URBREATH [101139711]

Systemic Integration of Transformative Technical and Nature-based Solutions to Improve Climate Neutrality of European Cities and Regions and tackle Climate Change: the URBreath Approach

D3.1 AI models for climate change vulnerability assessment and weather forecast - Version 1

Project Reference No	URBREATH – 101139711
Deliverable	D3.1 AI models for climate change vulnerability assessment and weather forecast – V1
Work package	WP3: URBREATH data strategy and tools
Type	OTHER
Dissemination Level	PU - Public
Date	28/02/2025
Status	Final
Editor(s)	Emma Gaitán, Toni Rubio, Darío Redolat, Carlos Prado, Lorena Galiano, César Paradinas, Robert Monjo (FIC)
Contributor(s)	Task 3.3 participants
Reviewer(s)	Stijn Vranckx (VITO) Ettore Etenzi (EXUS)
Document description	This document describes a comprehensive framework for complete time-scale atmospheric predictions and projections within the URBREATH project, integrating methodologies for short-term weather forecasting, seasonal predictions, and long-term climate projections. This deliverable establishes a robust predictive system that will be further refined in subsequent project phases. This deliverable is linked to T3.2 and T3.3, updates of this deliverable are foreseen in M24 (December 2025) and M36 (December 2026).

Document Revision History

Version	Date	Modifications Introduced	
		Modification Reason	Modified by
0.1	15/12/2024	Table of contents	FIC
0.2	18/12/2024	Draft ready for internal review	FIC
1.0	25/02/2025	Internal review	ICCS
1.0	26/02/2025	Internal review	EXUS
1.1	27/02/2025	Final version	FIC
2.0	27/02/2025	Submitted to Project Coordinator for quality check	LC
2.1	28/02/2025	Final version ready for submission	LC

Disclaimer

The URBREATH project is co-funded by the European Union under grant agreement ID 101139711. The information and views set out in this document are those of the URBREATH Consortium only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Executive summary

The URBREATH project is dedicated to improving urban climate resilience by integrating AI-driven forecasting and projections methodologies across three harmonized temporal scales: short-term weather forecasting, seasonal prediction, and long-term climate projections. This deliverable, D3.1-V1, presents the methodological framework and initial results for enhancing urban-scale meteorological forecasts, ensuring that the project provides a coherent and complementary forecasting system at all timescales. By leveraging ensemble prediction systems (EPS), statistical corrections, and AI-driven post processing, this work aims to optimize predictions for key climate variables such as temperature, precipitation, and snow accumulation, which are essential for urban planning and adaptation.

At the weather forecasting scale, this deliverable evaluates the performance of different ensemble-based probabilistic models to determine their accuracy for short-term predictions (0-15 days). Due to limited data availability, a full multi-model comparison over 2023-2024 could not be conducted, as outlined in Section 2.2.3. Instead, a comparative assessment for December 2024 - January 2025 was carried out, applying the same verification methodologies planned for the full dataset. Additionally, systematic bias correction techniques were applied to Global Ensemble Forecast System (GEFS) data for 2023-2024, demonstrating the potential for significant forecast improvements, which will be further developed in future phases.

For the seasonal forecasting scale, this deliverable outlines the methodology for providing reliable probabilistic forecasts for lead times of 1 to 6 months. The approach is based on multi-model ensembles, post-processing adjustments, and integration with key climate drivers such as sea surface temperatures (SSTs) and large-scale circulation patterns. This ensures that seasonal predictions align with both short-term weather forecasts and long-term climate projections, creating a seamless transition between different forecasting timescales.

At the climate projection scale, this deliverable integrates high-resolution downscaling techniques to adapt global climate model (GCM) projections to urban environments. The methodologies ensure that climate projections align with seasonal and short-term forecasts, providing a comprehensive climate service for the project's Front Runner (FR) cities. The projections assess long-term trends, extreme event probabilities, and vulnerability factors, ensuring that decision-makers receive the necessary insights for strategic climate adaptation planning.

By harmonizing these three forecasting/projection scales, this deliverable establishes a robust, multi-timescale framework for predicting urban meteorological conditions. Future project phases will expand the dataset for a complete multi-model comparison, optimize AI-based forecast postprocessing, and get ready to deliver to the operative decision-support tools of the project for the FR cities in the next phase and all cities in the last one. These advancements will ensure that URBREATH delivers a cutting-edge forecasting service, enabling URBREATH pilot cities to make data-driven decisions for climate resilience and adaptation.

Table of Contents

List of Figures 6
List of Tables 11
List of Terms and Abbreviations 12
1. Introduction 13
1.1 Context, purpose and scope 13
1.1.1 Timescale links and crucial synergies 15
1.2 Approach and relation to other Work Packages and deliverables 16
1.3 Structure of the deliverable 18
2. Data and study area 20
2.1 Study Area 20
2.2 Data requirements 21
2.2.1 Observed data 22
2.2.2 Weather forecast 22
2.2.3 Seasonal forecast 23
2.2.4 Climate projections 24
2.3. Data Sources Selected 25
2.3.1 Observed data 25
2.3.2 Weather forecast 40
2.3.3 Seasonal forecast 42
2.3.4 Climate projections 43
3. Methodology 48
3.1 Weather forecast 48
3.2 Seasonal forecast 56
3.3 Climate projections 60
4. Results 68
4.1 Weather forecast 69

4.2 Seasonal forecast 83
4.3 Climate projections 85
5. Conclusions 98
6. References 100
Annex 1. Observed data information 105
1. Information about Cluj-Napoca's observed data 105
2. Information about Leuven's observed data 106
3. Information about Madrid's observed data 112
4. Information about Tallinn's observed data 117

List of Figures

Figure 1. Map of Europe showing the Front Runner cities (in red) and follower cities (in purple) marked with a cross. 21

Figure 2. Outline of data required for the climatic study. 25

Figure 3. Resulting quality control for the minimum temperature series of the observatory with identifier '58-El Pardo' for the city of Madrid. Outliers of the extreme values type are detected. 29

Figure 4.Resulting quality control for the precipitation series of the observatory with identifier '15120099999' for the city of Cluj-Napoca. Outliers of the extreme values type are detected. 30

Figure 5.Resulting homogeneity check for the maximum temperature series of the observatory with identifier 'LC-011' for the city of Leuven. The test detects a homogeneity jump with a mean difference of less than 3°C. 32

Figure 6. Resulting homogeneity check for the precipitation series of the observatory with identifier '64510' for the city of Leuven. 34

Figure 7. The figure shows a) observatories available at the start of the study, as well as the observatories ultimately used in the study for b) weather forecast, c) seasonal forecast and d) climate projections for Cluj-Napoca. Observatories with temperature data are shown in red, those with precipitation data in blue, and those with both variables in green. Meteo Romania is the Romanian weather agency. The district of Cluj in Romania, whose capital is Cluj-Napoca, is shown in green, with the study area marked with a cross 36

Figure 8. The figure shows a) observatories available at the start of the study, as well as the observatories ultimately used in the study for b) weather forecast, c) seasonal forecast and d) climate projections for Leuven. Observatories with temperature data are shown in red, those with precipitation data in blue, and those with both variables in green. KMI-IRM is the Belgian weather agency. The city boundary of Leuven is shown in green, with the study area marked with a cross. 37

Figure 9. The figure shows a) observatories available at the start of the study, as well as the observatories ultimately used in the study for b) weather forecast, c) seasonal forecast and d) climate projections for Madrid. Observatories with temperature data are shown in red, those with precipitation data in blue, and those with both variables in green. AEMet is the Spanish weather agency, whereas City Council stands for Madrid's weather monitoring system. The municipality of Madrid is shown in black, and the district of Villaverde is shown in green, with the study area marked with a cross. 38

Figure 10. The figure shows a) observatories available at the start of the study, as well as the observatories ultimately used in the study for b) weather forecast, c) seasonal forecast and d) climate projections for Tallinn. Observatories with temperature data are shown in red, those with precipitation data in blue, and those with both variables in green. ILM is the Estonian weather agency. The borders of Estonia are shown in black, and the

county of Harju, which includes the city of Tallinn, is shown in green, with the study area marked with a cross. 39

Figure 11. Example of the spatial representation of ERA5-Land reanalysis, representing a 30-year return period event for daily maximum temperature. Source: C3S 44

Figure 12. Shared Socioeconomic Pathways (in the Figure, OECD stands for Organizations of Economic Co-operation and Development). Source: O'Neill et al., 2017 45

Figure 13. Scenarios defined in CMIP6. Level 1 and level 2 are defined as the scenarios to be provided as mandatory and optional, respectively, by all climate models that are part of each CMIP phase. Source: Figure adapted from O'Neill et al., 2016 46

Figure 14. Example ROC diagram: EPS vs. deterministic forecast. Source: ECMWF 50

Figure 15. Area under the ROC curves for two score classifiers, A and B. Source: Springer Nature website 51

Figure 16. Contingency table for a yes/no event. Source: CAWCR 51

.Figure 17. Scheme of seasonal forecast methodology 59

Figure 18. Key features of climate and Earth System Models (ESMs): ESMs include biological and chemical processes. Source: Heavens et al. 2013 61

Figure 19. Requirements for climate projections 63

Figure 20. FICLIMA statistical downscaling: key features 65

Figure 21. Maximum temperature BIAS in Madrid 70

Figure 22. MAE for corrected maximum temperature in Madrid 71

Figure 23. MAE for corrected minimum temperature in Madrid. GEFS 2023-2024 72

Figure 24. ROC curve for precipitation in Madrid. Thicker lines represent predictions for D+1 for each model, while thinner lines represent forecasts from D+2 to D+10 73

Figure 25. Minimum temperature BIAS in Leuven 74

Figure 26. MAE for corrected minimum temperature in Leuven 75

Figure 27. ROC curve for precipitation in Leuven. GEFS model 2023-2024 76

Figure 28. Sum of misses and false alarms for precipitation in Leuven. GEFS model 2023-2024 77

Figure 29. Maximum temperature BIAS in Tallinn 78

Figure 30. MAE for corrected maximum temperature in Tallinn 79

Figure 31. Area under the ROC Curve (AUC) for precipitation in Tallinn. GEFS 2023-2024 80

Figure 32. Minimum temperature BIAS in Cluj-Napoca 81

Figure 33. MAE for corrected minimum temperature in Cluj-Napoca 82

Figure 34. ROC curve for precipitation in Cluj-Napoca. GEFS 2023-2024 83

Figure 35. Figure of the CNN zero stage of predictor selection. On the left, the comparison between all the observations (for the 9 observatories in black) and the two best performed methods (MultiLinear in blue and LinearDeepCNN in green). On the right, a boxplot showing the Pearson Correlation for each method used 84

Figure 36. Expected evolution of maximum temperature for Madrid city throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Madrid. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results 85

Figure 37. Expected evolution of minimum temperature for Madrid city throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Madrid. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results 86

Figure 38. Expected evolution of annual cumulative precipitation for Madrid city throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Madrid. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results 87

Figure 39. Expected evolution of maximum temperature for Leuven throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Leuven. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results..... 88

Figure 40. Expected evolution of minimum temperature for Leuven throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Leuven. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results..... 89

Figure 41. Expected evolution of annual cumulative precipitation for Leuven throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Leuven. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results 90

Figure 42. Expected evolution of maximum temperature for Tallinn throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Tallinn. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results..... 92

Figure 43. Expected evolution of minimum temperature for Tallinn throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Tallinn. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results..... 92

Figure 44. Expected evolution of annual cumulative precipitation for Tallinn throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four

emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Tallinn. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results 93

Figure 45. Expected evolution of maximum temperature for Cluj-Napoca throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Cluj-Napoca. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results 95

Figure 46. Expected evolution of minimum temperature for Cluj-Napoca throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Cluj-Napoca. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results 96

Figure 47. Expected evolution of annual cumulative precipitation for Cluj-Napoca throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Cluj-Napoca. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results 97

List of Tables

Table 1. Compilation of available observatories with climatic series of temperature and precipitation and sources. 27

Table 2. Compilation of selected observatories with climatic series of temperature and precipitation. 33

Table 3. Compilation of not selected observatories with climatic series of temperature and precipitation. First value is referred to the weather forecast and the second is the seasonal forecast. 35

Table 4: Compilation of available raw EPS models and summary of issues 41

Table 5: Variables required for the seasonal forecast design of predictors 43

Table 6. Information on 10 CMIP6 climate models (IPCC AR6), retrieved from the ESGF portal 47

Table 7. Compilation of available raw EPS models 53

List of Terms and Abbreviations

Abbreviation	Definition
AUC	Area Under the Curve
IPCC6	Intergovernmental Panel on Climate Change, Sixth Assessment Report
FR	Front Runner
AI	Artificial Intelligence
GCM	Global Climate Model
EPS	Ensemble Prediction System
WP	Work Package
NBS	Nature-Based Solutions
ICT	Information and Communication Technology
LLL	Local Living Labs
CMIP6	Coupled Model Intercomparison Project Phase 6
SSP	Shared Socioeconomic Pathways
ERA5	Fifth generation ECMWF atmospheric reanalysis
ESM	Earth System Model
NOAA	National Oceanic and Atmospheric Administration
NCEI	National Centers for Environmental Information
GSOD	Global Surface Summary of the Day
AEMet	Agencia Estatal de Meteorología (España)
ECMWF	European Centre for Medium-Range Weather Forecasts
GEFS	Global Ensemble Forecast System
CDS	Climate Data Store (Copernicus)
CNN	Convolutional Neural Networks
KS Test	Kolmogorov-Smirnov Test
M	Month
MAE	Mean Absolute Error
MSE	Mean Squared Error
ROC	Receiver Operating Characteristic (ROC curve)
WMO	World Meteorological Organization

1. Introduction

1.1 Context, purpose and scope

The first conclusions presented by the sixth report of the Intergovernmental Panel on Climate Change (IPCC6, 2021a and b), focusing on the physical bases of climate and climate change, present an indisputable reality: the climatic emergency that the planet is experiencing is directly linked to human activity, being one of its main precursors. No region of the planet has not been affected, to a greater or lesser extent, by climate change-related effects such as melting ice, rising sea levels, or extreme weather events, among others (IPCC, 2021a, IPCC, 2021b). This report is categorical in its conclusions: 'It is unequivocal that human activity has warmed the atmosphere, ocean, and land surface' and that 'global surface temperature will continue to increase at least until mid-century under all possible emissions scenarios.' In other words, even in a best-case scenario, the temperature will continue to rise due to the actions already taken, and it will take at least 30 years to reverse the damage caused. This presents an unprecedented climate emergency, as many of the observed climate changes are without precedent over thousands, if not hundreds of thousands, of years (IPCC, 2021a, IPCC, 2021b). Consequently, the report advises: 1) large-scale, immediate, and forceful mitigation measures to reduce its impact, 2) adaptation measures appropriate to each sector and problem, and 3) using climate research as a tool for climate change adaptation.

Climate change poses a series of risks that affect the development and economy of all countries worldwide, which have been increasingly impacted by extreme weather events linked to climate change in recent decades. It is expected that in the coming decades, temperatures (both maximum and minimum) across Europe will rise by between 2 and 4 °C, with slight changes in the amount of annual precipitation, although with more concentrated precipitation events over time and increasingly intense extreme phenomena. In other words, events such as heatwaves, droughts, and extreme precipitation will occur more frequently, and the socio-economic risks associated with their occurrence will increase considerably.

In order to assess and estimate climate risk, it is necessary to define it according to the extent to which climate impacts affect the sector being assessed. It is important to note that the impact of climate change does not affect all regions or socio-economic sectors equally. Therefore, these impacts must be evaluated at the local scale. To estimate climate risk, climate information projected at the local level is required (future evolution of basic meteorological variables such as temperature, precipitation, or wind), to analyze the impacts and implications of these projections for the sector under study (using indicators adjusted to the needs of that sector) and to identify the potential risks that may arise in the coming decades. Once the impacts are identified, adaptation measures can be defined to address them, minimizing negative effects while attempting to capitalize on positive ones.

For this reason, this study aims to provide an overview of how climate change will affect the evolution of meteorological variables in the coming decades in the Living Labs. To ensure that this study serves as valuable input for local climate information, it should be considered as a support tool when designing adaptation measures in response to climate change. A set of key meteorological variables has been selected to characterize the region climatically. Thus, decisions should be made based on the new climatic conditions expected in the future, rather than relying on past climatic conditions.

In addition, it is essential to integrate both short- and medium-term meteorological forecasts with long-term climate projections. Short-term weather forecasts and seasonal predictions offer critical information for immediate and near-future planning, while climate projections provide an outlook for longer-term trends. The combination of these tools enables more effective adaptation strategies, helping to address both immediate impacts and long-term shifts in climate patterns. By utilizing both short-term and long-term data, decision-making can be more informed and responsive to the varying timescales of climate change.

This document presents the methodology, approaches, and preliminary results of downscaling climate models to the demo site level, effectively translating Global Climate Model (GCM) outputs to a local scale. Furthermore, this document includes a methodological assessment of the approaches considered for enhancing the accuracy and reliability through AI for each timescale model. This prospective methodological analysis will explore the techniques and innovations planned to improve forecasting performance, ensuring that the downscaling process aligns with the specific characteristics and risks of each demo site. It also captures the activities carried out within Tasks T3.2. *Climate modelling and assessment of vulnerability to climate change* and T3.3. *Short-term seasonal weather forecast*.

The report details the collection and analysis of climate data in Front Runner (FR) cities, focusing on the characterization of their current climate conditions by identifying extreme events and climate hazards. Additionally, in shorter timescales, on the one hand, it is outlined the process of selecting and adapting Weather Forecast Models to each FR city, ensuring that forecasting capabilities are optimized for local needs. On the other hand, seasonal forecast methodology and expected results are analyzed for each FR city.

A key aspect of this deliverable is the continuous adaptation and improvement of the operational model using AI-based approaches. The AI-driven methodology will enable a dynamic refinement of the models by integrating real-time observations and continuously learning from past weather patterns. While climate projections are performed as a one-time analysis, the weather forecasting model will be updated daily, ensuring adaptive improvements in predictive accuracy and responsiveness to evolving atmospheric conditions. In terms of seasonal scale, the AI will be implemented within the first stage, or stage-zero, when a set of prediction points will be calculated using convolutional neural networks

(CNN), which will help us to find the most statistically significant ones to build our custom forecast for each point.

Furthermore, this document includes a methodological assessment of the approaches considered for enhancing the accuracy and reliability through AI for each timescale model. This prospective methodological analysis will explore the techniques and innovations planned to improve forecasting performance, ensuring that the downscaling process aligns with the specific characteristics and risks of each demo site.

Around the scope of weather forecast timescale, the present version of this deliverable has not been able to include a complete comparison of all ensemble weather prediction models over the desired period of 2023-2024 due to delays and restrictions in accessing meteorological datasets from official providers. Despite efforts to retrieve historical ensemble forecasts from various national meteorological services, data availability issues and procedural delays in data delivery prevented a full comparative evaluation across the selected models. The specific challenges faced in data acquisition, including slow retrieval rates, lack of response from some providers, and technical API limitations, are detailed in Section 2.2.3 of this report.

To mitigate these limitations, the analysis in this deliverable has focused on the available forecast period spanning December 2024 and January 2025, applying the same statistical methodologies originally planned for the full 2023-2024 period. Additionally, a comprehensive evaluation of the bias correction methodology was conducted using the complete dataset from 2023 and 2024, allowing for an initial assessment of the potential improvement that can be achieved when AI-based postprocessing is applied to the raw ensemble model outputs. This serves as a crucial benchmark for future work, as the next version of this deliverable, scheduled for December 2025, aims to provide a fully justified selection of the optimal ensemble model based on a complete two-year historical dataset, alongside the implementation of optimized AI-based corrections to further enhance forecasting accuracy.

This deliverable is directly linked to Tasks T3.1, T3.2, and T3.3, contributing to the overarching goal of strengthening local climate resilience through high-resolution, AI-enhanced weather forecasting methodologies.

1.1.1 Timescale links and crucial synergies

Extreme events are occurring more intensely and frequently now (WMO,2023), with most of them linked to climate change through attribution studies. It is crucial to take advantage of the synergistic feedback that exists in decision-making processes when considering all timescales—climate projections, seasonal forecasting, and weather forecasting—to support optimal decision-making and enable more accurate protection and adaptation measures.

Through local downscaled climate change projections, pilot cities can identify warning signals that may pose risks to specific urban areas or planned Nature-Based Solutions (NBS) for ecosystem services in the long term, across different climate scenarios. These signals typically consist of co-designed indices explicitly tailored to each location.

Once long-term warning signals are identified, the next step is to translate and adapt these indices to shorter timescales—seasonal and weather forecasts—ensuring, in collaboration with pilot cities and stakeholders, that these indices align with the short-term alerts they expect to receive. This approach enables bidirectional work with these indices, both forwards and backwards in the timescale.

For strategic planning of adaptation measures, historical meteorological data can be analyzed to determine how often the defined index thresholds have been exceeded over the past five years. This provides an initial assessment of exposure to potential impacts that each city and its proposed measures might currently face—typically lower than what is projected for the coming years. Moreover, this threshold analysis can serve as a preliminary assessment to determine the most probable climate change scenario for the coming decades.

Focusing on adaptation and protection measures for NBS and urban infrastructure, meteorological and seasonal forecasts can be leveraged as an early warning system to anticipate impact risks using the co-designed indices. Advanced impact forecasting allows cities to develop proactive protection measures to safeguard key urban elements.

In the long term, this tool can be used to assess which climate change scenario reality is aligning with. By monitoring and recording local observations, cities can track how extreme values evolve over time for each index. Comparing climate simulations of different scenarios with daily meteorological observations will allow for regular updates and reviews, helping cities determine which scenario they are most likely experiencing and adjust their planning accordingly.

1.2 Approach and relation to other Work Packages and deliverables

The approach adopted in this deliverable is designed to show schematically the main information required for understanding the proposed methodologies, initial evaluations and preliminary results for each timescale. For this purpose, each timescale is compiling an overview of its current state of the art, including a summary of the key findings and gaps and an overview of the relevant competing market products. This document also compiles the data requirements for each timescale, with particular focus on execution data and model development data on one hand, while on the other, it highlights the limitations of available data sources. These limitations encompass both existing datasets and outline

the next steps. Around results, in this document are compiled both, the input data analysis and the expected outputs (preliminary results).

Based on all information gathered, conclusions are summarizing the findings, next steps and future development plans. This deliverable will be followed by another, which will present the final results and complete the methodology development.The methodology will then be applied successively to the rest of the pilot cities.

Relation to other work packages and deliverables

Naturally, this deliverable is strongly aligned with WP3 (URBREATH Data Strategy and Tools), particularly the following tasks:

- **T3.1 - AI-based algorithms and tools:**

  The AI methodologies presented in this deliverable contribute to the development of high-speed big data clustering, ML-based classification models, and high-level data fusion services. These components are critical for improving the downscaling of climate models and operational weather forecasting at the city level.

- **T3.2 - Climate modelling and assessment of vulnerability to climate change:**

  The document provides insights into localized climate modeling for FR cities, supporting the identification of vulnerability hotspots, extreme event trends, and long-term climate risks. This enables an integrated approach to climate adaptation strategies.

- **T3.3 - Short-term seasonal weather forecast:**

  The adaptation of weather forecast models using AI techniques aligns with T3.3, ensuring that seasonal and daily forecasts reflect the latest atmospheric conditions. The deliverable also provides a methodological framework for integrating climate projections, seasonal outlooks, and real-time weather forecasting into a cohesive operational model.

This deliverable also supports the objectives of other Work Packages:

WP4 - URBREATH decision-making framework:

- The improved climate and weather models will provide data-driven insights for the Local Digital Twin framework (T4.1), impact visualization tools catalogue (T4.3), and will be an important input for adjusting the URBREATH NBS's (T4.4).
- The predictive outputs will inform the URBREATH ICT integrated solution (D4.7, D4.8), ensuring that NBS decision-making is guided by high-quality climate projections and also for current extreme events risks by weather and seasonal forecasts.

WP5 - Local Living Labs (LLL):

- This deliverable contributes to the URBREATH KPIs multi-level impact assessment framework (D5.11, D5.12), ensuring that climate risk assessments align with pilot city-specific baselines and adaptation goals.
- The weather forecast combined indices will also be validated and iterated within the Living Labs, allowing for stakeholder-driven adjustments to forecast outputs.

WP6 - Urban greening and renaturing actions:

- Climate and weather predictions will support real-time NBS monitoring (D6.2, D6.3), ensuring that deployed solutions are optimized for changing environmental conditions.
- The threshold-based climate risk assessments in this deliverable will help define the institutional and technical frameworks for NBS implementation (D6.4, D6.5, D6.6).

WP7 - Scaling, knowledge transfer, and standardization:

- The standardization of climate and weather modeling approaches developed in this deliverable will contribute to the URBREATH Global Roadmap (D7.11), ensuring that forecasting methodologies can be replicated and scaled across European cities.
- Additionally, the AI-driven climate risk insights will be included in the Follower Cities Roadmap (D7.10), fostering knowledge transfer and best-practice sharing.

This deliverable bridges multiple WPs by providing high-resolution climate seasonal and weather insights, enabling improved risk assessments, better-informed urban planning, and proactive adaptation strategies. Its methodological innovations in AI-based forecasting not only refine local climate predictions but also support digital decision-making frameworks, Living Labs, and NBS deployment efforts.

1.3 Structure of the deliverable

This deliverable is structured in five main sections. Section two details the data requirements and sources for the study across various timescales, and outlines the available databases and final selections

for each front-runner city. The third section consists of a description of all the methods identified for addressing the requirements of URBREATH in terms of climate change projections, seasonal and weather forecasts. The fourth section presents the preliminary results of climate projections or how the best weather forecast model can be selected for each city. Finally, a section of Conclusions (5th) is presented to remark on the main findings and future steps. A single annex section is included to provide an in-depth discussion on the selection of observatories.

2. Data and study area

This section gathers, on one hand, the information related to the various databases required for the study across its different timescales, and on the other hand, it presents the available databases and those ultimately selected for each of the front cities.

2.1 Study Area

The URBREATH project plans to work in four climatic areas, selecting one main city (FR) and several secondary cities for each of them. The case studies are:

- The Mediterranean region is represented by Madrid (Spain), Parma (Italy) and Athens (Greece).
- The Atlantic region comprises Leuven (Belgium) and Aarhus (Denmark).
- The Boreal region is working in Tallinn (Estonia) and in Kajaani (Finland).
- The Continental region's Living Labs are Cluj-Napoca (Romania) and Pilsen (Czechia).

In this first phase, the studies have focused on the four FR cities: Madrid, Leuven, Tallinn, and Cluj-Napoca.

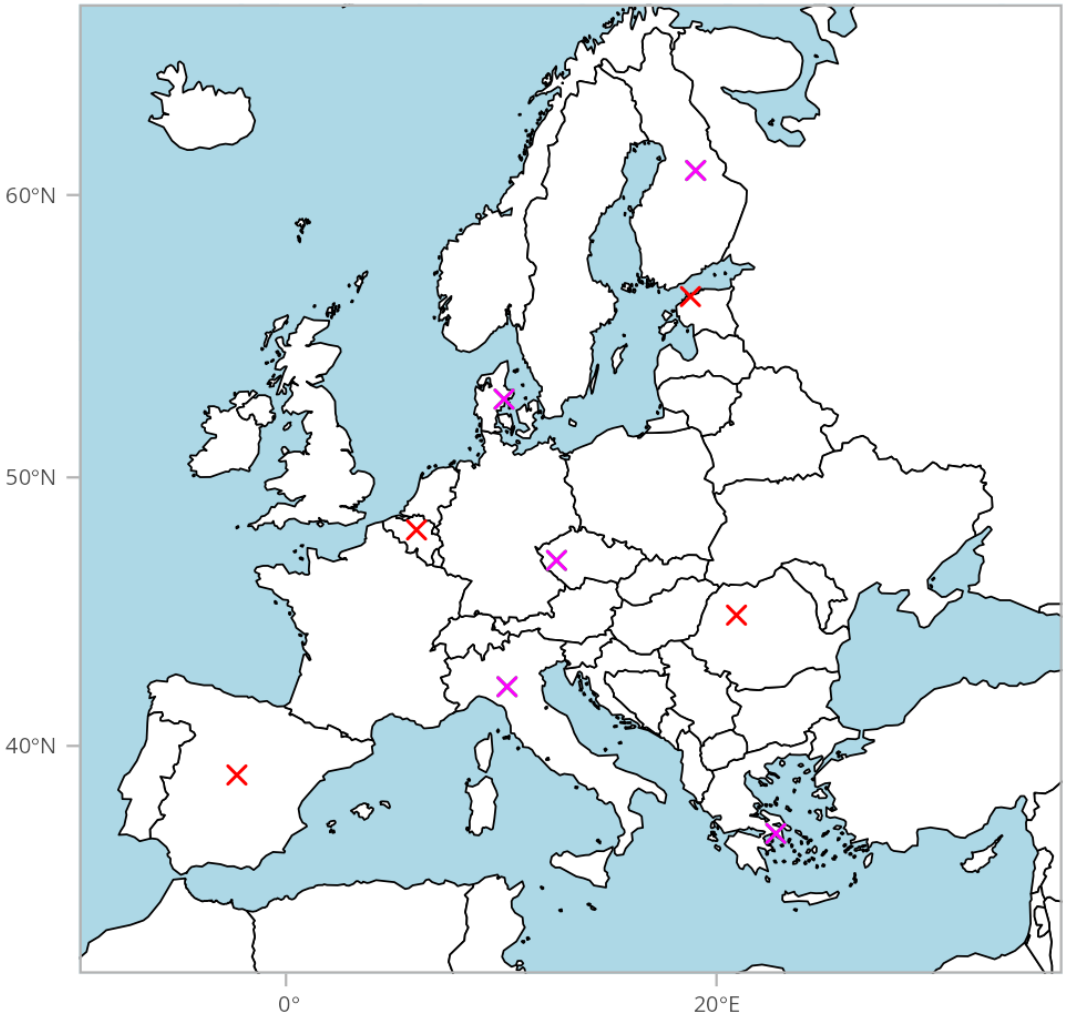

Figure 1. Map of Europe showing the Front Runner cities (in red) and follower cities (in purple) marked with a cross.

2.2 Data requirements

Section 2.2. outlines the meteorological and climatic data needs, both historical and simulated, for climate projections and short- to medium-term predictions.

2.2.1 Observed data

In the generation of meteorological and climatic information, it is essential to have historical meteorological data collected in the FR city areas. The need for this information arises from the necessity of incorporating microclimatic effects that are beyond the reach of models due to their spatial scale (approximately 5 km in weather forecast, and about 100-120 km in climate scale). Obtaining this information is, in most cases, a significant challenge, since access to this data is often complicated, limited in many cases by access to official sources, as well as the poor quality of the data itself.

The main required variables are maximum and minimum temperature, precipitation, wind and relative humidity and snowfall.

2.2.2 Weather forecast

The implementation of ensemble-based probabilistic forecasting within the URBREATH project requires a dual approach to data acquisition. On the one hand, a comprehensive historical dataset is essential for evaluating and selecting the optimal ensemble prediction system (EPS) for each city. On the other hand, continuous operational forecasts are needed to feed our AI-enhanced predictive framework, ensuring the integration of real-time probabilistic forecasts into the URBREATH decision-support system.

To determine which EPS model performs best for urban applications, it has been required historical ensemble forecasts from all available models with data covering at least 2 years. This dataset is used to conduct statistical evaluations against observed meteorological records, focusing on main weather variables with hourly data for optimised forecasts. The weather variables required are 2-meter temperature, maximum temperature, minimum temperature, relative humidity, and accumulated precipitation (including snow) forecasts.

These historical ensemble datasets are then statistically compared with observed data, which has also been successfully retrieved from official meteorological sources. The goal is to quantify bias, ensemble spread, and overall model skill for each available EPS. The results of this evaluation will determine which ensemble system is best suited for operational urban weather prediction in the framework of URBREATH.

Beyond historical forecast needs, the operational deployment of the URBREATH AI-enhanced prediction model requires real-time ensemble forecasts to drive the system daily. The selected ensemble system must provide daily updates to stand the probabilistic forecasts, serving as the input data for the AI post-processing methods outlined in the methodology section.

The real-time ensemble forecast data will be used to run our enhanced urban-scale probabilistic predictions, incorporating machine learning bias correction, uncertainty quantification, and high-resolution downscaling techniques. This ensures that the URBREATH model not only has been selected as the best available ensemble system but also improves its predictive capabilities through advanced AI methodologies.

The integration of observed meteorological data is crucial for both the evaluation of EPS and the enhancement of operational forecasts through artificial intelligence methodologies. Then, to assess and identify the most suitable EPS for urban weather forecasting, it is imperative to compare ensemble model outputs against reliable historical observational data. For the period spanning 2023 to 2024, it has successfully acquired observed data from official sources, which will serve as the benchmark for evaluating the performance of various ensemble models.

2.2.3 Seasonal forecast

The seasonal forecasting method requires exhaustive long data to work: a predictor or range of predictors that provide a way to describe the explained variance of predictands or surface variables that are expected to be forecasted. In a classical way, it has been explored the use of teleconnection indices to estimate the variability and predict the variables, but, in this case, and due to: (1) the wide range of climates in the European continent, (2) the incorporation of AI and CNN, it was decided to adopt a new way to define the indices used. Nevertheless, a minimum of 50 full years is highly recommended to train the CNN-based models. The predictors also have to cover the main sources (or sinks) of energy in the atmosphere and their transport along the Coriolis-driven currents to higher latitudes, so several types of predictors have to be selected (e.g. sea surface temperature or latitudinal and zonal wind). However, an excessive number of predictors must not be exceeded because of the risk of overtraining, and to achieve this a precise selection must be made before training the model.

Several statistics are used in the "zero stage" to select the coordinates that best describe the explained variance of the surface predictors at the monthly scale: Pearson correlation, KS or Kolmorov-Smirnof test, and median absolute error (MAE) and median square error (MSE), and they have to be applied within the training stage.

Once the zero step is done, the next steps will still require the best coordinate predictor data to train the second and third steps (predictor autoregressive approach and wavelet-arima approach, respectively). This step, unlike the CNN training, will be needed to perform the final and updated prediction every month. In this sense, it will also be necessary to have the updated data of the predictors on a monthly scale, which is why a stable platform such as the Climate Data Store of Copernicus (CDS) has been chosen to satisfy the seasonal forecasting requirement.

Regarding the observation needs, Global Surface Summary of the Day (GSOD) was chosen because of the ease of downloading and the wide availability of data. Although observations are crucial for training the neural networks in stage zero, they are also required for stages one and two, where they are assimilated into the autoregressive and wavelet-Arima approaches. These steps are performed every month to generate the latest updated seasonal model.

2.2.4 Climate projections

To conduct a detailed climate analysis with the proposed methodology (see section 3.3), it is essential to have a set of data which includes:

- Historical daily observed data of temperature and precipitation with the longest possible length. This information provides real and accurate information about past climate conditions.
- A reanalysis which is a reconstruction of meteorological data using both observations and numerical models, offering a continuous and consistent representation of past climate.
- A set of climate models which are mathematical representations of the climate system that simulate how climate parameters evolve based on various factors, such as greenhouse gas concentrations, solar radiation, and other components of the climate system.

The following figure (Figure 2) outlines the necessary data as well as the part of the process in which they are required. The final sources of the selected observed data are analyzed exhaustively in Section 2.3.1, as well as the reanalysis and climate models in Section 2.3.4.

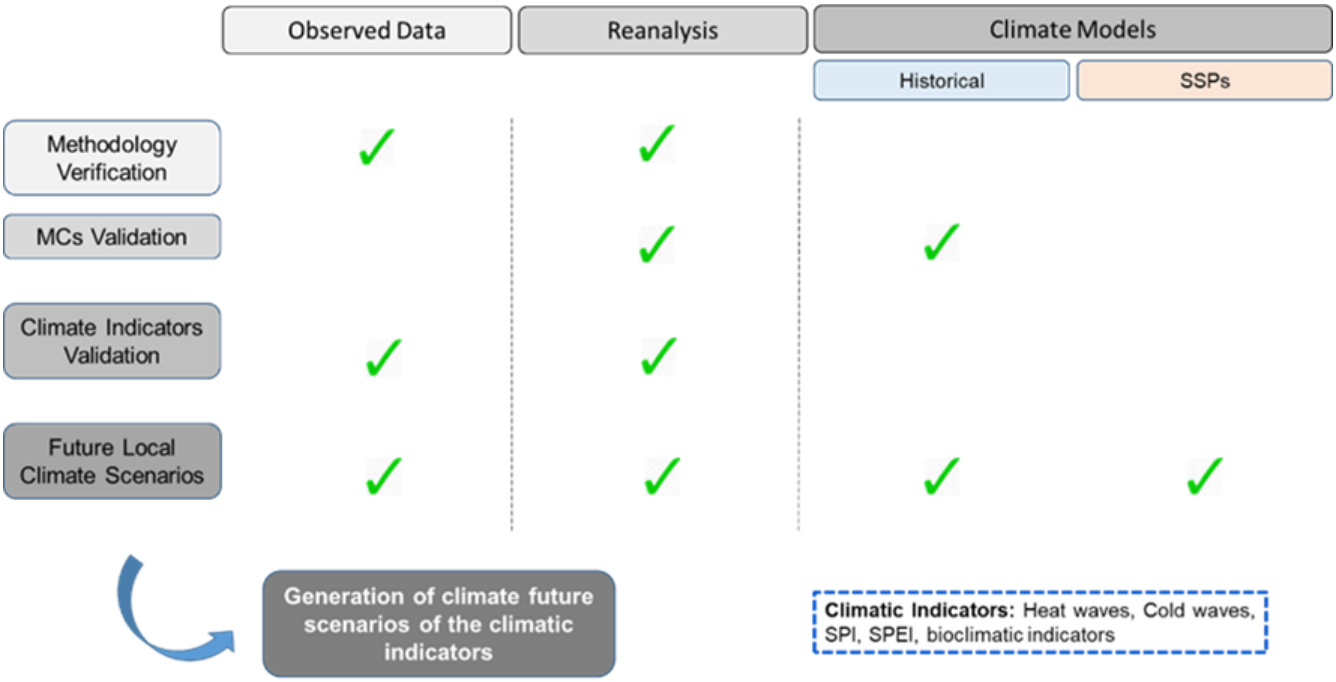

Figure 2. Outline of data required for the climatic study.

2.3. Data Sources Selected

Section 2.3. outlines the databases used in the study, from observed data to the predictive models, including both meteorological and climatic forecasts. It also describes the data treatments applied and the challenges encountered during the process.

2.3.1 Observed data

Beyond historical data, the implementation of AI-driven improvements in our forecasting methodology requires access to up-to-date observational data. These real-time observations are essential for:

- **Model training and calibration:** Continuously refining AI algorithms to enhance predictive accuracy.
- **Bias correction:** Adjusting ensemble forecasts to mitigate systematic errors.
- **Validation:** Assessing the performance of AI-enhanced forecasts in near real-time.

To fulfill this requirement, and after a thorough analysis of the data availability and other open data policies from several national weather agencies, it has been decided to establish a data acquisition

pipeline straight from the GSOD dataset, a complete repository of weather observations provided by the National Center for Environmental Information (NCEI) of the National Oceanic and Atmospheric Administration (NOAA) of the United States of America. The GSOD dataset collects weather data from the primary networks of weather stations ruled by the national (and also regional) weather agencies of most of the countries around the world, testing the data and storing it for public access. It offers **daily** (and in some cases even hourly) summaries of key meteorological variables such as mean temperature, precipitation and wind speed. Data is typically available within two to three days after observation, ensuring timely integration into our forecasting models.

A custom script has been developed to automate the daily retrieval of GSOD hourly data from the NCEI portal, accessible at its webpage[1]. This automated process ensures a consistent and up-to-date influx of observational data, which is critical for the operational application of our AI methodologies.

In addition to the GSOD network, data from other regional or local entities were considered, such as three municipal observatories managed by the Madrid City Council and one station operated by the Spanish Meteorological Agency (AEMET) for the Madrid case, or other extra stations from the weather agencies of Estonia not available at GSOD for the Tallinn case. These local data sources provide high-resolution observations specific to the urban area of the cities, thereby enhancing the granularity and relevance of the data used in our models. See the map identifying the location of each weather station in figures 7 to 10. Additionally, a table summarizing the observed data used for weather forecast modelization is attached in Annex 1.

For each case study, the available observatories with daily temperature and/or precipitation values are compiled here at Table 1, for more detailed analysis see on the Annex 1 (Table A.1, A5, A9 and A13).

[1] GSOD webpage at NOAA: https://www.ncei.noaa.gov/data/global-hourly/access/

Table 1. Compilation of available observatories with climatic series of temperature and precipitation and sources.

City	*Source*	*Temperature*	*Temperature/ precipitation*	*Precipitation*	*Total temperature*	*Total precipitation*
Cluj-Napoca	GSOD	-	4	-	4	4
Leuven	GSOD and Leuven.cool	-	92	11	92	103
Madrid	GSOD, AEMet and Madrid City Council	15	29	2	16	31
Tallinn	GSOD and ILM	1	5	1	5	6
Total	-	**16**	**130**	**14**	**117**	**141**

Maps showing the location of these observatories can be found in section (a) in Figures 7, 8, 9 and 10 for Cluj-Napoca, Leuven, Madrid and Tallinn, respectively.

The values of these observatories are analysed in terms of quality and homogeneity of their data to determine their possible use in the different applications: weather forecast, seasonal forecast and climate projections.

The **quality control** of an observed meteorological data series consists of developing a set of tests to ensure that the data are consistent within the study series itself. The two main automatic controls to be used in quality control are:

- **Basic consistency**. Direct rejection of manifestly erroneous values, e.g. negative precipitation.
- **Outliers**. Unusual outliers within a given data set, i.e. values that appear to come from different data sources or to have been generated in a different way from the rest of the data. In this case, their detection will come from our definition of 'atypical'.

In the case of **temperature**, the **basic consistency** has been to look for daily values where the maximum temperature was lower than the minimum temperature. This occurs when the temperature data makes sense on its own but not when compared to the other daily temperature series; in such a case, either

series could be the wrong one, which makes it necessary to reject the daily data for both, replacing them with NA values (no data). On the other hand, there are two types of outliers: **extreme values** that are not known to have been exceeded by official records and **values above the deviation value**.

The range of values set as valid for temperature values - and thus to be able to detect and correct for extreme outliers above or below - are, for each case study:

- **Cluj-Napoca**: between -25°C and 44.5°C for maximum temperature and between -39.6°C and 30°C for minimum temperature.
- **Leuven**: between -25°C and 42.5°C for maximum temperature and between -22°C and 30°C for minimum temperature.
- **Madrid**: between -5°C and 42.7°C for maximum temperature and between -17°C and 33°C for minimum temperature.
- **Tallinn**: between -30°C and 35.6°C for maximum temperature and between -43.5°C and 23°C for minimum temperature.

In order to detect the second type of outlier from a theoretical point of view, it is assessed how far the outlier deviates from the typical values of the series. The formal way to perform such a test consists in determining how many times a certain data deviates from the standard deviation of the total observed series in question, thus detecting the values above the deviation value. Therefore, the tests will have to determine:

- The standard deviation and the monthly mean of each observatory.
- A 'deviation factor' indicating the number of units of standard deviation from the mean above which a certain daily value can be flagged as an outlier. In this sense, a deviation factor of 8 has been applied.
- Calculation of a 'deviation value' that should not be exceeded for each monthly distribution:

$$deviation\ value = m \pm deviation.factor \times std$$

  Where m corresponds to the mean of the month of the value to be studied, std to the standard deviation of the same month and deviation. factor to the number of standard deviation units established in the second step.
- A direct study of the data flagged as outliers by the previous step to determine whether they are true and should therefore be discarded or whether a new threshold value should be reassigned and the study repeated.

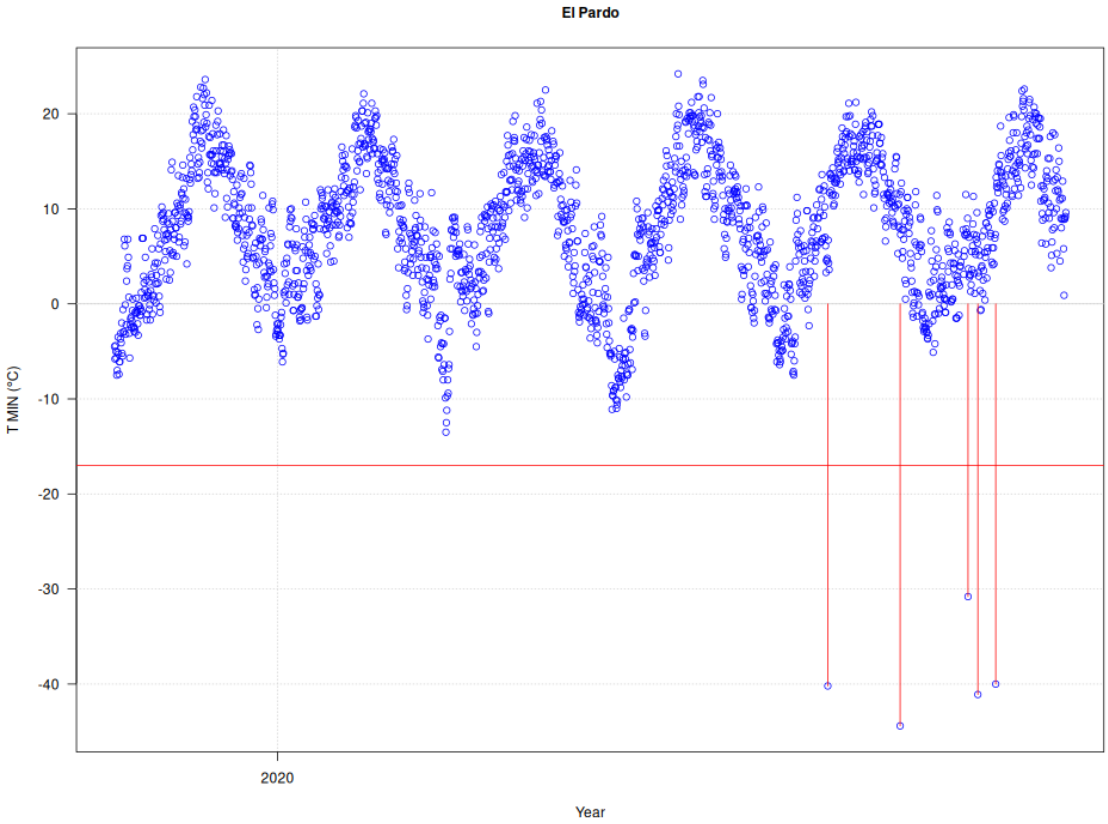

Figure 3. Resulting quality control for the minimum temperature series of the observatory with identifier '58-El Pardo' for the city of Madrid. Outliers of the extreme values type are detected.

In the case of **precipitation**, the basic consistency check involves searching for negative daily values. Given the obviously erroneous nature of these cases, all such detected cases are immediately replaced by NA (no data) values.

Attending outliers, as for the temperature variable, there are two types. In order to detect **extreme outliers**, precipitation observatories of official networks are analysed. Thus, the following extreme values of maximum daily accumulated precipitation are defined for each case study:

- **Cluj-Napoca**: 98 mm.
- **Leuven**: 242 mm.
- **Madrid**: 124 mm.
- **Tallinn**: 85 mm.

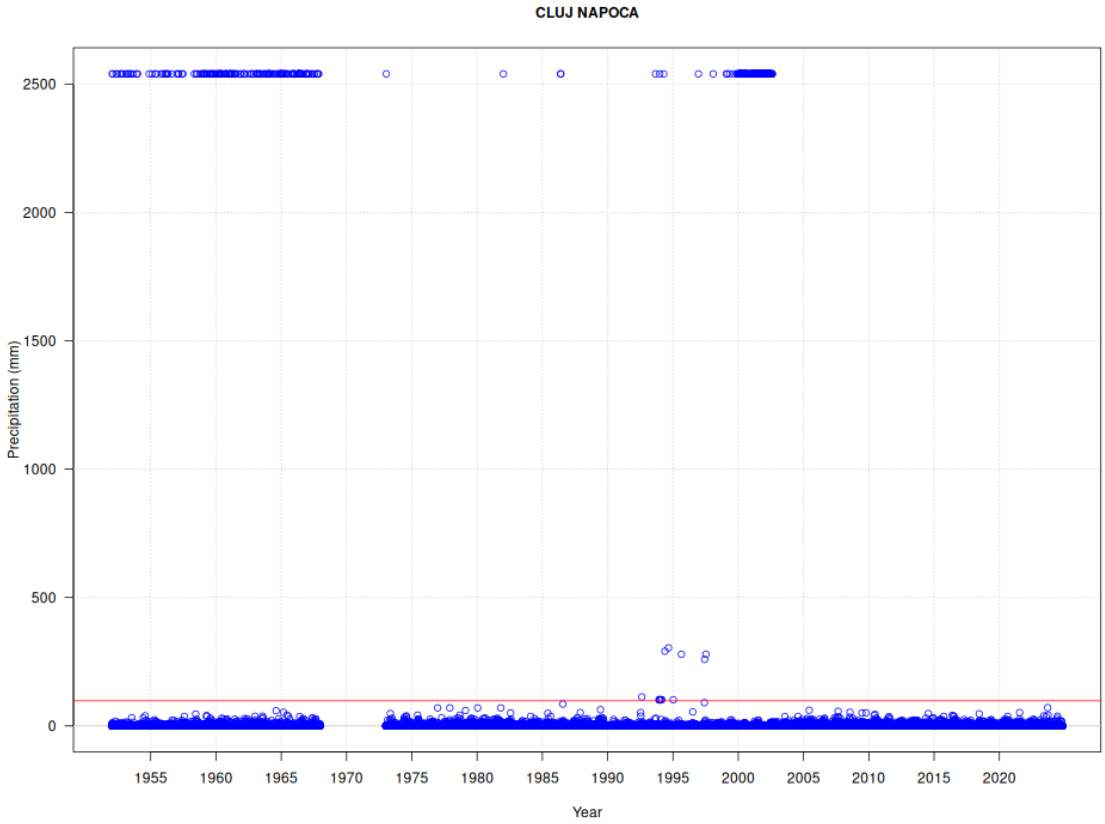

Figure 4.Resulting quality control for the precipitation series of the observatory with identifier '15120099999' for the city of Cluj-Napoca. Outliers of the extreme values type are detected.

On the other hand, values higher than the deviation value are detected by exceeding such data a pre-set number of times the standard deviation of the whole series (1), in this case the deviation factor is 2. However, and given the nature of this meteorological variable, its detection does not imply its automatic elimination, as is the case with temperature, since extreme precipitation is a phenomenon that, although unusual, is not necessarily impossible, and therefore requires a manual review that compares it with the climatology of the area.

Once quality control has been carried out, the resulting values are analysed for each observatory. Some are found to have very unrealistic records and/or too much missing data within the time length. In addition, observatories must have records of more than two years to be able to proceed to the next step.

The **homogenisation** of the data of a time series refers to the quality control of the data as elements of a time series, i.e. it studies the possible coherence of the data exactly in the order in which they are presented. Note that the above checks could be performed on the same series but in a disordered sequence; however, they do not provide information about the possible time variability of the data, which is almost always linked, at the very least, to annual cycles. The homogeneity test that has been used– applied only to observatories with at least two years of data - operates as follows:

- In order to measure how similar one year is to another, a Kolmogorov-Smirnov distribution comparison test, a non-parametric statistical test (which does not assume distributions of the variable to be studied, therefore) was performed, providing a p-value that can be used as a measure of the similarity between two years. Note that this first part only indicates similarities between consecutive years and is no more than a preliminary mark on the possibility of the existence of inhomogeneity.
- If a certain year has been flagged as a possible indicator of inhomogeneity then it is subjected to a more generic test. Once the cut-off year and the following year (known to be of different distribution and assumed to represent different periods) have been marked, the p-value of each of the years in the series is determined with respect to those two years. If there is a jump or a break between all these p-values in the years considered, then a true inhomogeneity for the whole series is established.

This approach makes it possible to identify the specific years in which inhomogeneity appears in a series. Since establishing how small a p-value must be to indicate potential inhomogeneity is a matter of judgement, the same test has been run several times with different p-value thresholds (from very negative values to those closer to 0). This helps eliminate the subjectivity of the criterion, since inhomogeneity, if true, should appear in most of the test runs.
Finally, once certain homogeneity gaps have been detected, a manual evaluation is necessary to correct the series, as there may be a great variety of results and, therefore, different criteria to act according to them.

In the case of **temperature**, the maximum and minimum temperature series are treated separately. In order to automate the process, the following situations and actions are taken into account for each homogeneity jump, in which the most recent records are considered as valid temporally: two subseries are differentiated, one from the end of the series to the most recent homogeneity jump and another from the most recent jump to the previous jump or, in case there is no jump, to the beginning of the series. The averages of the two sub-series are then compared (see Figure 5):

- If the mean difference is less than 3°C, the series is not modified*.
- If the mean difference is greater than or equal to 3°C:
  - If the length of the most recent sub-series is greater than half the total length of the series, everything before it is removed, leaving only that last sub-series in the observatory record.
  - If the length of the most recent subseries is equal to or less than half the total length of the series, the series is not modified[2].

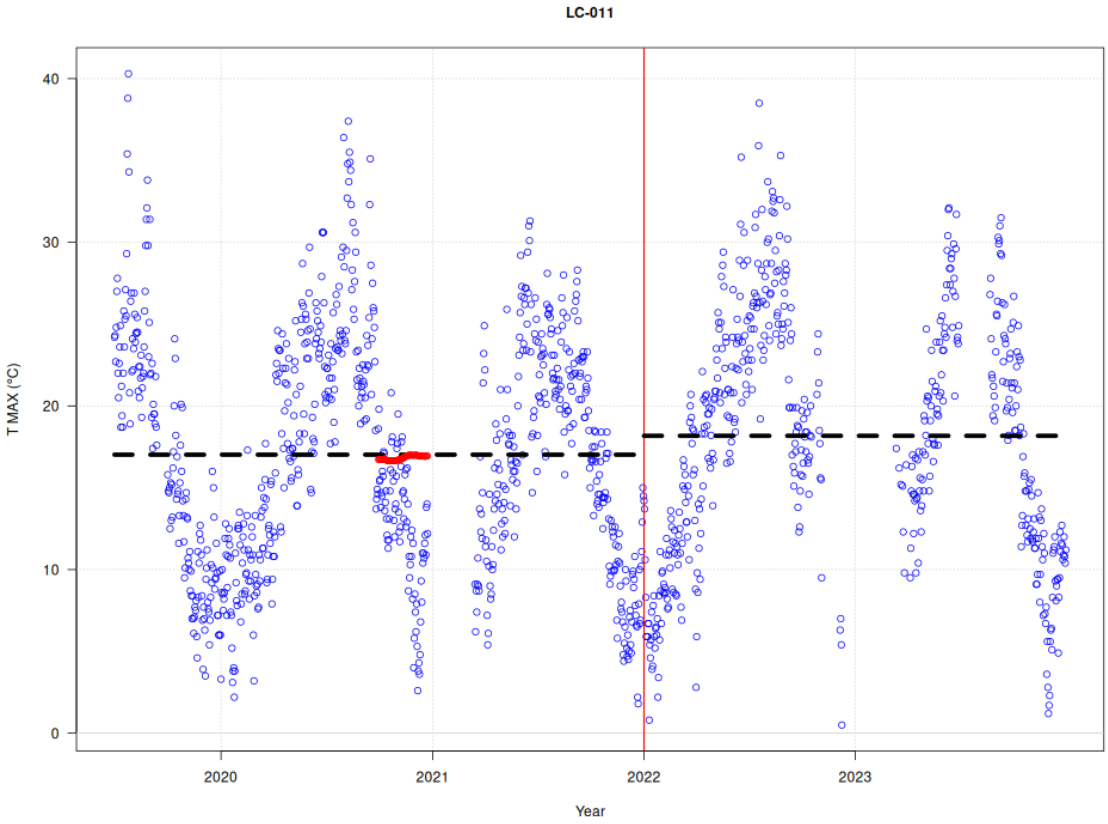

Figure 5.Resulting homogeneity check for the maximum temperature series of the observatory with identifier 'LC-011' for the city of Leuven. The test detects a homogeneity jump with a mean difference of less than 3°C.

[2] *In case of more than one homogeneity jump, the process would be repeated by considering two new sub-series, one from the end of the series to the jump before the most recent jump (temporarily) and one from that jump to the previous jump or, if there is no previous jump, to the beginning of the series*

In the case of **precipitation** (Figure 6), the correction is carried out completely manually due to its complexity. The process is not automated.

During the quality and homogenisation processes, certain observatories that do not meet the conditions required by the tests are eliminated. Finally, for each case study, the results are compiled here at Table 2, for more detail see on the Annex 1 (Table A.4-A8, A12 and A16).

Table 2. Compilation of selected observatories with climatic series of temperature and precipitation.

City	*Temperature*	*temperature/ precipitation*	*Precipitation*	*Total temperature*	*Total precipitation*
Cluj-Napoca	1	3	-	4	3
Leuven	-	3	-	3	3
Madrid	1	15	3	16	18
Tallinn	1	4	1	5	5
Total	**3**	**25**	**4**	**28**	**29**

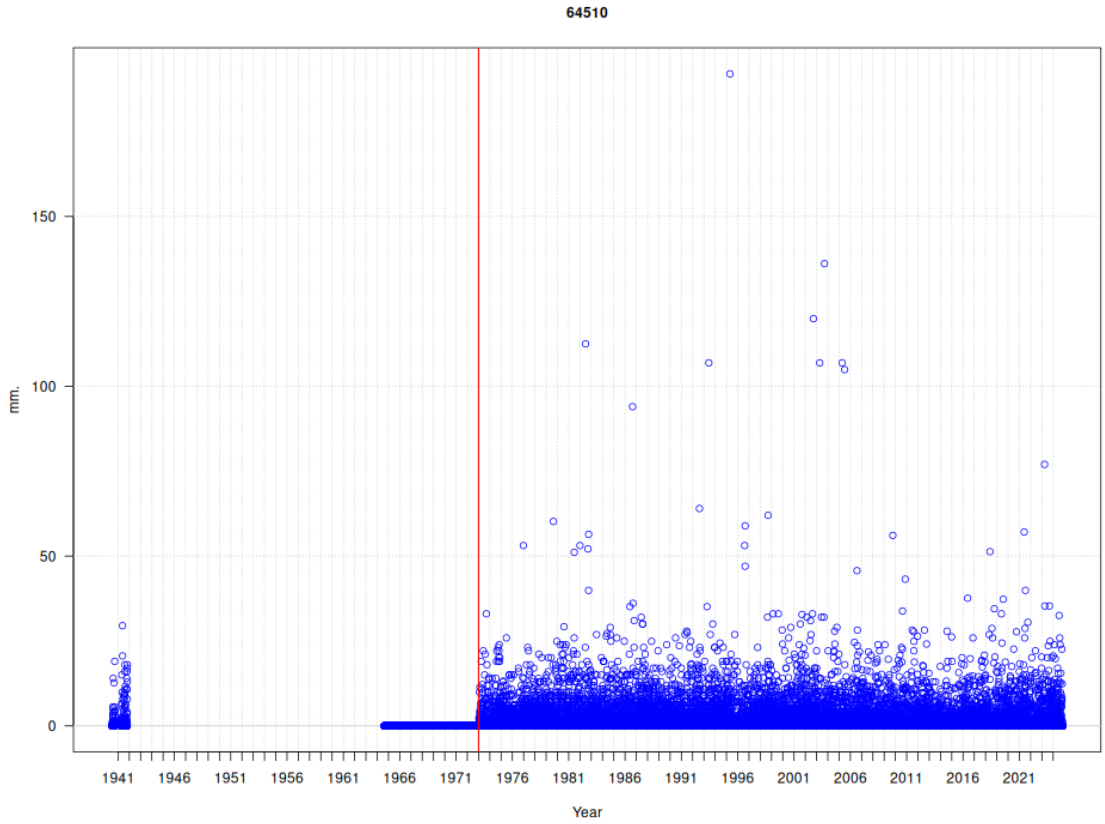

Figure 6. Resulting homogeneity check for the precipitation series of the observatory with identifier '64510' for the city of Leuven.

All observatories that have undergone the controls are employed in the study of climate projections (section (d) in figures 7, 8, 9 and 10). However, due to not updating data periodically enough, not all of them have been selected for the weather and seasonal forecast study (sections (b) and (c) in figures 7, 8, 9 and 10). For each case study, the results are compiled here at Table 3, for more detail see on the Annex 1 Tables A.2, A.3 A6, A7, A10, A11, A14, A15.

Table 3. Compilation of not selected observatories with climatic series of temperature and precipitation. First value is referred to the weather forecast and the second is the seasonal forecast.

City	*Temperature*	*temperature/ precipitation*	*Precipitation*	*Total temperature*	*Total precipitation*
Cluj-Napoca	-	1|1	-	1|1	1|1
Leuven	-	2|1	-	2|1	2|1
Madrid	-	2|1	-	2|1	2|1
Tallinn	-	2|1	-	2|1	2|1
Total	**0**	**7|7**	**0**	**7|7**	**7|7**

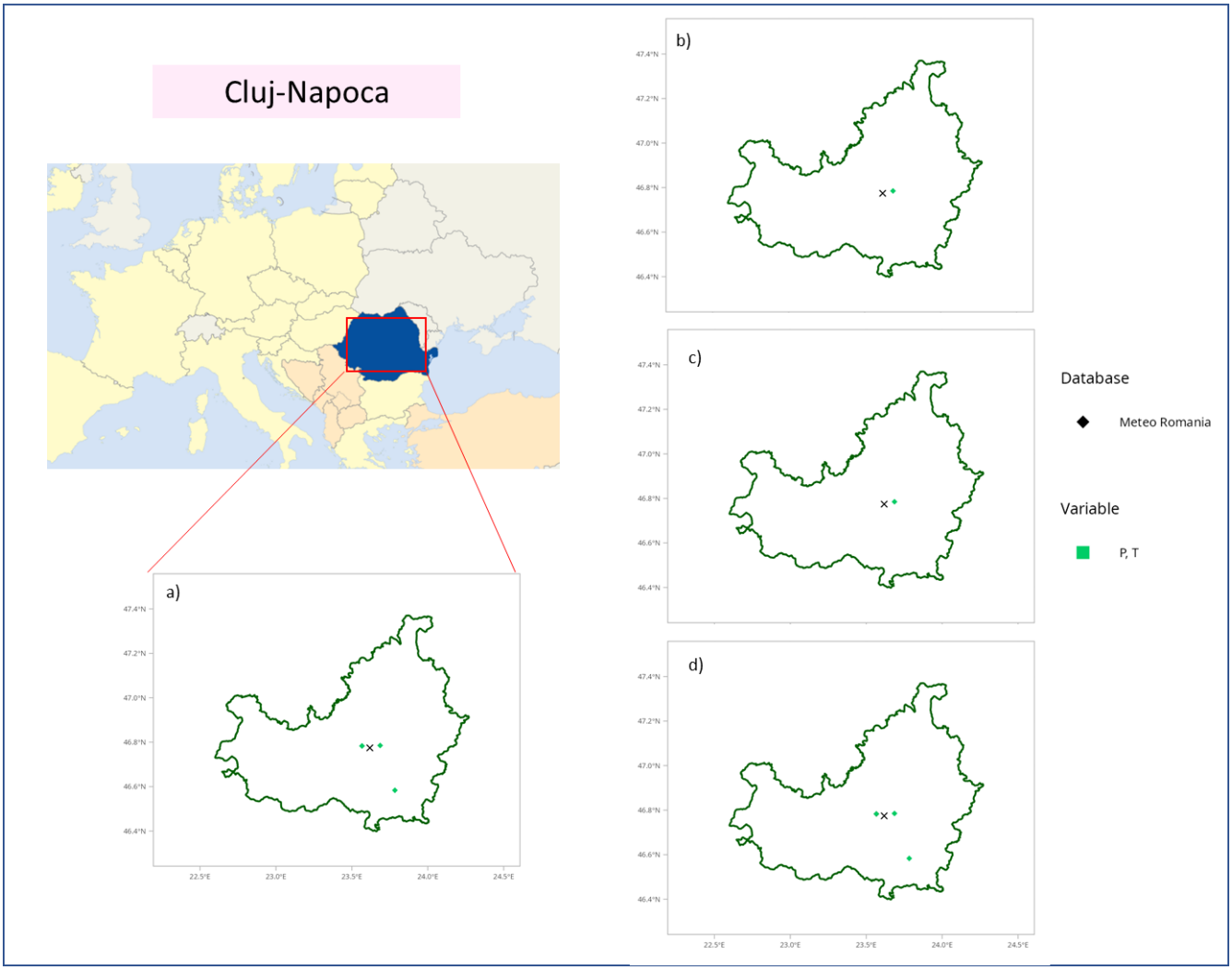

Figure 7. The figure shows a) observatories available at the start of the study, as well as the observatories ultimately used in the study for b) weather forecast, c) seasonal forecast and d) climate projections for Cluj-Napoca. **Observatories with temperature data are shown in red, those with precipitation data in blue, and those with both variables in green. Meteo Romania is the Romanian weather agency. The district of Cluj in Romania, whose capital is Cluj-Napoca, is shown in green, with the study area marked with a cross**

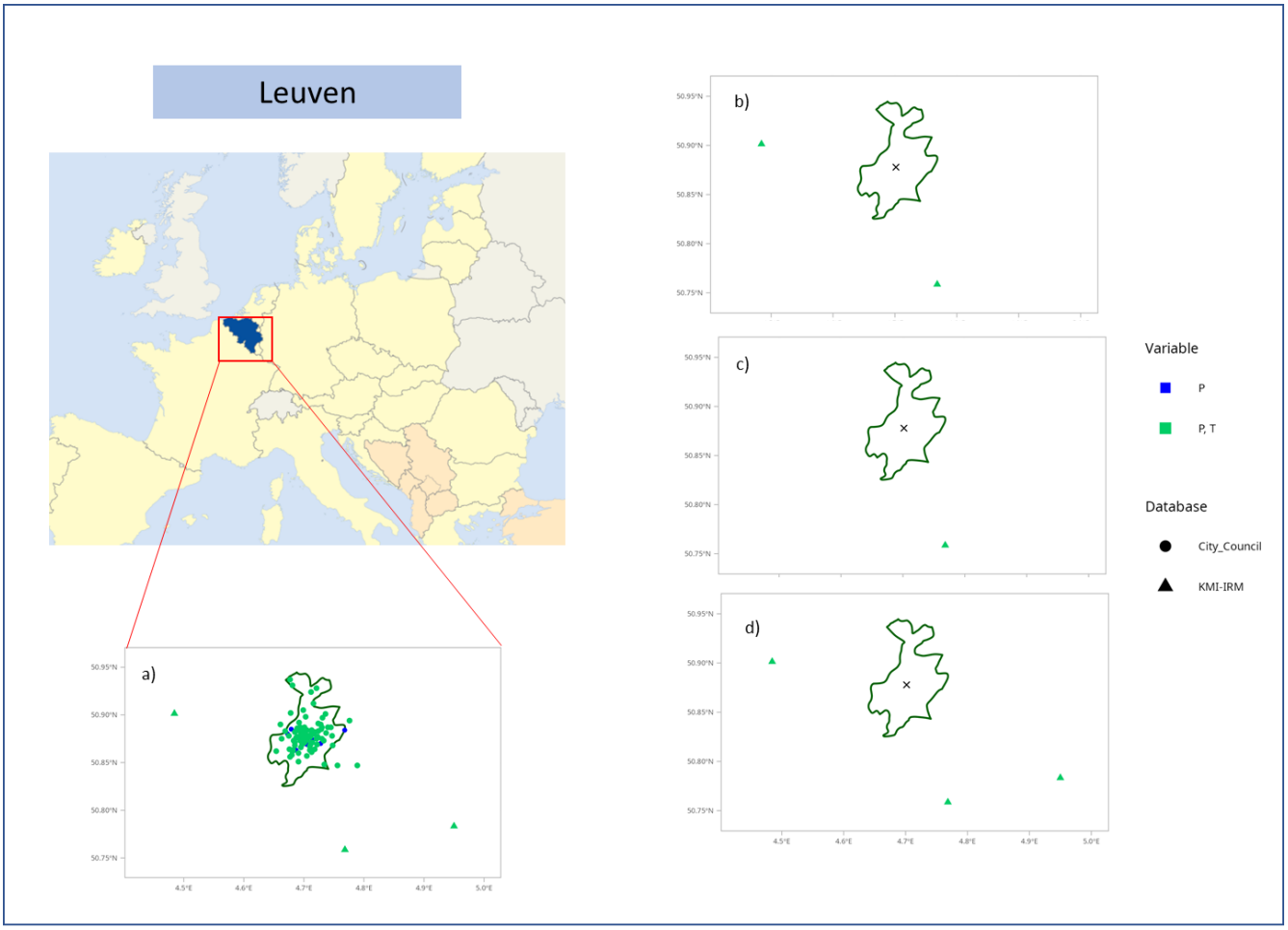

Figure 8. The figure shows a) observatories available at the start of the study, as well as the observatories ultimately used in the study for b) weather forecast, c) seasonal forecast and d) climate projections for Leuven. **Observatories with temperature data are shown in red, those with precipitation data in blue, and those with both variables in green. KMI-IRM is the Belgian weather agency. The city boundary of Leuven is shown in green, with the study area marked with a cross.**

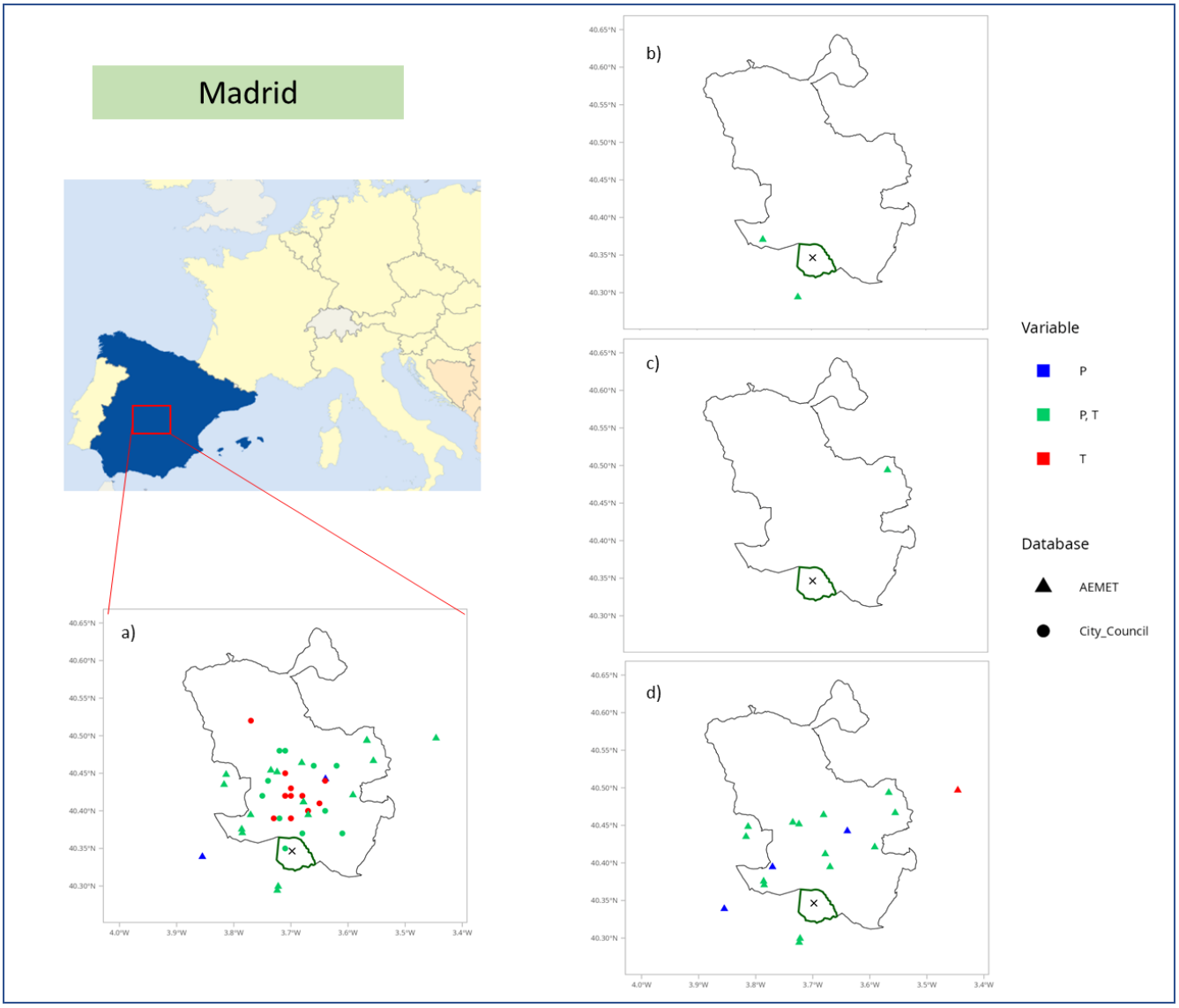

Figure 9. The figure shows a) observatories available at the start of the study, as well as the observatories ultimately used in the study for b) weather forecast, c) seasonal forecast and d) climate projections for Madrid. **Observatories with temperature data are shown in red, those with precipitation data in blue, and those with both variables in green. AEMet is the Spanish weather agency, whereas City Council stands for Madrid's weather monitoring system. The municipality of Madrid is shown in black, and the district of Villaverde is shown in green, with the study area marked with a cross.**

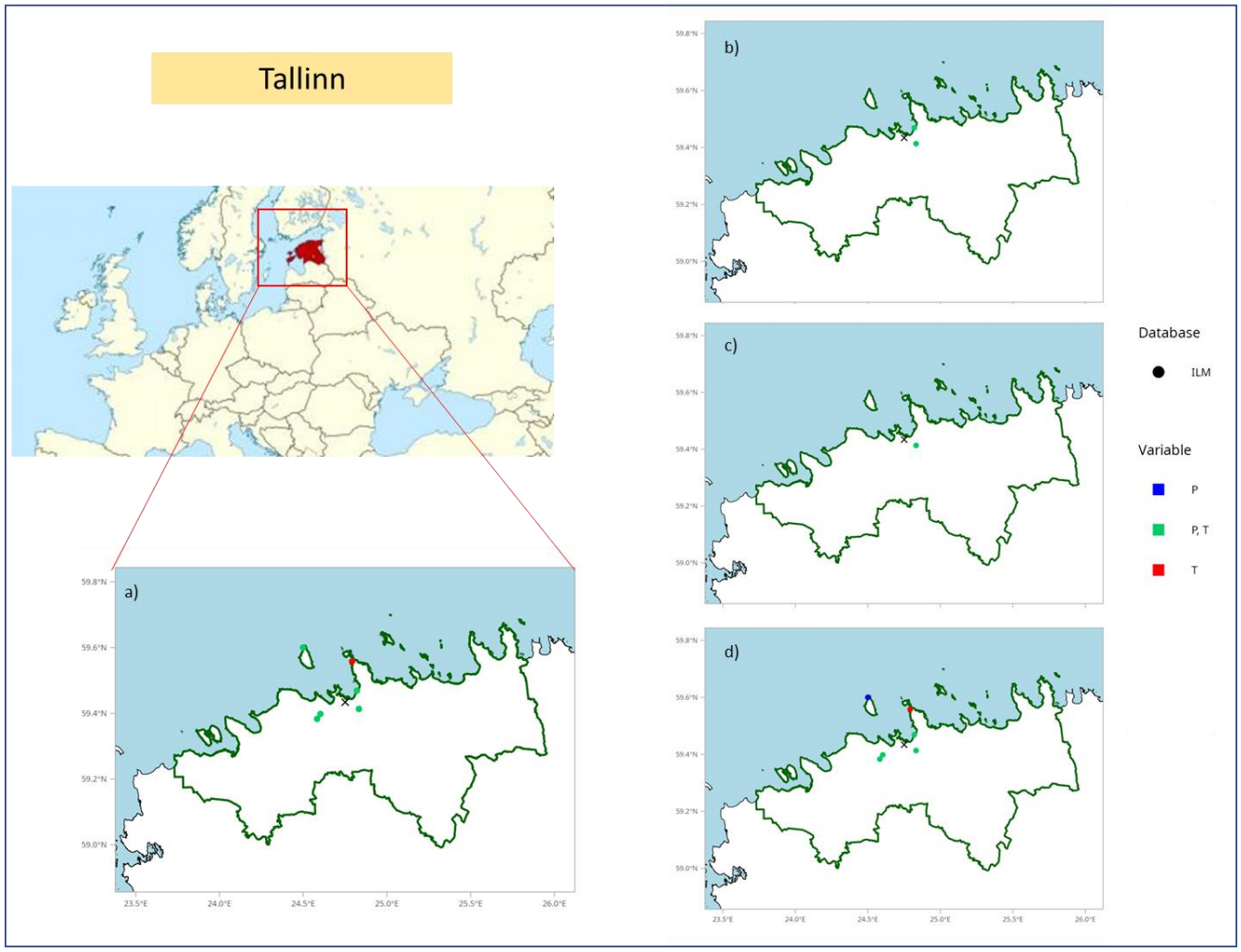

Figure 10. The figure shows a) observatories available at the start of the study, as well as the observatories ultimately used in the study for b) weather forecast, c) seasonal forecast and d) climate projections for Tallinn. **Observatories with temperature data are shown in red, those with precipitation data in blue, and those with both variables in green. ILM is the Estonian weather agency. The borders of Estonia are shown in black, and the county of Harju, which includes the city of Tallinn, is shown in green, with the study area marked with a cross.**

Exclusion of Leuven's City Council weather stations network

While expanding our observational data sources, it has been considered to incorporate data from Leuven's municipal weather network. However, upon evaluation, it has been identified significant reliability issues with this network. Reviewing the article titled "Quality control and correction method

for air temperature data from a citizen science weather station network in Leuven, Belgium" by Beele et al. (2022), it has been identified specific limitations within the Leuven.cool network, that impact the reliability of its data for our forecasting purposes.

These are the main data quality issues identified:

1. **Positive temperature bias:** The study found that the Leuven.cool network's low-cost weather stations exhibited a positive temperature bias, particularly under conditions of high incoming solar radiation. This bias is attributed to design flaws in the stations, such as inadequate shielding from direct sunlight, leading to overheating of the sensors.
2. **Variable sensor calibration:** Inconsistencies in sensor calibration across different stations resulted in inter-station temperature discrepancies. These variations compromise the uniformity of the dataset, making it challenging to compare temperature readings across different locations within the network.
3. **Environmental exposure variability:** Differences in station siting, including proximity to buildings, vegetation, and other environmental factors, introduced additional biases. Stations located in less ideal settings were more susceptible to recording inaccurate temperature data due to localized microclimatic effects.

While Beele et al. developed a quality control (QC) procedure to address these issues, the process is intricate and requires extensive calibration against official weather stations. The QC method involves multiple levels of data filtering and correction, including the application of a random forest model to adjust for identified biases. Despite these efforts, the corrected data still exhibit residual uncertainties, and the QC process necessitates a significant amount of high-quality reference data for effective implementation.

In conclusion, given these challenges, incorporating data from the Leuven.cool network without rigorous quality control could compromise the accuracy of our forecasts. The necessity for complex correction procedures and the remaining uncertainties post-QC render the dataset less suitable for our operational needs. Therefore, to maintain the integrity and reliability of our forecasting system, it has been decided to exclude data from the Leuven.cool network from our analysis.

2.3.2 Weather forecast

The selection of the optimal ensemble prediction system (EPS) for the URBREATH project relies on an extensive dataset of historical forecasts and real-time ensemble outputs to assess the statistical performance of different models under operational conditions. However, our efforts to acquire a comprehensive multi-model EPS dataset, in order to select the more skilled prediction for each city, were met various limitations, significantly affecting the scope of the comparison process. The table

below (Table 4) provides a summary of the status and issues encountered while attempting to download historical ensemble forecast data:

Table 4: Compilation of available raw EPS models and summary of issues

EPS Model	Organization	Period of historical data downloaded	Summary of the status and issues
ECMWF Ensemble Prediction System (ENS)	ECMWF	From 1st January 2024 to 30th April 2024 and from 18th December 2024 until 31 th January 2025	MARS, ECMWF's official downloading data platform, performs very slowly, downloading a single day of data per day.
Global Ensemble Forecast System (GEFS)	NOAA/NCEP (USA)	2023-2024	Required data perfectly downloaded
Met Office Global and UK Ensemble (MOGREPS-G & MOGREPS-UK)	UK Met Office	None	Only last month available https://app.snowflake.com/marketplace
Canadian Ensemble Prediction System (CEPS)	Environment Canada (CMC)	from 18th December 2024 until 31 th January 2025	Landing page for download data is unable: https://caspar-data.ca/ Only a month and few days more available for the project downloaded from other limited source
Météo-France ARPEGE Ensemble Prediction System	Météo-France	None	Major issues with API
ICON Ensemble Prediction System (ICON-EPS)	Deutscher Wetterdienst (DWD, Germany)	from 18th December until 31 th January	Major issues with API (Pamore) Only a month and few days more available for the project downloaded from other limited source

Final dataset used for model selection

Due to these constraints, the only ensemble model available for full historical comparison over the target period (2023-2024) was the GEFS (NOAA/NCEP) dataset with these weather variables: temperature, maximum temperature, minimum temperature, relative humidity, wind gust, and

accumulated precipitation (including snow). This dataset compared with available observations (see section 2.3) served as the primary reference for statistical analysis, allowing us to quantify forecast biases, ensemble spread, and overall skill.

To partially mitigate the limited availability of multi-model historical data, it has also incorporated IFS from ECMWF, ICON and GEM ensembles data from the 18th of December until 31st of January.

This dataset, albeit incomplete, enabled a partial comparative analysis between GEFS and the other ensemble models for temperature and precipitation using statistical methodologies. While the limited period of available multi-model data prevents a long-term performance assessment, it still provides initial insights into forecast differences and bias correction needs for urban-scale predictions.

The limitations in historical data acquisition significantly impacted the ability to conduct a full-scale multi-model comparison. However, despite these constraints, the GEFS dataset and the limited real-time EPS data collected allowed us to establish a baseline performance assessment for ensemble models in urban forecasting. Also to analyze forecast skill for temperature and precipitation within the available time frame, and finally to identify potential biases and model uncertainty that will guide future post-processing efforts. These findings have been summarized in the results section.

2.3.3 Seasonal forecast

The study area is represented by 9 reference stations covering the main pilot cities for the first stage of the project. Due to the wide range of climates (for instance, from Mediterranean climates to subpolar ones, containing continental and oceanic too), it was decided to work with self-generated predictors instead of employing pre-defined teleconnection indexes. These observatories, collected from Global Surface Summary of the Day (GSOD, NCEI 2022), have been selected on the basis of quality and geographical criteria and hold precipitation and 2-meter temperature data for the period 1982–2024. For the quality analysis, outliers and inhomogeneity filters were based on the Kolmogorov–Smirnov (KS) test (Monjo et al. 2013). The KS filters revealed no remarkable problems, since none of the significant inhomogeneities detected lie within the reference period considered, 1993–2024. (Minor inhomogeneities were detected and corrected in temperature time series for Aarhus and Parma, plus other observatories like Plzen Line were rejected and the next nearest was employed for the city -only for precipitation- which was replaced for the Prague International Airport observatory).

Regarding the ‘predictors’ selection and development, the Climate Data Store of Copernicus (CDS) was used for four variables, as specified in Table 5.

Table 5: Variables required for the seasonal forecast design of predictors

Variable	hPa level
Sea surface temperature	Surface
Specific humidity	700 hPa
U component of wind	1000-500 hPa
V component of wind	1000-500 hPa

The choice of these predictors is based on the seasonal patterns and characteristics between the lower and middle atmosphere and the oceans, and the transfer across the hemispheres (also known as 'atmospheric bridges').

2.3.4 Climate projections

As mentioned in section 2.2.4, in order to carry out a proper regionalization, it is necessary to have a solid database to work with. The observed data used have been explained in section 2.3.1. The ERA5 reanalysis has been used, along with a set of 10 climate models from the CMIP6 (Coupled Model Intercomparison Project Phase 6).

Reanalysis datasets

Different reanalysis data from the European Centre for Medium-Range Weather Forecasts (ECMWF): ERA5 (atmospheric) and ERA5-Land (surface) were selected for several reasons. First, since they are developed by the ECMWF, they offer superior information compared to other reanalyses, particularly in Europe. Second, they represent the most up-to-date versions of European reanalyses, providing enhanced spatial and temporal resolution compared to earlier versions. Lastly, they can be freely downloaded through the Copernicus Climate Change Service program.

ERA5-Land focuses on surface variables, covering only land areas, and provides data for up to 50 different variables. Its global grid has a spatial resolution of 0.073 x 0.073 degrees (approximately 9 km x 9 km), with vertical coverage from 2 meters above the surface to a depth of 289 cm, along with 4 levels of the ECMWF surface model. ERA5-Land offers hourly data from January 1951 to the present, with monthly updates and a delay of about three months.

ERA5, the latest atmospheric reanalysis from ECMWF, has been available since July 2019 and is the most accurate atmospheric reanalysis to date. It integrates a vast array of data, including observations from weather stations, soundings, satellites, and other reanalyses (such as oceanic data), across different atmospheric and marine levels. ERA5 aims to replicate the actual atmospheric conditions as accurately as possible for any given point in the past. Due to the high quality of its input data, only post-

release satellite data is included. ERA5 covers all of Europe with a regular spatial resolution of 0.25° (approximately 30 km).

The reanalysis data undergoes a quality control process to eliminate inconsistent values, such as negative precipitation, missing data areas, relative humidity exceeding 100%, or minimum temperatures that exceed maximum temperatures (figure 11 shows an example of ERA5-Land data).

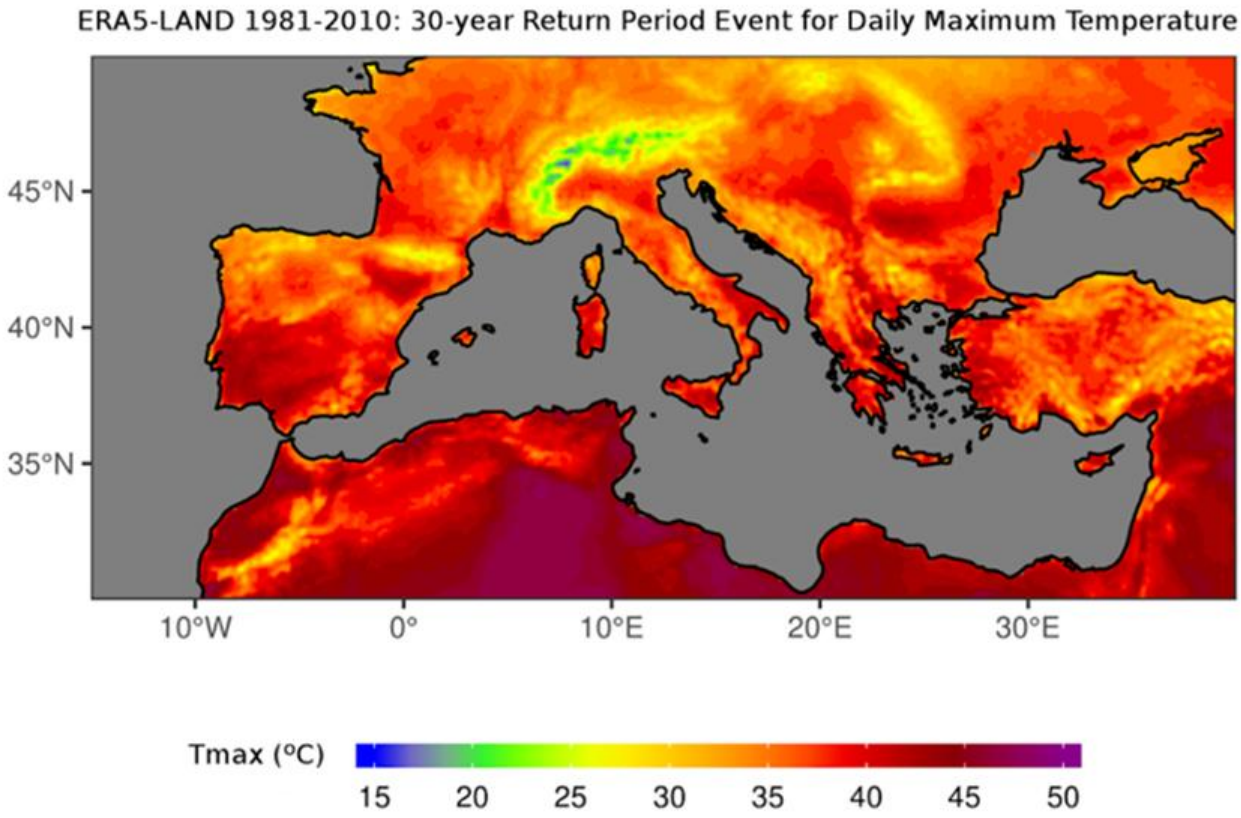

Figure 11. Example of the spatial representation of ERA5-Land reanalysis, representing a 30-year return period event for daily maximum temperature. Source: C3S

Climate models

The final data set required consists of climate models. Specifically, models from the 6th phase of the Coupled Model Intercomparison Project (CMIP6) were used, along with the emission scenarios defined in this phase, known as the Shared Socioeconomic Pathways (SSPs). Statistical techniques, due to their computational efficiency, enable the use of a large number of climate models (n) and SSPs (m), resulting in a set of (n x m) climate projections. These models are run continuously from the past into the future.

After simulating the control period, the simulations are divided into as many runs as there are SSPs. For each climate model, there is a control simulation labeled 'Historical' for the period 1951-2014, and four SSPs for the period 2016-2100 (SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5).

The importance of using these new climate models lies in the significant updates introduced in CMIP6, which are aligned with the World Climate Research Programme (WCRP) Grand Science Challenges. These include studies on Clouds, Climate Circulation and Sensitivity, Changes in the Cryosphere, Climate Extremes, Regional Sea Level Rise, Water Availability, Short-Term Climate Prediction, and Biogeochemical Cycles and Climate Change.

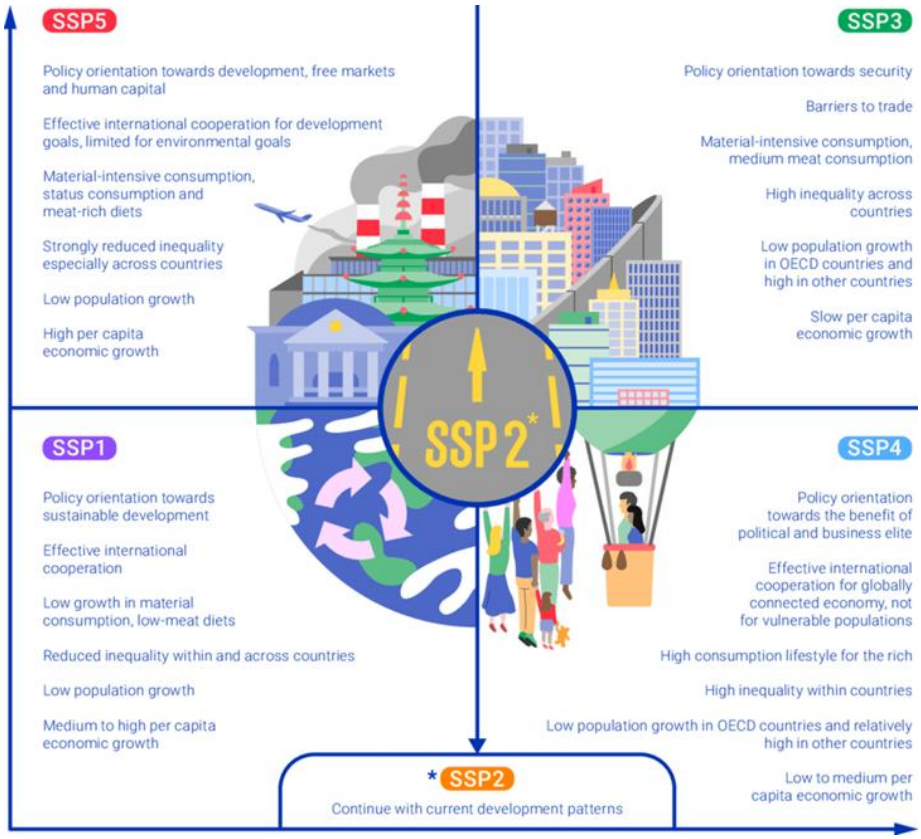

Figure 12. Shared Socioeconomic Pathways (in the Figure, OECD stands for Organizations of Economic Co-operation and Development). Source: O'Neill et al., 2017

Also new in CMIP6 is the approach to future greenhouse gas concentration scenarios, now represented by SSPs (Shared Socioeconomic Pathways) instead of the RCPs (Representative Concentration Pathways) used in CMIP5. The SSPs cover a similar range as the RCPs but address critical gaps, such as

the impact of certain forcings (land use, short-lived species, or air quality), the effects of peak and trough forcings, and the consequences of limiting global warming to below 2 °C (Eyring et al., 2016).

SSPs are projections of global socio-economic changes up to 2100, describing alternative socio-economic developments. Each SSP outlines a potential future trajectory (figure 12). CMIP6 has established SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5 as main scenarios (referred to as Tier 1). CMIP6 extends Tier 1 from 2 (in CMIP5) to 4 scenarios (figure 13).

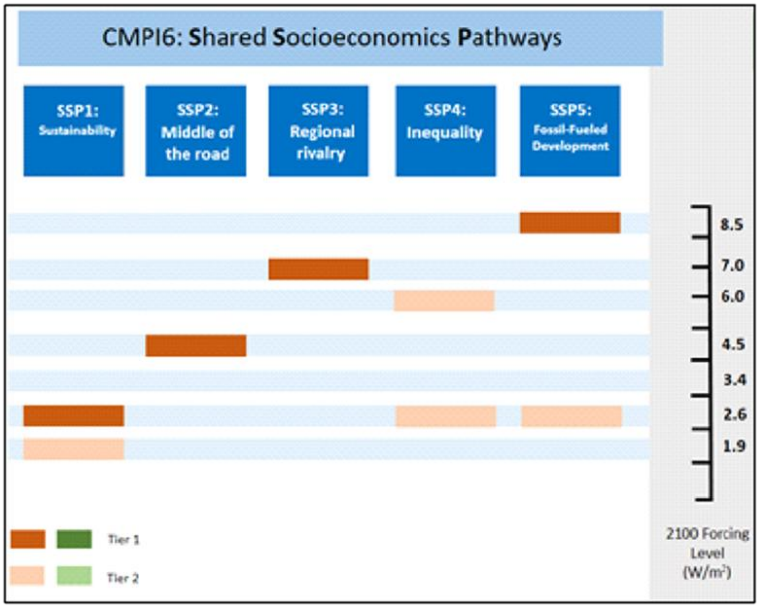

Figure 13. Scenarios defined in CMIP6. Level 1 and level 2 are defined as the scenarios to be provided as mandatory and optional, respectively, by all climate models that are part of each CMIP phase. Source: Figure adapted from O'Neill et al., 2016

In this study, 10 daily resolution climate models have been used, collected within the CMIP6 (Table 6). Throughout the different phases of this CMIP project, improvements have been made to the quality of the climate models up to the current Earth System Models (ESM). Furthermore, new emission scenarios have been defined in each of the phases, adjusting to the new adaptation and mitigation needs in the face of climate change. CMIP6 has a common set of future scenarios covering land use and emissions, as required for future SSPs (Eyring et al., 2016).

Table 6. Information on 10 CMIP6 climate models (IPCC AR6), retrieved from the ESGF portal

CMIP6 MODELS	Resolution	Responsible Centre	References
ACCESS-CM2	1,875º x 1,250º	Australian Community Climate and Earth System Simulator (ACCESS), Australia	Bi, D. et al (2020)
BCC-CSM2-MR	1,125º x 1,121º	Beijing Climate Center (BCC), China Meteorological Administration, China.	Wu T. et al. (2019)
CanESM5	2,812º x 2,790º	Canadian Centre for Climate Modeling and Analysis (CC-CMA), Canadá.	Swart, N.C. et al. (2019)
CMCC-ESM2	1,000º x 1,000º	Centro Mediterraneo sui Cambiamenti Climatici (CMCC).	Cherchi et al, 2018
CNRM-ESM2-1	1,406º x 1,401º	CNRM (Centre National de Recherches Meteorologiques), Meteo-France, Francia.	Seferian, R. (2019)
EC-EARTH3	0,703º x 0,702º	EC-EARTH Consortium	EC-Earth Consortium. (2019)
MPI-ESM1-2-HR	0,938º x 0,935º	Max-Planck Institute for Meteorology (MPI-M), Germany.	Müller et al., (2018)
MRI-ESM2-0	1,125º x 1,121º	Meteorological Research Institute (MRI), Japan.	Yukimoto, S. et al. (2019)
NorESM2-MM	1,250º x 0,942º	Norwegian Climate Centre (NCC), Norway.	Bentsen, M. et al. (2019)
UKESM1-0-LL	1,875º x 1,250º	UK Met Office, Hadley Centre, United Kingdom	Good, P. et al. (2019)

3. Methodology

Within the URBREATH project, various methodologies have been proposed to simulate the different time scales considered in the study: weather and seasonal forecasts, and climate projections. Each methodology has been carefully designed to address the specific needs of these time scales. The following sections provide a deeper insight into each of these methodologies.

3.1 Weather forecast

Weather or short-term forecasting is the result of numerical model simulations, where the fundamental Navier-Stokes equations governing fluid dynamics and thermodynamics are applied to describe the evolution of atmospheric conditions. These models rely on known initial conditions, typically derived from observational data sources such as ground-based weather stations, satellites, and radiosondes, to generate future atmospheric states. The output is visualized in the form of gridded maps, where key meteorological variables—such as temperature, wind speed, humidity, and precipitation—are represented spatially at different forecast lead times (Kalnay, 2003).

There are two primary methodologies for weather forecasting: deterministic and probabilistic. Deterministic forecasting provides a single best-estimate scenario based on the most accurate initial conditions available, while probabilistic forecasting, implemented through ensemble prediction systems (EPS), generates multiple realizations of the forecast by introducing perturbations in the initial conditions or model physics. The latter accounts for the chaotic nature of the atmosphere and provides uncertainty quantification, making it a more robust approach for decision-making in complex applications such as early warning systems, disaster risk management, and renewable energy planning (Bouallégue et al., 2023).

In the context of the URBREATH project, during the Madrid General Assembly's technical partners meeting, it was agreed that probabilistic forecasting would be the optimal approach for integrating weather prediction into end-user decision-support tools. This decision was made considering the added value of ensemble forecasting, particularly its ability to quantify forecast confidence, provide risk-based insights, and improve the communication of uncertainty in operational contexts.

Furthermore, the application of probabilistic weather forecasting in URBREATH is expected to offer several technical and operational advantages combined with AI methodologies, which will be explored throughout this document. These advantages include:

- Higher predictive skill at short-medium ranges (1-10 days) forecasts, particularly for extreme weather events (Price et al., 2023).

- Enhanced adaptability for urban resilience strategies, where forecast uncertainty plays a critical role in planning mitigation and adaptation measures (Benáček et al., 2023).
- Optimized usability for stakeholders, as end-user tools in related sectors such as air quality monitoring, health risk assessment, and infrastructure planning benefit from probabilistic risk communication rather than deterministic yes/no outputs (Bouttier et al., 2024).

Given these significant advantages, URBREATH's decision to adopt probabilistic forecasting represents a strategic shift toward more resilient, uncertainty-aware decision-making, aligning with the best practices recommended by leading meteorological institutions such as ECMWF, NOAA, and UK Met Office. This document will deepen into the methodologies, validation strategies, and implementation frameworks for integrating ensemble weather forecasting of weather variables like extreme precipitation or maximum temperature -and related indices to be codesigned- into URBREATH's operational workflow.

Precipitation verification

- **Ranked Probability Score (RPS):** Discrete Ranked Probability Score measures the deviation of the forecast values assigned to a category compared to the corresponding observations that actually fall within that category. The terms "discrete" and "ranked" refer to the discrete nature of the forecast categories.

$$RPS = \frac{1}{M-1}\sum \left(\left(\sum \quad \rho_k\right) - \left(\sum \quad o_k\right)\right)$$

where M is the number of forecast categories, p_k is the predicted probability in forecast category k, and o_k is an indicator (0=no, 1=yes) for the observation in category k. The RPS is a measure of how good forecasts are in matching observed outcomes. Where:

RPS = 0 the forecast is perfectly accurate.

RPS = 1 the forecast is completely inaccurate.

- **ROC Curve:** The Relative Operating Characteristics (ROC) diagram is a powerful way to verify probabilistic forecasts and, in particular, to compare their performance with deterministic forecast systems. These categorical forecasts will produce a set of pairs of "Hit Rate" and "False Alarm Rate" values to be entered into the ROC diagram: False Alarm Rate (FAR) on the x-axis and Hit Rate (HR) value on the y-axis (derived from a contingency table).

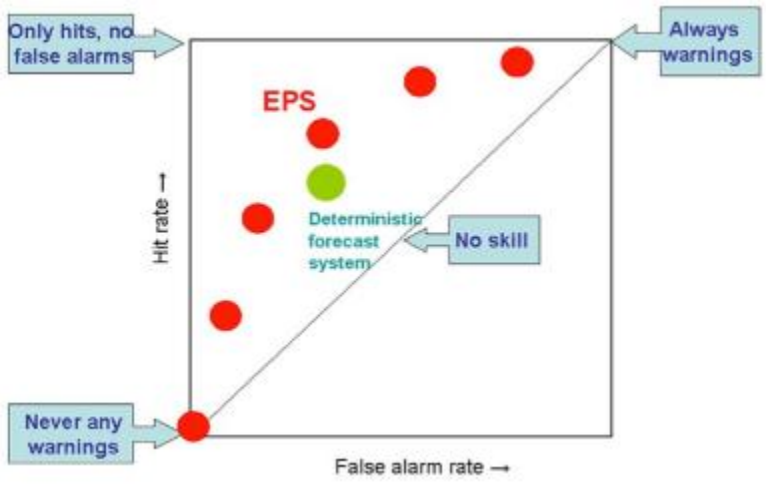

Figure 14. Example ROC diagram: EPS vs. deterministic forecast. Source: ECMWF

- The upper left corner of the ROC diagram represents a perfect forecast system (no false alarms, only hits). The closer any verification is to this upper left corner, the higher the performance. The lower left corner (no false alarms, no hits) represents a system which never warns of an event. The upper right corner represents a system where an event is always warned.

- **Area Under the ROC Curve (AUC):** The area under a ROC is a scalar value that measures the overall performance of a binary classifier (Hanley and McNeil 1982). The AUC value is within the range [0.5–1.0], where the minimum value represents the performance of a random classifier and the maximum value would correspond to a perfect classifier (with a classification error rate equivalent to zero).

  The AUC is a robust overall measure to evaluate the performance of score classifiers because its calculation relies on the complete ROC curve and thus involves all possible classification thresholds. The AUC is typically calculated by adding successive trapezoid areas below the ROC curve.

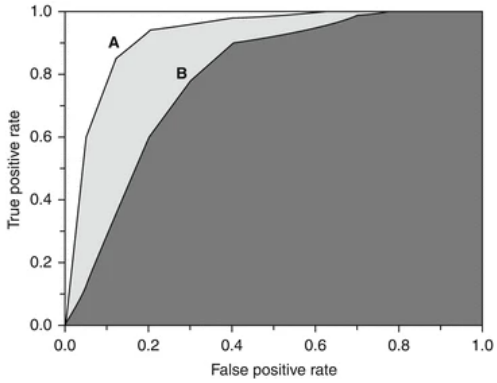

Figure 15. Area under the ROC curves for two score classifiers, A and B. Source: Springer Nature website

- **Contingency table statistics**:

  To verify a forecast of a yes/no event, for example a rainfall or a fog event, a contingency table that shows the frequency of "yes" and "no" forecasts and occurrences is a powerful tool. The four combinations of forecasts (yes or no) and observations (yes or no), called the *joint distribution*, are:

  hit - event forecast to occur, and did occur
  miss - event forecast not to occur, but did occur
  false alarm - event forecast to occur, but did not occur
  correct negative - event forecast not to occur, and did not occur

		Observed		
		yes	no	Total
Forecast	yes	*hits*	*false alarms*	*forecast yes*
	no	*misses*	*correct negatives*	*forecast no*
Total		*observed yes*	*observed no*	*total*

Figure 16. Contingency table for a yes/no event. Source: CAWCR

Given a contingency table for precipitation events, the following statistics have been calculated:

- Accuracy: the accuracy represents the proportion of correct predictions (both true positives and true negatives) among the total number of cases examined.

$$ACC = \frac{hits\ +\ correct\ negatives}{Total\ cases}$$

- Error rate: represents the proportion of wrong predictions.

$$Error\ Rate\ =\ 1 - ACC$$

- False positive rate (false alarm): proportion of false alarms among observed negative cases.

$$FPR = \frac{False\ alarms}{False\ alarms\ +\ correct\ negatives}$$

- False negative rate (miss rate): proportion of misses among observed positive cases

$$FNR = \frac{Misses}{Misses\ +\ hits}$$

- Brier Score: Brier Score (BS) is a measure, over a large sample, of the correspondence between each forecast probability against the frequency of occurrence of the verifying observations. Observation frequency is plotted against forecast probability as a graph. A perfect correspondence means the graph will lie upon the diagonal; the area between the graph and the diagonal measures the Brier Score - values lie between 0 (perfect) and 1 (consistently wrong).

Temperature verification

- BIAS: Average difference between the forecast and observed value in a weather forecast.
- Mean Absolute Error (MAE): The Mean Absolute Error is the average error of the absolute value between the difference of forecast and observed value:

$$MAE\ = \frac{1}{N}\sum \qquad (F_i - O_i)$$

Use of ensemble prediction systems for urban forecasting

EPS models from different meteorological agencies vary in their spatial resolution, forecast horizon, and ensemble size. Below (Table 7) is an overview of the major EPS systems available globally with at least 4 days of forecasting temporal range.

Table 7. Compilation of available raw EPS models

EPS Model	Organization	Temporal Resolution	Forecast Range	Ensemble Members
ECMWF Ensemble Prediction System (ENS)	European Centre for Medium-Range Weather Forecasts (ECMWF)	6 hours	15 days	51
Global Ensemble Forecast System (GEFS)	NOAA/NCEP (USA)	6 hours	16 days	31
Met Office Global and UK Ensemble (MOGREPS-G & MOGREPS-UK)	UK Met Office	6 hours	7 days	18
Canadian Ensemble Prediction System (CEPS)	Environment Canada (CMC)	6 hours	16 days	20
Météo-France ARPEGE Ensemble Prediction System	Météo-France	6 hours	4 days	35
ICON Ensemble Prediction System (ICON-EPS)	Deutscher Wetterdienst (DWD, Germany)	3 hours	7 days	40

These models form the backbone of operational weather prediction across different regions and forecast horizons. Their strengths lie in capturing uncertainty, but their raw outputs must be optimized for high-resolution urban applications through advanced AI-based postprocessing methods.

Optimising urban weather forecasting: from statistical optimisation treatment to integration of ensemble prediction systems with AI-based postprocessing

EPS are a key component of modern numerical weather forecasting, providing probabilistic predictions that allow for uncertainty quantification. These systems, developed by various national meteorological centers, generate multiple realizations of future atmospheric states by perturbing initial conditions and/or model physics. Their outputs are crucial for urban weather forecasting, where accurate predictions of temperature, precipitation, and snow accumulation are essential for infrastructure planning, public safety, and climate adaptation strategies (Bouallégue et al., 2023).

However, raw ensemble forecasts often contain biases, systematic errors, and under-dispersive predictions, particularly at high spatial resolutions relevant to urban environments. Therefore, the integration of AI-based postprocessing techniques is a promising solution to refine ensemble outputs, enhance forecast skill, and improve reliability for city-scale predictions (Rasp & Lerch, 2018).

To bridge the gap between coarse ensemble forecasts and high-resolution urban meteorology, AI-based post-processing techniques are being increasingly used. These methods help improve spatial precision, bias correction, and probabilistic reliability. Climate Research Foundation usually uses statistical methodologies to improve the raw weather forecast output. Here is a description:

For each forecast day and ensemble member, the values of the 10 immediately preceding days are selected and compared with the observed values. Thus, two parameters are calculated:

- Variability: $V = F_0 - F$
- Standard deviation: $\sigma = \sqrt{\frac{\sum (F_i - F)^2}{10}}$

where F_0 is the forecast for the problem day, F is the average forecast for the ten previous days and F_i is the forecast for each one of the ten previous days.

Then, two types of corrections are calculated:

- Climate correction: A simple linear regression of prediction vs observation is applied through the complete climate series.
- Operative correction: A simple linear regression of prediction vs observation is applied only to the data of the ten previous days selected.

To calculate the corrected value of the current forecast, one of these two corrections is applied:

- If the forecasts of the previous ten days present a very high variability ($V > 2\sigma$), the corrected value is calculated through the climate correction.
- Otherwise, if the forecasts of the previous ten days present a small variability ($V < 2\sigma$), the corrected value is calculated through the operative correction.

These statistical methods will be applied to all available models and take a complementary part of the process of optimisation of the weather forecasting with AI methods. With the aim of identifying best results modelisation, both methods will be explored with available models during 2025, comparing the required 2 years of observed data with the different modelisations and optimisation methods. These are the AI methods that have been documented to study and implement their development in the coming months:

- **Neural networks for bias correction**
  - Deep learning methods, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have demonstrated the ability to correct systematic errors in ensemble temperature and precipitation forecasts (Rasp & Lerch, 2018).

  - AI models can learn complex relationships between forecasted and observed weather variables, significantly improving short-term urban weather predictions.
- **Statistical post-processing for probabilistic forecasting**
  - Techniques such as quantile regression forests (QRF) and ensemble model output statistics (EMOS) provide enhanced probabilistic predictions of snow accumulation, precipitation extremes, and temperature variability (Évin et al., 2021).
  - These models allow for adjusting ensemble spread, ensuring more realistic probability distributions of extreme weather events in urban settings.
- **Generative AI for precipitation forecasting**
  - Generative AI models, such as Latent Diffusion Models (LDMs) and Vision Transformers (ViTs), have recently been developed to enhance ensemble precipitation forecasts (Sha et al., 2024).
  - These techniques generate realistic precipitation scenarios, improving the detection of extreme rainfall events that are critical for urban flood management.
- **Hybrid AI-physical models for snow forecasting**
  - The integration of analog ensemble techniques with parametric AI models has improved snowfall predictions by combining historical analogs with real-time ensemble outputs (Scheuerer, 2018).
  - These hybrid methods refine urban snow accumulation estimates, supporting transportation management and infrastructure resilience.

Given that the official meteorological sources used for historical data retrieval also offer real-time ensemble forecasts, it has been ensured that the operational data pipeline is sustainable and reliable for operational weather forecast purposes supported with AI. This includes using:

- Daily downloads of ensemble model outputs from selected EPS providers: Weather predictions from three different models are being automatically downloaded on a daily basis, with the aim of selecting and training an optimal AI algorithm and then building an operational system that corrects each model's predictions to produce an optimal forecast.
- Integration of near-real-time ensemble predictions into our AI-based framework: The data downloaded from the forecast models on a daily basis will join the models' historical database, allowing the AI model to automatically adjust to the new data it will receive.
- Operative calibration and bias correction through AI, including the statistical findings from the historical model (2023-2024) evaluation phase: The historical data from the prediction model will be used to train the AI algorithm in order to obtain a prediction tailored to each study case. A neural network will be considered as a priority algorithm for implementation: neural networks are highly powerful mathematical tools that can be used to recognize images, design artificial intelligences capable of surpassing the best chess players, or build predictive models. In the

context of weather forecasting, neural networks can be utilized to enhance the performance of post-processing in numerical models.

3.2 Seasonal forecast

The ECMWF Strategy 2016-2025 (Alexe M. et al. 2023) includes the upgrading of the Integrated Forecasting System (IFS). The IFS has brought ocean-atmosphere coupling to all ECMWF forecasts (up to 1 year ahead). This is particularly noteworthy given the importance of ocean data in the forecasting process. However, the most recent versions have integrated artificial intelligence into the IFS itself, becoming the AIFS (Artificial Integrated Forecasting System) from 2023. ECMWF's decision-makers opted for graph neural networks, allowing them to move away from lat-long grids and use reduced Gaussian grids (which have almost the same distance between grid points wherever you are on the globe).

Other neural network applications have recently been developed (Kumari et al. 2021). One example is the prediction of solar irradiance, which is very useful for solar farms and power plants. A combination of LSTM-CNN models has been used in the state of California. The results show a skill score of 35 to 55% over traditional dynamic seasonal forecasts.

Other studies have been carried out, for example for predicting temperature (Guo et al. 2024). In this example, many methods were compared to see which one was better at predicting temperature in Northeast China on a monthly scale: ANN, RNN, LSTM, CNN and the CNN-LSTM combination were used. The CNN and CNN-LSTM methods performed best, with the lowest BIAS.

Energy balance studies have long highlighted the critical role of the global oceans in influencing seasonal atmospheric anomalies through coupling mechanisms. However, weather prediction models still struggle with long-term forecasts, largely due to the nonlinear sensitivity to initial conditions in the deep ocean. Given the significant unpredictability of seasonal variability in the Mediterranean climate, this study adopts a statistical approach inspired by Redolat et al. (2024). This method assumes the following hypothesis: 1) delayed teleconnection patterns provide information about ocean-atmosphere coupling on subseasonal time scales through the lens of 2) partially predictable quasi-periodic oscillations since 3) forecast signals can be extracted by smoothing noise in a continuous lead-time horizon.

The novel method consists of combining lag-correlated automatically designed teleconnection indices, with self-predictability techniques of residual quasi-oscillation based on wavelet (cyclic) and autoregressive integrated moving average (ARIMA) (linear) analyses, including Convolutional Neural Networks to generate the predictors (indices). The prediction skill of this teleconnection–wavelet–

ARIMA (TeWA) combination will be cross-validated and compared to that of the ECMWF's Seasonal Forecast System 5 (SEAS5)–ECMWF model (3 months ahead). To validate the method, monthly predictability of temperature and precipitation will be analyzed at 9 reference stations of the URBREATH regions for the 1980-2024 period.

This study modifies and extends the Teleconnection Wavelet-ARIMA Approach (TeWA) by incorporating CNNs to improve the selection of teleconnection indices and adapt the methodology to a monthly scale. The framework combines advanced machine learning for teleconnection analysis with statistical methods to enhance the predictability of temperature, precipitation, and wind anomalies in the case study regions across a forecast horizon of 0 to 6 months.

The original TeWA method relies on predefined teleconnection indices (like ENSO, PDO, ULMO, etc) and self-predictability metrics to estimate subseasonal anomalies. This adaptation introduces CNNs to dynamically identify teleconnection patterns from a global dataset of monthly atmospheric and oceanic variables.

Predictor selection and TeWA approach

CNNs analyze spatial fields of monthly anomalies in key climate variables, including: Sea Surface Temperature (SST), Humidity at 700 hPa, zonal wind (u) at 1000 hPa and 500 hPa, meridional wind (v) at 1000 hPa and 500 hPa.

The CNNs identify regions and combinations of these variables that serve as optimal predictors for case study surface anomalies. This step ensures that teleconnection indices are not predefined but emerge from the data, dynamically adapting to varying climate influences across different forecast horizons and regions (for example, the ENSO index may be highly correlated in the Netherlands and Estonia, but not in inland observatories).

Once CNN-selected areas are identified, teleconnection indices are derived by aggregating the spatiotemporal patterns of the global predictors. These indices are then integrated with self-predictability measures, calculated from the temporal persistence of local monthly anomalies, to form a hybrid statistical model. The combined model estimates the likelihood of surface anomalies across the region.

Model training and prediction

Wavelet transforms are applied to the identified teleconnection indices and observed anomalies to extract multi-scale temporal patterns.

Firstly, each time series is decomposed into distinct frequency bands, isolating high-frequency (sub-monthly) and low-frequency (seasonal to interannual) components. The relationship between indices or predictors and observed anomalies is analyzed for each frequency band to identify the dominant scales contributing to variability. Finally, by discarding noise-dominated scales, the wavelet step ensures that only meaningful components are passed to the next phase.

ARIMA models are used to model and predict the residual components of the anomalies after removing scale-specific influences.

The time series of anomalies is deseasonalized and detrended, isolating purely anomaly-based signals. The ARIMA models are fit to the remaining variability, capturing autocorrelated behavior that cannot be explained by the teleconnection patterns alone. ARIMA predictions are generated independently for each horizon (1–6 months) and combined with the teleconnection-based forecasts to produce a final prediction.

In the last step, the outputs from the teleconnection AI-based (CNN-driven) and ARIMA (time-series-driven) phases are combined into a hybrid forecast: The contributions of the teleconnection and ARIMA components are weighted dynamically based on their performance in hindcast experiments for each lead time. The combined forecast reflects the best of both methods: the spatial skill of CNNs in identifying teleconnection patterns and the temporal skill of ARIMA in modeling local dynamics.

Validation Metrics

Model performance is evaluated using hindcast experiments spanning several decades. Validation metrics include:

- Pearson Correlation Coefficient: Measures the correlation between predicted and observed monthly anomalies, assessing spatial and temporal skill.
- Standardised Mean Absolute Error (SMAE): Quantifies the magnitude of error between predictions and observations.

$$SMAE_H = \frac{\sum \left(\sum P_{h,i} - \sum O_{h,i}\right)}{\sum \left(\sum O_{h,i}\right)}$$

- Root Mean Absolute Error (RMAE): Provides a straightforward average of absolute prediction errors.

$$RMAE_H = \frac{\sum \left(\sum P_{h,i} - \sum O_{h,i}\right)}{\sum \left(\sum B_{h,i} - \sum O_{h,i}\right)}$$

- Skill Scores (e.g., Brier Skill Score, Continuous Ranked Probability Score): Compare the hybrid model's performance against baseline models like climatology and persistence.

As can be seen in the schematic figure 17, by combining the flexibility of CNNs in detecting teleconnection patterns with the robustness of TeWA's self-predictability metrics, this methodology enables improved subseasonal forecasting at a monthly scale. The automated selection of predictors at the initial stage ensures adaptability to complex climate dynamics, while validation against historical records confirms the reliability and accuracy of the approach.

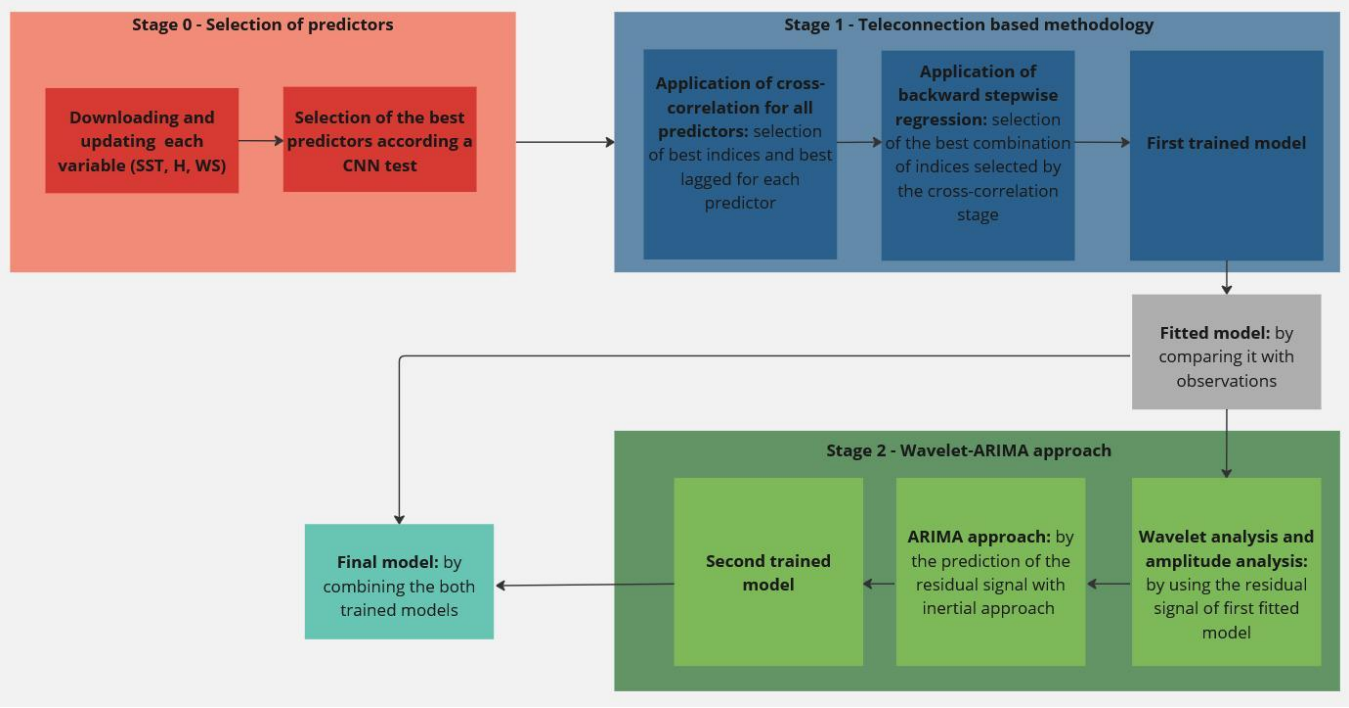

Figure 17. Scheme of seasonal forecast methodology

Key Gaps and Findings

The seasonal method based on Tele-Wavelet-Arima (in combination with convolutional neural networks) is very sensitive to the stationarity (coupling) between teleconnection indices (predictors) and surface variables (predictands). In other words, there are years or decades in which the method may not predict correctly (uncoupling, signal change, etc.). In addition, it is sensitive to the aforementioned stationarity of the couplings (for example, it sometimes only works for a few weeks in winter or autumn, especially for southern Europe and the Mediterranean, because it "decouples" from the northern jet stream during the summer months and therefore depends more on regional patterns

with a higher dependence on the movement of the ITCZ or Intertropical Convergence Zone and cut-off depressions).

Another weakness is the computational cost, since it requires a fairly long training period.

On the other hand, the method, even without incorporating CNN and AI techniques, was better if compared with the dynamical analogue seasonal forecast model (SEAS5) of the main centre for medium-range forecasting (ECMWF). It achieves up to 70% improvement with respect to SEAS5 for precipitation and up to 60% if compared with temperature. The first estimations by using the CNN in the first stages reach up to 50% improvement in the short-term forecast.

3.3 Climate projections

To assess and/or estimate climate risk, it is essential to define it based on the extent of climate impacts on the sector being analyzed. It is important to consider that the effects of climate change do not impact all regions or socio-economic sectors uniformly. Therefore, these impacts need to be evaluated at the local scale. To estimate climate risk, it is necessary to have climate projections at the local level (such as the future evolution of basic meteorological variables like temperature, precipitation, or wind), analyze the impacts and implications of these projections for the sector under study (using indicators tailored to its needs), and identify the potential risks that may arise in the coming decades. Once the impacts are identified, adaptation measures can be defined to address them, minimizing negative effects as much as possible while trying to take advantage of the positive ones.

For this reason, this study aims to provide an overview of how climate change affects the evolution of meteorological variables in the coming decades in the Living Labs. To ensure that this study serves as input for local climate information, it is intended to be used as a support tool when considering adaptation measures in response to climate change. A set of basic meteorological variables has been selected to characterize the region climatically. Thus, decisions should be based on the new climate conditions expected for the future, rather than past climate conditions.

State of the art in climate modelling

The most advanced tool for climate simulation today is the Climate Model (CM). A CM is a numerical representation of the climate system, based on the physical, chemical, and biological properties of its components, their interactions, feedback processes, and including some or all of their known characteristics. Climate systems can be represented by models of varying complexity, meaning that for each component or set of components, there is a spectrum or hierarchy of models that differ in aspects such as the number of spatial dimensions, the level of explicit representation of physical, chemical, or biological processes, or the extent of empirical adjustments used.

Earth System Models (ESMs) are the most modern climate models, which describe the processes within and between the atmosphere, ocean, cryosphere, marine and terrestrial biospheres. Process equations capture the physical, chemical, and biological mechanisms that govern the elements of the Earth system, including volcanic eruptions and variations in incoming solar radiation. ESMs also represent the carbon cycle, enabling the interactive calculation of atmospheric CO2 emissions or similar gases. They may include other components, such as atmospheric chemistry, ice sheets, dynamic vegetation, the nitrogen cycle, and urban or crop models, among others. The key improvement over the previously used General Circulation Models (GCMs) is that ESMs allow the interaction of the system with the carbon cycle and consider marine biochemistry and biogeology (figure 18).

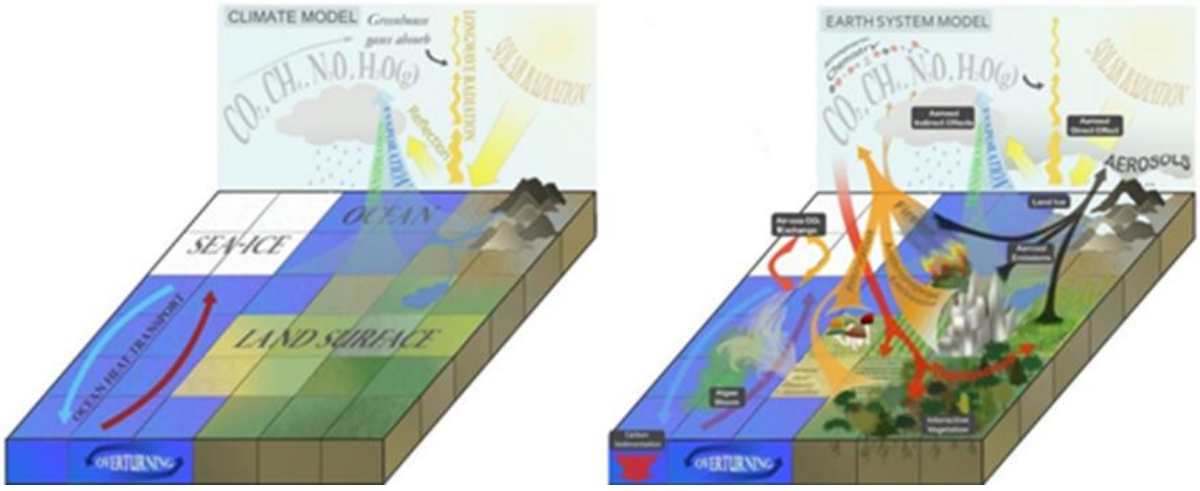

Figure 18. Key features of climate and Earth System Models (ESMs): ESMs include biological and chemical processes. Source: Heavens et al. 2013

Climate Models (CMs) are run based on different assumptions about future Greenhouse Gas (GHG) concentrations, reflecting plausible scenarios for socio-economic, technological, and environmental changes, as well as emissions of GHGs and pollutants. These models produce climate scenarios or projections that represent potential future climates, helping to analyze the impacts of human-induced climate change and inform impact models.

The Intergovernmental Panel on Climate Change (IPCC) coordinates global climate change efforts, providing scientific, technical, and socio-economic reports on climate risks, impacts, and adaptation strategies. These reports support the UN Framework Convention on Climate Change (UNFCCC), which aims to stabilize GHG concentrations to prevent dangerous interference with the climate system. The

IPCC, recognized as an authoritative body on climate change, releases periodic reports based on global consensus, with the Sixth Assessment Report published in August 2022.

In parallel, the Coupled Model Intercomparison Project (CMIP), established in 1995 by the World Climate Research Programme (WCRP), aims to improve understanding of climate change through multi-model frameworks. CMIP has enhanced climate models over the years and developed new emission scenarios to address adaptation and mitigation needs. The models from the sixth phase of CMIP (CMIP6) are currently the most advanced and state-of-the-art, incorporating the latest improvements and developments in climate simulations. These models represent the most up-to-date technology in climate modeling and are used in the latest IPCC assessments and other national climate change studies.

The need for regionalization or downscaling

Climate Models (CMs) can effectively simulate large-scale atmospheric circulation, but due to their spatial resolution (around 100 km), they are unable to capture smaller-scale atmospheric phenomena that are crucial for local climatology. To address this and other limitations, regionalization or downscaling techniques are applied. These methods adapt the outputs from CMs to local scales by using predictors—reliable information from CMs, such as low-resolution atmospheric configurations, typically at higher altitudes. This is because CMs perform poorly near the surface due to their low resolution, which limits their ability to accurately represent topography, land use, etc. These predictors are then used to translate low-resolution atmospheric data into localized surface effects (such as rainfall, temperature, and wind), known as predictands.

There are two primary regionalization approaches: statistical and dynamic, each with its own strengths and limitations. The choice of method depends on factors like the number of climate projections expected, the availability of observed data, and the orographic complexity of the study area. For this study, a statistical downscaling approach called FICLIMA (Ribalaygua et al., 2013) will be employed. This methodology, developed by the Climate Research Foundation (FIC), has been validated in numerous national and international projects (Gutierrez et al., 2019).

To ensure the reliability and robustness of climate projections, these projections must meet specific technical requirements, as outlined in figure 19.

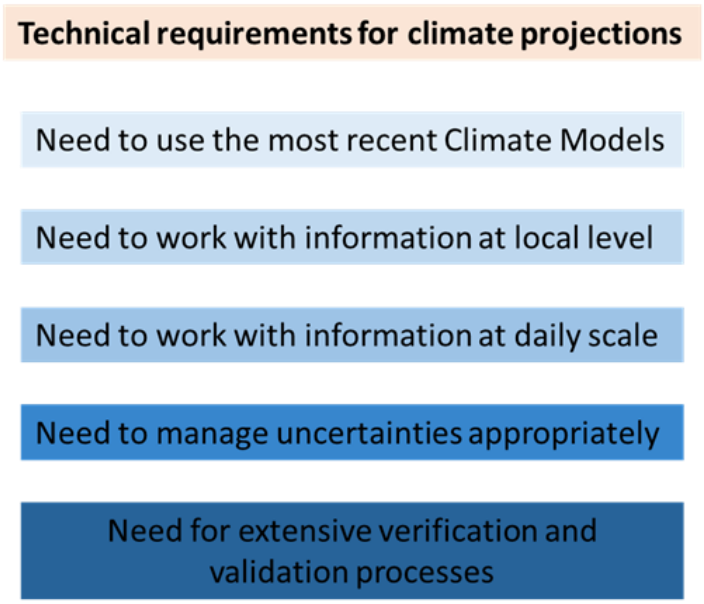

Figure 19. Requirements for climate projections

Meeting these requirements ensures that the set of climate projections generated is robust and based on verified and validated techniques that allow uncertainties to be quantified and taken into consideration when interpreting future scenarios.

Two-step statistical downscaling methodology: FICLIMA

URBREATH project has used a two-step statistical regionalisation methodology developed by the Foundation for Climate Research. This methodology has been applied to the variables: temperature (both maximum and minimum) and precipitation.

The methodology requires a prior selection of fields to be used as predictors and, once these have been selected, it performs a treatment based on the analogue's methodology. In general terms, the methodology follows the scheme below: for a specific problem day 'X', where the low-resolution atmospheric fields (geopotentials, temperatures at different pressure levels...) are known (through the outputs of the Climate Models for day 'X'), the aim is to estimate the value of the surface meteorological variables (maximum and minimum temperatures, precipitation i.e) for day 'X' at a specific location (observatory) from these known fields. The method works in two successive steps:

- The first step, called analogue stratification, consists of selecting, from a reference data bank, those n days with atmospheric configurations most similar to those of the problem day 'X'. The

similarity measure used compares the resemblance between the predictor variables used to characterise the atmospheric synoptic situations (as for example, geopotential or wind); these variables determine the synoptic forcing causing the descents and ascents of air, which generate cloudiness and precipitation. The aim is also to provide information on the surface wind direction, which allows us to study the effects that topography has on the rising air masses and, therefore, on the spatial distribution of cloudiness and precipitation.

- The second step applies different methods depending on the variable to be calculated:

  To estimate daily minimum and maximum temperatures, a multiple linear regression with automatic forward and backward predictor selection is performed for each variable. The data population will consist of the n (analogue) days selected in the previous step. As potential predictors, on the one hand, the values of atmospheric variables related to temperature forcings (such as the thickness of the air layers) in the vertical of the point for which the surface temperature is to be estimated, and on the other hand, predictors related to other temperature forcings, such as an indicator of the duration of the day in question (related to the radiative warming potential) and a weighted average of the temperatures of the previous days (which considers the effect of the thermal inertia of the ground) are offered. Once the linear relationship between the selected predictors and the predictand (minimum, maximum temperature) has been established, this relationship is applied to the predictor values for day 'X' to estimate the value of the predictand on that day.

  In the case of precipitation, precipitation is obtained by averaging the k analogous days most similar to 'X'. In addition to estimating the amount of rainfall, this method allows obtaining the probability of rain or dry weather.

Figure 20 illustrates the methodology described graphically.

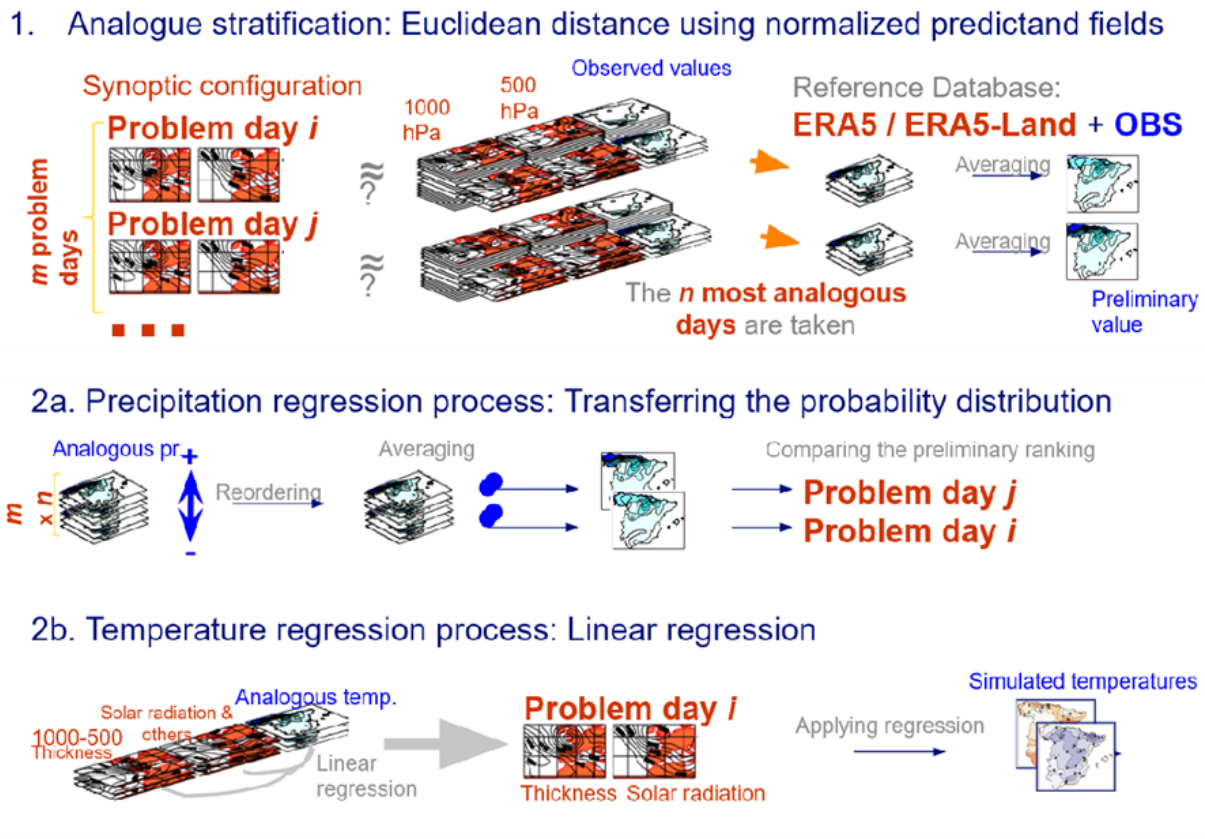

Figure 20. FICLIMA statistical downscaling: key features

Need for extensive verification and validation studies

Before simulating future climate, it's crucial to verify that the tools used can accurately simulate the observed past climate in the study area. This involves conducting thorough analyses to verify the downscaling methodologies and validate the Climate Models. Understanding the errors in simulating the observed climate is key because if these errors are too high, the methodology or CM should not be used for that specific point. Smaller errors lead to lower uncertainties in future simulations.

The **verification process** compares observed data with simulated data from applying the downscaling methodology to a reanalysis (ERA5). This helps assess if the methodology can accurately simulate the current and recent past climate. The methodology was applied to the ERA5 reanalysis for 1979-2020, adjusting the high spatial and temporal resolution data to the lower resolution of the CMs used later. This ensures more realistic verification results. Comparisons between observations and simulated data

should ideally be made on a day-by-day basis, which helps confirm that the methodology captures the physical relationships between predictors and predictands, minimizing the stationarity issue common in statistical methods.

It is essential to assess each climate model's performance in the study area to determine its reliability in simulating the region's present climate and whether it is suitable for use. This evaluation process, known as **validation**, involves comparing the simulated series from regionalising the ERA5 reanalysis with the series from regionalising the model's Historical scenario (its control simulation) for a common past period (1979-2015). The Historical scenario, based on past data, helps determine if the climate model accurately reflects the study area's climate, including average values and natural climate variability.

AI approach on current FICLIMA methodology

A statistical downscaling method based on analogs and transfer functions is a technique used to obtain climate or weather information at smaller spatial or temporal scales from larger-scale models (such as global or regional climate models). This type of method can be considered a use of artificial intelligence in downscaling techniques due to the way it applies principles related to pattern recognition and knowledge transfer between scales. Therefore, the methodology used in the simulation of climate projections falls within the range of AI methodologies.

Here's how artificial intelligence fits into this type of technique:

1. **Analog identification**: The method relies on finding similar historical patterns, known as "analogs," from large climate data sets. This is where artificial intelligence comes into play, as a machine learning algorithm (such as a k-nearest neighbors model, neural networks, or similar techniques) can be used to identify the most relevant analogs from historical data by looking for similarities between global/regional climate conditions and local conditions.

2. **Transfer functions**: Once the analogs are identified, transfer functions are used to adjust the predictions from the larger-scale model (e.g., temperature and precipitation from a global climate model) to the local scale. Transfer functions are typically mathematical formulas or statistical relationships that project large-scale conditions (global or regional) to a local scale. At this step, artificial intelligence techniques are key in learning and adjusting these transfer functions, using approaches like neural networks or advanced regression models that optimize the relationship between input data (large scales) and expected results at the local level.

Within the URBREATH project, it is foreseen to use machine learning for accuracy improvement. AI will be employed to optimize statistical models for predicting local conditions, such as through deep neural

networks that can learn complex nonlinear patterns in historical data. These models will be continuously updated as new data becomes available, enhancing the accuracy of the downscaling process. This approach will enable the refinement of predictions over time, enhancing the precision and adaptability of local climate projections.

In summary, artificial intelligence is applied in these downscaling techniques by using machine learning algorithms to recognize patterns, adjust transfer functions, and optimize the relationship between global and local scales. This integration of AI allows the methods to be more accurate, flexible, and better able to adapt to the variability of climate data.

4. Results

The expected results of the project for each of the front cities are as follows:

- A database consisting of the observed meteorological time series that have passed quality control and homogenization checks for each FR city.
- Operational weather forecast. It has performed a comparative analysis of temperature and precipitation variables of four ensemble weather prediction models—GEFS, ECMWF, ICON, and GEM—against observed meteorological data from December 2024 to January 2025. The results show systematic biases in temperature and precipitation forecasts, with errors varying by location and forecast range. Some models exhibited temperature overestimations of up to 2.5°C, with ECMWF generally providing the most accurate raw forecasts, while GEFS had the largest discrepancies, particularly in minimum temperature predictions. Additionally, a two-year comparison (2023-2024) between GEFS (as the only available model for this period) and a bias-corrected version of GEFS was performed for temperature. The applied correction focused on systematic biases detected in the raw forecasts, leading to notable improvements in this variable prediction accuracy. These results establish a baseline for further refinements, guiding the final model selection process and future improvements to enhance urban weather forecasting within URBREATH.
- Seasonal forecast simulations will include a total of 6 months (7 if counting the month from which the forecast is made for each observatory). Each month will have five values: percentile 05, percentile 25, median, percentile 75 and percentile 95. All data will be shown in the expected anomaly and in the corresponding units (ºC, mm, m/s for temperature, precipitation and wind respectively). It is expected that in the final stages of the project, when statistical methodology will be fully implemented, a combination in the final results could include a comparison between our statistically based forecast and a dynamically based forecast to strengthen the final seasonal forecast.
- A set of between 28 and 40 daily future climate projections (10 models x 4 SSPs) at the local scale, covering both the historical period (1951-2015) and the future period (2016-2100), for maximum and minimum temperature and precipitation. Future climate scenarios have been generated using a robust regionalization methodology, based on the most up-to-date climate information and surpassing all verification and validation processes that ensure the quantification of associated uncertainties.

The results obtained during the process of generating this information are presented below, considering the databases outlined in section 2 and the methodologies described in section 3.

4.1 Weather forecast

The results of the analysis of the ensemble weather forecast models (GEFS, ECMWF, ICON and GEM) for the common period for which data were available (18 December 2024 to 31 January 2025) show different results depending on the variable and location. These results are briefly discussed below for each Living Lab.

Results for the Living Lab of Madrid

In the case of the city of Madrid, it is observed that the temperature forecasts, especially the maximum temperature, are highly biased in all models. This could be explained by the fact that the area under study has thermal inversions in winter and large intraday thermal oscillations, which are a challenge to achieve optimised corrections using AI. The figure below shows how, in the case of the maximum temperature, the BIAS is around 2°C.

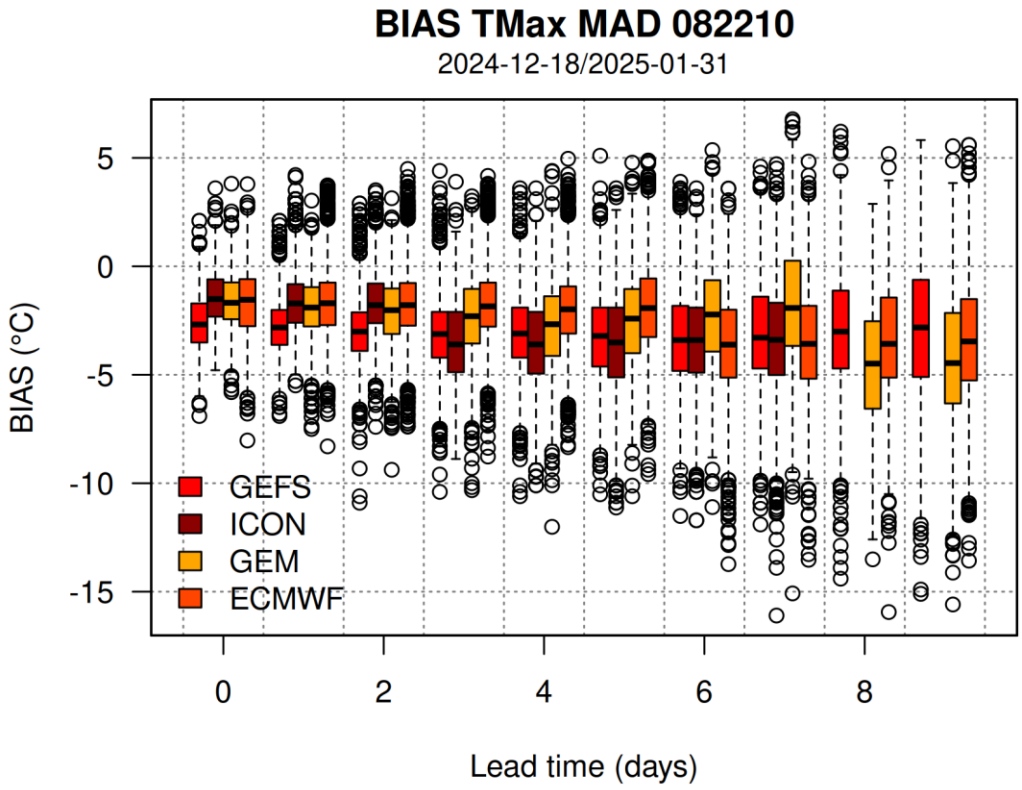

Figure 21. Maximum temperature BIAS in Madrid

However, by applying our simple quantile-quantile correction to temperature, very encouraging results have been obtained, which open the door to the possibility of obtaining very accurate weather forecasts in the next steps. In this case, as shown in the following figure, the MAE was reduced by half, from 2.5°C to around 1°C for all models and up to medium-range time steps (9 days).

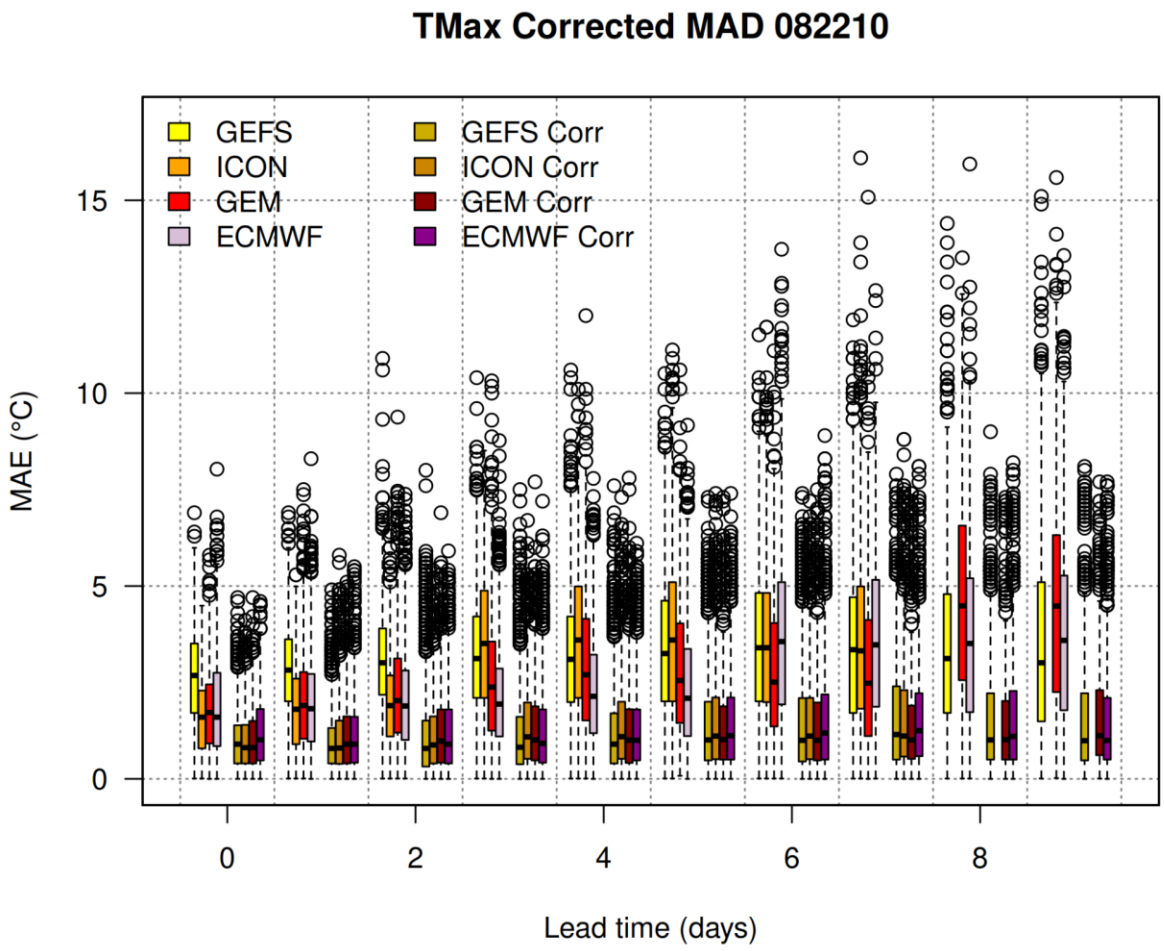

Figure 22. MAE for corrected maximum temperature in Madrid

In addition to the multi-model comparison for December 2024 and January 2025, a temperature error analysis has been performed for the two years for which GEFS model data are available (2023 and 2024). These predictions have also been corrected using our bias correction method, obtaining encouraging results:

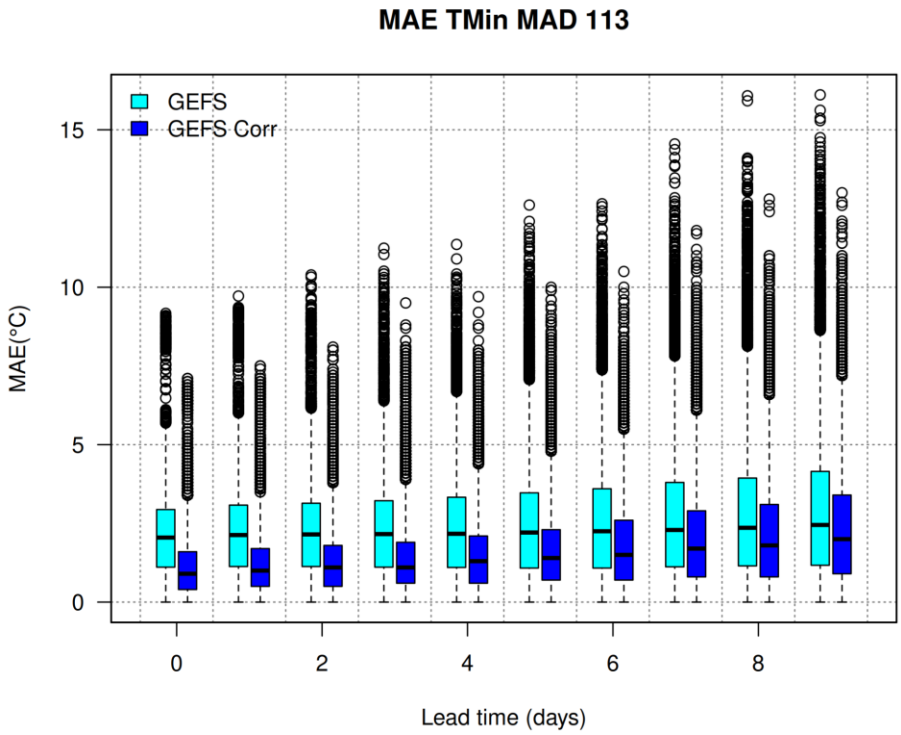

Figure 23. MAE for corrected minimum temperature in Madrid. GEFS 2023-2024

As a result of this correction, it can be seen in the figure that the minimum temperature error - in an observatory that has thermal inversions in winter - is reduced by half in almost all forecast times. Similar results are obtained for other observatories and also for maximum temperature, which is a very encouraging starting point for obtaining more accurate weather forecasts by training more complex AI models in next steps.

Regarding precipitation, in the case of Madrid, the vast majority of days have atmospheric stability and, when analysing the contingency matrix of predicted and observed rainy days, the vast majority of days (more than 80%) correspond to correct negatives (days when rain is not predicted and it does not rain), so the statistics are conditioned by this fact. Looking at the ROC curve, the performance is generally very good for both GEFS and ICON, while the GEM model has a higher error (around 30% adding up misses and false alarms).

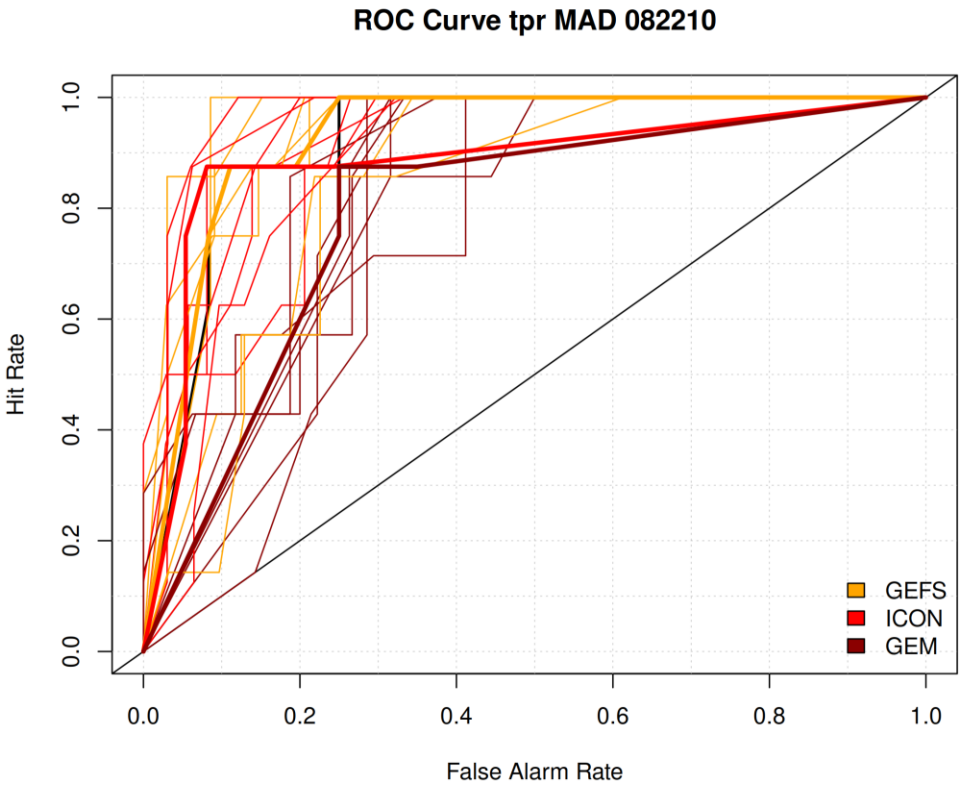

Figure 24. ROC curve for precipitation in Madrid. Thicker lines represent predictions for D+1 for each model, while thinner lines represent forecasts from D+2 to D+10

Results for the Living Lab of Leuven

In the case of the Living Lab in Leuven, the maximum temperature shows very little bias and consequently the mean errors are also very low, in the region of 1.5°C. All models show a similar performance for all time steps analysed. Corrections do not lead to significant improvements.
As for the minimum temperature, the predictions are slightly biased towards positive values in all models except GEM. In the case of GEFS, this bias increases to about 2.5°C, probably due to a significant difference in the altitude of its grid point with respect to the observatory altitude.

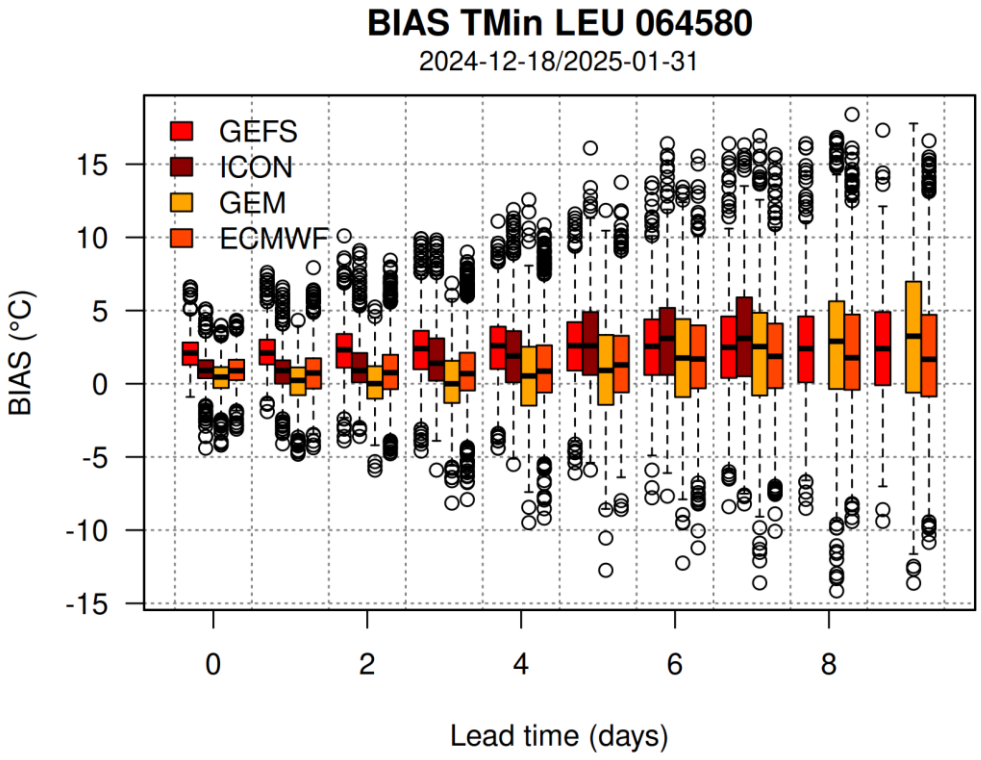

Figure 25. Minimum temperature BIAS in Leuven

Looking at the MAE of the minimum temperature, it can be seen that the GEFS performs significantly worse than the other models, especially in the first days of the forecast. However, once our bias correction method is applied, all the forecasts improve to a point where the MAE is within 1°, with the ECMWF being the model with the least error in its forecasts after the correction.

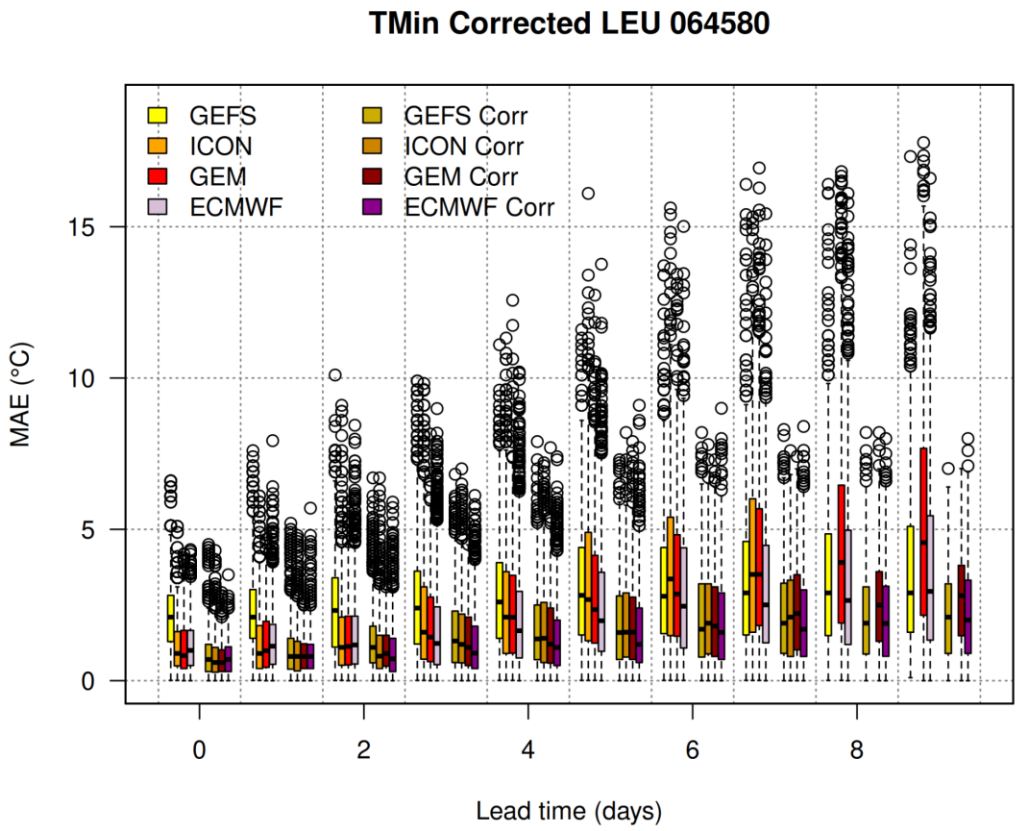

Figure 26. MAE for corrected minimum temperature in Leuven

Looking at the precipitation of the GEFS model for the years 2023 and 2024, it can be seen that, in general, the model has good predictive ability during the first few days by looking at its ROC curve. On the ROC curve, it can be clearly seen how, as the days pass, the area under the curve decreases, indicating that the prediction errors increase.

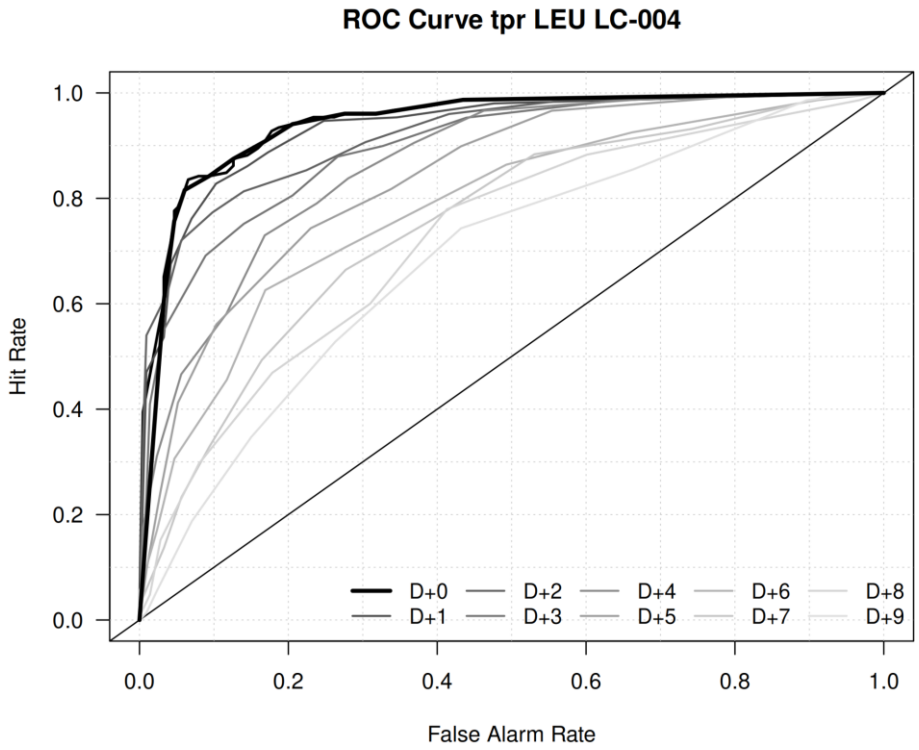

Figure 27. ROC curve for precipitation in Leuven. GEFS model 2023-2024

This loss of predictive capacity is reflected in the increase in misses and false alarms over the days. The following figure shows how the error on the first day of forecasting is only 15% and how this error increases over the days to around 40% for a forecasting horizon of 9 days.

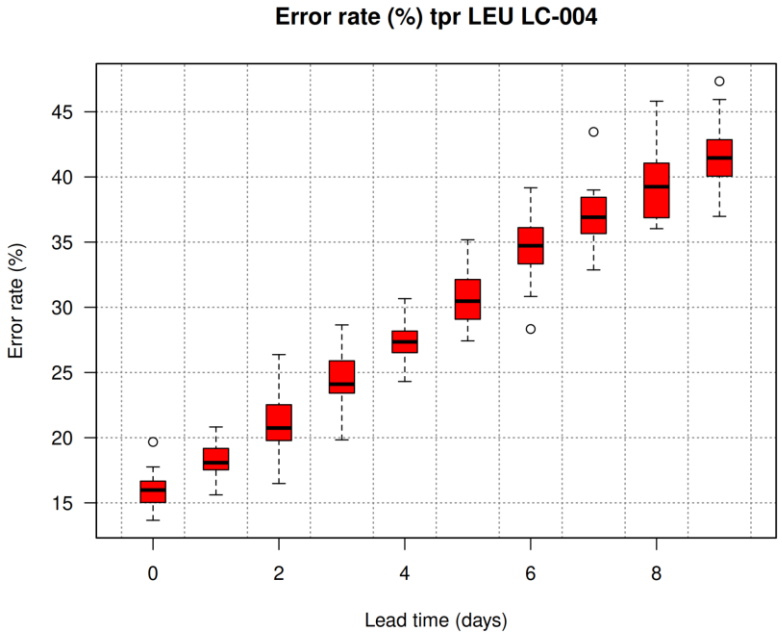

Figure 28. Sum of misses and false alarms for precipitation in Leuven. GEFS model 2023-2024

Results for the Living Lab of Tallinn

The results for the Tallinn case show small biases in both the maximum and minimum temperatures, making these differences significantly larger in the case of the ICON model.

A striking detail is that the BIAS tends to be more negative for longer forecast horizons, which means that in the long term the models tend to predict lower temperatures than what actually occurs.
It has been observed that there can be large temperature differences in this city, depending on whether the stations are closer or further away from the coast. This will be a variable to take into account when optimising weather forecasts with our AI models.

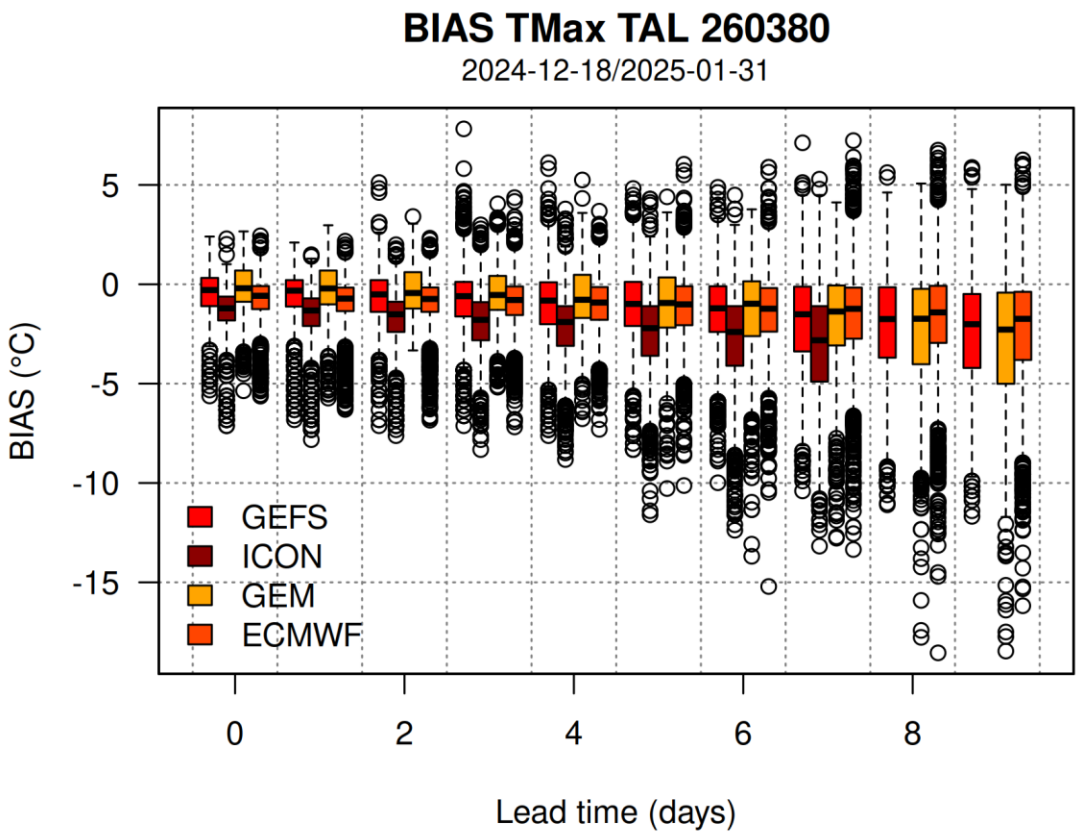

Figure 29. Maximum temperature BIAS in Tallinn

When our simple correction is applied, it is observed that the most biased forecasts are corrected to the level of the other models, reducing the MAE even below 1°C. This improvement in the forecasts is more noteworthy the further ahead in time, so that with this method it is possible to obtain forecasts with an error of less than 1.5°C for time horizons beyond 5 days.

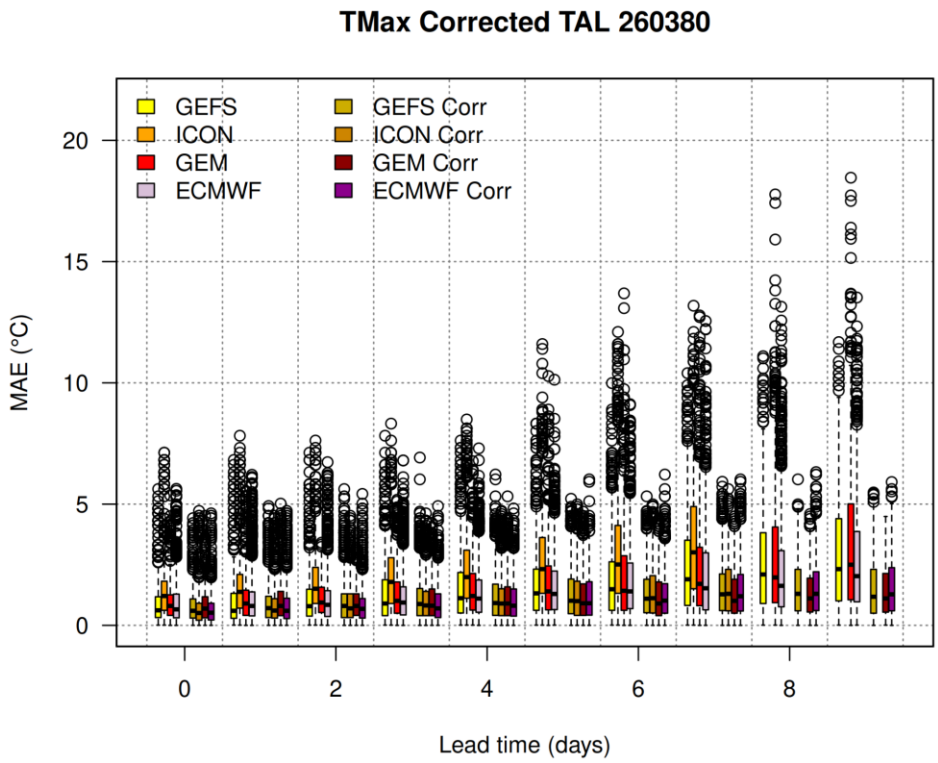

Figure 30. MAE for corrected maximum temperature in Tallinn

With regard to precipitation, and in line with the previous cities, the GEFS for the years 2023 and 2024 also performs well for Tallinn, with values very similar to those of Leuven in the case of type 1 and 2 errors (misses and false alarms). In this case, the area under the ROC curve (AUC) is shown, where it can be seen that its value decreases from 0.93 to 0.72 as the days go by, but it is still well above 0.5 even on the 9th day of the forecast.

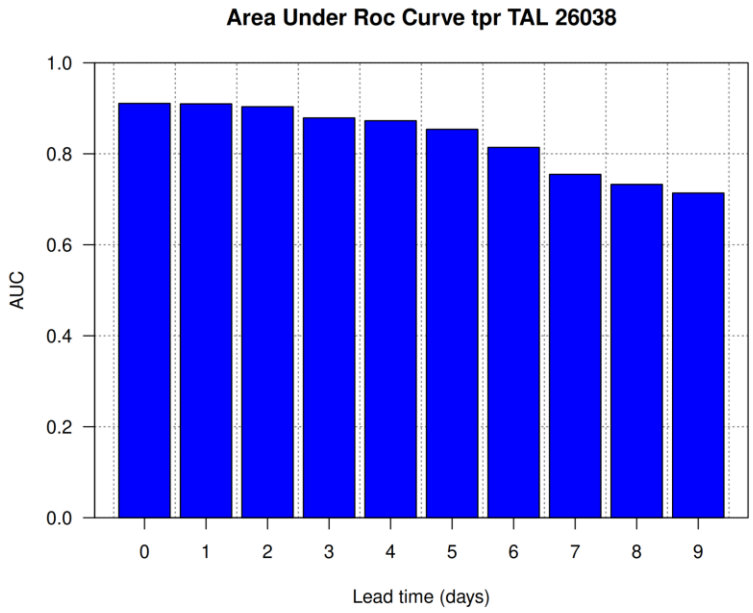

Figure 31. Area under the ROC Curve (AUC) for precipitation in Tallinn. GEFS 2023-2024

Results for the Living Lab of Cluj-Napoca

In the Cluj-Napoca case study, the raw temperature predictions of the meteorological models show a high error compared to the previous cases. This is because the predictions show a high bias for both maximum and minimum temperature, although the bias is very uneven between models. For example, in the case of the minimum temperature, the GEFS model tends to overestimate the temperature by 2.5°C, while the GEM model tends to underestimate it by about 2°C. The ICON and ECMWF forecasts are also slightly biased by about 1°C.

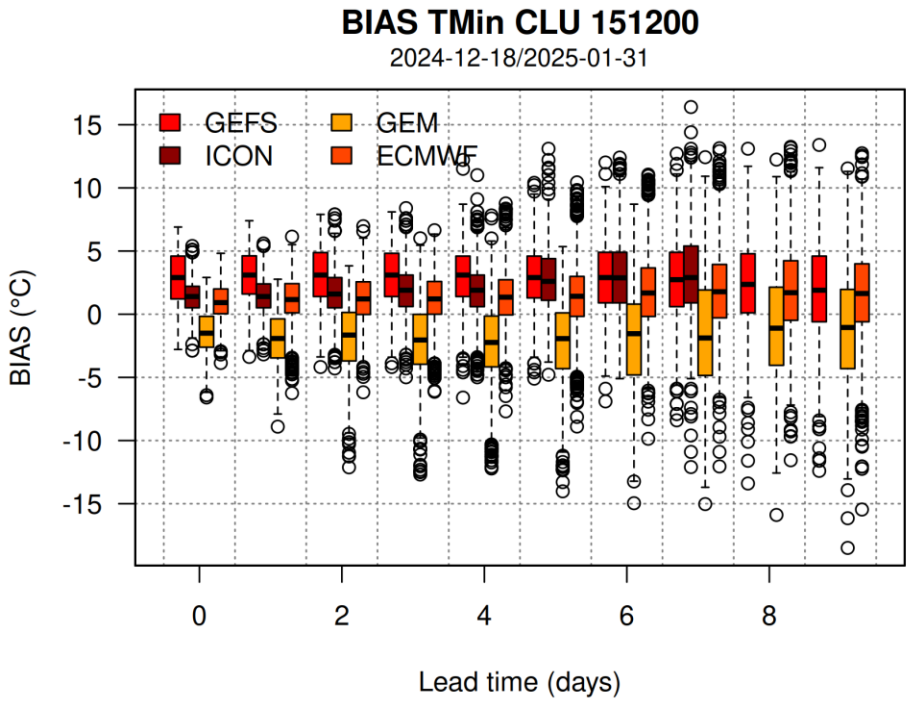

Figure 32. Minimum temperature BIAS in Cluj-Napoca

As can be seen in the figure below, the MAE of the uncorrected forecasts is very high, especially in the case of the GEFS model, which exceeds 2.5°C on average. The model which performs best is the ECMWF, whose errors are less than 2°C even at horizons beyond 5 days. The predictions of the models with higher BIAS can be largely corrected by our simple method to bring the average error below 1.5°C, being the ECMWF model the best one after correction too.

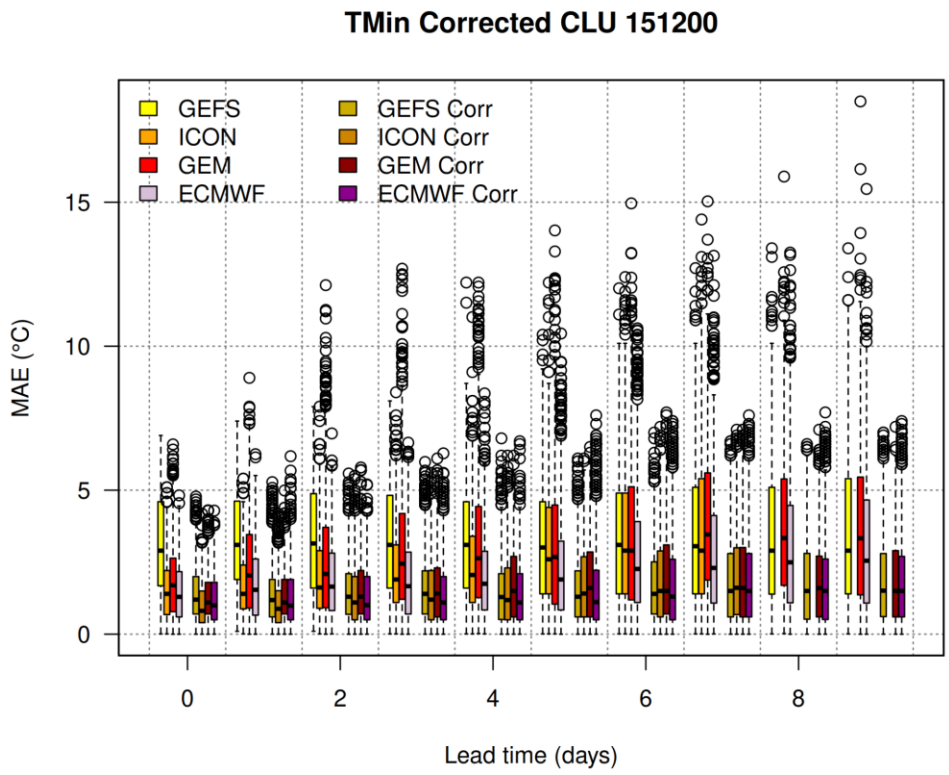

Figure 33. MAE for corrected minimum temperature in Cluj-Napoca

The case of Cluj-Napoca differs from the other cities in the GEFS predictions for the years 2023 and 2024. In this case, the predictive capacity of the model is much more limited, especially with a very high number of false positives (days when the model predicts rain, but in the end there is no rain). The values of the area under the curve (AUC) are between 0.7 and 0.8 for the first forecast horizons, and for a forecast 9 days ahead this value drops to 0.6, which is very close to the critical value of 0.5 that would indicate that the model has no skill. This fact can be seen on the ROC curve, where it can be seen that the curves for the D+8 and D+9 forecasts are close to the diagonal line:

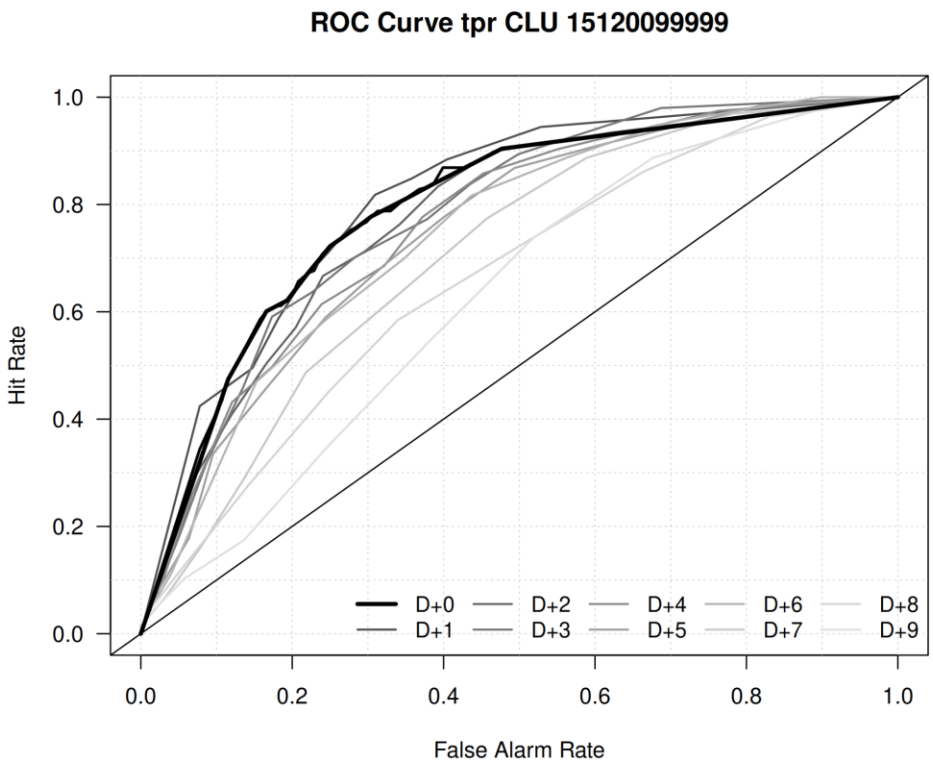

Figure 34. ROC curve for precipitation in Cluj-Napoca. GEFS 2023-2024

4.2 Seasonal forecast

During the zero stage, several CNN methods were compared, for example: multilayer (which does not use convolutional layers), CNN with only a shallow depth of analysis, a "deeper" CNN, also called DeepCNN, a multilinear method which does not also use convolutional layers but implies several fully connected linear layers, and a Linear Deep CNN which includes the deepest analysis with several layers. In this sense, in all scales, this method represents the best performance for the selection of predictors and the first stages of seasonal forecasts (Figure 35).

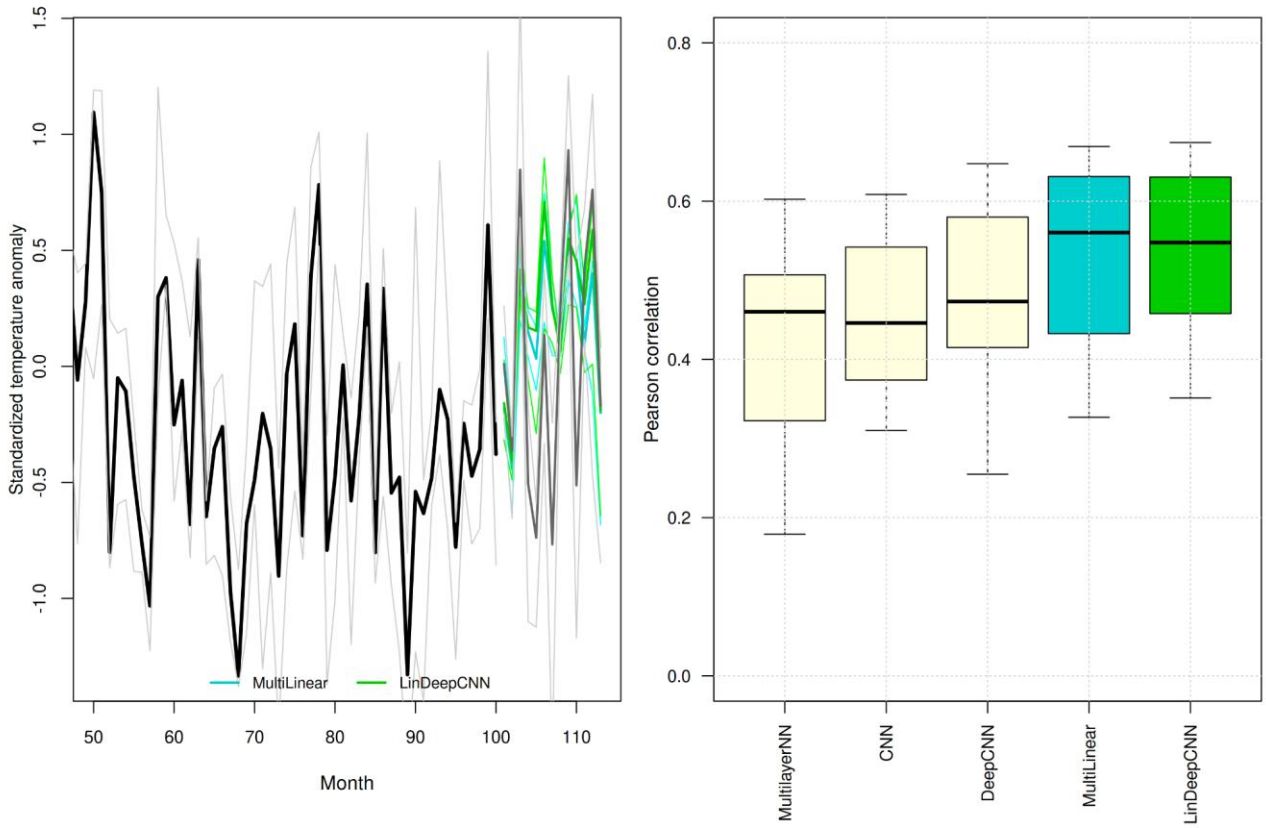

Figure 35. Figure of the CNN zero stage of predictor selection. On the left, the comparison between all the observations (for the 9 observatories in black) and the two best performed methods (MultiLinear in blue and LinearDeepCNN in green). **On the right, a boxplot showing the Pearson Correlation for each method used**

Preliminary results on seasonal forecasting show that the TeWA approach improves the predictability of first-month temperature and precipitation anomalies by 50%–70% compared with the forecast of SEAS5. On a moving-averaged daily scale, the optimum prediction window is 30 days for temperature and 16 days for precipitation. The predictable ranges are consistent with atmospheric bridges in teleconnection patterns [e.g., Upper-Level Mediterranean Oscillation (ULMO)] and are reflected by spatial correlation with sea surface temperature (SST). Our results suggest that combinations of the TeWA approach and numerical models could boost new research lines in subseasonal-to-seasonal forecasting.

4.3 Climate projections

Results for the Living Lab of Madrid

Future climate scenarios for temperature

According to the climate change local scenarios generated for Madrid (see in the figure 9 the location of the observatories used), it can be expected an increase in temperatures across all scenarios, time horizons, and periods of the year. Maximum temperature increases (figure 36) are expected to range from 2.6 to 3.2 degrees Celsius by mid-century and from 3.3 to 6.0 by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 20.9ºC.

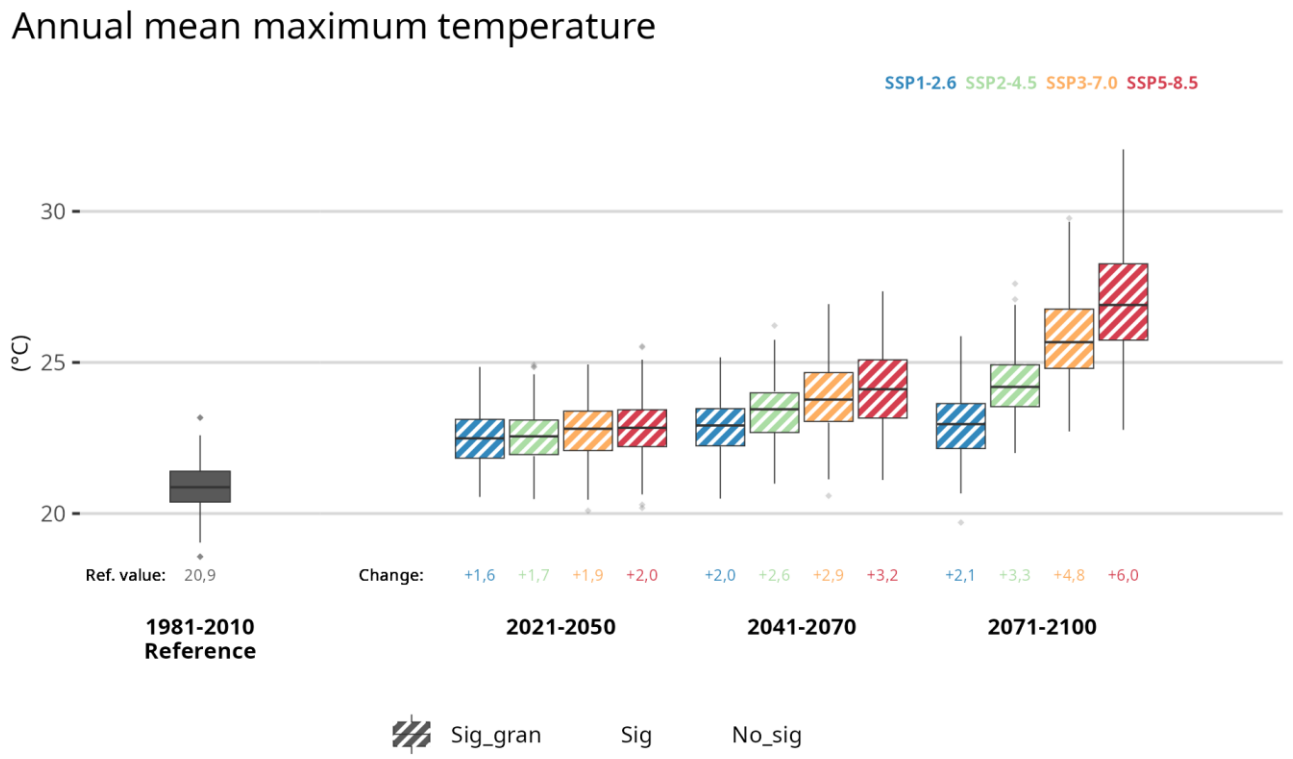

Figure 36. Expected evolution of maximum temperature for Madrid city throughout the 21st century **for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Madrid. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and**

25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results

Similar results are obtained for minimum temperature (figure 37), although they are somewhat less pronounced. This variable is expected to range from 2.6 to 3.2 degrees Celsius by mid-century and from 3.3 to 6.0 by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively, with respect to the Historical value of 8.6ºC.

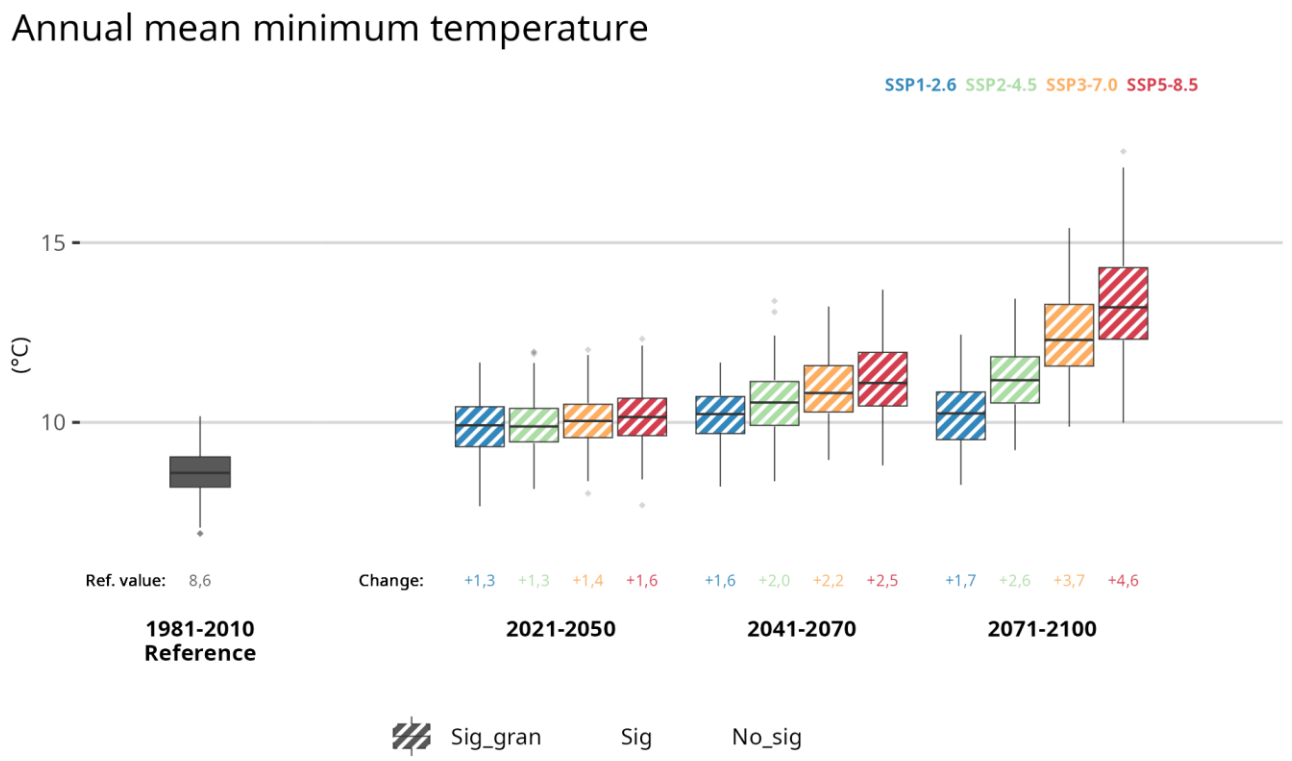

Figure 37. Expected evolution of minimum temperature for Madrid city **throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Madrid. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results**

These results will lead to an increase in extreme temperature-related events, such as the intensification of heatwaves and tropical nights, as well as a decrease in frost episodes.

Future climate scenarios for precipitation

Regarding the precipitation results, no significant changes are expected in the annual accumulated precipitation, as shown in figure 38. Regarding the precipitation results, no significant changes are expected in the annual accumulated precipitation, as shown in the figure. However, it is expected that the way precipitation occurs will be altered, changing the monthly distribution of precipitation and, consequently, the contributions to water resources.

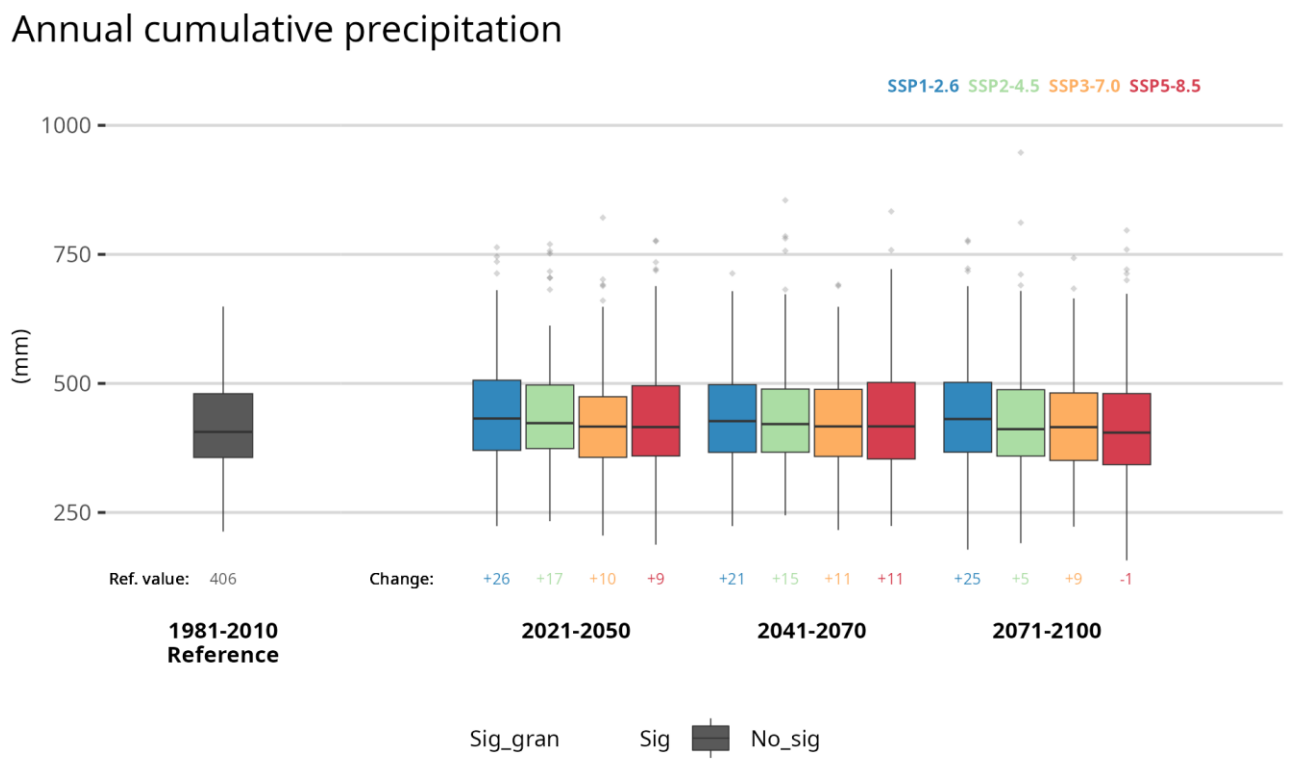

Figure 38. Expected evolution of annual cumulative precipitation for Madrid city **throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Madrid. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results**

The precipitation results, combined with the effect of rising temperatures, will impact factors such as evapotranspiration, drought episodes, and more. Therefore, it will be important to analyze these variables in subsequent phases of the URBREATH project.

Results for the Living Lab of Leuven

Future climate scenarios for temperature

According to the climate change local scenarios generated for Leuven (see in the figure 8 the location of the observatories used), it can be expected an increase in temperatures across all scenarios, time horizons, and periods of the year.

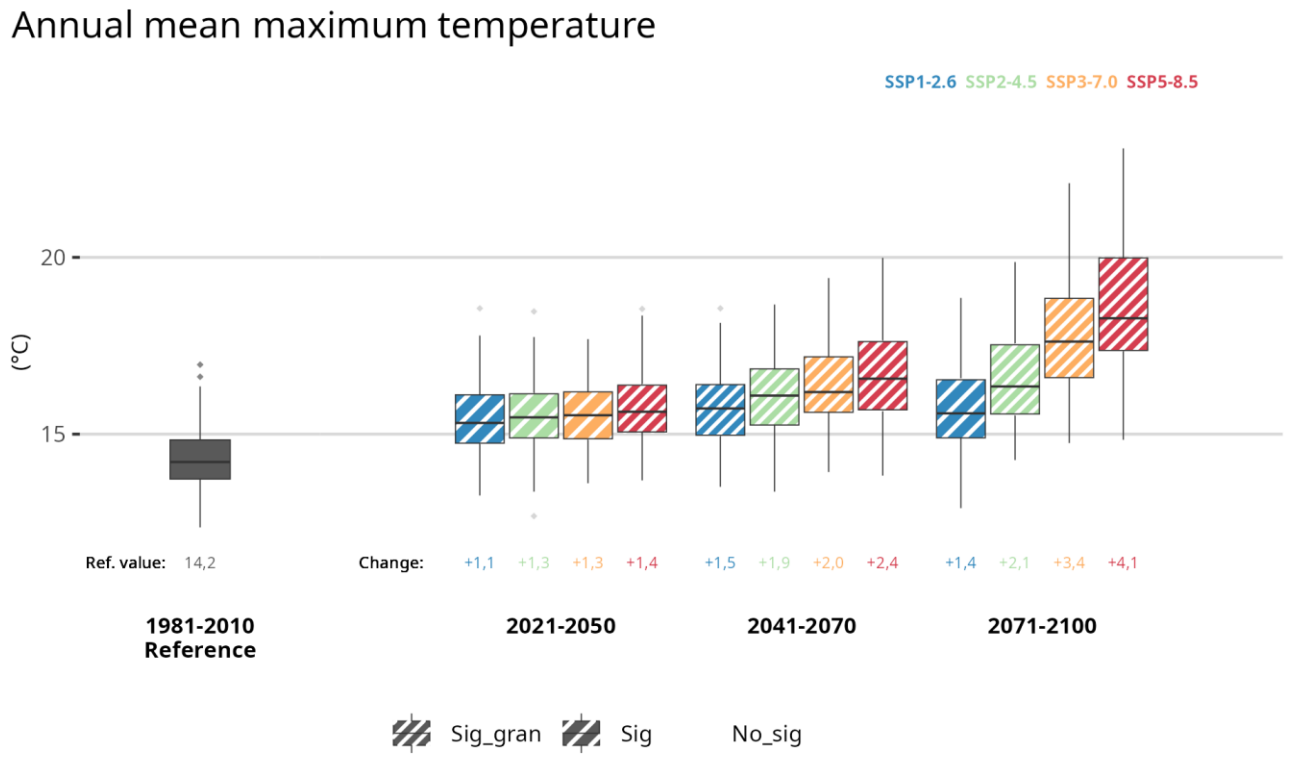

Figure 39. Expected evolution of maximum temperature for Leuven **throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Leuven. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and**

25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results

Maximum temperature increases (figure 39) are expected to range from 1.9 to 2.4 degrees Celsius by mid-century and from 2.1 to 4.1 by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 14.2ºC. Similar results are obtained for minimum temperature (figure 40), although they are somewhat less pronounced. This variable is expected to range from 1.4 to 1.8 degrees Celsius by mid-century and from 1.7 to 3.3 by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively, with respect to the Historical value of 7.0ºC.

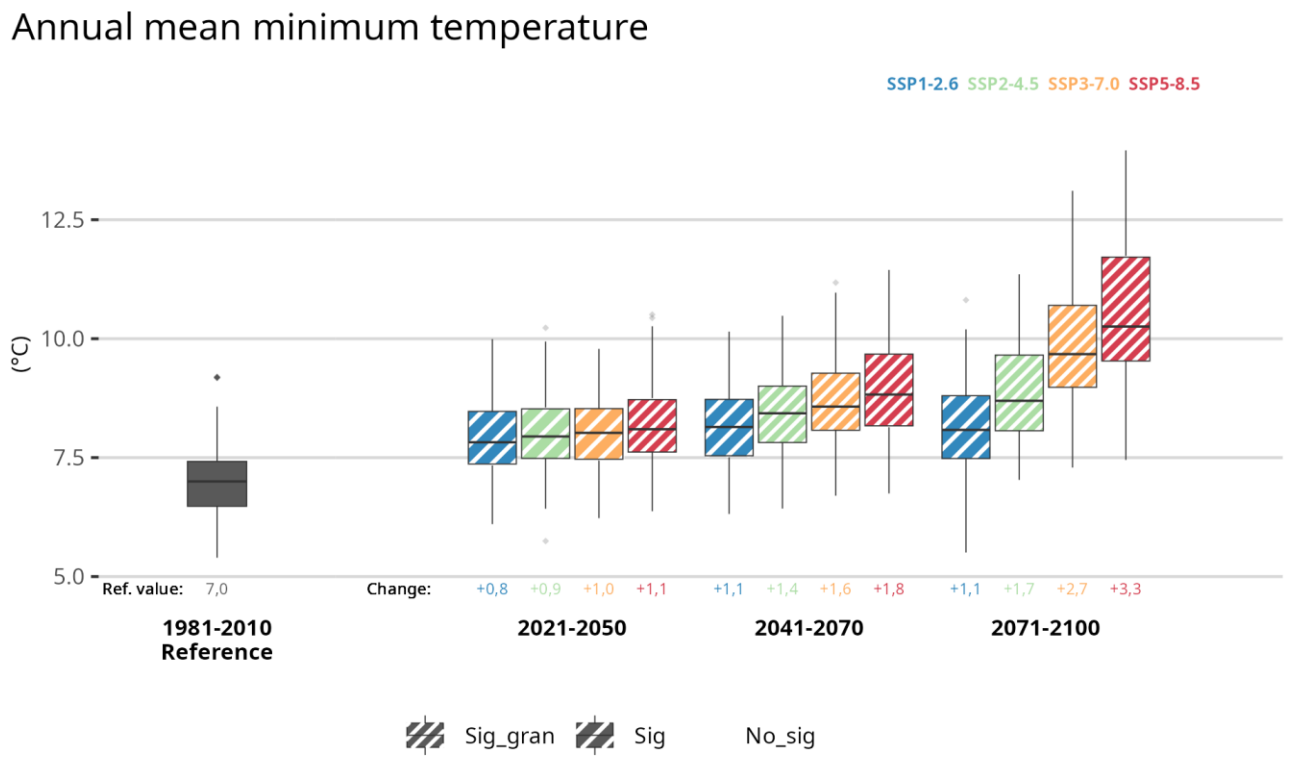

Figure 40. Expected evolution of minimum temperature for Leuven **throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Leuven. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results**

These results will lead to an increase in extreme temperature-related events, such as the intensification of heatwaves and tropical nights, as well as a decrease in frost episodes.

Future climate scenarios for precipitation

Regarding the precipitation results, no significant changes are expected in the annual accumulated precipitation, as shown in figure 38. Regarding the precipitation results, no significant changes are expected in the annual accumulated precipitation, as shown in the figure 41, being significant the expected slight increases by the end of the century. On the other hand, it is expected that the way precipitation occurs will be altered, changing the monthly distribution of precipitation and, consequently, the contributions to water resources.

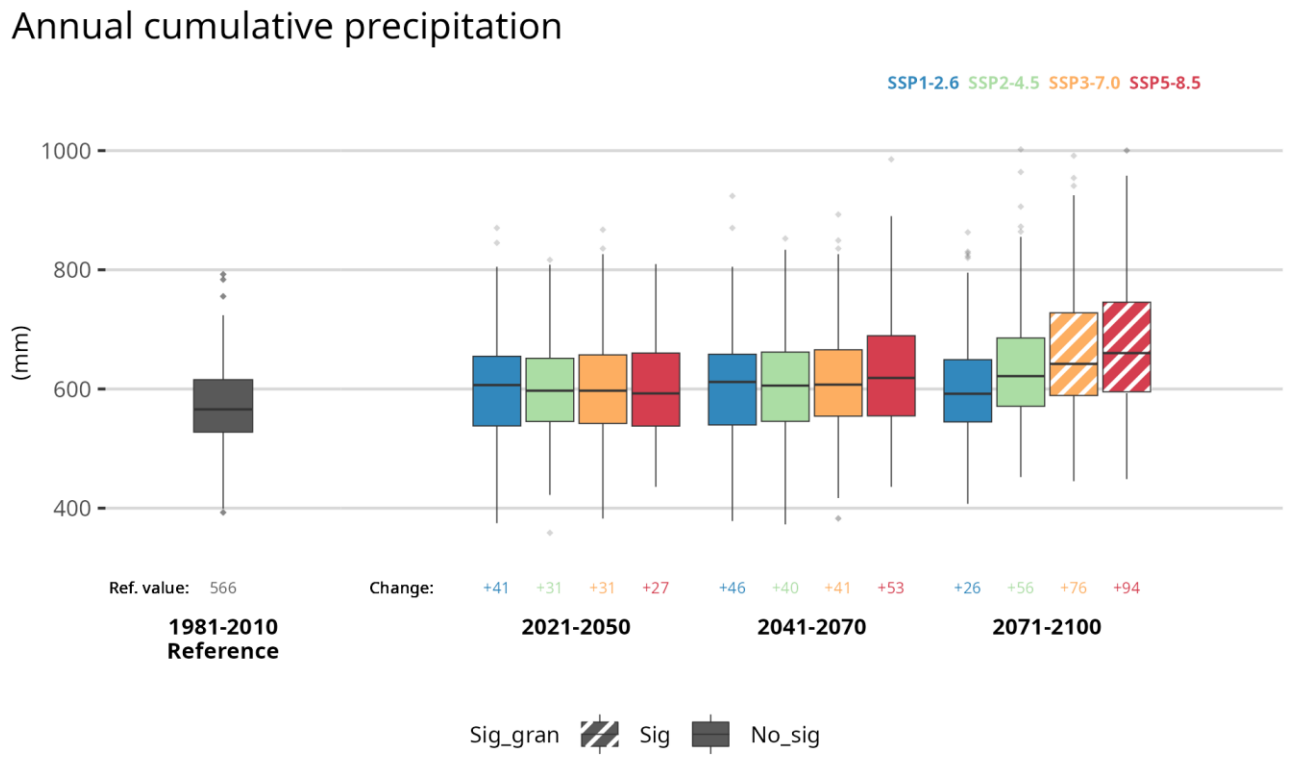

Figure 41. Expected evolution of annual cumulative precipitation for Leuven **throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the**

observatories employed in Leuven. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results

The precipitation results, along with the influence of rising temperatures, will affect factors like evapotranspiration, drought events, and other related aspects. As a result, it will be crucial to examine these variables in the following phases of the URBREATH project.

Results for the Living Lab of Tallinn

Future climate scenarios for temperature

According to the climate change local scenarios generated for Tallinn (see in the figure 10 the location of the observatories used), It can be expected an increase in temperatures across all scenarios, time horizons, and periods of the year.

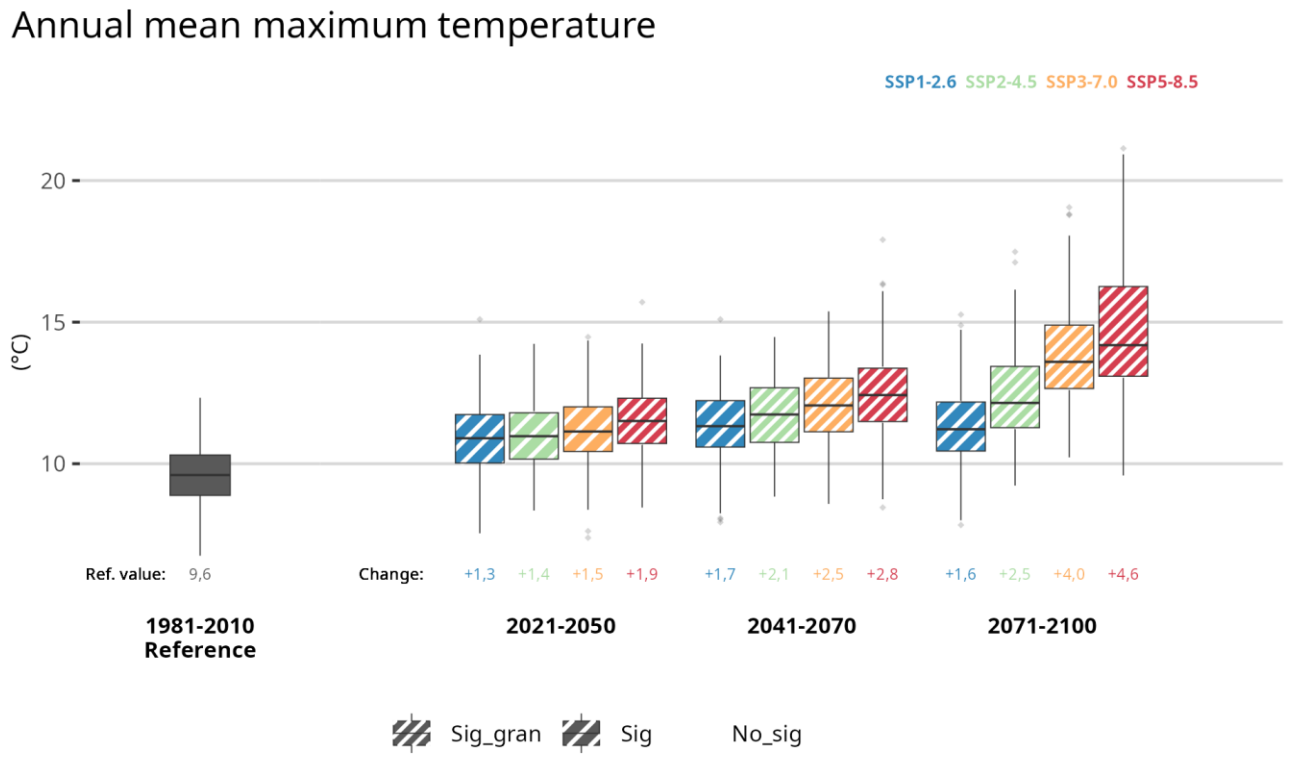

Figure 42. Expected evolution of maximum temperature for Tallinn **throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Tallinn. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results**

Maximum temperature increases (figure 42) are expected to range from 2.1 to 2.8 degrees Celsius by mid-century and from 2.5 to 4.6 by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 9.6ºC. Similar results are obtained for minimum temperature (figure 43), although they are somewhat less pronounced. This variable is expected to range from 2.1 to 2.7 degrees Celsius by mid-century and from 2.6 to 4.6 by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively, with respect to the Historical value of 4.3ºC.

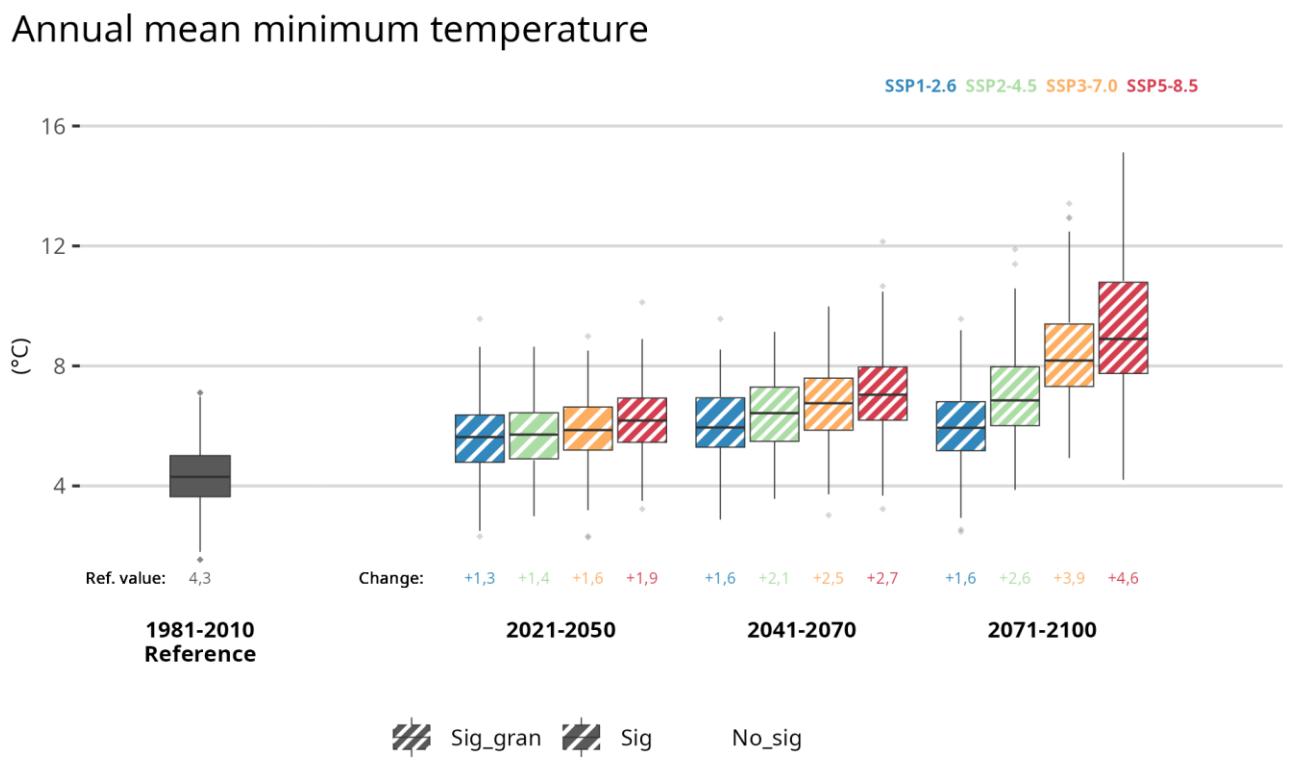

Figure 43. Expected evolution of minimum temperature for Tallinn **throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories**

employed in Tallinn. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results

These results will lead to an increase in extreme temperature-related events, such as the intensification of heatwaves and tropical nights, as well as a decrease in frost episodes.

Future climate scenarios for precipitation

No significant changes are expected in the annual accumulated precipitation (figure 44), while by the end of the century, an increase of between 141 and 284 mm is expected which represents an increase of 20.7 % and 41.8 % compared to the reference value of 679 mm.

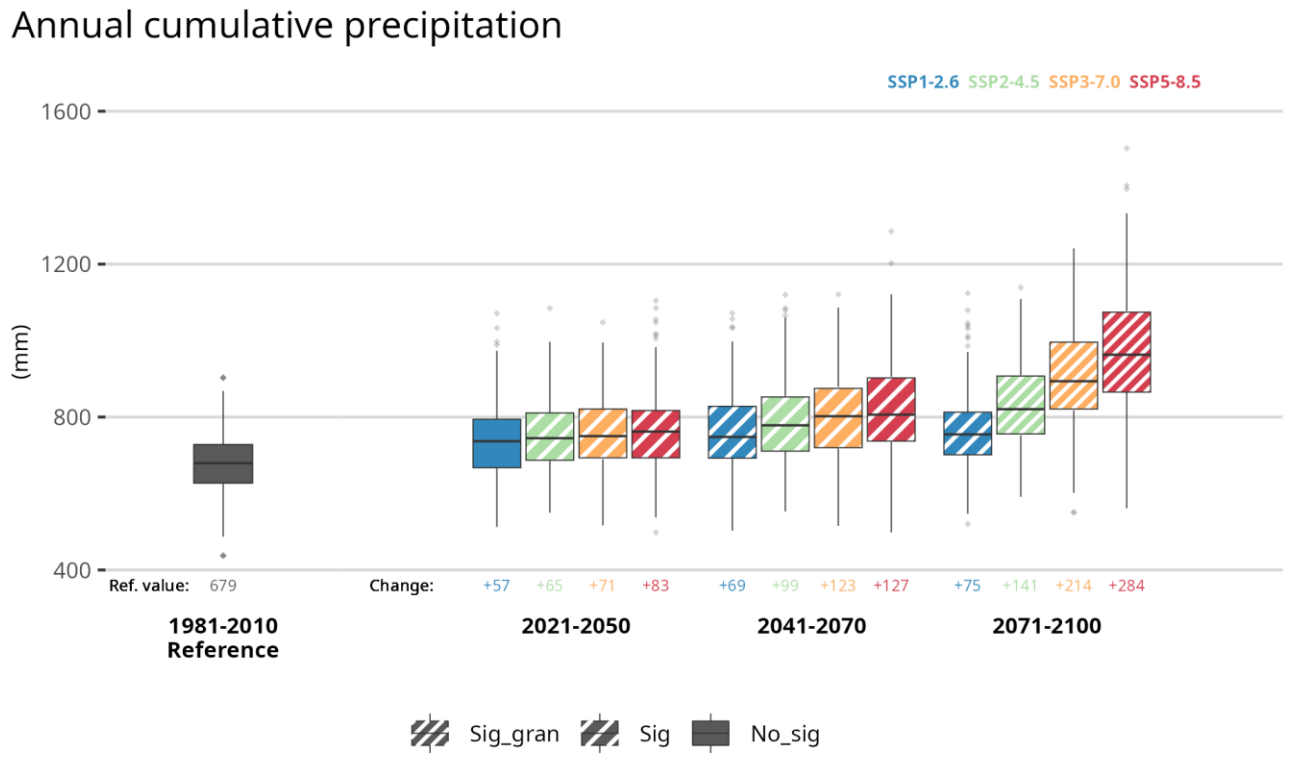

Figure 44. Expected evolution of annual cumulative precipitation for Tallinn **throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100)**

under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Tallinn. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results

The precipitation results, combined with the effect of rising temperatures, will impact factors such as evapotranspiration, drought episodes, and more. Therefore, it will be important to analyze these variables in subsequent phases of the URBREATH project.

Results for the Living Lab of Cluj-Napoca

Future climate scenarios for temperature

According to the climate change local scenarios generated for Cluj-Napoca (see in the figure 7 the location of the observatories used), it can be expected an increase in temperatures across all scenarios, time horizons, and periods of the year. Maximum temperature increases (figure 45) are expected to range from 2.3 to 3.1 degrees Celsius by mid-century and from 2.9 to 5.3 by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 14.2ºC.

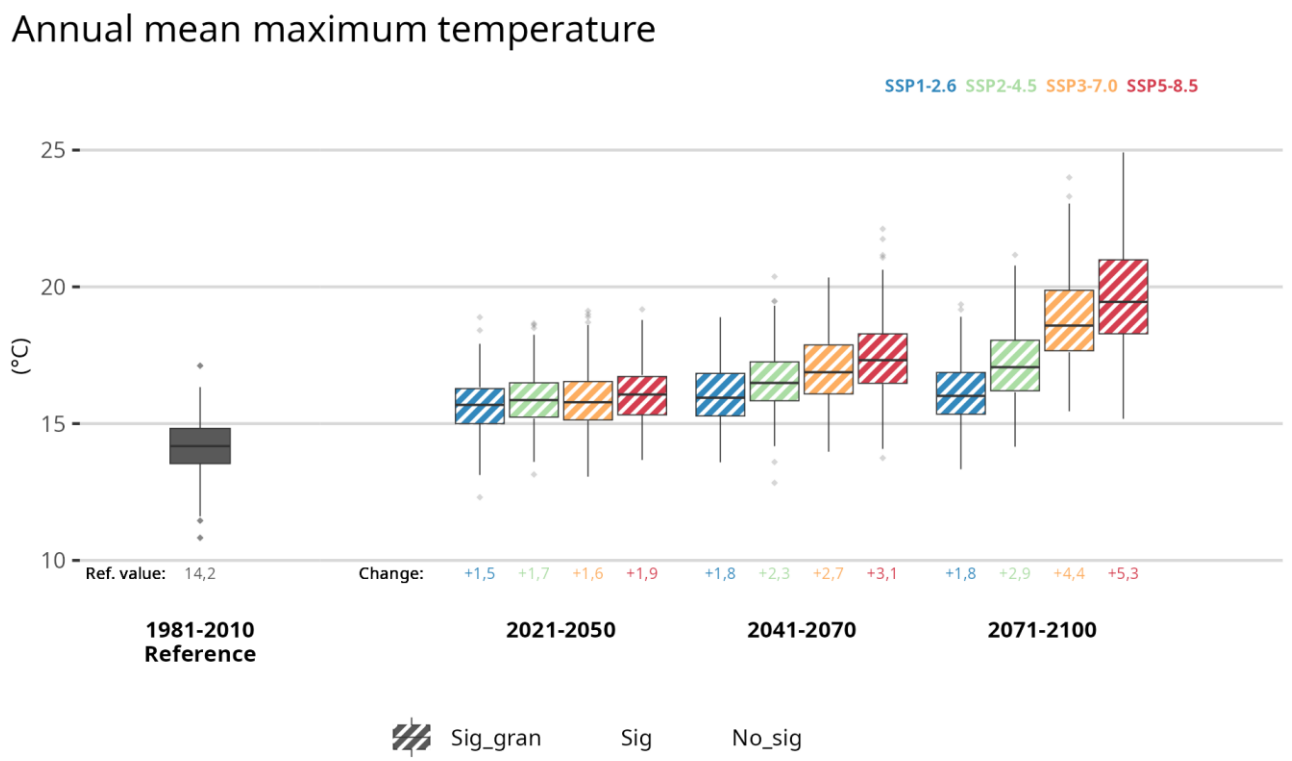

Figure 45. Expected evolution of maximum temperature for Cluj-Napoca **throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Cluj-Napoca. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results**

Similar results are obtained for minimum temperature (figure 46), although they are somewhat less pronounced. This variable is expected to range from 1.8 to 2.5 degrees Celsius by mid-century and from 2.3 to 4.1 by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively, with respect to the Historical value of 4.9ºC.

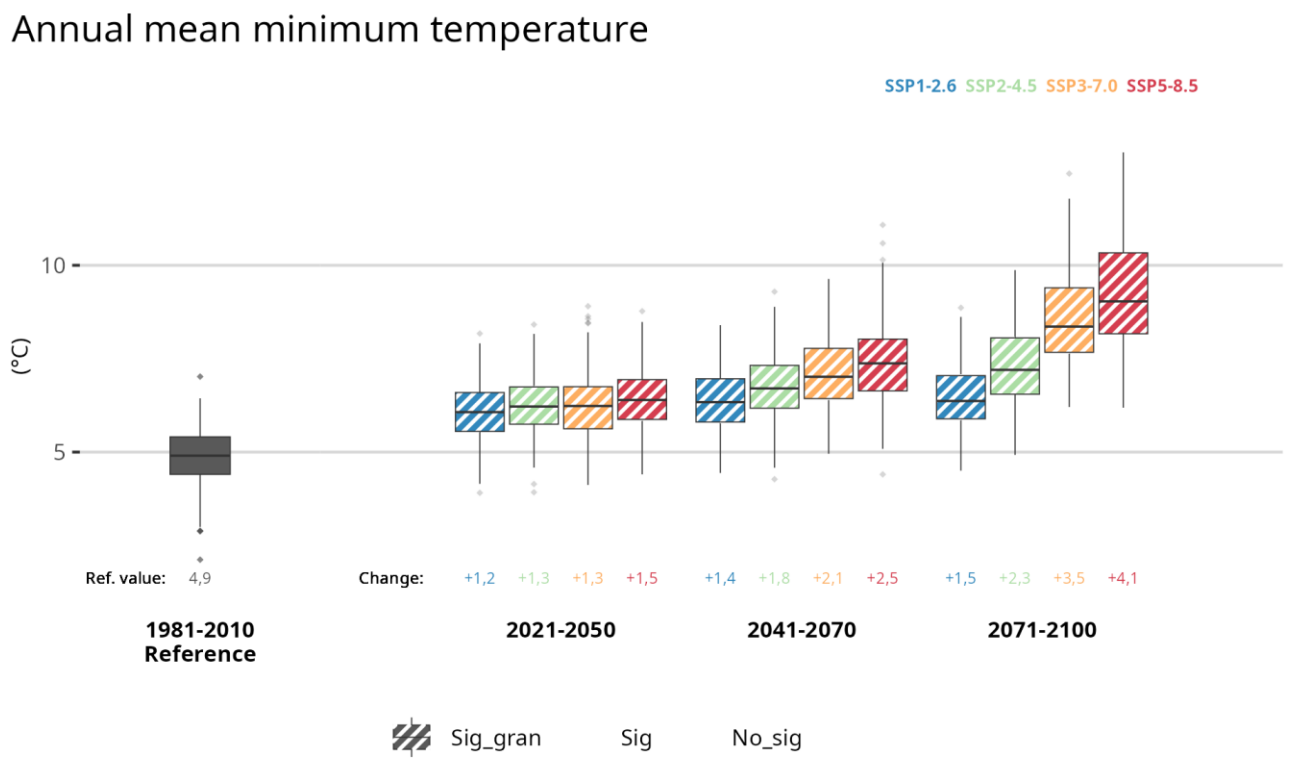

Figure 46. Expected evolution of minimum temperature for Cluj-Napoca **throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Cluj-Napoca. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results**

These results will lead to an increase in extreme temperature-related events, such as the intensification of heatwaves and tropical nights, as well as a decrease in frost episodes.

Future climate scenarios for precipitation

Light changes are expected in the annual accumulated precipitation by the end of the century, an increase of between 73 and 89 mm is expected, which represents an increase of 15,2% and 18,5% compared to the reference value of 480 mm (Figure 47).

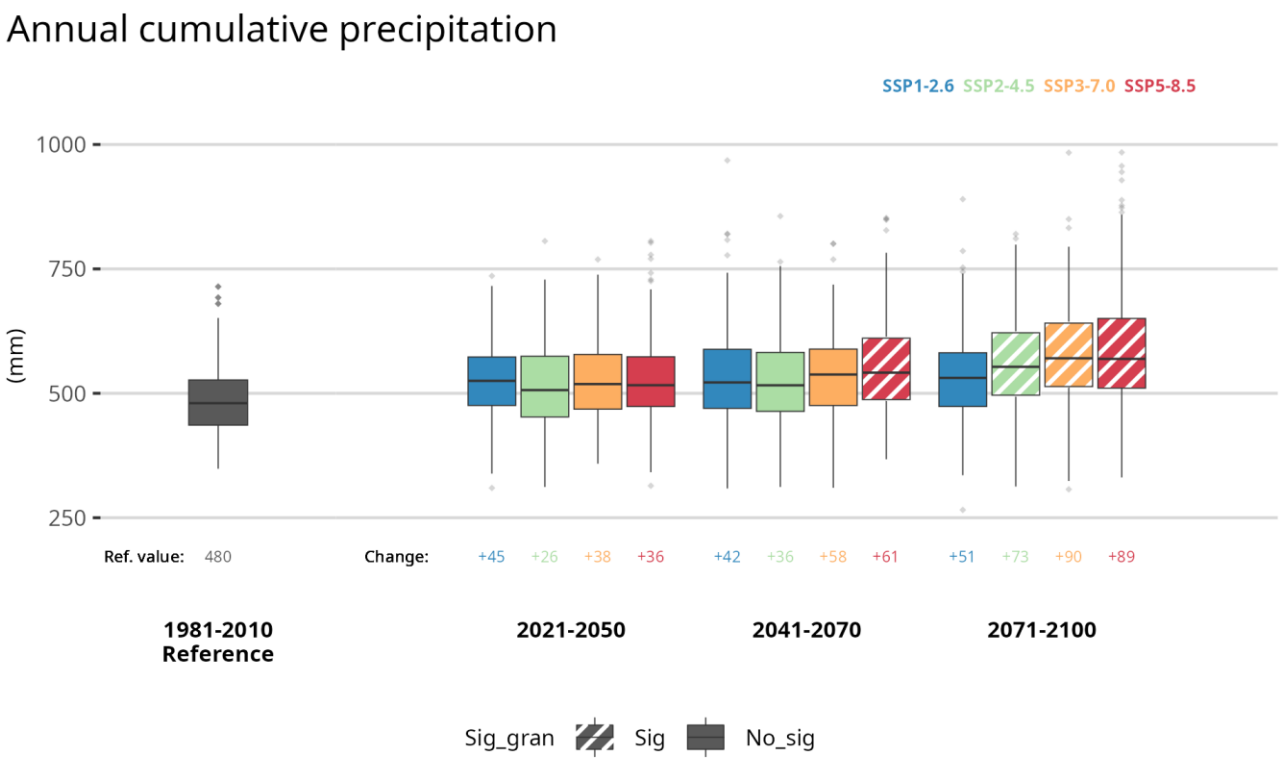

Figure 47. Expected evolution of annual cumulative precipitation for Cluj-Napoca

***Figure 47. Expected evolution of annual cumulative precipitation for Cluj-Napoca* throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories employed in Cluj-Napoca. The boxplots represent the results from 10 climate models, with the central horizontal line corresponding to the median and the upper and lower extremes representing the 75th and 25th percentiles, respectively. The lower value indicates the increase compared to the absolute value for the Historical period. The fill codes specify the statistical significance of the results**

The precipitation results, combined with the effect of rising temperatures, will impact factors such as evapotranspiration, drought episodes, and more. Therefore, it will be important to analyze these variables in subsequent phases of the URBREATH project.

5. Conclusions

The conclusions presented in this report reflect the urgent need for local-scale climate assessments, particularly in light of the profound implications of climate change as highlighted by the latest IPCC report. Human activity has clearly driven the climate emergency, and its impact is already being felt globally, with increasing temperatures, rising sea levels, and extreme weather events affecting regions around the world. This study emphasizes the necessity of adapting to these changes by utilizing both short-term weather forecasts and long-term climate projections. By applying advanced methodologies, including artificial intelligence, the accuracy and reliability of climate and weather models can be continually enhanced, providing more tailored insights for local decision-making. Despite some data access limitations, the preliminary findings offer valuable directions for refining forecasting tools and adapting strategies to mitigate the impacts of climate change in the coming decades.

The evaluation of ensemble weather prediction models in this deliverable has provided valuable insights into the systematic biases and predictive skill of different models for urban-scale forecasting. Due to data availability constraints, the full historical comparison of all ensemble models over the desired period (2023-2024) was not feasible, as outlined in Section 2.3.2. However, the assessment conducted for the December 2024 – January 2025 period using available models (GEFS, ECMWF, ICON, and GEM) revealed significant differences in forecast accuracy. ECMWF demonstrated higher overall skill in predicting temperature variations, while GEFS exhibited notable biases, particularly in minimum temperature forecasts. The results highlight the importance of model selection and post-processing methodologies to improve forecasting performance at the urban level. The final results presented on the next iteration of the deliverable (December 2025) will consider all this gathered valuable information, assessing the best base model final choice.

In addition to the comparative model evaluation, this deliverable also assessed the potential for systematic bias correction in ensemble forecasts using the full 2023-2024 dataset for GEFS and more limitedly for a shorter period with all EPS models. The application of statistical correction techniques showed notable improvements in forecast accuracy, establishing a baseline for future enhancements testing all available models. The findings from this phase will guide the next deliverable (December 2025), where the integration of AI-driven post-processing techniques are expected for operational results. These improvements aim to optimize urban-scale probabilistic forecasting, ensuring more reliable predictions that will ultimately support climate resilience efforts in the project's Front Runner cities.

Seasonal forecasting is a rapidly growing discipline. The development of new techniques and the application of neural networks have had a major impact on its growth. An improvement of more than 50% has been achieved in the first "windows" of the forecast (the first 2 months or the first 60 days) by

using a linear deep CNN combination of neural networks. It should also be noted that, considering our techniques, a better performance in seasonal forecasting has been observed in the Mediterranean cities of Athens and Madrid, but this is preliminary and can be modified as the project progresses. A final combination of this methodology with dynamic seasonal forecasting could strengthen the final prediction, since the moment the two signals could agree and reduce the uncertainty at these long-time scales.

The applied climate methodology is a two-step analog method that enables the generation of localized climate information by incorporating microclimatology through the use of observed data. The results obtained are unanimous regarding the expected temperature variations, showing progressive increases throughout the 21st century in alignment with global warming trends. Regarding precipitation, no significant changes are expected until mid-century. However, by the end of the century, Cluj-Napoca and Madrid are projected to experience only minor alterations, while Leuven and Tallinn are expected to see an increase in accumulated precipitation.

Adjusting both weather prediction methodologies and climate projection techniques allows for more precise planning, helping to avoid unnecessary risks. For example, it can prevent inaction in the face of short-term flooding while also guiding long-term decisions, such as infrastructure development and green space planning.

In conclusion, the findings of this project underscore the pressing need for localized climate assessments and adaptation strategies in response to the ongoing climate crisis, across all time scales. Whether addressing short-term weather patterns or long-term climate projections, the results emphasize the importance of tailoring solutions to specific local needs and ensuring that both immediate and future impacts are effectively considered in decision-making processes.

6. References

Abbott-Halpin, E. and Rankin, C., 2020. Introduction: Public Library Governance and Wicked Problems. In Public Library Governance (pp. 5-16). De Gruyter Saur.

Alexe M., Ben Bouallegue Z. (2023). ECMWF unveils alpha version of new ML model. doi: 10.21957/d29513c51f

Armour-Gemmen, M.G., 2020. Innovation for the Engaged Librarian, American Society for Engineering Education, 2020, Faculty & Staff Scholarship. 2945. https://peer.asee.org/34831.pdf.

Barras, R. (1990). Interactive innovation in financial and business services: the vanguard of the service revolution. Research Policy, 19, 215 – 237.

Beele, E. et al, (2022), "Replication Data for: Quality control and correction method for air temperature data from a citizen science weather station network in Leuven, Belgium", https://doi.org/10.48804/SSRN3F, KU Leuven RDR, V6

Benáček, P. et al. "Postprocessing of Ensemble Weather Forecast Using Decision Tree–Based Probabilistic Forecasting Methods". Weather and Forecasting 38.1 (2023): 69-82. https://doi.org/10.1175/WAF-D-22-0006.1

Bentsen, M., Bethke, I., Debernard, J.B., Iversen, T., Kirkevag, A., Seland, O., et al., 2013. The Norwegian Earth System Model, NorESM1-M - part 1: description and basic evaluation of the physical climate. Geosci. Model Dev. 6, 687–720.

Bentsen, Mats; Oliviè, Dirk Jan Leo; Seland, Øyvind; Toniazzo, Thomas; Gjermundsen, Ada; Graff, Lise Seland; Debernard, Jens Boldingh; Gupta, Alok Kumar; He, Yanchun; Kirkevåg, Alf; Schwinger, Jörg; Tjiputra, Jerry; Aas, Kjetil Schanke; Bethke, Ingo; Fan, Yuanchao; Griesfeller, Jan; Grini, Alf; Guo, Chuncheng; Ilicak, Mehmet; Karset, Inger Helene Hafsahl; Landgren, Oskar Andreas; Liakka, Johan; Moseid, Kine Onsum; Nummelin, Aleksi; Spensberger, Clemens; Tang, Hui; Zhang, Zhongshi; Heinze, Christoph; Iversen, Trond; Schulz, Michael (2019). NCC NorESM2-MM model output prepared for CMIP6 ScenarioMIP. Version YYYYMMDD[1]. Earth System Grid Federation. doi: 10.22033/ESGF/CMIP6.608.

Bi, Daohua, Dix, Martin, Marsland, Simon, O'Farrell, Siobhan, Sullivan, Arnold, Bodman, Roger, Law, Rachel, Harman, Ian, Srbinovsky, Jhan, Rashid, Harun A., Dobrohotoff, Peter, Mackallah, Chloe, Yan, Hailin, Hirst, Anthony, Savita, Abhishek, Dias, Fabio Boeira, Woodhouse, Matthew, Fiedler, Russell,

Heerdegen, Aidan, 2020: Configuration and spin-up of ACCESS-CM2, the new generation Australian Community Climate and Earth System Simulator Coupled Model. Journal of Southern Hemisphere Earth Systems Science, 70, 225, doi:10.1071/es19040.

Bouallègue, B. et al (2023). The rise of data-driven weather forecasting. doi: 10.48550/arXiv.2307.10128.

Bouttier, F. et al. (2024). Probabilistic short-range forecasts of high-precipitation events: optimal decision thresholds and predictability limits. Natural Hazards and Earth System Sciences. https://doi.org/10.5194/nhess-24-2793-2024

Cawcr - Joint Working Group on Forecast Verification Research. Forecast Verification - Issues, Methods and FAQ. Methods for dichotomous (yes/no) forecasts. Contingency table. Link: https://www.cawcr.gov.au/projects/verification/verif_web_page.html

Cherchi, A., Fogli, P.G., Lovato, T., Peano, D., Iovino, D., Gualdi, S., Masina, S., Scoccimarro, E., Materia, S., Bellucci, A. and Navarra, A., 2019. Global mean climate and main patterns of variability in the CMCC-CM2 coupled model. Journal of Advances in Modeling Earth Systems, 11(1), pp.185-209. https://doi.org/10.1029/2018MS001369.

Chylek, P., Li, J., Dubey, M.K., Wang, M., Lesins, G., 2011. Observed and model simulated 20th century Arctic temperature variability: Canadian Earth System Model CanESM2. Atmos. Chem. Phys. Discuss. 11, 22893–22907. https://doi.org/10.5194/acpd-11- 22893-2011

Collins, W.J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Hinton, T., Jones, C.D., Liddicoat, S., Martin, G., O'Connor, F., Rae, J., Senior, C., Totterdell, I., Woodward, S., Reichler, T., Kim, J., Halloran, P., 2008. Evaluation of the HadGEM2 model. Hadley Centre Technical Note HCTN. vol 74. Met Office Hadley Centre, Exeter, UK.

Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [MassonDelmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press. In Press.

Dunne, J.P., John, J.G., Adcroft, A.J., Griffies, S.M., Hallberg, R.W., Shevliakova, E., et al., 2012. GFDL's ESM2 global coupled climate-carbon Earth System Models. Part I: physical formulation and baseline simulation characteristics. J. Clim. 25, 6646–6665.

EC-Earth Consortium (EC-Earth) (2019). EC-Earth-Consortium EC-Earth3-Veg model output prepared for CMIP6 Scenario MIP. Earth System Grid Federation. doi:10.22033/ESGF/CMIP6.727.
Evin, G., et al. (2021): Calibrated ensemble forecasts of the height of new snow using quantile regression forests and ensemble model output statistics, Nonlin. Processes Geophys., 28, 467–480, https://doi.org/10.5194/npg-28-467-2021.

ECMWF web page, section 12.B Statistical Concepts - Probabilistic Data. Link: https://confluence.ecmwf.int/display/FUG/Section+12.B+Statistical+Concepts+-+Probabilistic+Data#:~:text=The%20Ranked%20Probability%20Score%20

Good P., Sellar A., Tang Y., Rumbold S., Ellis R., Kelley D., Kuhlbrodt T. (2019). MOHC UKESM1.0-LL model output prepared for CMIP6 ScenarioMIP ssp245. 1]. Earth System Grid Federation. doi:10.22033/ESGF/CMIP6.6339.

Guo, Q., He, Z., & Wang, Z. (2024). Monthly climate prediction using deep convolutional neural network and long short-term memory. Scientific Reports, 14(1), 17748.

Gutiérrez, J. M., Maraun, D., Widmann, M., Huth, R., Hertig, E., Benestad, R., Ribalaygua, J., Pórtoles, J., … & Pagé, C. (2019). An intercomparison of a large ensemble of statistical downscaling methods over Europe: Results from the VALUE perfect predictor cross-validation experiment. International journal of climatology, 39(9), 3750-3785

IPCC, 2021a: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press. In Press
IPCC, 2021b: Summary for Policymakers. In: Climate Change 2021: The Physical Science Basis.

Kalnay, E. (2002). Atmospheric Modeling, Data Assimilation and Predictability. Cambridge: Cambridge University Press.

Kumari P., Toshniwal D. (2021). Long short term memory–convolutional neural network based deep hybrid approach for solar irradiance forecasting, Applied Energy, Volume 295,117061,ISSN 0306-2619, https://doi.org/10.1016/j.apenergy.2021.117061.

Monjo R, Pórtoles J, Ribalaygua J. 2013. Detection of inhomogeneities in daily data: a test based in the Kolmogorov-Smirnov goodness-of-fit test. 9th Data Management Workshop of EUMETNET, El Escorial (Madrid), 6th-8th November.

Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., ... & Marotzke, J. (2018). A higher-resolution version of the max planck institute earth system model (MPI-ESM1. 2-HR). Journal of Advances in Modeling Earth Systems, 10(7), 1383-1413.

O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., & Sanderson, B. M. (2016). The scenario model intercomparison project (ScenarioMIP) for CMIP6. Geoscientific Model Development, 9(9), 3461-3482.

O'Neill, B. C., Kriegler, E., Ebi, K. L., Kemp-Benedict, E., Riahi, K., Rothman, D. S., ... & Solecki, W. (2017). The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. Global environmental change, 42, 169-180. DOi: https://doi.org/10.1016/j.gloenvcha.2015.01.004

Price, Ilan et al. (2023) "GenCast: Diffusion-based ensemble forecasting for medium-range weather." Doi: 10.48550/arXiv.2312.15796

Rasp, S., & Lerch, S. (2018). Neural networks for post-processing ensemble weather forecasts. ArXiv, abs/1805.09091. https://doi.org/10.1175/MWR-D-18-0187.1.

Ribalaygua J, Torres L, Portoles J, Monjo R, Gaitan E, Pino MR. Description and validation of a two-step analogue/regression downscaling method. Theoretical and Applied Climatology 2013b; 114: 253-269.

Scheuerer, M. (2018). Probabilistic Forecasting of Snowfall Amounts Using a Hybrid between a Parametric and an Analog Approach. Monthly Weather Review. https://doi.org/10.1175/MWR-D-18-0273.1.

Seferian R. (2019). CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 AerChemMIP hist-1950HC. Version YYYYMMDD[1].Earth System Grid Federation. doi:10.22033/ESGF/CMIP6.4041.

Sha, Y., Sobash, R., & Gagne, D. (2024). Improving ensemble extreme precipitation forecasts using generative artificial intelligence. ArXiv, abs/2407.04882. https://doi.org/10.48550/arXiv.2407.04882.

Springer Nature web page, section of Enciclopedia of System Biology, Area under the ROC Curve, Figure 110. Link: https://link.springer.com/referenceworkentry/10.1007/978-1-4419-9863-7_209/figures/110

Swart N. C., Cole J. N. S., Kharin V. V., Lazare M., Scinocca J. F., Gillett N. P., Anstey J., Arora V., Christian J. R., Hanna S., Jiao Y., Lee W. G., Majaess F., Saenko O. A., Seiler C., Seinen C., Shao, A., Sigmond M., Solheim L., von Salzen K., Yang D. and Winter B. (2019). The Canadian Earth System Model version 5 (CanESM5.0.3), Geosci. Model Dev., 12, 4823–4873, doi:10.5194/gmd-12-4823-2019.
World Meteorological Organization (WMO). (2023). 2023 shatters climate records, with major impacts. Retrieved from:
https://public.wmo.int/en/media/press-release/2023-shatters-climate-records-major-impacts

World Meteorological Organization (WMO), 2023. Provisional State of the Global Climate 2023. Available at: https://wmo.int/news/media-centre/2023-shatters-climate-records-major-impacts

Wu T., Lu Y., Fang Y., Xin X., Li L., Li W., Jie W., Zhang J., Liu Y., Zhang L., Zhang F., Zhang Y., Wu F., Li J., Chu M., Wang Z., Shi X., Liu X., Wei M., Huang A., Zhang Y. and Liu, X. (2019). The Beijing Climate Center Climate System Model (BCC-CSM): the main progress from CMIP5 to CMIP6 , Geosci. Model Dev. 12, 1573–1600. doi:10.5194/gmd-12-1573-2019, 2019.

Xiao-Ge, X., Tong-Wen, W., Jie, Z., 2013. Introduction of CMIP5 experiments carried out with the climate system models of Beijing Climate Center. Adv. Clim. Chang. Res. 4, 41–49. https://doi.org/10.3724/SP.J.1248.2013.041

Yukimoto S., Koshiro T., Kawai H., Oshima N., Yoshida, K., Urakawa S.; Tsujino H., Deushi M., Tanaka T., Hosaka M., Yoshimura H., Shindo E., Mizuta R., Ishii M., Obata A., Adachi, Y. (2019). MRI MRI-ESM2.0 model output prepared for CMIP6 CMIP. Earth System Grid Federation. doi:10.22033/ESGF/CMIP6.621.

Yukimoto, S., Yoshimura, H., Hosaka, M., Sakami, T., Tsujino, H., Hirabara, M., Tanaka, T.Y., Deushi, M., Obata, A., Nakano, H., Adachi, Y., Shindo, E., Yabu, S., Ose, T., Kitoh, A., 2011. Meteorological research institute- earth system model v1 (MRI-ESM1)— model description. Technical Report of MRI. vol 64.

Annex 1. Observed data information

1. Information about Cluj-Napoca's observed data

Table A.1: *Initial observatories for the Cluj-Napoca case study*

ID	Latitude (º)	Longitude (º)	Database	Variable
ROE00100902	46.78	23.57	Meteo Romania	P, T
15120099999	46.79	23.69	Meteo Romania	P, T
15120599999	46.78	23.57	Meteo Romania	P, T
15143099999	46.58	23.78	Meteo Romania	P, T

Table A.2: *Final observatories of the weather forecast study for the Cluj-Napoca case study.*

ID	Latitude (º)	Longitude (º)	Database	Variable
15120099999	46.79	23.69	Meteo Romania	P, T

Table A.3: *Final observatories of the seasonal forecast study for the Cluj-Napoca case study.*

ID	Latitude (º)	Longitude (º)	Database	Variable
15120099999	46.79	23.69	Meteo Romania	P, T

Table A.4: *Final observatories of the climate projections study for the Cluj-Napoca case study.*

ID	Latitude (º)	Longitude (º)	Database	Variable
ROE00100902	46.78	23.57	Meteo Romania	P, T
15120099999	46.79	23.69	Meteo Romania	P, T
15143099999	46.58	23.78	Meteo Romania	P, T
15120599999	46.78	23.57	Meteo Romania	T

2. Information about Leuven's observed data

__Table A.5__: Initial observatories for the Leuven case study

ID	Latitude (º)	Longitude (º)	Database	Variable
LC-002	50.85	4.76	City_Council	P, T
LC-004	50.87	4.69	City_Council	P, T
LC-005	50.88	4.71	City_Council	P, T
LC-006	50.91	4.72	City_Council	P, T
LC-008	50.88	4.66	City_Council	P, T
LC-009	50.87	4.71	City_Council	P, T
LC-010	50.88	4.75	City_Council	P, T
LC-011	50.86	4.69	City_Council	P, T
LC-012	50.89	4.73	City_Council	P, T
LC-013	50.93	4.72	City_Council	P, T
LC-014	50.88	4.71	City_Council	P, T
LC-016	50.88	4.74	City_Council	P, T
LC-019	50.92	4.71	City_Council	P, T
LC-021	50.89	4.70	City_Council	P, T
LC-022	50.89	4.69	City_Council	P, T
LC-023	50.91	4.70	City_Council	P, T
LC-028	50.85	4.69	City_Council	P, T
LC-029	50.88	4.69	City_Council	P, T
LC-030	50.86	4.65	City_Council	P, T
LC-031	50.86	4.71	City_Council	P, T
LC-032	50.88	4.70	City_Council	P, T
LC-033	50.88	4.72	City_Council	P, T
LC-034	50.89	4.69	City_Council	P, T
LC-039	50.87	4.69	City_Council	P, T
LC-040	50.89	4.66	City_Council	P, T
LC-042	50.88	4.69	City_Council	P, T

LC-043	50.87	4.75	City_Council	P, T
LC-044	50.88	4.71	City_Council	P, T
LC-045	50.93	4.68	City_Council	P, T
LC-046	50.87	4.70	City_Council	P, T
LC-047	50.87	4.68	City_Council	P, T
LC-048	50.89	4.73	City_Council	P, T
LC-049	50.89	4.70	City_Council	P, T
LC-052	50.89	4.73	City_Council	P, T
LC-053	50.88	4.70	City_Council	P, T
LC-054	50.89	4.74	City_Council	P, T
LC-056	50.88	4.72	City_Council	P, T
LC-057	50.90	4.70	City_Council	P, T
LC-059	50.90	4.73	City_Council	P, T
LC-060	50.86	4.71	City_Council	P, T
LC-061	50.88	4.73	City_Council	P, T
LC-062	50.88	4.67	City_Council	P, T
LC-064	50.85	4.79	City_Council	P, T
LC-065	50.87	4.70	City_Council	P, T
LC-066	50.86	4.72	City_Council	P, T
LC-067	50.88	4.70	City_Council	P, T
LC-068	50.88	4.71	City_Council	P, T
LC-070	50.88	4.68	City_Council	P, T
LC-071	50.90	4.68	City_Council	P, T
LC-072	50.88	4.69	City_Council	P, T
LC-073	50.85	4.73	City_Council	P, T
LC-074	50.87	4.73	City_Council	P, T
LC-075	50.88	4.72	City_Council	P, T

LC-076	50.89	4.69	City_Council	P, T
LC-077	50.86	4.71	City_Council	P, T
LC-078	50.89	4.75	City_Council	P, T
LC-080	50.94	4.68	City_Council	P, T
LC-081	50.86	4.68	City_Council	P, T
LC-082	50.87	4.72	City_Council	P, T
LC-084	50.89	4.72	City_Council	P, T
LC-085	50.88	4.72	City_Council	P, T
LC-087	50.87	4.70	City_Council	P, T
LC-088	50.90	4.74	City_Council	P, T
LC-089	50.89	4.78	City_Council	P, T
LC-091	50.88	4.70	City_Council	P, T
LC-092	50.87	4.71	City_Council	P, T
LC-094	50.88	4.70	City_Council	P, T
LC-095	50.87	4.70	City_Council	P, T
LC-097	50.87	4.71	City_Council	P, T
LC-099	50.88	4.72	City_Council	P, T
LC-100	50.86	4.68	City_Council	P, T
LC-102	50.88	4.70	City_Council	P, T
LC-104	50.88	4.69	City_Council	P, T
LC-105	50.88	4.70	City_Council	P, T
LC-106	50.88	4.71	City_Council	P, T
LC-107	50.87	4.71	City_Council	P, T
LC-110	50.87	4.70	City_Council	P, T
LC-111	50.88	4.71	City_Council	P, T
LC-112	50.88	4.70	City_Council	P, T
LC-113	50.88	4.70	City_Council	P, T

LC-116	50.88	4.69	City_Council	P, T
LC-117	50.88	4.70	City_Council	P, T
LC-118	50.88	4.70	City_Council	P, T
LC-119	50.87	4.69	City_Council	P, T
LC-120	50.86	4.68	City_Council	P, T
LC-121	50.86	4.68	City_Council	P, T
LC-122	50.86	4.68	City_Council	P, T
LC-123	50.86	4.68	City_Council	P, T
LC-126	50.88	4.72	City_Council	P, T
64580	50.76	4.77	KMI-IRM	P, T
64510	50.90	4.48	KMI-IRM	P, T
64630	50.78	4.95	KMI-IRM	P, T
LC-003	50.87	4.73	City_Council	P
LC-007	50.87	4.71	City_Council	P
LC-041	50.88	4.70	City_Council	P
LC-096	50.89	4.68	City_Council	P
LC-108	50.87	4.70	City_Council	P
LC-109	50.87	4.70	City_Council	P
LC-114	50.88	4.70	City_Council	P
LC-124	50.86	4.69	City_Council	P
LC-125	50.88	4.67	City_Council	P
LC-127	50.87	4.72	City_Council	P
LC-128	50.88	4.77	City_Council	P

***Table A.6**: Final observatories of the weather forecast study for the Leuven case study.*

ID	Latitude (º)	Longitude (º)	Database	Variable
64580	50.76	4.77	KMI-IRM	P, T
64510	50.90	4.48	KMI-IRM	P, T

***Table A.7:** Final observatories of the seasonal forecast study for the Leuven case study.*

ID	Latitude (º)	Longitude (º)	Database	Variable
64580	50.76	4.77	KMI-IRM	P, T

***Table A.8:** Final observatories of the climate projections study for the Leuven case study.*

ID	Latitude (º)	Longitude (º)	Database	Variable
64580	50.76	4.77	KMI-IRM	P, T
64510	50.90	4.48	KMI-IRM	P, T
64630	50.78	4.95	KMI-IRM	P, T

3. Information about Madrid's observed data

Table A.9: *Initial observatories for the Madrid case study*

ID	Latitude (º)	Longitude (º)	Database	Variable
8223099999	40.37	-3.79	AEMET	P, T
8224099999	40.29	-3.72	AEMET	P, T
8221099999	40.49	-3.57	AEMET	P, T
8227099999	40.50	-3.45	AEMET	P, T
SP000003195	40.41	-3.68	AEMET	P, T
3128C	40.42	-3.59	AEMET	P, T
3129	40.47	-3.56	AEMET	P, T
3194A	40.44	-3.82	AEMET	P, T
3194I	40.45	-3.74	AEMET	P, T
3194U	40.45	-3.72	AEMET	P, T
3194Y	40.45	-3.81	AEMET	P, T
3195	40.41	-3.68	AEMET	P, T
3195A	40.46	-3.68	AEMET	P, T
3195Q	40.39	-3.67	AEMET	P, T
3196	40.38	-3.79	AEMET	P, T
3196A	40.39	-3.77	AEMET	P, T
3200	40.30	-3.72	AEMET	P, T
82210	40.49	-3.57	AEMET	P, T
24	40.42	-3.75	City_Council	P, T
39	40.48	-3.71	City_Council	P, T
54	40.37	-3.61	City_Council	P, T
56	40.39	-3.72	City_Council	P, T
59	40.46	-3.62	City_Council	P, T
102	40.40	-3.64	City_Council	P, T
103	40.35	-3.71	City_Council	P, T
104	40.37	-3.68	City_Council	P, T

106	40.44	-3.74	City_Council	P, T
107	40.46	-3.66	City_Council	P, T
108	40.48	-3.72	City_Council	P, T
3195T	40.44	-3.64	AEMET	P
3197	40.34	-3.86	AEMET	P
4	40.42	-3.71	City_Council	T
8	40.42	-3.68	City_Council	T
16	40.44	-3.64	City_Council	T
18	40.39	-3.73	City_Council	T
35	40.42	-3.70	City_Council	T
36	40.41	-3.65	City_Council	T
38	40.45	-3.71	City_Council	T
58	40.52	-3.77	City_Council	T
109	40.43	-3.70	City_Council	T
110	40.42	-3.71	City_Council	T
111	40.42	-3.68	City_Council	T
112	40.40	-3.67	City_Council	T
113	40.40	-3.67	City_Council	T
114	40.39	-3.70	City_Council	T
115	40.39	-3.70	City_Council	T

Table A.10: *Final observatories of the weather forecast study for the Madrid case study.*

ID	Latitude (º)	Longitude (º)	Database	Variable
8223099999	40.37	-3.79	AEMET	P, T
8224099999	40.29	-3.72	AEMET	P, T

Table A.11: *Final observatories of the seasonal forecast study for the Madrid case study.*

ID	Latitude (º)	Longitude (º)	Database	Variable
82210	40.49	-3.57	AEMET	P, T

Table A.12: *Final observatories of the climate projections study for the Madrid case study.*

ID	Latitude (º)	Longitude (º)	Database	Variable
3128C	40.42	-3.59	AEMET	P, T
3129	40.47	-3.56	AEMET	P, T
3194A	40.44	-3.82	AEMET	P, T
3194I	40.45	-3.74	AEMET	P, T
3194U	40.45	-3.72	AEMET	P, T
3194Y	40.45	-3.81	AEMET	P, T
3195	40.41	-3.68	AEMET	P, T
3195A	40.46	-3.68	AEMET	P, T
3195Q	40.39	-3.67	AEMET	P, T
3196	40.38	-3.79	AEMET	P, T
3200	40.30	-3.72	AEMET	P, T
82210	40.49	-3.57	AEMET	P, T
8223099999	40.37	-3.79	AEMET	P, T
8224099999	40.29	-3.72	AEMET	P, T
8221099999	40.49	-3.57	AEMET	P, T
3195T	40.44	-3.64	AEMET	P
3196A	40.39	-3.77	AEMET	P
3197	40.34	-3.86	AEMET	P
8227099999	40.50	-3.45	AEMET	T

4. Information about Tallinn's observed data

***Table A.13**: Initial observatories for the Tallinn case study*

ID	Latitude (º)	Longitude (º)	Database	Variable
EN000026038	59.38	24.58	ILM	P, T
ENE00175051	59.40	24.60	ILM	P, T
26038	59.41	24.83	ILM	P, T
26034	59.60	24.50	ILM	P, T
Pirita	59.47	24.82	ILM	P, T
EN000026034	59.60	24.50	ILM	P
Rohuneeme	59.56	24.79	ILM	T

***Table A.14:** Final observatories of the weather forecast study for the Tallinn case study.*

ID	Latitude (º)	Longitude (º)	Database	Variable
26038	59.41	24.83	ILM	P, T
Pirita	59.47	24.82	ILM	P, T

***Table A.15:** Final observatories of the seasonal forecast study for the Tallinn case study.*

ID	Latitude (º)	Longitude (º)	Database	Variable
26038	59.41	24.83	ILM	P, T

Table A.16: *Final observatories of the climate projections study for the Tallinn case study.*

ID	Latitude (º)	Longitude (º)	Database	Variable
EN000026038	59.38	24.58	ILM	P, T
ENE00175051	59.40	24.60	ILM	P, T
26038	59.41	24.83	ILM	P, T
Pirita	59.47	24.82	ILM	P, T
EN000026034	59.60	24.50	ILM	P
Rohuneeme	59.56	24.79	ILM	T