MISSIONE 4
ISTRUZIONE
RICERCA

CLIMATE RESILIENT STRATEGIES BY
ARCHETYPE-BASED URBAN ENERGY MODELLING

Generation of future UHI weather data

DELIVERABLE 2.2

This study was carried out within the «Climate Resilient Strategies by Archetype-based Urban Energy Modelling (CRiStAll)» project – funded by European Union – Next Generation EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca), M4C2, I 1.1. This manuscript reflects only the authors' views and opinions and the Ministry cannot be considered responsible for them.

www.cristall.polito.it

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
Politecnico di Torino

Index

1. INTRODUCTION

1.1 Purpose

The purpose of this report is to explain the methods adopted for the generation of future weather data relative to the cities identified in the proposal. The deliverable describes the sources of weather data projections, the methods implemented to bias correct the future weather data, and the generation of future weather data for building simulation.

1.2 Structure of the report

This deliverable is structured into four sections, covering the collection and evaluation of future weather data from international databases, the bias correction of the datasets, and the generation of input data for building simulation and for the input to generators for incorporating the Urban Heat Island effect. Section 1 serves as the introduction, delineating the purpose (1.1), deliverable structure (1.2), and partner contributions to Task 2.2 development (1.3). Section 2 describes the characteristics of the data files used to generate future weather conditions. Section 3 presents the methods used to reduce bias errors by exploiting the data measured data previously obtained and described in deliverable D2.1. Section 4 addresses the problem of generating the test reference years for the future, ready to be modified for incorporating the Urban Heat Island effect.

1.3 Contribution of partners

Units obtained future projections from the CORDEX database, developed a script to retrieve weather data relative to specified locations, the team worked on the projected weather files corresponding to the stations identified in deliverable 2.1 for the locations of Rome and Bari and generated the relative TMY files corresponding to the selected stations. Units developed a general Python code to perform the bias correction of data using the measured and modelled data, implementing different univariate and multivariate bias correction methods. POLITO focused on Turin, obtaining future projections from the CORDEX database, performing multivariate bias correction on model projections using R scripts, and producing the relevant TMY files. The unibz team took care of developing a procedure to integrate in current and future weather files the UHI effect as modelled with UWG by MIT.

2. COLLECTION GCM-RCM projections for the area

Future weather projections were obtained from the CORDEX database, which provides data for different models and timesteps. However, it can be obtained with a time resolution of one day, three hours, and hourly. Hourly data was selected, avoiding the necessity of time interpolating data for the three-hour interval or using morphing methods with daily intervals. An additional issue is relative to the spatial coverage of the data. The data available refers to a wide area comprising the whole of Europe; therefore, projections or historical values were obtained only for the locations of interest to avoid unmanageable large data. For instance, the size of the files is strictly related to the interval time used to store the data, and in the case of hourly data, the sizes increase considerably with respect to the daily ones. The models selected to generate the future weather files are: GERICS_CNRM-CM5, designed as Model 1, and GERICS_IPSL-CM5A-MR, as Model 2.

Both models provide hourly data with a resolution of 0.11 degrees, corresponding to a grid of about 12.5 km. Modelled data are provided for two periods: a historical period from 1971 to 2005 and a projection period from 2006 to 2100. However, the available measured data for Rome and Bari do not cover the historical period; therefore, only the projection period corresponding to the RCP 8.5 has been considered for these two locations.

For Turin, the GERICS_M-MPI-ESM-LR model was selected, providing hourly data at a spatial resolution of 0.11 degrees (approximately 12.5 km), as it was closest to the median temperature of all climate model projections [1]. Unlike the other cities, measured data for Turin was also available for years before 2005; therefore, files related to the historical period were obtained as well.

For each station location identified in Deliverable 2.1, the variables relevant to the building energy simulation were obtained, as reported in Table 1.

Table 1: Variables used for the generation of future weather files

Variable	**Name**	**unit**
Near-Surface Air Temperature	tas	K
Near-Surface Specific Humidity	huss	kg_v/kg_a
Near-Surface Relative Humidity	hurs	%
Near-Surface Wind Speed	sfcWind	m/s
Surface Air Pressure	ps	Pa
Surface Downwelling Shortwave Radiation	rsds	W/m^2
Eastward Near-Surface Wind	uas	m/s
Northward Near-Surface Wind	vas	m/s

3. BIAS CORRECTION OF PROJECTIONS

Model generated future weather files usually show a bias error if compared with the measured files; therefore, before generating weather data for simulations, the values were bias corrected to avoid unreal behaviour.

3.1 Methodology

Simulations from global and regional climate models (GCMs and RCMs) may differ from local measurements, generating a bias error that will influence the expected climatic impact [2]. The difference can be more evident in long term scenarios, as highlighted by [3]. To correct the bias in models' outputs from simulations, available measurements for typical weather variables collected at the weather stations described in Deliverable 2.1 can be used. Bias correction can be applied singularly for each variable with a univariate bias correction method or dealing with multiple variables together to perform a multivariate bias correction. In this report, both approaches have been adopted. The univariate Bias Correction used is the Quantile Delta Mapping proposed by Cannon et al. [4], while the multivariate Correction has been performed with the MBCn method [5].

For both univariate and multivariate versions of the corrections, three periods can be identified. A historical period where both measurements and model data are present, a check period to check the effect of correction where both measurements and model data are present, and a future period where only the model is present. Figure 1 presents an example of the subdivision of periods using the data available for Rome. For other locations, different dates, depending on the measured data available, were used.

		recorded	recorded	
		2013 2019	2020 2023	2024
Historical model	Future model			
1971 2005	2006	2013 2019	2020 2023	2024 2100
		calibration	correction check	correction

Figure 1, weather data distribution between calibration, check and correction periods

3.2 Univariate model

A common methodology for correcting biases in future weather files follows a statistical approach. In this report, Delta Quantile Mapping (DQM) was used for the cities of Rome and Bari. DQM accounts for models' errors, assuming that biases in historical observations will be repeated during the projections. An additional feature of the method is that can automatically deal with projected values outside the range of the historical period. To check the performance of the procedure, the corrected and original model values have been compared to the measurements for a common period. However, the short period of recorded data required the splitting of the data between a training dataset and a check dataset with few years. Model data usually comprises a historical period ranging from 1971 up to 2005 and a future part from 2006 to 2100, however, the measured data from Rome and Bari required the use of modelled data as a historical period.

3.3 Multivariate model

The Multivariate Bias Correction with the N-dimensional probability density function (MBCn) method enhances the application of the Quantile Delta Mapping (QDM) method in a multivariate context. Initially, individual climate variables are corrected using the QDM method. Subsequently, the dependence structure among these climate variables is adjusted through an iterative reshuffling process. In each iteration, climate data are rotated by multiplying them with random orthogonal matrices. The QDM is then corrected and re-correlated using inverse random matrices.

While all climate variables undergo bias correction via the MBCn method, the QDM method specifically corrects global solar irradiance. This is because reshuffling marginally corrected global solar irradiance values, as performed in the MBCn method, can disrupt the diurnal structure of global solar irradiance. Such disruption may result in unrealistic values not only for global solar irradiance but also for the direct and diffuse solar irradiance components derived from it.

The calibration of the MBCn and QDM methods, along with the subsequent prediction of bias-corrected values, was conducted separately for each month of the year to preserve month-to-month variability in the bias-corrected climate data. The methods assume that the current bias will remain consistent in the future. In this report, the MBCn/QDM method is applied to the city of Turin. To evaluate the performance of the procedure, the corrected model values were compared with the original model values and measurements for a common period (check period).

4. Evaluation of weather data correction

4.1 Rome

Available weather data for Rome ranges from 2013 to 2023, so the Correction of model data cannot be carried out using the historical model; therefore, in the present work, the future model data for RCP 8.5 has been used since this model was considered the most adherent to the current situation. Figure 1 presents the splitting of data into a training period and a check period to test the efficiency of the method and the future projection.
Table 2 reports the data of the stations from which the measured data were collected for each coordinate presented in the Table 1 model, or both models were obtained from the netcfd files downloaded from the CORDEX database. The values were extracted using the *cdo* tool at the coordinates of the stations. Only RCP 8.5 data were used, and the historical parts of the model were disregarded.

Table 2: Weather stations of the Rete micrometeorological of Regione Lazio

Station	Code	City	Latitude	Longitude	height
Tor Vergata	AL001	Roma	41.84153	12.64792	104
Cavaliere	AL003	Roma	41.92889	12.65832	57
Castel di Guido	AL004	Roma	41.88942	12.2665	61
Boncompagni	AL007	Roma	41.9096	12.49657	72

According to Figure 1, the model and measured data were split into two periods, from 2013 to 2019 and from 2020 to 2023. The first period was used to train model data through the measurements available from the same period, and then the second, the test period, was used to evaluate the effect of the Correction on different data. For the test period, the monthly average for a generic parameter has been computed for the observed subscript "*o*", model, subscript "*m*", and corrected subscript "*q*", values. Then, the RMS error of monthly values was computed using the model and corrected values using Eq. 1, where the subscript "*j*" can be "*m*" for the model and "*q*" for corrected values.

$$RMS_j = \sqrt{\frac{\sum_k^{n_m} \left(v_o - v_j\right)^2}{n}} \tag{1}$$

Table 3 reports the obtained values for temperature, relative humidity, and global radiation. The Correction reduces the bias error for every model and variable; however, the reduced number of years of measured data does not allow for a large correction of the files.

Table 3: RMS for different variables and models. Station AL001

	GERICS_CNRM-CM5			GERICS_IPSL-CM5A-MR		
RMS	tas	hurs	rsds	tas	hurs	rsds
no correction	2.527	9.888	25.27	2.056	12.46	27.98
correction	1.864	8.714	21.25	1.831	7.873	24.74

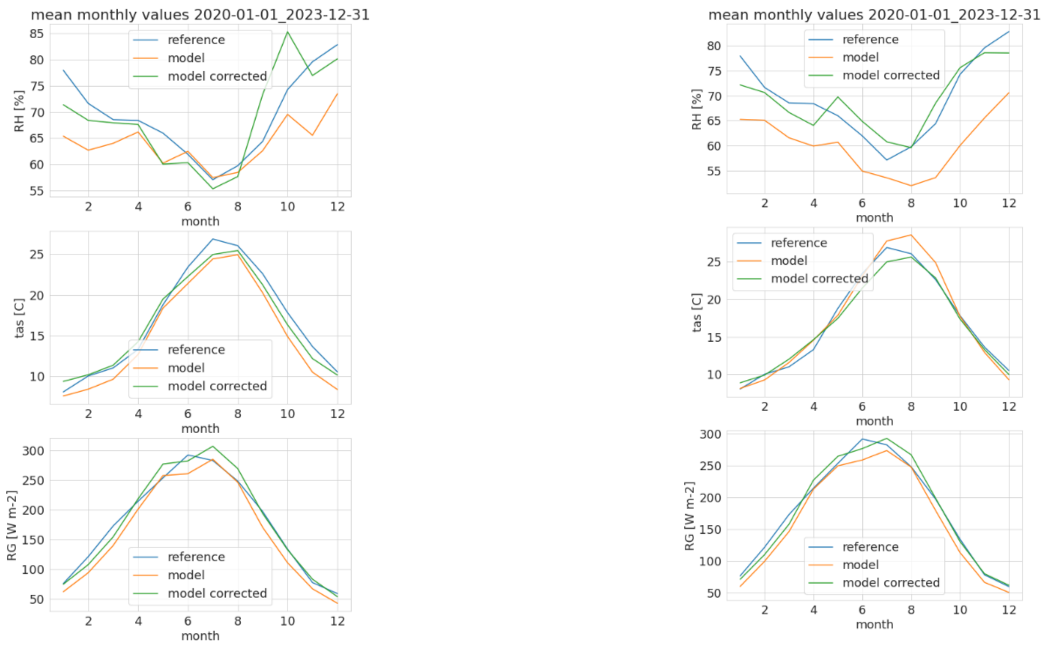

Figure 2: mean monthly values of parameters, left model 1, right model 2

Figure 2 presents the comparison between mean monthly values of temperature, humidity, and radiation for the two models. Quantile correction tries to match the statistical distribution of the model data with one of the measured ones. Figure 3 presents a qqplot for the station AL001 and temperature distribution for the calibration period on the left and check period on the right and Model 1, similarly, Figure 3 presents the same results for Model 2; analogous distributions can be obtained for other parameters. From both figures it appears that the quantile correction is able to correct the bias errors between measured data and the model since the dots representing singular values show that the quantile correction reduces the difference in quantile distribution, since the dots representing the comparison between measures and modelled data, as can be seen the dots are placed after Correction almost near the medium line of the plot. Figure 3 presents the comparison of measured, modelled, and modelled with quantile correction for the AL001 Station and temperature for Model 1, while Figure 4 presents the same distribution for Model 2. The efficiency of a model in replicating the statistical distribution of values is the density plot, which represents the probability density function of the variable. Figure 5 and Figure 6 show the density distribution of the temperature for the reference, modelled, and quantile corrected model for the training and check period of temperature and relative humidity for station AL001 and Model 1. Again, it is possible to identify how the quantile correction tries to replicate the statistical distribution of the reference data.

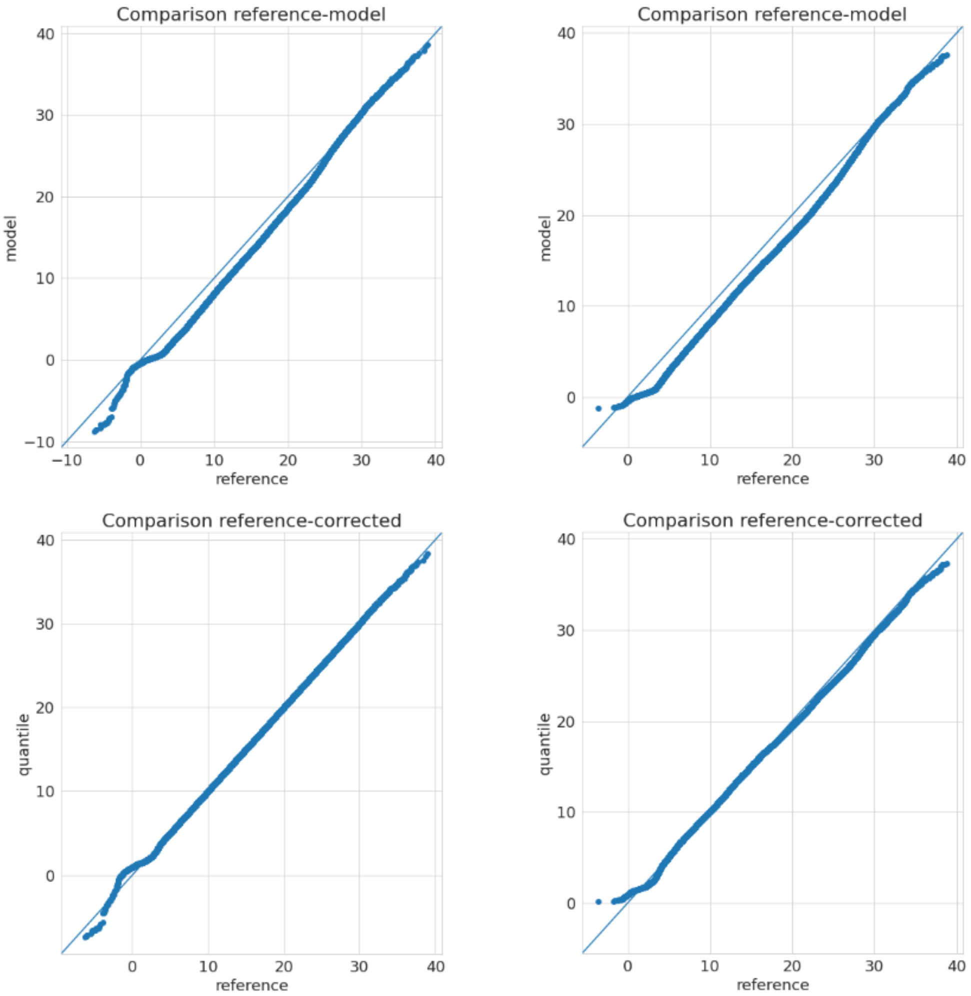

Figure 3: quantile-quantile distribution for the training period, left and test period right for model 1

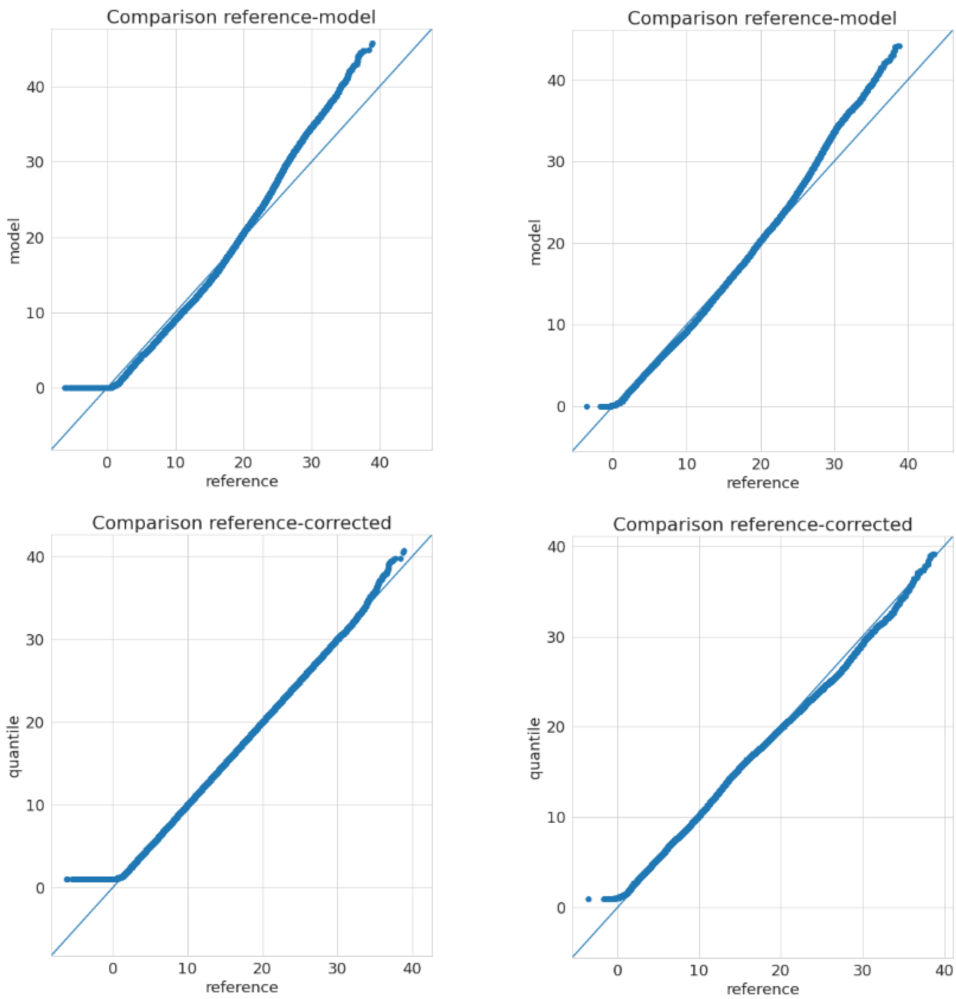

Figure 4: quantile-quantile distribution for the training period, left and test period right for model 2

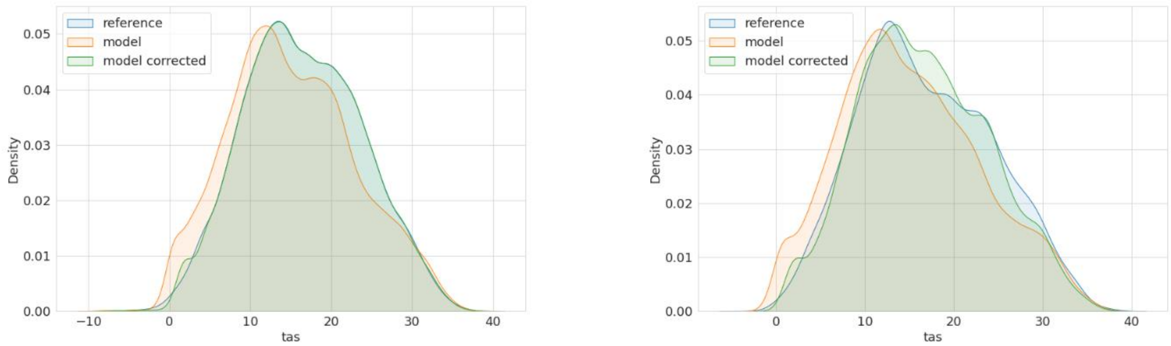

Figure 5: density plot for temperature, training period left and check period right

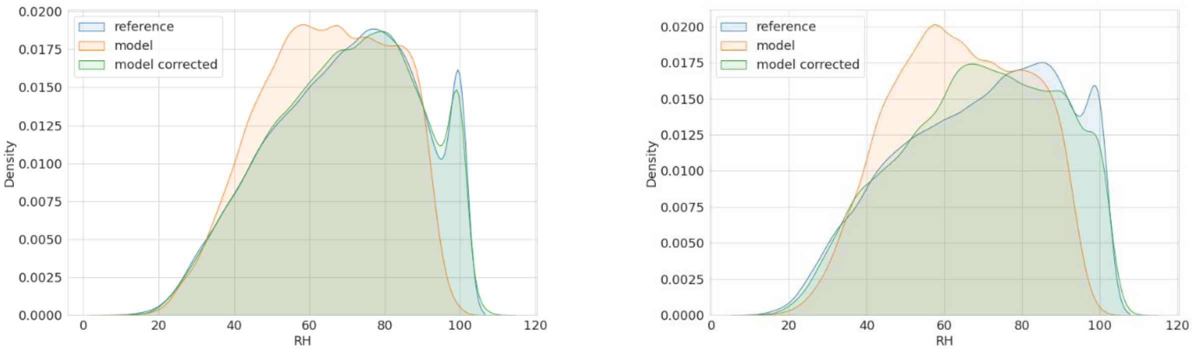

Figure 6: density plot for relative humidity, training period left and check period right

4.2 Bari

The same approach for Rome was followed for Bari. Table 4 reports the RMS errors obtained for temperature, relative humidity and global radiation. However, the small number of years of measured data did not allow a large correction of the data, especially for relative humidity, in this case Model 2 performed better than Model 1

Table 4: RMS for different variables and models. Station Bari Corso Trieste

	GERICS_CNRM-CM5			**GERICS_IPSL-CM5A-MR**		
RMS	tas	hurs	rsds	tas	hurs	rsds
no correction	2.224	6.193	24.707	2.052	8.536	23.47
correction	1.739	7.761	21.613	1.914	7.877	18.94

4.3 Turin

For Turin, the data from the “Bauducchi” station, from which the measured data were collected, were obtained from the NetCDF files downloaded from the CORDEX database for the selected model (GERICS_MPI-M-MPI-ESM-LR). The station coordinates are latitude 44.96° N, longitude 7.70° E, and height 226 m. A Python script was used to extract the nearest point to these coordinates and assemble the various weather variables and years into a single dataset.

Two sets of measured data were collected from the “Bauducchi” station: one from 1994 to 2003 (the training period) for performing bias correction using historical models, and another from 2014 to 2023 (the check period) to test the efficiency of the method and assess future projections.

In the first step of evaluating the weather data correction, the same procedure for calculating the RMS Error of monthly values (Eq. 1) was applied, and the results are presented in Table 5 for temperature, relative humidity, global irradiation, and wind speed. Similar to the findings in Rome, the correction effectively reduces the bias error for all variables.

The next step involves comparing the probability density functions (PDFs) of temperature, wind speed, and relative humidity from observations (grey), the raw model (blue), and the bias-corrected model (red) for both the training and check periods. Figures 7, 8, and 9 illustrate that the PDFs of the raw model are closely aligned with the observed data during the training period. For the check period, the bias-correction procedure effectively adjusts the PDFs of the raw model to more closely mimic those of the observations. This adjustment holds true not only for temperature but also for more complex variables such as wind speed, underscoring the effectiveness of the bias correction step in producing realistic estimates across a range of climate variables.

Table 5: RMS for different variables. Bauducchi Station

	GERICS_MPI-M-MPI-ESM-LR			
RMS	tas	hurs	rsds	sfcWind
no correction	0.748	12.684	26.73	0.519
correction	0.624	6.337	20.89	0.089

Further validation of this effectiveness is provided by the comparison of mean monthly values for temperature, humidity, global solar irradiation, and wind speed, as presented in Figure 10. The results indicate that the bias-corrected model not only enhances individual variables but also maintains their interrelationships, thereby ensuring a more accurate representation of the underlying climatic conditions.

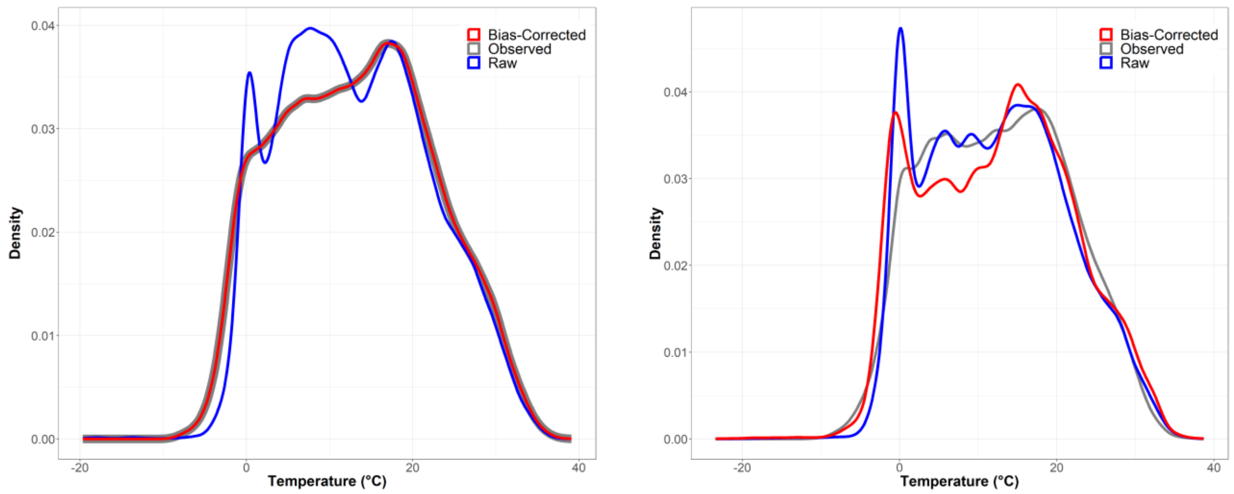

Figure 7: density plot for temperature, Bauducchi Station, training period left and check period right

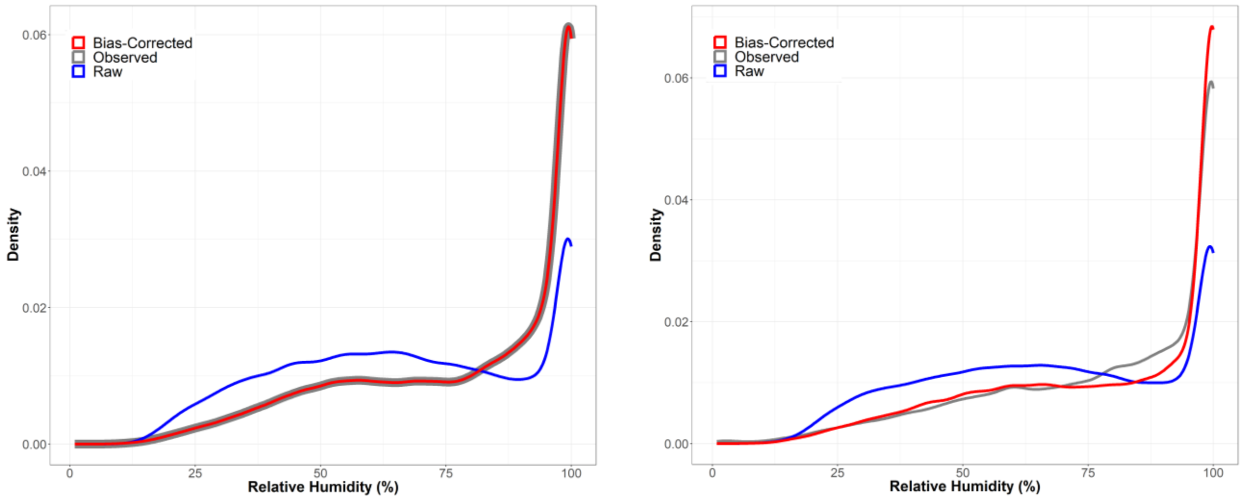

Figure 8: density plot for relative humidity, Bauducchi Station, training period left and check period right

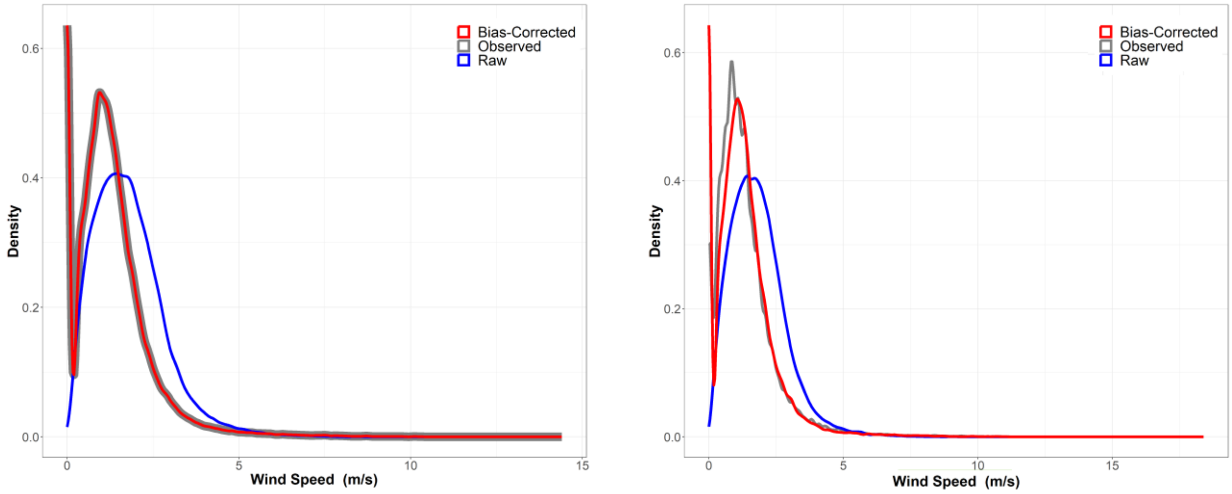

Figure 9: density plot for wind speed, Bauducchi Station, training period left and check period right

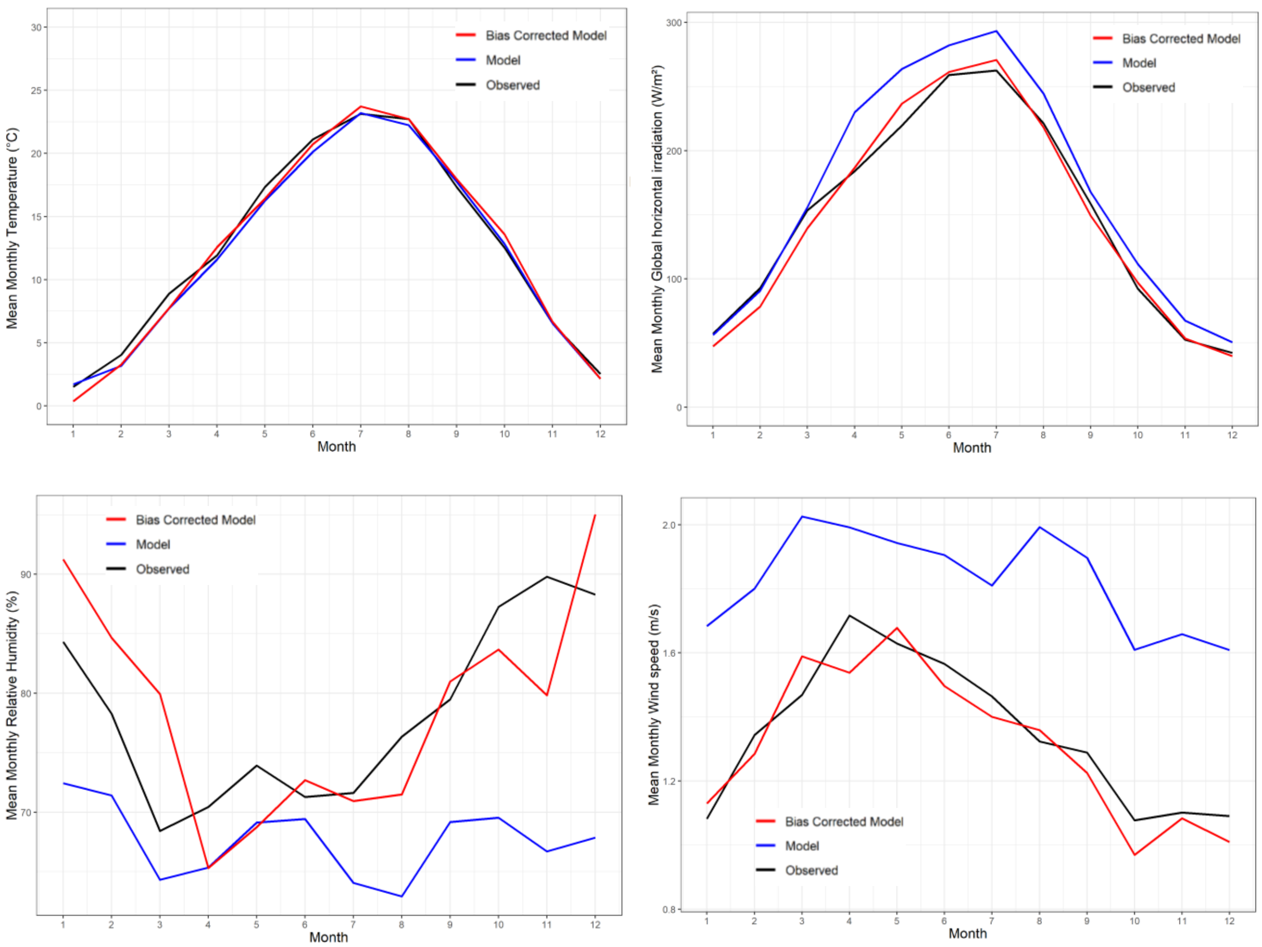

Figure 10: mean monthly values of parameters for Bauducchi Station-2014-2023

5. Generation of future reference years

The input files for simulations have been generated following the UNI-EN-ISO 15927–4 standard. The files were created for the present period and the future as well. For the future, three periods were considered: a present weather file computed using the corrected modeled data, but with reference to the dates of historical measured data available for each station, a mid-term future file 2081-2100, and a long term one 2076-2010 using the quantile corrected projections previously obtained. Heating and Cooling degree days using a base temperature of 18 °C were used to synthetically describe the characteristics of the generated TMYs as reported in the following tables.

5.1 Rome

The quantile corrected future files can be used to generate future TMY using the standard methodologies. This is possible since future weather data is on an hourly basis. Other models present only daily future weather data; therefore, to generate future TMY, a morphing method (Manzan et al., 2024) should be used. For each station and model, three files are generated for three different periods. A historical one from 2013 to 2023, a middle future period from 2041 to 2061, and a long term period file 2081 to 2100. Table 6 reports the heating and cooling days for the generated files:

Table 6: Heating and cooling degree days for Rome and in different periods

Station	period	Model 1		Model 2	
		HDD_{18} [°C·d]	CDD_{18} [°C·d]	HDD_{18} [°C·d]	CDD_{18} [°C·d]
AL001	2013_2023	1144	838	982	874
	2041_2060	838	961	825	1171
	2081_2100	579	1265	488	1606
AL003	2013_2023	1021	898	1100	943
	2041_2060	939	1042	823	1281
	2081_2100	585	1426	496	1790
AL004	2013_2023	886	871	1157	802
	2041_2060	823	965	749	1037
	2081_2100	561	1198	373	1438
AL007	2013_2023	909	934	939	1003
	2041_2060	797	1073	755	1337
	2081_2100	543	1522	435	1710

Table 6 shows the general behaviour of the generated file. First, as a general rule, heating degree days decrease when considering future generated simulation data; instead, cooling degree days increase at the same time. This highlights a general trend due to climate change. An additional observation is that Model 2 presents a scenario more affected by climate change. This behaviour is also confirmed by the inspection of Figure 11, which presents the daily mean distribution of temperature for the reference periods. The temperature increases between the vase value of 2013-2023 to the future periods of 2041-2060 and 2081 2100, and the temperatures generated using Model 2 are higher than the ones of Model 1.

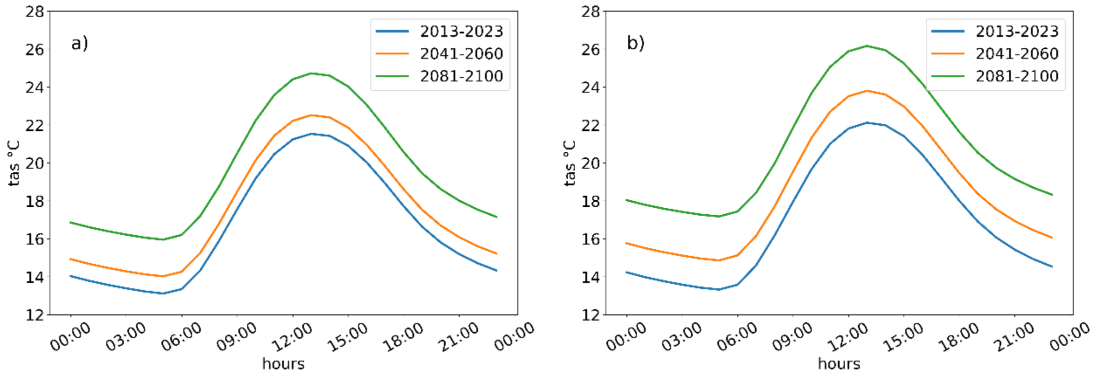

Figure 11: mean daily temperatures for station AL001 for different future periods, a) with model 1 b) with model 2

5.2 Bari

The same approach was followed for Bari, in this case, with the values of the weather station of Corso Trieste. Future weather files show the increase of temperatures in the short and long term, as can be inferred from the values of heating and cooling days reported in Table 7. Figure 12 presents the mean daily temperatures with the two models, and the increase of mean temperatures in future periods is again visible. The results of Bari show, as noticed with Rome, that model 2 provides temperature values higher than the ones obtained with model 1.

Table 7: Heating and cooling degree days for Bari and in different periods

Station	period	Model 1		Model 2	
		HDD_{18} [°C·d]	CDD_{18} [°C·d]	HDD_{18} [°C·d]	CDD_{18} [°C·d]
AL001	2013_2023	759	947	717	1102
	2041_2060	582	1178	500	1365
	2081_2100	387	1516	291	1861

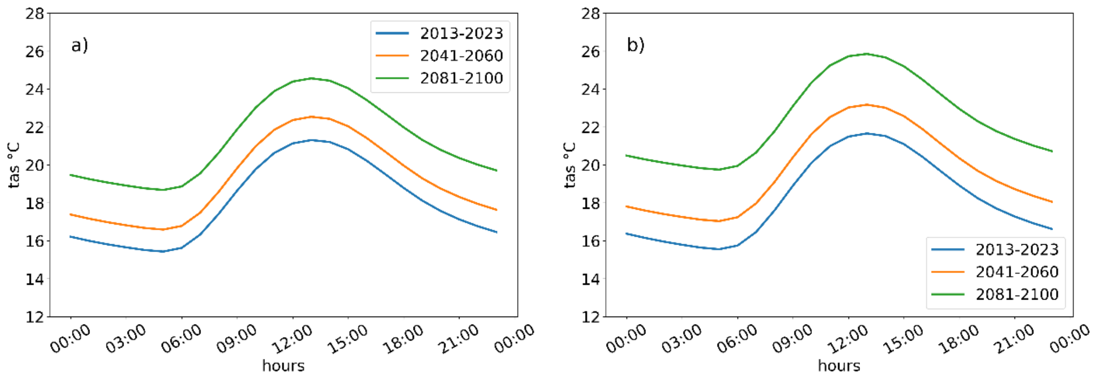

Figure 12: mean daily temperatures for Bari for different future periods, a) with model 1 b) with model 2

5.3 Turin

The same method for creating reference years was applied to Turin. Future weather files indicate a temperature increase in both the short and long term, as reflected in the heating and cooling days presented in Table 8. Figure 13 illustrates mean daily temperatures, further emphasizing the rise in mean temperatures in future periods, particularly for the long-term future.

Table 8: Heating and cooling degree days for Turin in different periods

Station	period	GERICS_MPI-M-MPI-ESM-LR	
		HDD_{18} [°C·d]	CDD_{18} [°C·d]
Bauducchi	2013_2023	2143	550
	2041_2060	1918	720
	2081_2100	1518	1120

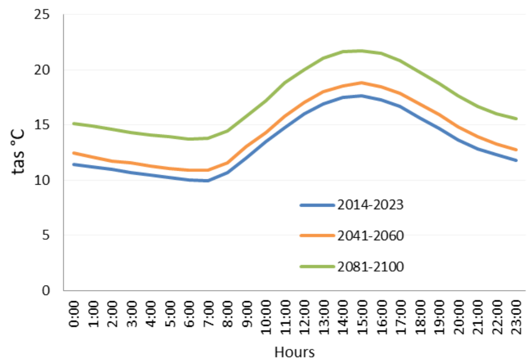

Figure 13: mean daily temperatures for Turin for different future periods

6. IMPLEMENTING UHI EFFECT

The UWG was used as a tool to model and account for the Urban Heat Island effect in weather data. A dedicated procedure was developed by the Free University of Bolzano to select representative urban blocks and perform for those a detailed modelling with the UWG by the MIT. The results were used to develop statistical correlations based on the urban morphology to quantify in a simple way the impact of the UHI effect. Such correlations were later tested and validated with an additional set of urban blocks, different from those selected for the generation of the statistical correlations, in order to assess their robustness and representativeness. Once the correlations proved to be sufficiently representative, they were used for a quick and effective visual representation of the UHI for the entire city under consideration. The proposed approach, combining detailed modelling with the UWG, statistical analyses and GIS mapping, is described in Figure 14.

An application for Turin was presented in the framework of the conference Building Simulation Applications BSA 2024, which took place in Bolzano, June 26th – 28th 2024 (Borelli G., Ballarini I., Corrado V., Gasparella A., Pernigotto G. 2024. “*Assessment and mapping of the urban heat island* *effect: a preliminary analysis on the impact on urban morphology for the city of Turin, Italy*”). In the case of Turin, 170 urban blocks were used for training the statistical correlations and 20 for their test. As it can be seen in Figure 15, the impact of UHI can be significant with respect to the rural weather data (Torino Bauducchi), with deviations of HDD_{18} ranging from slightly more than 450 K d to about 300 K d, respectively for current and future scenarios. Likewise, the deviations for CDD_{18} can range from the current 200 K d to more than 300 K d in the future.

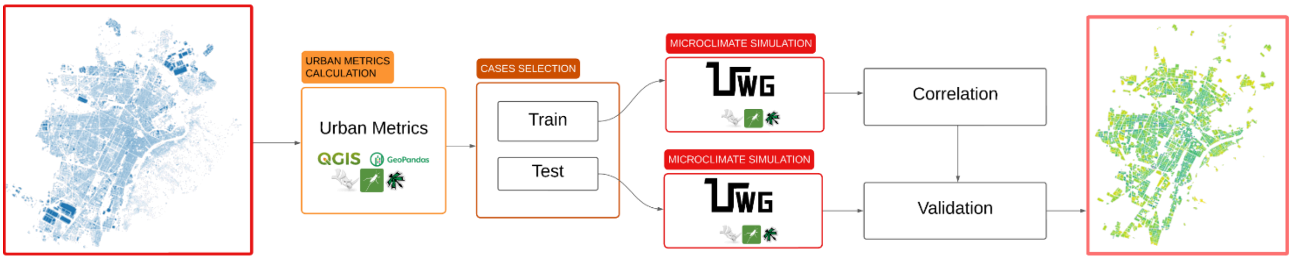

Figure 14: Proposed procedure for the inclusion of UHI effects in current and future weather files

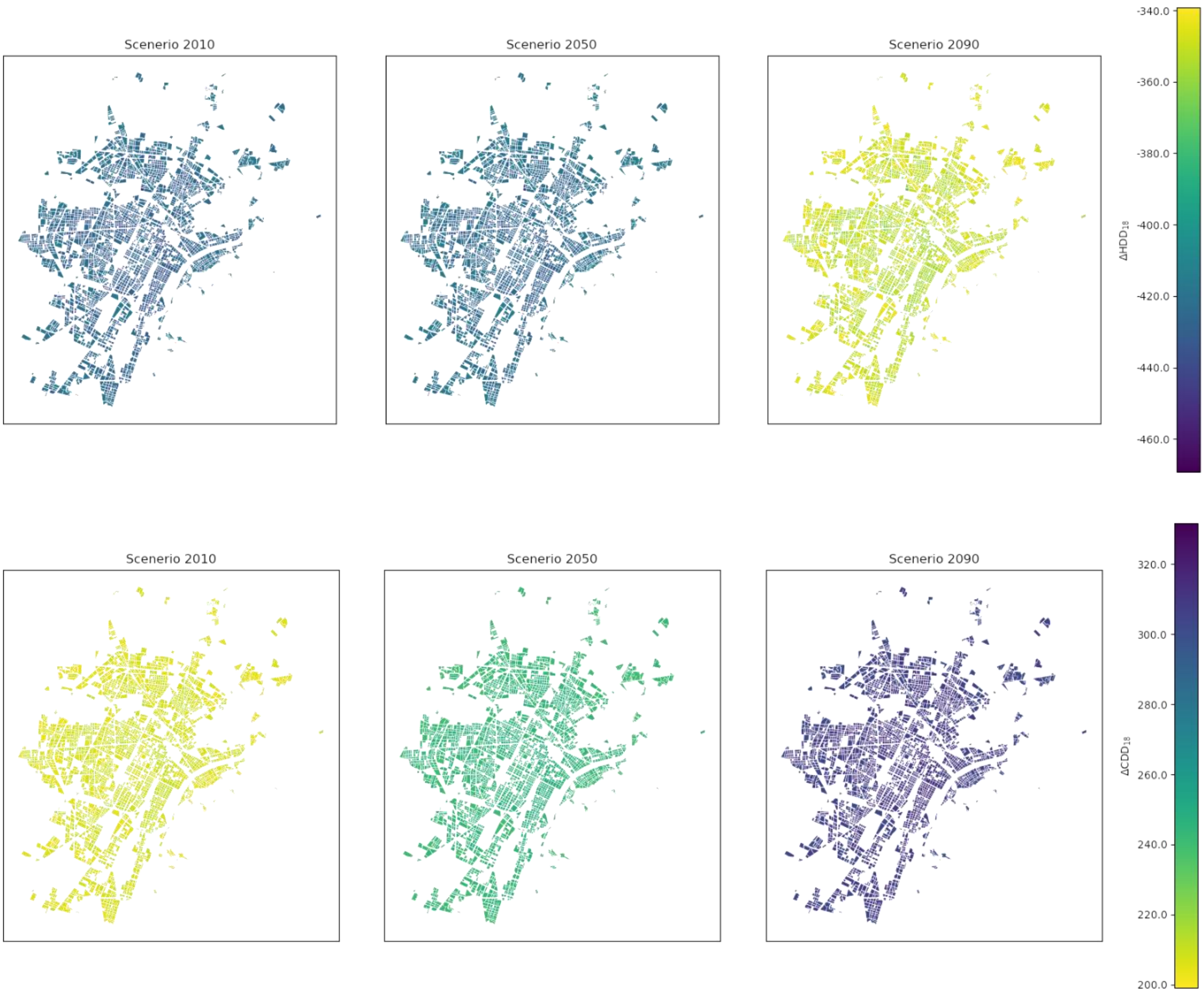

Figure 15: Impact of UHI effects on Turin for current and future weather files, respectively in terms of variation for the HDD18 (top) and CDD18 (bottom)

Finally, the proposed approach allowed to explore the variability of the UHI effect within the different areas of the city, allowing for an easier identification of districts to prioritize in the next work packages.

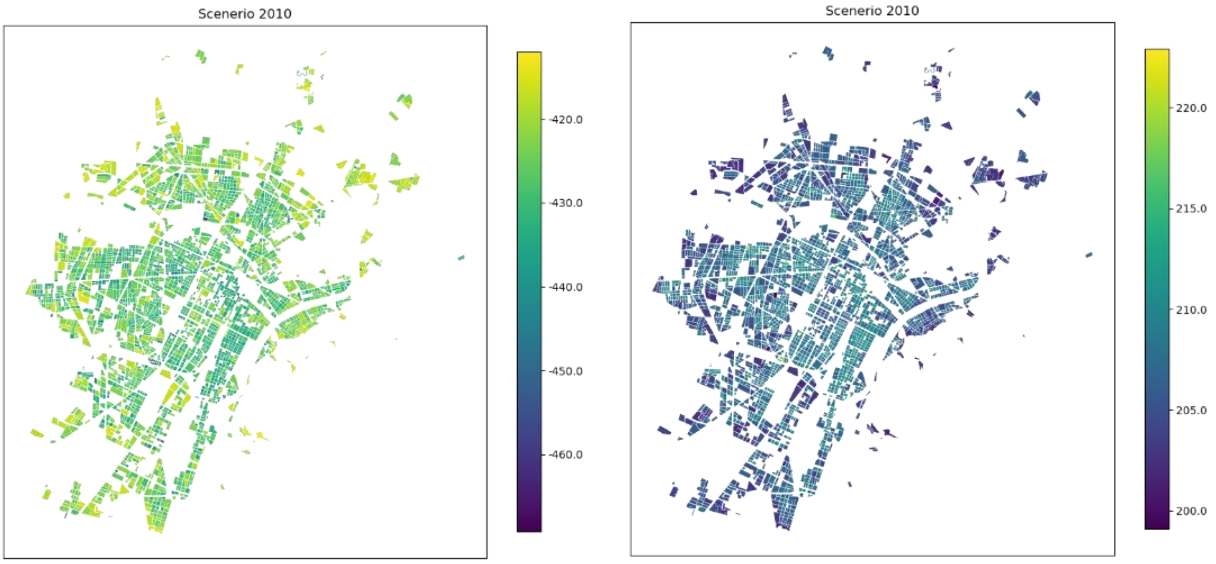

Figure 16: Impact of UHI effects in Turin for the current climate scenario (2010), respectively in terms of variation for the HDD18 (left) and CDD18 (right)

7. CONCLUSIONS

In this research, the availability of data obtained from GCM-RCM projections was exploited to obtain weather files describing the future behaviour of climate. This is an important feature to design buildings that will be used in future periods, but it is also important to analyse urban areas, since they will be particularly affected by climatic change due to the anthropic impact. However, as shown in this report, some steps are mandatory prior to implement models capable to project data up to year 2100. First, the data from models must be treated and bias corrected before they can be used; in literature, there are different bias correction methods, so different approaches are available. Therefore, it is crucial to assert how the bias correction modifies the original model data. For this step, a check period different from the one used to train the bias correction, but with measured and modelled data, must be identified and the relative measured and model data compared using appropriate metrics. The data obtained from methods come from numerical simulations and, therefore, it is affected by intrinsic uncertainties; for this, it is desirable to use different models so to be able to incorporate the uncertainty into results. An additional issue is the time resolution of the modelled data. Using the available hourly data, the final generation of Typical Meteorological Years (TMY) was straightforward. If the modelled data had been available at different temporal resolutions, usually every three hours or daily, interpolation or the application of morphing methods would have been necessary. The outcome of the presented procedure is the generation of TMYs files ready to be used in building simulation or used as inputs for tools such as UWG to generate future weather files incorporating urban heat island effect. The results described in this report demonstrate that the now available future projections are an important tool for building and urban area simulations.

8. REFERENCES

[1] Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S.C. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov CR and MR. Evaluation of Climate Models. Climate Change 2013 – The Physical Science Basis, Cambridge University Press; 2014, p. 741–866. https://doi.org/10.1017/CBO9781107415324.020.

[2] Hoffmann H, Rath T. Meteorologically consistent bias correction of climate time series for agricultural models. Theoretical and Applied Climatology 2012;110:129–41. https://doi.org/10.1007/s00704-012-0618-x.

[3] Agbazo MN, Grenier P. Characterizing and avoiding physical inconsistency generated by the application of univariate quantile mapping on daily minimum and maximum temperatures over Hudson Bay. International Journal of Climatology 2020;40:3868–84. https://doi.org/10.1002/joc.6432.

[4] Cannon AJ, Sobie SR, Murdock TQ. Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes? Journal of Climate 2015;28:6938–59. https://doi.org/10.1175/JCLI-D-14-00754.1.

[5] François B, Vrac M, Cannon AJ, Robin Y, Allard D. Multivariate bias corrections of climate simulations: Which benefits for which losses? Earth System Dynamics 2020;11:537–62. https://doi.org/10.5194/ESD-11-537-2020.