw.pl/?page=BIOMET. Accesed 17 Aug 2022

IPMA (2022) UTCI (Universal Thermal Climate Index) às 00UTC. https://www.ipma.pt/pt/oclima/biometeo/utci/. Accesed 17 Aug 2022

Kingma BRM, Steenhoff H, Toftum J et al (2021) ClimApp-integrating personal factors with weather forecasts for individualised warning and guidance on thermal stress. Int J Environ Res Public Health 18(21):11317. https://doi.org/10.3390/ijerph182111317. https://www.mdpi.com/1660-4601/18/21/11317

Kingma BRM, Steenhoff H, Toftum J et al (2021) ClimApp-integrating personal factors with weather forecasts for individualised warning and guidance on thermal stress. Int J Environ Res Public Health 18(21):11317. https://doi.org/10.3390/ijerph182111317. https://www.mdpi.com/1660-4601/18/21/11317

Kozjek K, Dolinar M, Skok G (2017) Climate classification of slovenia. Int J Climatol. https://doi.org/10.1002/joc.5042

Krüger E (2017) Impact of site-specific morphology on outdoor thermal perception: a case-study in a subtropical location. Urban Clim 21. https://doi.org/10.1016/j.uclim.2017.06.001

Krzyżewska A, Dobek M, Domżał-Drzewicka R et al (2017) Emergency interventions due to weather-related hypothermia. Weather 72. https://doi.org/10.1002/wea.2926

Kuchcik M, Błażejczyk K, Halaś A (2021) Long-term changes in hazardous heat and cold stress in humans: multi-city study in poland. Int J Biometeorol 65. https://doi.org/10.1007/s00484-020-02069-7

Lindner-Cendrowska K, Błażejczyk K (2018) Impact of selected personal factors on seasonal variability of recreationist weather perceptions and preferences in warsaw (poland). Int J Biometeorol 62. https://doi.org/10.1007/s00484-016-1220-1

Liu J, Zhou G, Linderholm H et al (2022) Optimal strategy on radiation estimation for calculating universal thermal climate index in tourism cities of china. a Int J Environ Res Public Health 19. https://doi.org/10.3390/ijerph19138111

Lukić M, Milovanović J (2020) UTCI based assessment of urban outdoor thermal comfort in Belgrade, Serbia. Sinteza 2020 - international scientific conference on information technology and data related research. https://doi.org/10.15308/Sinteza-2020-70-77

Maas AL (2013) Rectifier nonlinearities improve neural network acoustic models. https://api.semanticscholar.org/CorpusID:16489696

Maracchi CLG (2017) Comfort termico. http://www.lamma.rete.toscana.it/meteo/comfort-termico. Accesed 17 Aug 2022

Milewski P (2013) Application of the UTCI to the local bioclimate of Poland's Ziemia Kłodzka region. Geogr Pol 86. https://doi.org/10.7163/GPol.2013.6

Morabito M, Grifoni D, Crisci A et al (2014) Might outdoor heat stress be considered a proxy for the unperceivable effect of the ultraviolet-induced risk of erythema in florence? J Photochem Photobiol B Biol 130. https://doi.org/10.1016/j.jphotobiol.2013.12.009

Nastos P, Matzarakis A (2011) The effect of air temperature and human thermal indices on mortality in athens, greece. Theor Appl Climatol 108. https://doi.org/10.1007/s00704-011-0555-0

NOAA National Centers for Environmental Information (2022) ETOPO 2022 15 Arc-Second Global Relief Model. https://doi.org/10.25921/fd45-gt74

Novak M (2013) Use of the UTCI in the Czech Republic. Geogr Pol 86(1):21–28. https://doi.org/10.7163/GPol.2013.3. http://rcin.org.pl/igipz/publication/46819

Novák M (2021) UTCI as the NWP model ALADIN (CHMI) output - first experiences. Geogr Pol 94(2):237–249. https://doi.org/10.7163/GPol.0203. https://rcin.org.pl/igipz/publication/229971

Nowosad BM, Rodzik Wereski MS, Dobek, (2013) The UTCI Index in Lesko and Lublin and its circulation determinants. Geogr Pol 86. https://doi.org/10.7163/GPol.2013.4

Palmer T (2020) A vision for numerical weather prediction in 2030. http://arxiv.org/abs/2007.04830

Pappenberger F, Jendritzky G, Staiger H et al (2015) Global forecasting of thermal health hazards: the skill of probabilistic predictions of the Universal Thermal Climate Index (UTCI). Int J Biometeorol 59(3):311–323. https://doi.org/10.1007/s00484-014-0843-3. http://link.springer.com/10.1007/s00484-014-0843-3

Pedregosa F, Varoquaux G, Gramfort A et al (2011) Scikit-learn: Machine learning in python. J Mach Learn Res 12. https://doi.org/10.48550/arXiv.1201.0490

Reichstein M, Camps-Valls G, Stevens B, et al (2019) Deep learning and process understanding for data-driven Earth system science. Nature 2019 566:7743 566(7743):195–204. https://doi.org/10.1038/s41586-019-0912-1. https://www.nature.com/articles/s41586-019-0912-1

Romaszko J, Cymes I, Dragańska E et al (2017) Mortality among the homeless: causes and meteorological relationships. PLoS ONE 12. https://doi.org/10.1371/journal.pone.0189938

Rozbicki T, Rozbicka K (2018) Variability of UTCI index in South Warsaw depending on atmospheric circulation. Theor Appl Climatol 133. https://doi.org/10.1007/s00704-017-2201-y

Schultz MG, Betancourt C, Gong B, et al (2021) Can deep learning beat numerical weather prediction? https://doi.org/10.1098/rsta.2020.0097. https://doi.org/10.1098/rsta.2020.0097

Skutecki R, Jalali R, Dragańska E et al (2019) UTCI as a biometeorological tool in the assessment of cold-induced stress as a risk factor for hypertension. Sci Total Environ 688. https://doi.org/10.1016/j.scitotenv.2019.06.280

Termonia P, Fischer C, Bazile E et al (2018) The aladin system and its canonical model configurations arome cy41t1 and alaro cy40t1. Geosci Model Dev 11. https://doi.org/10.5194/gmd-11-257-2018

Urban A, Kyselý J (2014) Comparison of UTCI with other thermal indices in the assessment of heat and cold effects on cardiovascular mortality in the Czech Republic. Int J Environ Res Public Health 11. https://doi.org/10.3390/ijerph110100952

Urban A, Di Napoli C, Cloke H et al (2021) Evaluation of the ERA5 reanalysis-based Universal Thermal Climate Index on mortality data in Europe. Environ Res 198. https://doi.org/10.1016/j.envres.2021.111227

Vinogradova V (2019) Using the universal thermal climate index (utci) for the assessment of bioclimatic conditions in Russia. Int J Biometeorol 65. https://doi.org/10.1007/s00484-020-01901-4

von Storch H, Zwiers FW (1999) Statistical analysis in climate research. Cambridge University Press. https://doi.org/10.1017/CBO9780511612336

Wang Y, Bellus M, EHRLICH A, et al (2018) 27 years of regional cooperation for limited area modelling in central Europe (rc lace). Bull Am Meteorol Soc 99. https://doi.org/10.1175/BAMS-D-16-0321.1

Weihs P, Staiger H, Tinz B et al (2011) The uncertainty of UTCI due to uncertainties in the determination of radiation fluxes derived from measured and observed meteorological data. Int J Biometeorol 56. https://doi.org/10.1007/s00484-011-0416-7

Wilks D (2006) Statistical methods in the atmospheric sciences. Academic Press

Wu F, Yang X, Shen Z (2010) Regional and seasonal variations of outdoor thermal comfort in china from 1966 to 2016. Sci Total Environ 665. https://doi.org/10.1016/j.scitotenv.2019.02.190

Zare S, Hasheminejad N, Shirvan HE et al (2018) Comparing Universal Thermal Climate Index (UTCI) with selected thermal indices/environmental parameters during 12 months of the year. Weather Clim Extremes 19. https://doi.org/10.1007/s00484-011-0453-2