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

D3.2 AI models for climate change vulnerability assessment and weather forecast - V2

Project Reference No	URBREATH – 101139711
Deliverable	D3.2 AI models for climate change vulnerability assessment and weather forecast – V2
Work package	WP3: URBREATH data strategy and tools
Type	OTHER
Dissemination Level	PU - Public (fully open)
Date	22/12/2025
Status	Final
Editors	Emma Gaitán, Toni Rubio, Darío Redolat, Carlos Prado, Lorena Galiano (FIC)
Contributors	Task 3.3 participants
Reviewers	Georgios Kopsiaftis (ICCS) Ettore Etenzi (EXUS)
Document description	This document describes the validated methodologies and results of the comprehensive framework for complete time-scale atmospheric predictions and projections within the URBREATH project. It details the implemented systems for short-term weather forecasting, seasonal prediction, and long-term climate projections in the four Frontrunner cities (FRCs). This deliverable establishes a robust predictive system that is currently operational and actively delivering data to the project's decision-support tools. This work is primarily linked to Tasks T3.2 and T3.3. The final

	project update, D3.3, is foreseen in M36 (December 2026), which will focus on the final validation, scale-up to all nine cities and the co-design of user-specific indices.

Document Revision History

Version	Date	Modifications Introduced	
		Modification Reason	Modified by
0.1	1/09/2024	Table of contents	ICCS
1.0	12/12/2025	Draft ready for internal review	FIC
1.1	15/12/2025	Internal review and comments	ICCS
1.2	16/12/2025	Internal review and comments	EXUS
1.3	17/12/2025	Addressing suggestions: Final version	FIC
2.0	18/12/2025	Final quality check	LC
2.1	19/12/25	Minor review completed	FIC, LC
Final	22/12/2025	Version ready for submission to the EC Portal	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, D3.2, marks the completion of the core methodological development, presenting the validated approach and results for all three scales in the four Frontrunner cities. By leveraging FICLIMA's own methodologies, including rigorous statistical corrections, and implemented AI-driven postprocessing, this work successfully optimizes predictions for key climate variables such as temperature and precipitation. The focus on D3.2 shifts from methodology design to tackling to the verification, validation and operational delivery of these forecasts, which are essential inputs for urban planning, risk assessment, and informed adaptation strategies.

Regarding the short-term weather forecast scale (0–10 days), this report evaluates the performance of two ensemble-based probabilistic models. A comparison of the Global Ensemble Forecast System (GEFS) and the European Centre for Medium-Range Weather Forecasts (ECMWF) models, analyzed for three key daily variables: maximum temperature, minimum temperature, and precipitation across the four Frontrunner cities (FRCs) (Section 2.2.3), confirms that the ECMWF model consistently outperforms GEFS across all metrics. Furthermore, AI correction algorithms (Random Forest) were successfully applied to ECMWF temperature predictions, demonstrating satisfactory improvements, particularly in correcting systematic errors. While these algorithms show strong potential, future project phases will continue to explore optimization, enabling potential model upgrades and the delivery of even more accurate forecasts to all cities.

For the seasonal forecast, this report shows that the FICLIMA’s method provides the most robust and persistent station-scale skill over 1–6-month lead times, generally matching or outperforming SEAS5 and CFSv2, especially beyond the first 1–3 months. Dynamical systems still offer valuable complementary information, with SEAS5 adding large-scale seasonal guidance and CFSv2 contributing subseasonal skill at shorter horizons. Overall, FICLIMA’s methodology strengthens the seasonal signal that can be extracted from ocean–atmosphere teleconnections, while the combined use of statistical and dynamical sources lays the foundation for more reliable, operational climate services in the URBREATH regions.

At the climate projection scale, this deliverable develops and validates an ensemble of locally downscaled CMIP6 projections for the FRCs across four SSP pathways, combining the FICLIMA two-step statistical downscaling for temperature and precipitation with parametric quantile-quantile mapping (referenced to ERA5-Land) for wind, humidity and radiation, and extending the outputs to daily 1 km grids; it complements mean-climate trends with actionable extremes analysis (heatwaves and meteorological drought via SPI/SPEI), delivering robust, decision-ready indicators of long-term change, extreme-event likelihood and sector-relevant risk to support strategic climate adaptation planning.

By successfully implementing and validating the methodologies across all three forecasting and projection scales, this deliverable establishes a robust, multi-timescale Framework for predicting urban meteorological conditions in the Frontrunner cities. Operational delivery is already in place, with validated short-term and seasonal forecasts being actively transferred to the URBREATH decision-support tools (WMS/MinIO) for use by the FRCs. The final phase of the project, Deliverable D3.3, will focus on the final optimization of AI-based forecast postprocessing, and the scale-up of the service to all nine cities, including the co-design of user-specific indices. These advancements ensure that URBREATH delivers a cutting-edge, data-driven forecasting service, enabling all pilot cities to make informed decisions for climate resilience and adaptation.

Table of Contents

Executive summary 3
Table of Contents 5
List of Figures 7
List of Tables 9
List of Terms and Abbreviations 9
1. Introduction 11
  1.1 Background, purpose and scope 11
  1.2 Approach and relation to other Work Packages and deliverables 13
  1.3 Structure of the deliverable 15
2. Data and study area 16
  2.1 Study Area 16
  2.2 Data requirements 17
    2.2.1 Observed data (updates from D3.1) 17
      2.2.1.1 Weather forecast 17
  2.3. Data Sources Selected 19
    2.3.1 Weather forecast 19
    2.3.2 Seasonal forecast 21
    2.3.3 Climate projections 22
3. Methodology 23
  3.1 Weather forecast 23
  3.2 Seasonal forecast 24
  3.3 Climate projections 26
    3.3.1 State of the art in climate modelling 27
    3.3.2. Simulation of wind, relative humidity and radiation future climate scenarios 28
    3.3.3. Simulation of Heat weaves and Drought episodes 30
    3.3.4. AI approach on current downscaling methodology 30
4. Results 32
    4.1.1 Weather forecast 32
      Results for Madrid 32
      Results for Leuven 35
      Results for Tallinn 36

Results for Cluj-Napoca.....37
Verifications of AI models to correct predictions.....38
4.1.2 Seasonal forecast.....41
4.1.2.1 Seasonal forecast - SEAS5 model validation.....42
Pearson Correlation (Ensemble Mean).....42
Mean Absolute Error (MAE).....43
Root Mean Square Error (RMSE).....44
4.1.2.2 Seasonal forecast - Statistical-based seasonal forecast models results.....45
Key findings summary.....50
4.1.2.3 Seasonal forecast - CFSv2 models results.....51
Key findings summary.....53
4.1.2.4 Seasonal forecast summary and conclusions.....54
4.1.3 Climate projections.....54
Results for Madrid.....54
Results for Leuven.....57
Results for Tallinn.....61
Results for Cluj-Napoca.....65
5. Displaying results in the final tool.....69
5.1 Weather forecast.....69
5.2. Seasonal forecast.....72
5.3. Climate projections.....74
6. Conclusions.....77
7. References.....80
Annex 1. Verification metrics for weather and seasonal forecast.....83
Annex 2. Simulation of Heat weaves and Drought episodes.....92
Annex 3. Weather forecast results.....95
Annex 4. Seasonal Forecast Full Results.....117
Annex references.....223

List of Figures

Figure 1. Map of Europe showing the Frontrunner cities (in red) and Follower cities (in purple) marked with a cross 16
Figure 2. Scheme of seasonal forecast methodology........ 25
Figure 3. Example downscaling processes of a daily field output from CMIP6 Earth System Models, applying transfer functions: a) Original gridded model output of a particular day, b) Simple bilinear interpolation of (a) to ERA5-land resolution, c) Quantile correction of (b) by mapping the modelled ECDF (from CMIP6) and the reference ECDF (from ERA5-land) for each grid point of the ERA5-land reanalysis, taking all days of the common period........ 29
Figure 4. Daily average temperature BIAS in Madrid........ 32
Figure 5. MAE for minimum temperature in Madrid. 33
Figure 6. ROC curve for precipitation forecast in Madrid. 34
Figure 7. Area under the ROC curve (AUC) for precipitation forecast in Madrid........ 35
Figure 8. Minimum temperature BIAS of raw and corrected ECMWF ensemble predictions for Madrid........ 39
Figure 9. Minimum temperature MAE of raw and corrected ECMWF ensemble predictions for Madrid........ 39
Figure 10. Minimum temperature BIAS of raw and corrected ECMWF ensemble predictions for Cluj-Napoca. . 40
Figure 11. Minimum temperature MAE of raw and corrected ECMWF ensemble predictions for Cluj-Napoca.. 41
Figure 12. Pearson correlation for TMean, classified for lead month, observatory and model. 42
Figure 13. Mean Absolute Error for TMean, classified by lead month and model. 43
Figure 14. Root Mean Square Error for TMean, classified by lead month and model........ 44
Figure 15. Combined Pearson Correlation for different Statistical based models for each lag and for Mean Temperature. 47
Figure 16. Combined Pearson Correlation for the observatories used for each lag (Mean Temperature). 48
Figure 17. Combined SMAE for different Statistical based models for each lag and for Mean Temperature...... 49
Figure 18. Combined SMAE for the observatories used for each lag using all models (Mean Temperature). 50
Figure 19. Pearson correlation for maximum temperature across the entire dataset of observations for each lead day. 52
Figure 20. SMAE for maximum temperature across the entire dataset of observations for each lead day......... 53
Figure 21. The graph shows the evolution of annual number of Heat Waves per year 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. [Figure A5.6 in Annex 5] 56
Figure 22. The graph shows the evolution of annual number of Heat Waves per year 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. [Figure A5.25 in Annex 5] 60
Figure 23. The graph shows the evolution of annual number of Heat Waves per year 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. [Figure A5.44 in Annex 5] 63
Figure 24. The graph shows the evolution of annual number of Heat Waves per year 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. [Figure A5.63. in Annex 5] 67
Figure 25. Screenshot of the FICLIMA prototype webpage for weather forecast by VCS. Are displayed results for Madrid city for maximum wind gust probability of overpassing 40kmk/h. 70
Figure 26. Screenshot of the FICLIMA prototype webpage for weather forecast by VCS. Are displayed results for Leuven city for probabilistic daily accumulated precipitation at the threshold of appreciable precipitation (>0mm). 70
Figure 27. Screenshot of the FICLIMA prototype webpage for weather forecast by VCS. Are displayed results for Tallinn city for probabilistic for minimum temperature probability of overpassing below 0ºC. 71
Figure 28. Screenshot of the FICLIMA prototype webpage for weather forecast by VCS. Are displayed results for Cluj-Napoca city for probabilistic for maximum temperature probability of overpassing 5ºC 71
Figure 29. Screenshot of the FICLIMA prototype webpage for weather forecast by Municipia. In the picture can see displayed results for Leuven city for probabilistic for minimum temperature probability of overpassing every available threshold. 72
Figure 30. Screenshot of the FICLIMA prototype webpage for seasonal forecast by VCS. Are displayed results for Madrid city for relative monthly precipitation anomaly. 73
Figure 31. Screenshot of the FICLIMA prototype webpage for seasonal forecast by VCS. Are displayed results for Madrid city for relative monthly precipitation anomaly. 73
Figure 32. Screenshot of the FICLIMA prototype webpage for seasonal forecast by VCS. Are displayed results for Madrid city for relative monthly precipitation anomaly. 74

Figure 33. Example of a visualisation of climate projections for average annual temperature within the URBREATH visualization tool, along with an explanation of how to interpret the results shown in the graph for the city of Cluj-Napoca. 75
Figure 34. Example of the visualisation of climate projections for annual mean wind speed and annual mean shortwave radiation in the city of Cluj-Napoca within the URBREATH visualization tool 76

List of Tables

Table 1. Final observatories of the weather forecast timescale for the FRCs. 18
Table 2. Compilation of available raw EPS models and summary of issues. 20
Table 3. Information on 10 CMIP6 climate models (IPCC AR6), retrieved from the ESGF portal. 22
Table 4. Metrics used for model verification 24

List of Terms and Abbreviations

Abbreviation	Definition
AEMet	Agencia Estatal de Meteorología (España)
AI	Artificial Intelligence
AUC	Area Under the Curve
CDS	Climate Data Store (Copernicus)
CNN	Convolutional Neural Networks
CMIP6	Coupled Model Intercomparison Project Phase 6
ECMWF	European Centre for Medium-Range Weather Forecasts
EPS	Ensemble Prediction System
ERA5	Fifth generation ECMWF atmospheric reanalysis
ERA5-Land	Fifth generation ECMWF atmospheric reanalysis Land version
ESM	Earth System Model
GCM	Global Climate Model
GEFS	Global Ensemble Forecast System
GSOD	Global Surface Summary of the Day
FLC	Follower City
FRC	Frontrunner City
ICT	Information and Communication Technology
IPCC6	Intergovernmental Panel on Climate Change, Sixth Assessment Report

KS Test	Kolmogorov-Smirnov Test
MAE	Mean Absolute Error
MSE	Mean Squared Error
NBS	Nature-Based Solutions
NCEI	National Centers for Environmental Information
NOAA	National Oceanic and Atmospheric Administration
ROC	Receiver Operating Characteristic (ROC curve)
RPS	Ranked Probability Score
SMAE	Scaled Mean Absolute Error
SOTA	State-of-the-Art
SPEI	Standardized Precipitation Evapotranspiration Index
SPI	Standardized Precipitation Index
SSP	Shared Socioeconomic Pathways
TeWA	Teleconnection Wavelet (Self-predictability approach)
WMO	World Meteorological Organization
WP	Work Package

1. Introduction

1.1 Background, purpose and scope

The conclusions of the sixth report of the Intergovernmental Panel on Climate Change (IPCC6, 2021a and b) unequivocally establish that human activity is the primary driver of the current climate emergency. No region remains unaffected by effects such as droughts, ice melting, and increasing frequency of extreme weather events. The report confirms that global surface temperature will continue to rise until at least mid-century under all emissions scenarios, requiring immediate, large-scale mitigation efforts alongside adaptation measures tailored to specific sectors and problems.

Climate change poses significant economic and developmental risks. European projections indicate that average temperatures (maximum and minimum) will likely increase by 2 to 4ºC in the coming decades. Although annual precipitation changes may be slight, precipitation events are expected to become more intense and concentrated, leading to higher frequencies of heatwaves, droughts, and extreme rainfall.

Consequently, assessing and estimating climate risk must be executed at the local scale, as impacts are not uniform across regions or socio-economic sectors. Accurate risk estimation requires climate information projected at the local level, specifically the future evolution of meteorological variables (temperature, precipitation, wind), to analyze sector-specific implications and inform adaptation strategies based on expected future conditions, rather than relying on past climate records.

It is essential to integrate short- and medium-term weather forecasts with long-term climate projections. This combined approach is vital for implementing effective adaptation strategies, allowing decision-makers to address both immediate impacts (via forecasts) and long-term climate shifts (via projections).

This document presents the methodology and results of the definitive approach adopted for downscaling climate models to the demo site level, effectively translating Global Climate Model (GCM) outputs to a local scale. The core of this deliverable is the implementation through results and its evaluation of the AI-enhanced methodologies across all timescales, which validates the performance and ensures that the downscaling process aligns with the specific characteristics and risks of each demo site, which are already receiving this information through different URBREATH tools. To sum up, the deliverable just shows the results of 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 Frontrunner cities, focusing on the characterization of their current climate conditions by identifying extreme events and climate hazards. Additionally, on shorter timescales, the process of selecting and adapting Weather Forecast Models to each FRC city is outlined. This ensures that forecasting capabilities are optimized for local needs. On the other hand, the implemented seasonal forecast methodology, TeWA-CNN, and its expected results are analyzed for each FRC 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 enables 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 is updated daily, ensuring adaptive improvements in predictive accuracy and responsiveness to evolving atmospheric conditions. On a seasonal scale, which is updated monthly, the TeWA-CNN methodology is implemented. It utilizes Convolutional Neural Networks (CNN) to calculate a set of prediction points, which helps in identifying the most statistically significant ones for building FICLIMA's custom forecast for each point.

Regarding the weather forecast timescale, the successful acquisition of historical data acquisition from the ECMWF model, spanning the entire 2023-2024 period, represents a major achievement for this deliverable. This crucial access has allowed for a comprehensive statistical evaluation and the application of AI-based bias correction using a two-year historical dataset. However, it must be noted that a complete comparative evaluation across all ensemble models considered (GEFS, ECMWF, ICON, and GEM) remains limited. Specifically, the ICON dataset only covers a limited period, and access to the GEM model has not been possible.

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

1.2 Approach and relation to other Work Packages and deliverables

The approach adopted in this deliverable is designed to transition from preliminary assessments to the presentation and validation of the definitive AI-enhanced methodologies and their results across all required timescales.

This document compiles the methodologies, execution details, and results of the models developed for the four Frontrunner cities. Specifically, each sub-section outlines the justification for the chosen AI approach (such as the TeWA-CNN for seasonal forecast), detailing the necessary data requirements and addressing potential data limitations (such as the limited coverage of the ICON dataset).

Crucially, the focus of this deliverable is not limited to internal analysis, but includes the operational output required for other Work Packages (WPs) and deliverables. The validated short-term and seasonal forecasts are actively delivered to partners via the standard WMS protocol for integration into the URBREATH tools.

The main conclusions of this report summarize the performance of the implemented methodologies and results, and define the next steps. This deliverable will be followed by Deliverable D3.3, which will focus on the adaptation and fine-tuning of the models for all cities, including Follower cities (FLCs). This fine-tuning phase will encompass the co-design and development of derived climate/weather indices from the core variables already presented (for all scales), aligned with the needs expressed by the cities. This final version will complete the project’s current results and facilitate exploitation and scaling up across all nine cities within the project.

Relation to other work packages and deliverables

This deliverable establishes the high-resolution climate and weather forecasting methodologies that serve as the fundamental data layer for the URBREATH project.

The core contributions of this deliverable are directly linked to the following tasks within WP3:

- **T3.1 - AI-based algorithms and tools:** The methodologies detailed sequentially in D3.1 & D3.2 contribute to the final development of the AI-based clustering, classification, and data fusion models required for operational city-level forecasting.
- **T3.2 - Climate modelling and assessment of vulnerability to climate change:** This document provides the implemented localized climate modeling and downscaling detailed results for the

Frontrunner cities, supporting the assessment of vulnerability hotspots and long-term climate risk.

- **T3.3 - Short-term seasonal weather forecast:** The validated adaptation of weather forecast models aligns with T3.3, ensuring that cohesive seasonal outlooks and weather forecasts are delivered monthly and daily respectively.

The validated short-term, seasonal, and long-term outputs are critical inputs for decision-making and implementation in other Work Packages, ensuring the transfer of methodological innovations into tangible results:

WP4 (Decision-making):

- Provides the data-driven basis for the Local Digital Twin Framework (T4.1) and informs the Impact Visualization Tools (T4.3).
- The predictive outputs (weather/seasonal/climate) are essential for adjusting the Nature-Based Solutions (NBS) (T4.4) and feeding the URBREATH ICT integrated solutions (D4.7, D4.8).

WP5 (Local Living Labs):

- Supports the multi-level impact assessment Framework (D5.11, D5.12).
- The co-design and validation of derived climate/weather indices with the Living Labs ensure that forecast outputs are specifically tailored to stakeholder adaptation goals.

WP6 (NBS Implementation):

- Delivers the climate and weather predictions necessary to support real-time NBS monitoring (D6.2, D6.3).
- Ensures deployed solutions are optimized for evolving environmental conditions.
- The risk assessments help define the technical Frameworks for NBS implementation (D6.4, D6.5, D6.6).

WP7 (Scaling and Transfer):

- The standardization of the AI-based climate and weather modeling approaches contributes directly to the URBREATH Global Roadmap (D7.11).
- Facilitate replication and knowledge transfer across Europe through the Follower Cities Roadmap (D7.10).

In summary, this deliverable bridges multiple WPs by providing high-resolution, AI-enhanced weather and climate insights, enabling improved risk assessments, and supporting digital decision-making Framework and NBS deployment efforts.

1.3 Structure of the deliverable

This deliverable is organized into seven main sections. Following the introduction, in Section One, Section Two summarizes the corresponding section content from D3.1 and details the latest updates regarding data requirements and sources for the study across various timescales, outlining changes to available datasets and the final selections for each FRC city. The third section summarizes the methodologies initially described in D3.1, highlighting the updates made to all the methods identified for addressing URBREATH’s requirements for climate change projections, seasonal, and weather forecasts. The fourth section presents the results for each timescale and for each FRC city. This section allows the reader to evaluate how FICLIMA's models improve upon generic operational models for each city through AI methods. Section Five is dedicated to demonstrating how FICLIMA’s model outputs for each timescale are displayed to the Frontrunner cities via the project tools. Finally, a Conclusion section (Section Six) is presented, summarizing the main findings and outlining 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 summarizes the information related to the various databases required for the study across different timescales and presents the available and ultimately selected databases for each of the Frontrunner cities, highlighting the updates from D3.1.

2.1 Study Area

The URBREATH project plans to work in four climatic areas, selecting one main city (FRC) 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 (The Czech Republic).

In this second document, the studies are focused on the four FRCs: Madrid, Leuven, Tallinn, and Cluj-Napoca.

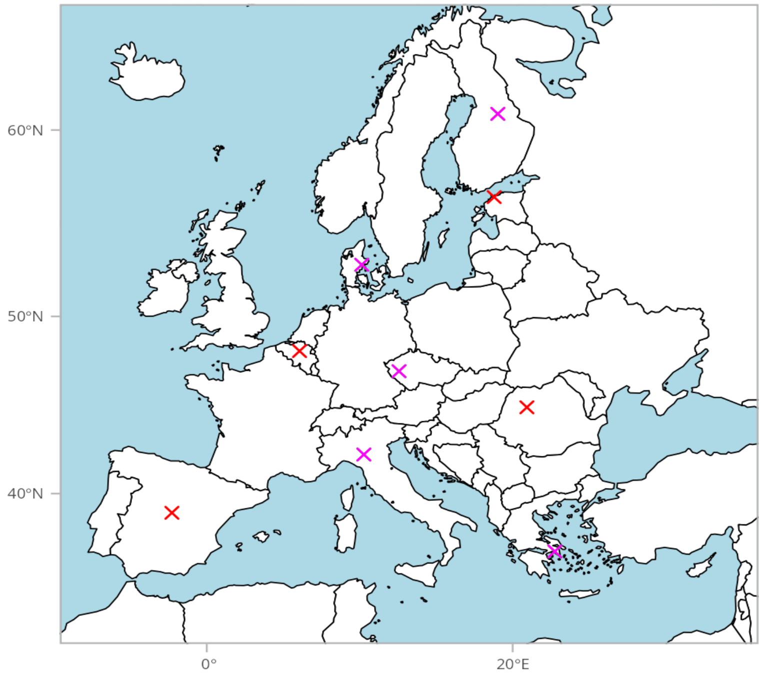


Figure 1. Map of Europe showing the Frontrunner cities (in red) and Follower cities (in purple) marked with a cross

2.2 Data requirements

Section 2.2. outlines the data needs, both historical and simulated, for climate projections and short- to medium-term predictions. It summarizes the advances in data collection and use made since the last deliverable. Regarding climate projections and seasonal forecasts, there are no updates; therefore, no corresponding information will be included in this document.

2.2.1 Observed data (updates from D3.1)

The new weather stations are not applicable for seasonal forecast timescale as methodologically it is appropriate to use only a single weather station for reduced area extensions like the URBREATH FRCs. Regarding climate projections timescales, these additional stations are not valid for this calculation, as their historical recorded period is not long enough for the modelling methodology.

2.2.1.1 Weather forecast

Weather forecast observed data has faced several challenges during the last months. As the modeling methodology requires as much data as possible, all available weather stations with online data from each Frontrunner city have been sought. Furthermore, some sources like GSOD have been discontinued recently; in this case, the last available day was August 29, 2025. This section will summarize all updates and how each issue has been managed.

In the previous deliverable, it has been outlined the decision to exclude data from the Leuven municipal network (Leuven.cool) due to significant quality and reliability concerns. As previously evidenced by the study by Beele et al. (2022), the network's low-cost stations exhibited key issues, including a positive temperature bias due to inadequate shielding, inconsistent sensor calibration, and high variability in environmental exposure. Although quality control (QC) procedures exist, the complexity and residual uncertainties associated with the necessary corrections rendered the dataset unsuitable for FICLIMA’s initial, purely scientific forecasting needs.

Following the advancement of the project and the detailed analysis of the FRC city requirements, a strategic decision has now been made to incorporate selected stations from the Leuven.cool network that meet the minimum data volume criterion (i.e., continuous data available since 2023). This choice is driven by a critical shift from seeking absolute scientific accuracy to prioritizing operational relevance for the Nature-based Solutions (NBS). The deployed NBS will operate under the actual microclimatic conditions measured by the municipal network, not the highly controlled environment of the airport observatory, which was the only reliable source previously available. Although the municipal data exhibits known biases, it provides unique information on microclimatic effects within the urban environment; peculiarities that are crucial for training and validating the performance of the localized NbS. By selectively incorporating only those stations that offer a sufficient volume of data, the project ensures that the developed forecasting and solution deployment models are grounded in the real-world conditions that the end-users and the NBS will encounter, accepting a necessary trade-off between standard scientific precision and the vital operational applicability at the microscale. Those observatories reaching quality control success have been compiled at Table 1.

Table 1. Final observatories of the weather forecast timescale for the FRCs.

Name ID	Latitude (º)	Longitude (º)	Database	Variable	FRC city
Avram Iancu	46.78	23.57	NOAA-WMO	P, T	Cluj-Napoca
Cluj Napoca	46.79	23.69	NOAA	P, T	Cluj-Napoca
Brussels NATL	50.90	4.48	NOAA	P, T	Leuven
Beauvechain	50.76	4.77	NOAA	P, T	Leuven
Wijgmaal	50.93	4.70	City council	P, T	Leuven
Wilsele Dorp	50.89	4.70	City council	P, T	Leuven
Heverlee	50.86	4.69	City council	P, T	Leuven
Kessel-Lo	50.88	4.74	City council	P, T	Leuven
Saint-Pieterskerk Church	50.88	4.70	City council	P, T	Leuven
Villaverde	40.35	-3.71	City council	P, T	Madrid
Ensanche de Vallecas	40.37	-3.61	City council	P, T	Madrid
Plaza Elíptica	40.39	-3.72	City council	P, T	Madrid
Cuatro vientos	40.37	-3.79	WMO	P, T	Madrid
Getafe	40.29	-3.72	WMO	P, T	Madrid
Retiro	40.41	-3.68	AEMET	P, T	Madrid
Tallin-Harku	59.41	24.83	Estonian Environment Agency	P, T	Tallinn
Pirita	59.47	24.82	Estonian Environment Agency	P, T	Tallinn
Pajupea	59.38	24.97	Estonian Environment Agency	P	Tallinn
Kloostrimetsa	59.46	24.88	Estonian Environment Agency	P	Tallinn
Hürüü	59.38	24.54	Estonian Environment Agency	P	Tallinn

The search for observed data from the other FRCs focused on Cluj-Napoca and Tallinn (as Madrid's observed data already has sufficient density for the project's scope). For Cluj-Napoca, an additional weather station was identified in the downtown area at the WMO section of the NOAA webpage. Additionally, as a result of the interaction with Tallinn city representatives, access to another weather station was obtained from the Estonian weather service. Those observatories reaching the quality control success have also been compiled at Table 1.

A major issue experienced during this period was the discontinuation of the GSOD observed data repository. Initially, a custom script was developed for the automated daily retrieval of this source, ensuring a consistent and timely data flow. However, the service interruption coincided with a US government shutdown, and because NOAA depends on this service, the discontinuation was noted some weeks later. To solve the issue, the connection was reestablished by an API connection to the NOAA servers.

2.3. Data Sources Selected

The analysis and commitments presented in this section incorporate the essential progress made in data acquisition since the previous deliverable. Structured by timescale, each sub-section summarizes the advancements in data acquisition from the selected sources. Where no major progress has occurred, a brief summary will still be provided to contextualize this deliverable, thus avoiding the need to revisit the previous one.

2.3.1 Weather forecast

Building on the efforts to refine the AI improvement methodologies for raw weather forecasts, the primary advancement achieved during this period is the successful acquisition of the data downloading process from the most promising model, the ECMWF weather forecast model. This crucial input of historical data has now been processed, enabling the subsequent development and training of the AI methodology intended to enhance raw weather forecasts. In this section, more detailed feedback is schematically provided regarding all raw weather forecast modeling.

A significant breakthrough has been achieved in the data acquisition phase, successfully resolving the key limitation noted in the previous version of this deliverable. The full historical ensemble forecast datasets for the two leading prediction models, GFS and ECMWF, were successfully acquired during 2025. However, the other downloaded model, ICON, could not provide the entire historical period required (2023-2024); consequently, it has not participated in the verification processes. This accomplishment is crucial, as it provides the necessary data to finally proceed with the comprehensive statistical evaluation.

By successfully retrieving this ensemble data and combining it with the local meteorological observation records from official sources (which serve as a benchmark for the 2023–2024 period), we can now move forward with a systematic comparison taking into account the available historical ensemble data. The primary accomplished goal of this thorough analysis is to quantify bias, ensemble spread, and overall skill for each model against observed urban conditions, which will allow us to the definitive selection of the most suitable Ensemble Prediction System (EPS) as the best base model for operational urban weather forecasting within the URBREATH Framework. This progress ensures that FICLIMA's AI-enhanced predictive model is built upon the most accurate baseline model input.

To demonstrate traceability with the previous deliverable, in the next Table (2) compiles all weather forecast historical data sources and how the issues have been solved or the source has been discarded.

Table 2. Compilation of available raw EPS models and summary of issues.

EPS Model	Organization	Period of historical data downloaded	Summary of the status and issues (D3.1)	Current solution
ECMWF Ensemble Prediction System (ENS)	ECMWF	from 1st January 2023 until now	MARS, ECMWF's official downloading data platform, performs very slowly, downloading a single day of data per day.	Downloaded
Global Ensemble Forecast System (GEFS)	NOAA/NCEP (USA)	from 1st January 2023 until now	Required data perfectly downloaded	Downloaded
Met Office Global and UK Ensemble (MOGREPS-G & MOGREPS-UK)	UK Met Office	None	Only last month available https://app.snowflake.com/marketplace	Discarded (data totally unreachable)

Canadian Ensemble Prediction System (CEPS)	Environment Canada (CMC)	None	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	Discarded (data not accessible)
Météo-FRCance ARPEGE Ensemble Prediction System	Météo-FRCance	None	Major issues with API	Discarded (data totally unreachable)
ICON Ensemble Prediction System (ICON-EPS)	Deutscher Wetterdienst (DWD, Germany)	from 12th December 2024 until now	Major issues with API (Pamore) Only a month and few days more available for the project downloaded from other limited source	Downloaded, but not the complete period required

2.3.2 Seasonal forecast

The primary advancement in FICLIMA's seasonal forecasting methodology lies in the strategic evolution of data inputs. Previously, the approach relied on predefined teleconnection indices and statistical methods (like Pearson correlation) to select predictors from a general pool of long-term global climate variables. Observational needs were largely met using data from GSOD.

The updated TeWA-CNN Framework introduces a dynamic, data-driven selection process by utilizing Convolutional Neural Networks (CNNs). This system analyzes specific spatial fields of monthly anomalies from global climate datasets to allow predictors to emerge from the data, rather than being fixed. The core inputs now focus explicitly on detailed atmospheric and oceanic variables: Sea Surface Temperature (SST), Humidity at 700 hPa, and the zonal (u) and meridional (v) wind components at both 1000 hPa and 500 hPa.

Crucially, the final forecast output is now strengthened by incorporating external model input. The final prediction is a combination of the internal CNN-driven forecast with anomaly data from best-performing SEAS5 models (CMCC, DWD, ECCC, ECMWF) and a post-processed CFS-90 days model. This shift ensures the internal forecast is validated and augmented by highly reliable external systems, creating a more robust and accurate prediction.

2.3.3 Climate projections

Climate risk assessment requires defining local-scale impacts based on the future evolution of key meteorological variables. To perform the necessary climate regionalization (downscaling), the study relies on a robust database comprising both reanalysis datasets and climate models.

Two reanalysis datasets from the ECMWF, accessible via the Copernicus Programme, were selected: ERA5 (for atmospheric variables, with 0.25° spatial resolution) and ERA5-Land (focused on surface variables over land areas, with higher 0.073° resolution and hourly updates). These data undergo a quality control process to ensure consistency.

The required set of climate projections utilizes ten climate models from the CMIP6 (Coupled Model Intercomparison Project Phase 6). The computational efficiency of the statistical downscaling techniques allows for the integration of these models with the four main emission scenarios (Tier 1) defined in CMIP6: the Shared Socioeconomic Pathways (SSPs). These scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5, represent a significant advancement over previous models (CMIP5/RCPs) by offering a more comprehensive coverage of future socio-economic developments and their resulting climatic forcings. The compilation at table 3.

Table 3. Information on 10 CMIP6 climate models (IPCC AR6), retrieved from the ESGF portal.

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

3. Methodology

Within the URBREATH project, a robust set of methodologies has been developed and successfully implemented to simulate the different time scales considered in the study: short-term weather forecasts, seasonal predictions, and long-term climate projections. Each methodology has been carefully designed and validated to address the specific predictive and operational needs of these time scales. The following sections provide a deeper insight into the implementation, performance evaluation, and current status of each methodology. For a detailed technical discussion on the data selection and processing underlying these methods, readers should consult the Annexes of this document or the previous deliverable methodology section.

3.1 Weather forecast

Weather or short-term forecasting relies on Numerical Weather Prediction (NWP) models that simulate atmospheric evolution based on known initial conditions derived from observational data. Forecasting methodologies are categorized as deterministic (single best-estimate scenario) or probabilistic, the latter implemented through Ensemble Prediction Systems (EPS). EPS generates multiple forecast realizations to account for atmospheric chaotic nature and provides essential uncertainty quantification.

In the context of the URBREATH project, probabilistic forecasting was adopted as the optimal approach for integration into end-user decision-support tools. This strategy provides crucial advantages, including:

- Higher predictive skill at short-medium ranges (1–10 days), particularly for extreme weather events.
- Enhanced adaptability for urban resilience strategies by quantifying forecast uncertainty.
- Optimized usability for stakeholders, as risk communication benefits from probabilistic insights over deterministic outputs.

By adopting probabilistic forecasting, URBREATH aligns with the best practices of leading meteorological institutions and establishes a robust Framework for integrating ensemble weather forecasts into its operational workflow.

A series of metrics, described in Annex 1, was used to analyse the performance of the ensemble models described in Section 2.3.1 against observations for each city. The following table (4) summarises these metrics:

Table 4. Metrics used for model verification

Variable	Metrics
Precipitation	Ranked Probability Score (RPS), ROC Curve, Area Under the ROC Curve (AUC), Contingency table, Accuracy, Error Rate, Brier Score
Temperature	BIAS, Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Correlation

3.2 Seasonal forecast

In this study it has been extended the Teleconnection Wavelet–ARIMA (TeWA) Framework into a CNN-based hybrid system for subseasonal–seasonal prediction. Convolutional neural networks are first applied to global monthly fields of SST, humidity at 700 hPa and upper/lower-level winds to identify the regions that act as optimal teleconnection predictors for each case-study station, instead of relying on predefined indices (e.g. ENSO, PDO). From these CNN-selected areas it has been derived data-driven teleconnection indices, which are combined with local self-predictability obtained via multi-scale wavelet decomposition and ARIMA time-series modelling of station anomalies. The two components are then dynamically weighted, based on multi-decadal hindcast performance, to generate monthly forecasts of temperature, precipitation and wind anomalies at 0–6-month lead times, which are benchmarked against SEAS5 and CFSv2. The final prediction also incorporates dynamic models such as SEAS5 and CFSv2 in the most appropriate format, along with the models that have obtained the best performance in the validation steps. A full technical description of the TeWA–CNN implementation is provided in Annex 4.

This methodology combines the flexibility of CNNs (Convolutional Neural Networks) for detecting teleconnection patterns with the robustness of TeWA's self-predictability metrics, enabling improved subseasonal forecasting at a monthly scale. The automated selection of predictors and validation against historical records ensure the approach's adaptability to complex climate dynamics and its reliability and accuracy. See figure 2.

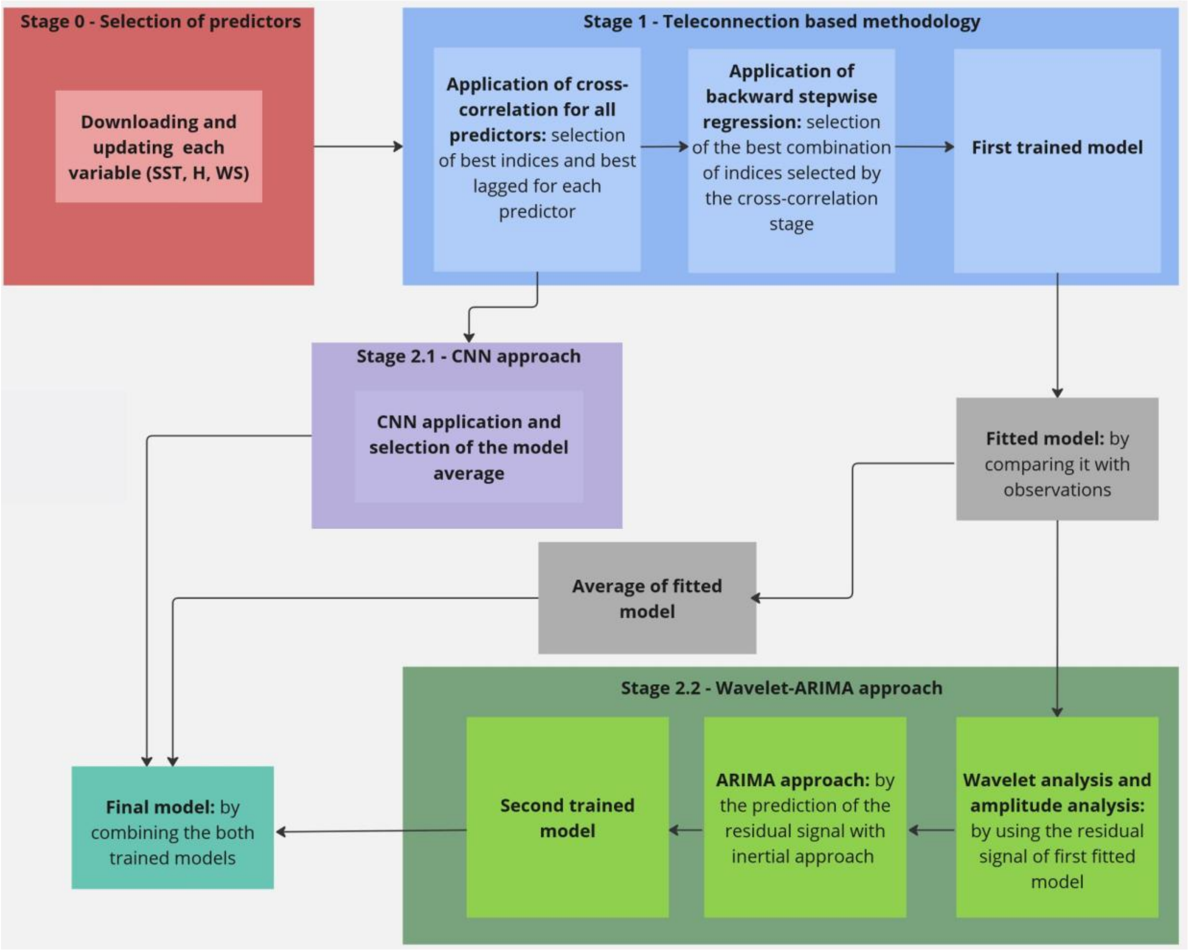


Figure 2. Scheme of seasonal forecast methodology.

Key Gaps and Findings

The seasonal method based on the Teleconnection–Wavelet–ARIMA Framework, extended here with CNN-derived teleconnection indices, remains sensitive to changes in the statistical coupling between large-scale predictors and local surface anomalies. Periods of non-stationarity in these teleconnections (e.g. shifts in the jet stream or in Mediterranean circulation regimes) can lead to locally reduced skill, and performance varies across regions and seasons, with wind generally more challenging than temperature and precipitation. Nevertheless, the multi-decadal hindcast validation shows that the TeWA-based models (CNN_Ens, CNN_TeWa and TeWa_Lin) provide station-scale skill that is comparable to, and often better than, state-of-the-art dynamical systems such as SEAS5 and CFSv2. At 1–3-month leads they typically match the best SEAS5 members in terms of correlation while achieving competitive or lower SMAE, and in several stations they retain small but positive correlations and controlled errors up to 4–6 months, when dynamical models tend to lose most of their predictive power.

Although the approach is computationally more demanding than simpler statistical baselines and the overall seasonal signal remains modest, the TeWA-based Framework demonstrably extracts additional predictable information from ocean–atmosphere teleconnections and local persistence. In this sense, it does not remove the fundamental limitations of seasonal forecasting, but it clearly strengthens the predictive skill at the subseasonal–seasonal range and represents a pragmatic step forward towards more reliable station-level climate services for the URBREATH regions.

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.

In Deliverable D.3.1 "*AI models for climate change vulnerability assessment and weather forecast-V1*", the study focused on the simulation of the meteorological variables maximum temperature, minimum temperature and precipitation using a two-step statistical downscaling methodology (analogue selection and transfer functions) known as FICLIMA (Ribalaygua et al., 2013). To complete this set of meteorological variables, climate projections of wind, relative humidity and precipitation have been produced. For the generation of these variables, a different methodology was applied; in this case, a statistical downscaling approach based on parametric quantile–quantile mapping was used.

Once the generation of the basic climate projections was completed, this work was expanded with the simulation of extreme events related to maximum temperature, such as heatwaves, and with the analysis of drought episodes linked to precipitation. In this way, the Frontrunner cities will have not only an overview of the future average climate but also insights into the occurrence of extreme events that are becoming increasingly frequent.

3.3.1 State of the art in climate modelling

Climate models are the most advanced tools available for simulating the climate system, ranging from simple representations to sophisticated Earth System Models (ESMs). ESMs incorporate physical, chemical and biological processes, including interactions between the atmosphere, oceans, cryosphere and biosphere, as well as the carbon cycle and other components such as atmospheric chemistry or dynamic vegetation. These models are run under different greenhouse gas concentration pathways defined by the CMIP6 Framework, known as Shared Socioeconomic Pathways (SSPs), which describe plausible futures combining socio-economic development, technological advances and emissions trajectories. Using these scenarios, climate models generate projections that help assess future climate conditions. International efforts like the IPCC and the CMIP6 modelling Framework coordinate and advance global climate research, are providing the most up-to-date and comprehensive climate simulations.

Because global models operate at coarse spatial resolutions, they cannot accurately capture local climate features, especially in complex terrain. To overcome this, regionalization or downscaling techniques are applied to translate large-scale climate information into local variables such as rainfall, wind or temperature. Two main approaches exist: dynamic and statistical. For this study, the validated statistical methods FICLIMA will be used. Ensuring that downscaled projections meet strict technical criteria guarantees that the resulting climate scenarios are robust, reliable and suitable for assessing future climate impacts at local scales.

To conduct the climate analysis, a set of ten climate models (see Table 3) and four SSP scenarios (SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5) were used. This means that for each climate variable and each extreme-event scenario, a total of 40 climate projections (model–SSP combinations) were generated for both the historical period 1985–2014 and the future period 2015–2100.

Rigorous verification and validation procedures are required before generating future climate scenarios. This ensures that the statistical downscaling methodology and the selected Climate Models can reliably reproduce past climate conditions in the study area. To quantify methodological and model errors, the downscaling Framework was first applied to ERA5 reanalysis data (1970-2020), comparing downscaled results with observations on a daily basis. This step confirms whether the methodology captures the physical relationships between predictors and predictands and reduces stationarity-related biases. In parallel, each climate model was validated by comparing its downscaled Historical simulation (1979–2015) with the corresponding ERA5-based downscaling, assessing its capacity to reproduce mean climate conditions and natural variability. All climate variables and derived indicators — including temperature, precipitation, wind, relative humidity, radiation, heatwave metrics and drought indices — are subjected to the same verification and validation processes to ensure the robustness of the full set of climate projections.

3.3.2. Simulation of wind, relative humidity and radiation future climate scenarios

In the first phase of generating downscaled climate future scenarios for wind, relative humidity and radiation (short and large) a statistical downscaling methodology based on parametric quantile-quantile mapping (Benestad 2010; Monjo et al. 2014) by using ERA5-Land as a reference (about 9km of spatial resolution) has been used.

A two-step method was commonly applied to all the variables and projections (10 climate models x 4 SSPs, when available) with the following steps:

1. Given a specific historical experiment, each (daily) climate variable is downscaled to each (reference) grid point of the ERA5-Land located within the Frontrunner cities, by comparing the Empirical Cumulative Distribution Function (ECDF) of the reanalysis and the ECDF of the historical experiment of the climate model, previously bilinearly interpolated (Figure 3).
2. From this comparison, a quantile-quantile mapping is obtained and used to fit parametric functions. The selected functions are polynomial curves to reduce the computational cost compared to the exponential distributions of Monjo et al. 2014 and Monjo et al., 2016.
3. Finally, parametric functions are applied to correct the SSP projections on a daily scale. These downscaled projections have been obtained for a set of grids covering the area occupied by the selected Wetlands, with the same spatial resolution as the ERA5-Land (9 km x 9 km).

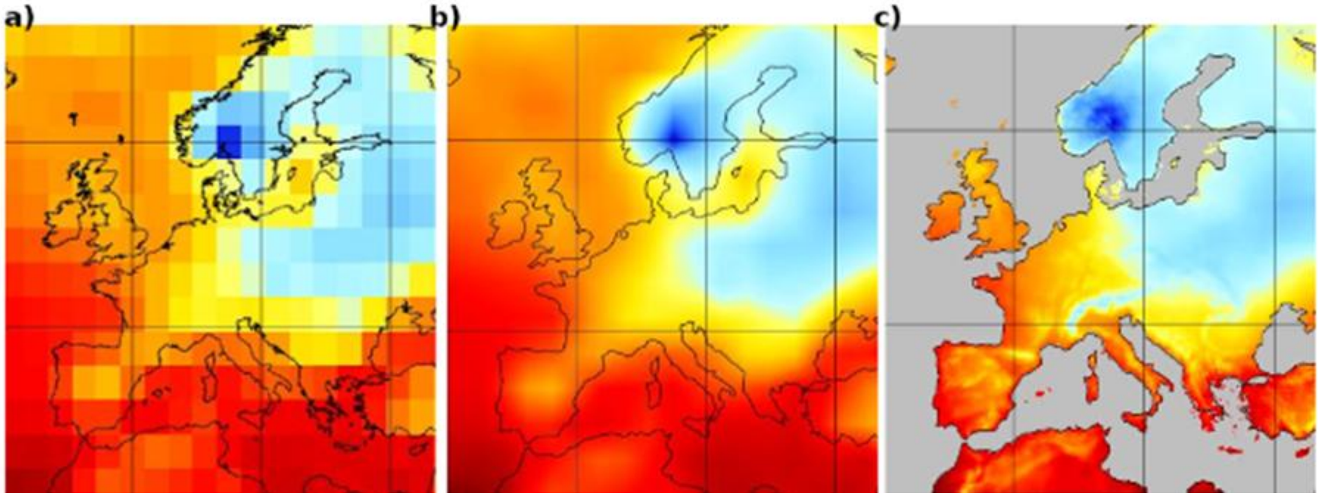


Figure 3. Example downscaling processes of a daily field output from CMIP6 Earth System Models, applying transfer functions: a) Original gridded model output of a particular day, b) Simple bilinear interpolation of (a) to ERA5-land resolution, c) Quantile correction of (b) by mapping the modelled ECDF (from CMIP6) and the reference ECDF (from ERA5-land) for each grid point of the ERA5-land reanalysis, taking all days of the common period.

Depending on whether the climate variable is more similar to (f) smooth (linear uniform, Gaussian) distributions (e.g. temperature, humidity) or more similar to (g) exponential (log-normal, Weibull, gamma…) distributions, the transfer functions (polygonal curves) are fitted from (f) increments or (g) factors obtained by comparison of each target quantile value (from HIST) and the corresponding reference quantile value (from ERA5). That is:

$$f: = mappin(sort(ERA5), sort(ERA5) - sort(HIST))$$

$$g: = mappin(sort(ERA5), sort(ERA5)/sort(HIST))\ for\ ERA5 > 0\ and\ HIST > 0$$

$$MOD\_smooth\_var = MOD\_smooth\_var + f(MOD\_smooth\_var)$$

$$MOD\_exponential\_var = MOD\_exponential\_var * g(MOD\_smooth\_var)$$

That means that extreme values (when MOD is outside the distribution tails of HIST) are transformed with the same increment (f) or factor (g) as the empirical tails of the training sample (HIST). This method can be directly applied for classical scalar variables (temperature, precipitation, humidity, pressure ...), with smooth or exponential distributions, according to up 6th degree polynomial fits applied to linear or logarithmic factors. As climate change will produce values located outside the historical tails, these values are corrected with the last truncated factor of the tailed mapping (f or g).

To achieve this, a topography-based transfer function was applied to the 9km-resolution downscaled model outputs, trained with the time series from the ERA5-Land reanalyses. Particularly, for each climate variable simulated by the CMIP6 models, the statistical downscaling was applied according to three steps:

1. For all the variables and each projected day (with 0.073°×0.073° resolution), a purely geo-statistical analysis was performed to a 1km×1km grid by using AIC-based stepwise regression with topological and geographical parameters (altitude, latitude, longitude and distance to the Atlantic Ocean and the Mediterranean Sea), and bilinear model for the residual errors.
2. Similarly, for each targeted day, residual errors of the AIC-based stepwise regression were interpolated from the original grid (0.073°×0.073°) to the final 1km×1km grid by using a bilineal model.
3. Finally, the projected value for each day is obtained by adding the result of Step (1) and Step (2), generating climate scenarios in a 1km×1km grid at a daily timescale.

3.3.3. Simulation of Heat weaves and Drought episodes

The concepts of heat waves and meteorological drought episodes are defined below. The criteria used to identify and calculate these events are described in detail in Annex 2.

A **heat wave** is a period of abnormally high air temperatures that persists for several consecutive days or longer and is significantly above the normal climate conditions of a given region. Heat waves often have negative impacts on human health, ecosystems, agriculture, and energy demand, especially when combined with high humidity or limited nighttime cooling.

A **meteorological drought episode** refers to a prolonged period of below-average precipitation compared to the long-term climatic mean for a specific area. It is typically identified based on rainfall deficits over weeks, months, or even years and represents the earliest stage of drought development, potentially leading to agricultural, hydrological, and socio-economic droughts if dry conditions persist.

3.3.4. AI approach on current downscaling methodology

The quantile–quantile (Q–Q) mapping methodology used in this study shares conceptual principles with modern artificial intelligence (AI) techniques, particularly in the way both approaches learn statistical relationships between input and output datasets. Q–Q mapping establishes a transfer function by analyzing the empirical distributions of climate model outputs and reference observations, generating a statistically informed correction function that can be applied to future projections. Similarly, AI-based downscaling models, such as neural networks or machine-learning regression architectures, learn non-linear mappings between predictors and predictands through data-driven training. While Q–Q mapping explicitly derives these relationships through distributional comparisons and parametric fitting, AI methods infer them implicitly through optimization algorithms. This parallel highlights how traditional statistical downscaling and AI-driven approaches both aim to model complex dependencies, with Q–Q mapping offering a transparent, computationally efficient Framework and AI providing potential advantages for capturing higher-order, non-linear interactions when sufficiently large and representative datasets are available.

4. Results

4.1.1 Weather forecast

This section presents the results of verifying the ensemble predictions of the ECMWF and GEFS models for all Frontrunner cities. The verified variables are minimum temperature, maximum temperature and precipitation. Data for these variables is available for the full years of 2023 and 2024. With the aim of reducing the length of the document, the explanation of metrics used in this study and its interpretation are presented in Annex 1.
The main difference from the same-name Section at D3.1 is that, whereas the previous deliverable only allowed the performance of the models to be analysed using data corresponding to two months, two full years of data are currently available.

<u>Results for Madrid</u>

In the case of Madrid, and more specifically the Villaverde study area during the control period, both models predict higher average temperatures than were subsequently recorded. The figure shows the average daily temperature for all days in the 2023–2024 period. The interquartile range represents the P10 and P90 percentiles of the mean error of the ensembles during the analysed period, meaning that the shaded area corresponds to 80% of days. The prediction bias is therefore between 1.5 °C and 3 °C on the first day of the forecast, with a median ranging from 2 to 2.5 °C, depending on the model. After nine days, the medians remain stable, but on some days these biases can increase to 5 °C. For all forecast horizons, the bias is approximately 0.5 °C higher for GEFS.

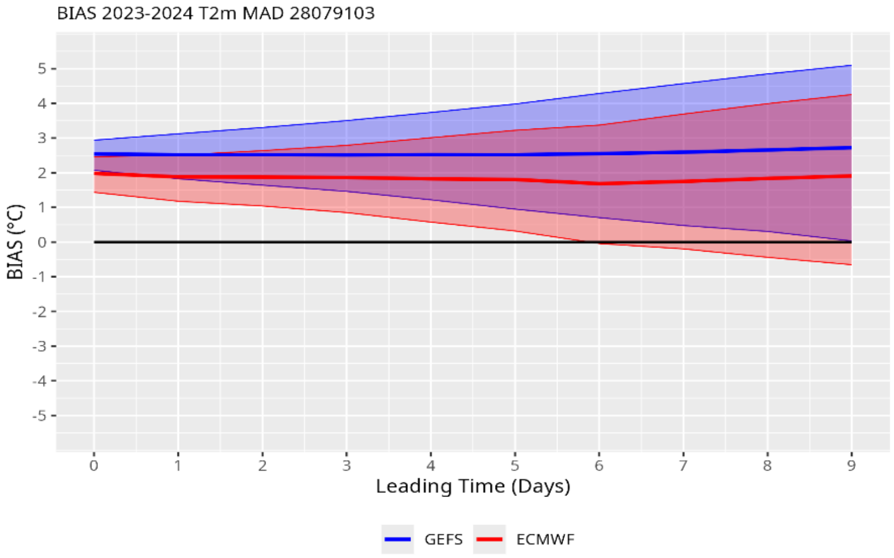


Figure 4. Daily average temperature BIAS in Madrid

This deviation in temperature prediction is primarily caused by errors in the predictions for night-time hours. According to the results, neither model can accurately predict thermal inversions, a phenomenon that is common in the study area. This lack of resolution means that the mean absolute error (MAE) for minimum temperature is very high, exceeding 3°C for GEFS, while ECMWF performs slightly better.

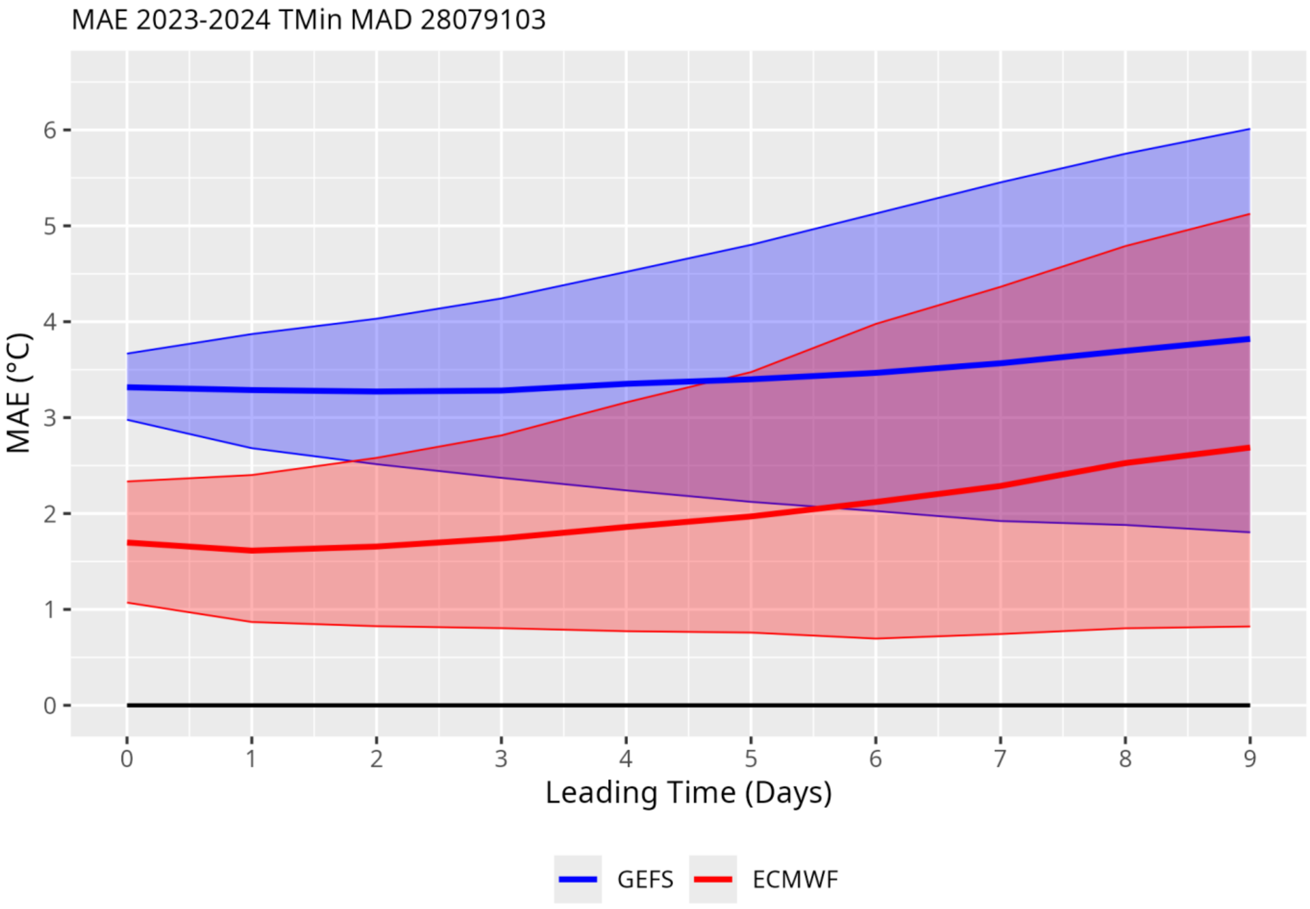


Figure 5. MAE for minimum temperature in Madrid.

In terms of precipitation, the results show that both models are highly accurate. As can be seen in the ROC curve graph, the lines approach the upper left corner -representing total accuracy- and move away from the diagonal line representing random prediction. Three prediction horizons have been selected: D+0 (the prediction made on the same day), D+4 and D+9. A similar curve can be observed in all of them, with the expected decrease in accuracy as the forecast period increases. At the top of the graph, the lines corresponding to the ECMWF model exceed those corresponding to the GEFS, indicating that the ECMWF model is more accurate for the selected time periods.

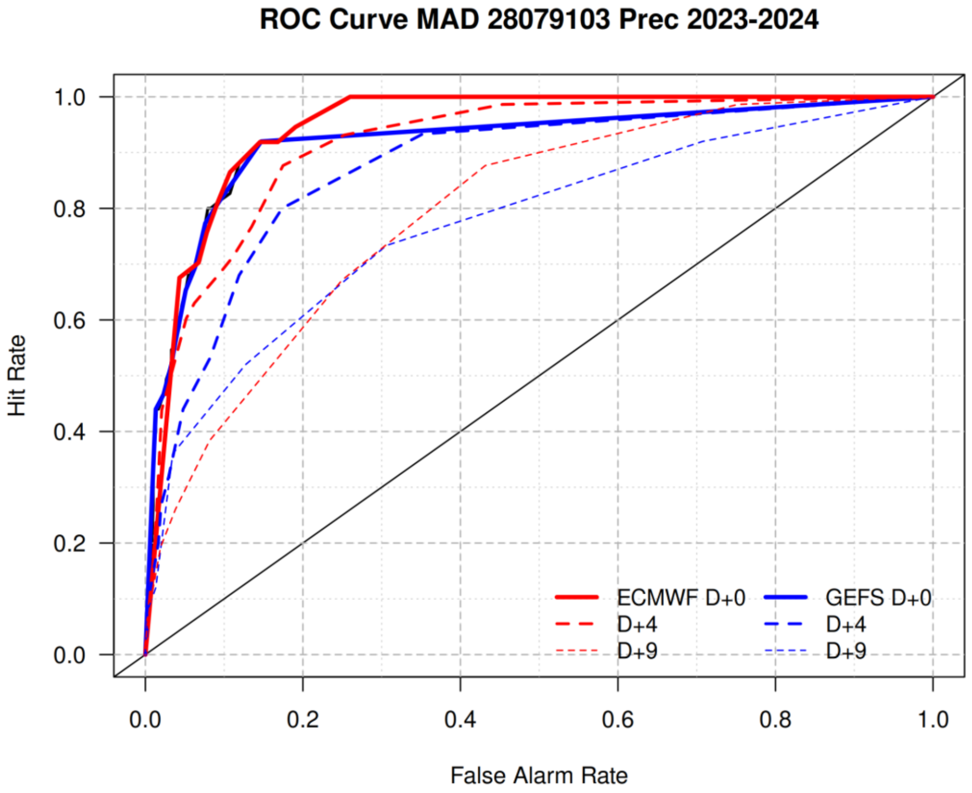


Figure 6. ROC curve for precipitation forecast in Madrid.

This result is best reflected by the area under the ROC curve (AUC) parameter: a value of 1 indicates an error-free prediction, while a value of 0.5 indicates a random prediction. The precipitation threshold considered in this study is 1 mm. As can be seen, the ECMWF model values are better than those of the GEFS for all prediction horizons. The results for the cases represented in the ROC curve graph are as follows: AUC 0.95 vs. 0.92 at D+1, 0.92 vs. 0.88 at D+4, and 0.8 vs. 0.78 at D+9, which indicates that the prediction is very accurate in both cases.

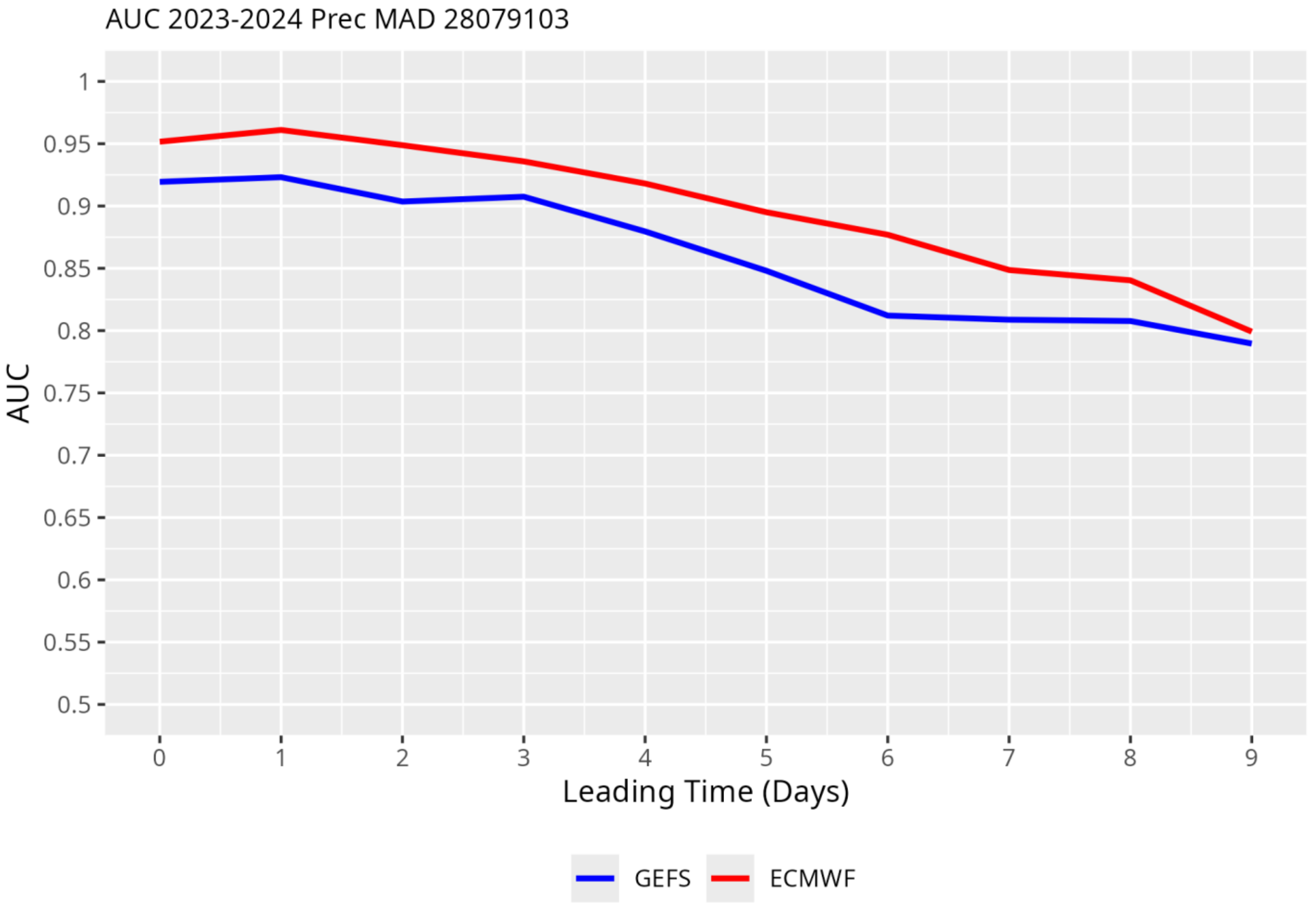


Figure 7. Area under the ROC curve (AUC) for precipitation forecast in Madrid.

<u>Results for Leuven</u>

In the case of the city of Leuven, the corresponding results for the Heverlee observation point are shown in Annex A3. Unlike the Madrid case study, the temperature predictions in Leuven are unbiased, for both minimum and maximum temperatures (figures A3.1 and A3.2).

This is probably because Leuven is in a very flat geographical area that is under constant influence from Atlantic winds. These factors make temperature prediction easier, as phenomena such as thermal inversions or adiabatic warming due to orographic barriers are uncommon.

In the case of maximum temperature, the mean absolute error (MAE) for ECMWF model forecasts is below 1°C on 90% of days on the first day of the forecast (figure A3.3). This error gradually increases as the forecast period extends, reaching a median error of around 2.5°C after nine days. This is a very low figure considering the chaotic nature of the atmosphere, indicating the low climatic variability of this area. For the GEFS model, the errors are also low, but around 0.5°C higher than those of the ECMWF.

This low climate variability and high accuracy in weather forecasts is also reflected in statistics such as correlation (figure A3.4). The correlation between observation and prediction for maximum temperature is also shown below, with values above 0.85 even at 9 days' notice and very close to 1 in the first days of prediction, indicating that the prediction is very good. For minimum temperature, the BIAS, MAE and correlation figures are very similar to those obtained for maximum temperature.

For precipitation -as in the case of Madrid-, the ECMWF model performs better than GEFS for the three forecast horizons selected in the ROC curve (figure A3.5). However, over a longer forecast period, the hit rate decreases more than in Madrid, probably due to the more oceanic climate, where the number of days of precipitation and the probability of model failure increase significantly. Nevertheless, the AUC parameter (figure A3.6) remains above 0.7 for all prediction horizons.

An important parameter for assessing the accuracy of a precipitation prediction model -in terms of amount- is the RPS. In this case, four ranges of precipitation amounts have been defined: less than 0.5 mm, between 0.5 and 5 mm, between 5 and 20 mm, and more than 20 mm. The greater the difference between the predicted and observed amounts, the greater the contribution to the index. Therefore, a prediction is considered better the closer its RPS value is to 0.
In this case, the RPS value for the ECMWF model is half that of the GEFS model for the first days (figure A3.7). This suggests that not only does the ECMWF model better discern days when there is a probability of rain, it also has a very good accuracy ratio in terms of precipitation amounts.

Results for Tallinn

In the case study of Tallinn, the environmental conditions are different from the previous cases. Certain meteorological situations have been identified in this city in which there are significant temperature differences between coastal areas and those further inland. These situations mainly occur in summer, when synoptic winds from the south cause the temperature to rise significantly inland, while the influence of the sea breeze keeps the temperature much lower on the coast. This effect is most notable in maximum temperatures, resulting in a negative bias in the models for this variable.

The verification results for the observation point in Pirita are shown in Annex 3. At this location, the maximum temperature shows a negative bias (Figure A3.8), which is quite close to zero given its proximity to the sea. Both the GEFS and ECMWF prediction models show an average negative bias of 1°C for all prediction horizons, which increases slightly as the prediction range increases.

This means that, for the maximum temperature, the root mean square error (RMSE, figure A3.9) ranges from 1 to 3°C throughout the analysed forecast period, although on some days this error can exceed 6°C.

For minimum temperature (figures A3.10 and A3.11), the figures are significantly better in all cases, with errors around 0.5°C lower than for maximum temperature and correlations between observations and predictions above 0.85. For both minimum and maximum temperatures, the ECMWF model consistently outperforms the GEFS, with an error margin around 0.3°C lower for all forecast horizons.

In the case of precipitation, the ROC curve (figure A3.12) is very similar to those of previous case studies. The only difference is that the ECMWF model produces a higher number of false positives for the initial forecast horizons, which negatively impacts its performance statistics compared to those of other analysed locations.

This means that, in this case, the results for both ECMWF and GEFS are practically tied for several parameters, such as RPS and error rate. The AUC (figure A3.13) remains above 0.7 in both cases, which is considered acceptable.
The models' error rate (false positives + false negatives) is shown in figure A3.14. Both models perform similarly, with error rates ranging from around 17% on the first day of prediction to 35% after nine days.

Results for Cluj-Napoca

In the case study of Cluj-Napoca, both models show a higher error rate in their weather forecasts than in previous cases. This may be because the city has a continental climate with greater variability and is in an area prone to convective and orographic phenomena, which are more difficult to predict.
Therefore, as is shown in figure A3.15, both models tend to underestimate the average daily temperature by approximately 1°C.
This BIAS is significantly greater for maximum temperature (figure A3.16), which is underestimated by around 2°C on average, whereas for minimum temperature it is underestimated by slightly less than 0°C on average. This means that average errors in maximum temperature are higher than in other case studies, ranging from 2 to 3.5°C, even for a one-day forecast. The mean absolute error (MAE, figure

A3.17) increases gradually over all forecast horizons, reaching an average of 4°C over a nine-day horizon. However, on 10% of days, nine-day forecasts have errors greater than 7°C.

Despite these errors, a rather striking result is that the correlation between observations and predictions remains very high, above 0.87 even for long prediction horizons (figure A3.18). This is encouraging because, although the predictions are biased, the models capture temperature changes in different weather situations well. Once this bias is eliminated, the predictions will be of high quality.
Mean absolute errors are significantly lower in the case of minimum temperature (figure A3.19): an average of 1.5–2.0°C for the first day of the forecast, increasing to 3°C for the 9-day forecast.

As in previous case studies, the ECMWF model outperforms the GEFS forecast in all analysed parameters and variables, achieving improvements of between 0.3°C and 0.5°C in minimum and maximum temperatures.

Regarding precipitation, it is evident that the models perform much worse here than in the other study areas. This may be due to the city's location. As it is in an area surrounded by mountain ranges, the prediction models may fail to account for processes that enhance or weaken precipitation due to orographic factors, such as the Föhn effect.

As can be seen in figure A3.20, the curves are far from the top left corner, which represents perfect prediction. As in all the other cases, the ECMWF model's predictions are better than those of the GEFS for all the analysed parameters, although deterioration in the predictions affects both models. In the case of the RPS (figure A3.21), the average value is approximately 0.1; the error rate exceeds 25% from the first prediction horizon; and the AUC (figure A3.22) values are below 0.85 in all cases, dropping below 0.7 for predictions with a horizon of more than one week.

Verifications of AI models to correct predictions

A model that uses artificial intelligence to correct predictions has been tested and is currently undergoing optimisation.
Based on the random forest algorithm, the model's training data corresponds to ECMWF ensemble predictions for 2023 and 2024. It has been validated for a period independent of the training test: May to October 2025. The predictor variables used to train the model are maximum and minimum temperature, temperature range, and month of the year.
The model has successfully corrected predictions that were originally highly biased, such as minimum temperature in Madrid and maximum temperature in Cluj-Napoca.
Results for these two cities are shown below:

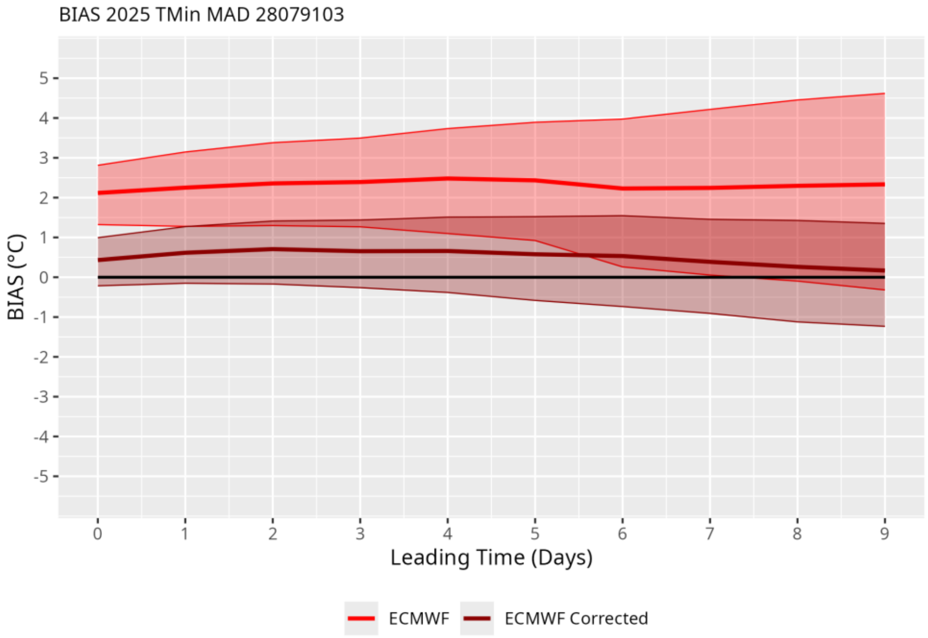


Figure 8. Minimum temperature BIAS of raw and corrected ECMWF ensemble predictions for Madrid.

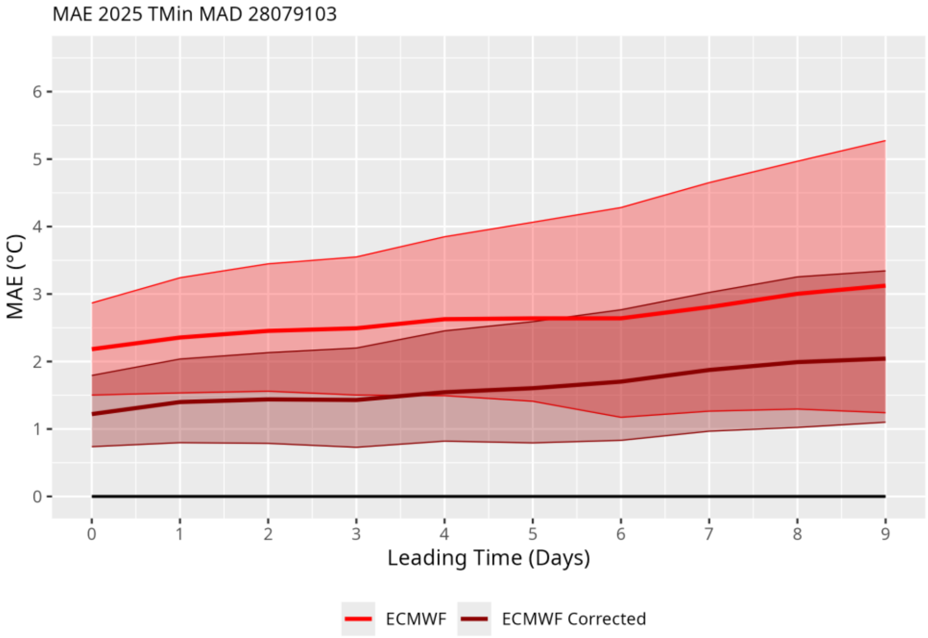


Figure 9. Minimum temperature MAE of raw and corrected ECMWF ensemble predictions for Madrid.

As can be seen in the figures, the bias of the minimum temperature in Madrid has decreased significantly, although there is still a slight positive overestimation of the temperature. This reduction in bias means that the mean absolute error (MAE) is halved, from an average of 2°C to 1°C in the initial forecast periods, while in forecasts with a time FRCame of more than one week, the error is less than 3°C on 90% of days, compared to 5°C in raw forecasts.

Similarly, when it comes to the maximum temperature in Cluj-Napoca, the same pattern emerges: the bias is reduced from an average of 2°C to just a few tenths of a degree. This means that the average error is significantly reduced, particularly in the initial days of the forecast. For instance, the average error on the first day drops from 2.3°C to 1.0°C. However, for longer forecast horizons, the improvement is more modest. For example, at nine days, the error reduces from an average of 3.7°C to 2.8°C.

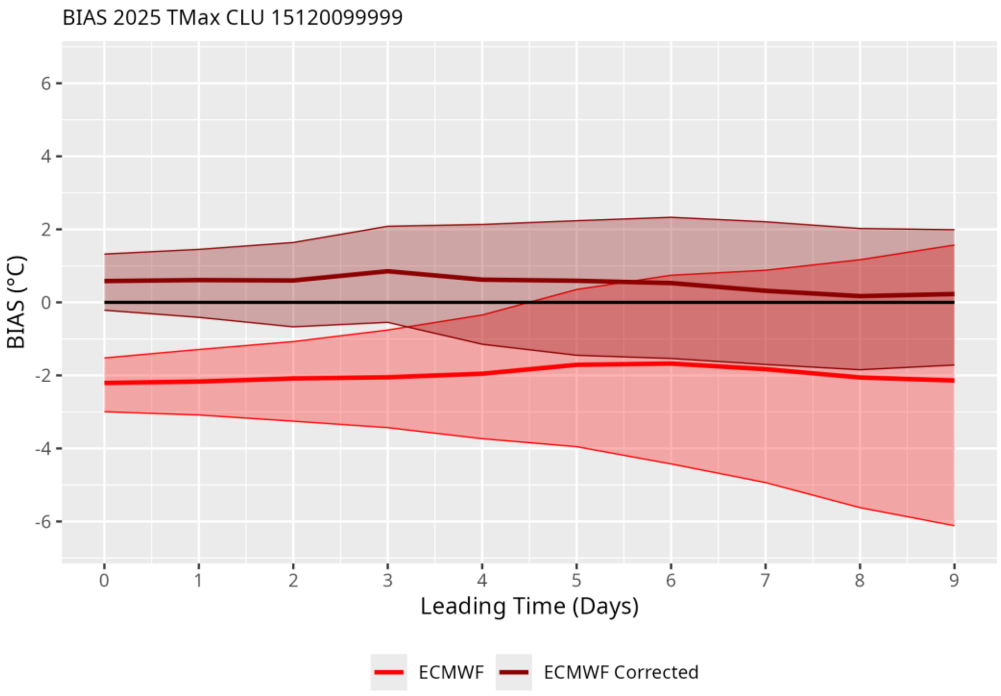


Figure 10. Minimum temperature BIAS of raw and corrected ECMWF ensemble predictions for Cluj-Napoca.

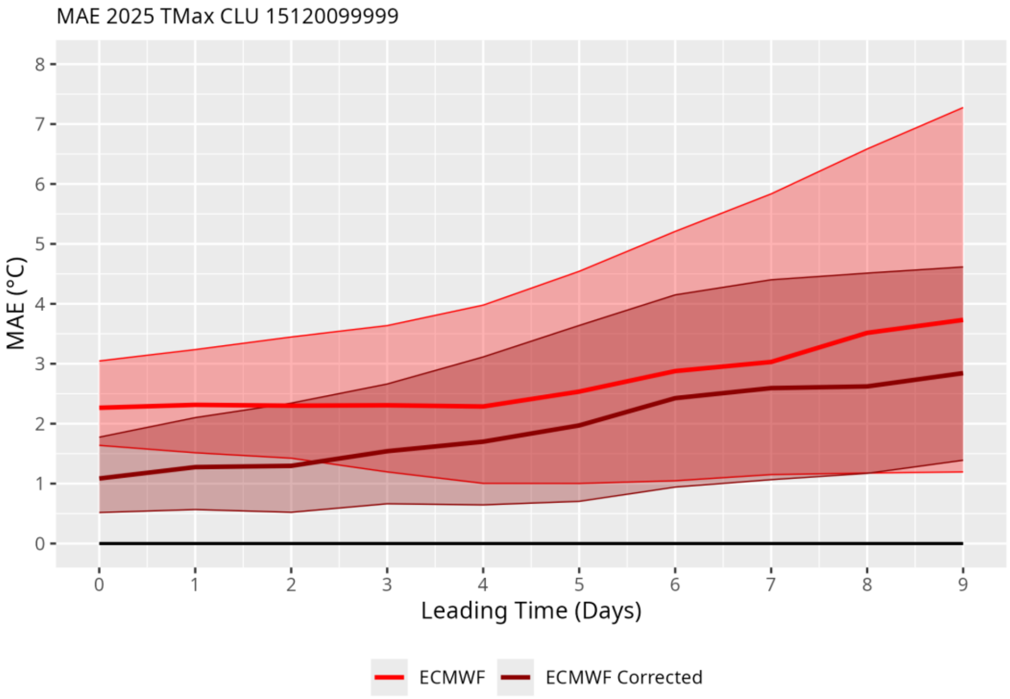


Figure 11. Minimum temperature MAE of raw and corrected ECMWF ensemble predictions for Cluj-Napoca.

4.1.2 Seasonal forecast

In this section, we assess the performance of the three seasonal prediction sources considered in the study: the TeWA–CNN statistical Framework, the ECMWF SEAS5 system, and the CFSv2 model. To avoid overwhelming the reader with information, results for the three sources will be illustrated using monthly mean temperature as a representative example, while the full set of figures for all variables and stations is provided in Annex 4. For each system, we summarise the main findings in terms of correlation and error metrics, followed by an overall synthesis that highlights their relative strengths and limitations for operational seasonal forecasting. Follower cities have been included in this analysis to provide clearer statistical signals using the aforementioned methodologies and improving results for all cities.

4.1.2.1 Seasonal forecast - SEAS5 model validation

<u>Pearson Correlation (Ensemble Mean)</u>

Mean Temperature (TMean)

The correlation analysis for mean temperature shows that forecast skill is strongest at the one-month lead, with positive correlations across most models and observation sites. Median values around 0.2 to 0.3 and upper extremes above 0.4 indicate meaningful predictability for short lead times. Beyond two months, correlations decline sharply toward zero, reflecting the challenges of longer-range forecasting. Models such as ECCC and ECMWF stand out with higher median correlations and greater consistency across stations, underscoring their relative strength for temperature prediction.

As an example, the validation results for the average temperature of the models that constitute the Copernicus seasonal prediction system are shown. For precipitation and wind, see Annex 4.

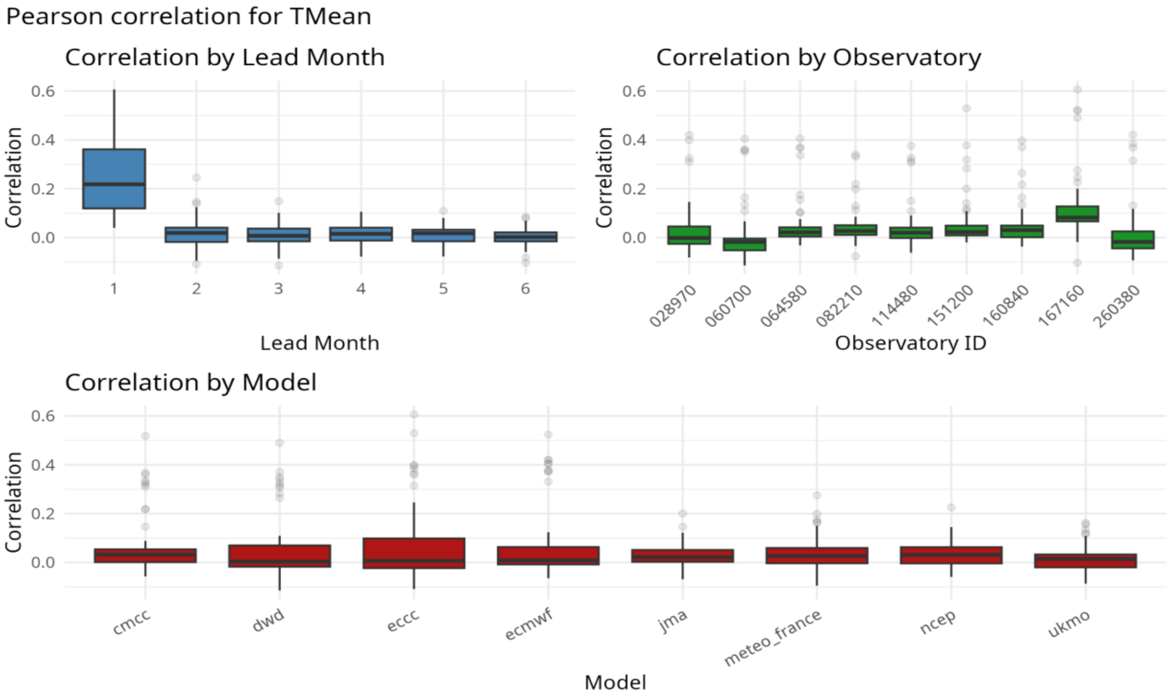


Figure 12. Pearson correlation for TMean, classified for lead month, observatory and model.

Mean Absolute Error (MAE)

Mean Temperature (TMean)

The MAE analysis for temperature forecasts shows the lowest errors at one-month lead, with steady increases as the forecast month progresses, reflecting diminishing predictability. Improvement over climatology is most pronounced at short lead times, reinforcing the added value of dynamical forecasts early in the season. To avoid saturating the graphs with redundant information, only statistics up to the second month are shown, since the difference between the results obtained in the second and sixth months was negligible. Among the models, ECCC and ECMWF consistently achieve lower MAE values, highlighting their better performance, while differences among models narrow beyond a two-month lead.

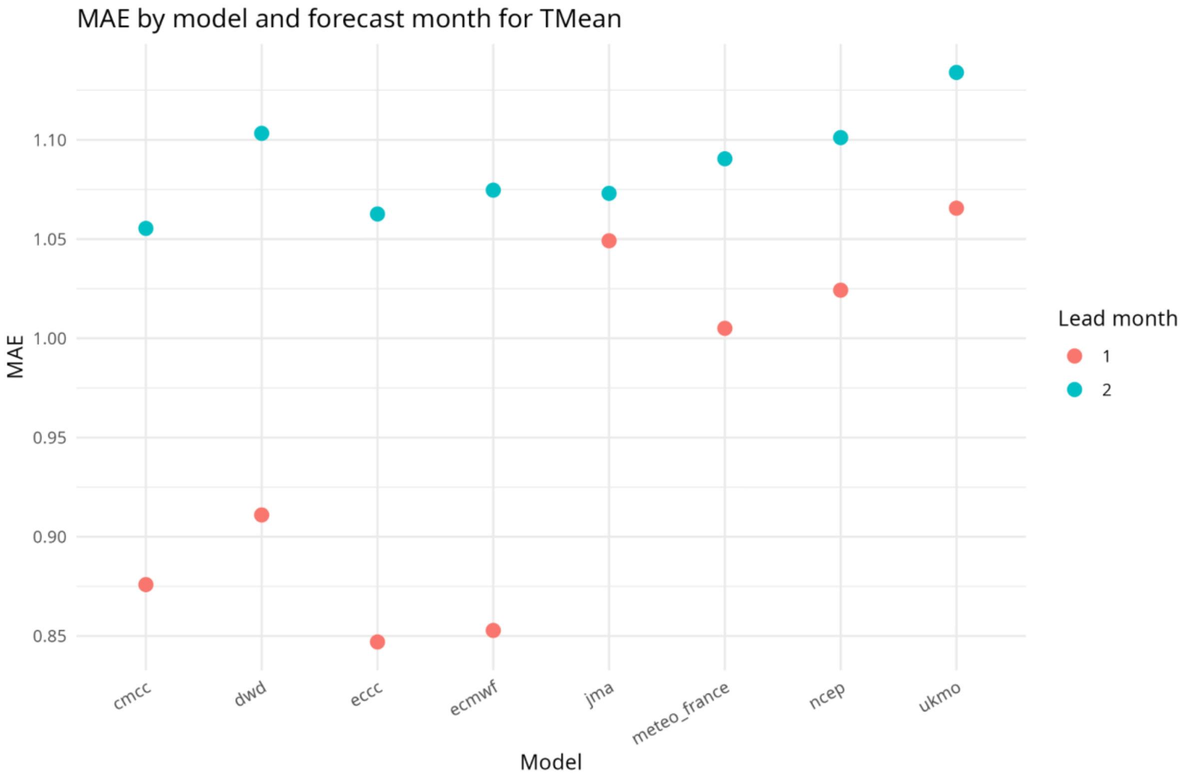


Figure 13. Mean Absolute Error for TMean, classified by lead month and model.

Root Mean Square Error (RMSE)

Mean Temperature (TMean)

RMSE patterns for temperature align closely with MAE results, showing the lowest errors at lead 1 and steady increases with lead time. This underscores the fading predictive skill beyond short leads. ECCC and ECMWF again emerge as the models with the lowest RMSE values at early leads, confirming their comparative strength in temperature forecasting.

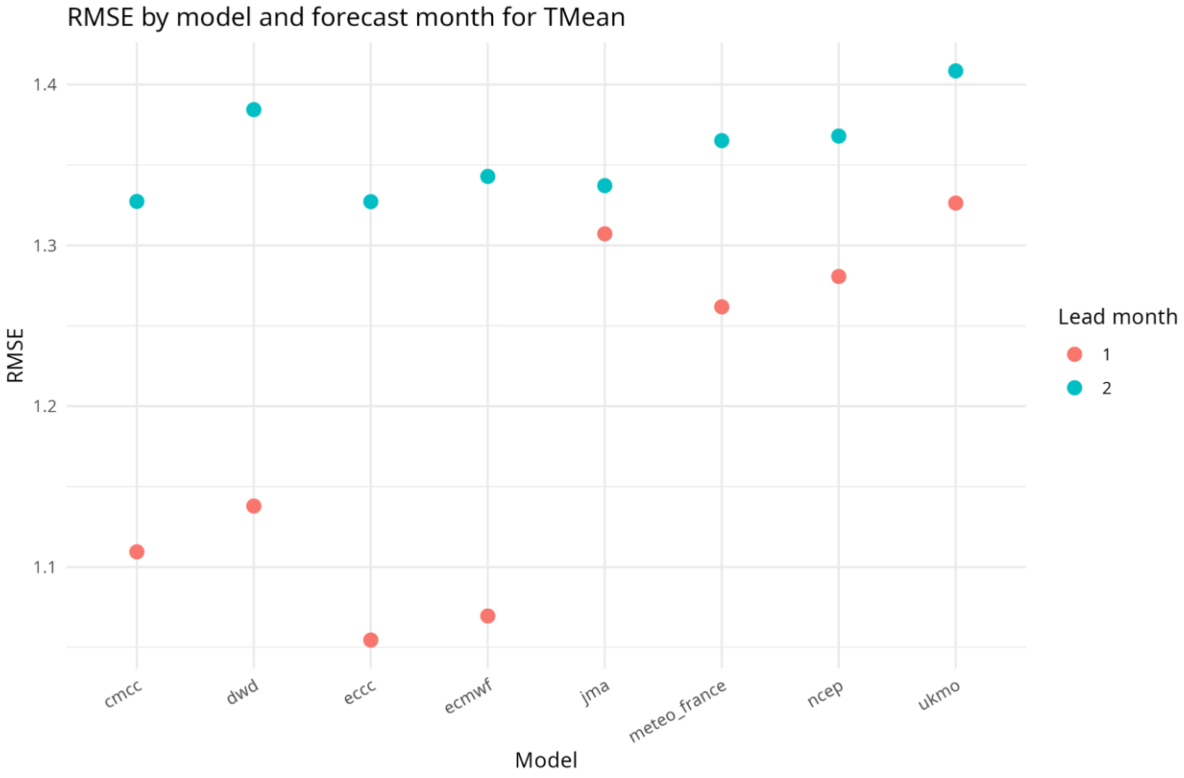


Figure 14. Root Mean Square Error for TMean, classified by lead month and model.

Key findings

This analysis demonstrates that seasonal forecast models consistently achieve their highest skill at one-month lead times across temperature, precipitation, and wind speed. Correlation values and error metrics alike show clear advantages for short-lead predictions, while skill declines sharply beyond two months. Models such as ECCC and ECMWF tend to perform better than others, particularly for temperature and wind forecasts. The results suggest that while dynamical seasonal forecasts provide meaningful guidance at short leads, cautious interpretation is needed for longer-range predictions, and model selection or weighting strategies can further improve forecast value.

Therefore, in the final model, only CMCC, DWD, ECCC and ECMWF will be selected to integrate in the final combined forecast due to their better skill than the rest of the models.

4.1.2.2 Seasonal forecast - Statistical-based seasonal forecast models results

To determine which model performs better, several statistical-based models were run for each observatory, lag and variable.

Regarding the models used, the seasonal forecasting system originally employed nine proposed models, each combining different computational approaches to predict meteorological variables. These models represented a comprehensive ensemble of modern statistical and machine learning techniques tailored for long-range weather prediction.

The CNN_ens model represents the core deep learning approach, utilizing three different random seeds to create an ensemble of Convolutional Neural Networks trained on the gridded ERA5 predictor fields. This ensemble approach helps capture spatial patterns in atmospheric data while providing robustness through model averaging. The CNN_Lin variant implements a simplified linear version within the Keras Framework, offering a more interpretable neural network architecture that maintains the convolutional structure but with linear activation functions.

For traditional statistical modeling, the MultiLinear model employs Ridge regression, a regularized linear regression technique that prevents overfitting while handling the high-dimensional predictor space. This serves as a baseline linear approach against which more complex models can be compared. The Teleconnection model implements a specialized R-based approach that identifies and utilizes atmospheric teleconnection patterns - large-scale atmospheric circulation patterns that influence weather across vast distances.

The TeWa model introduces wavelet-based time series analysis, combining temporal decomposition with atmospheric predictors to capture both FRCequency and time-domain relationships. This approach is particularly effective for identifying multi-scale temporal patterns in meteorological data. The hybrid models begin with CNN_Tele, which blends the CNN_ens predictions with Teleconnection-based forecasts, creating a synthesis of deep learning and atmospheric physics-based approaches.

Similarly, CNN_Tewa combines CNN_ens with TeWavelet predictions, merging spatial pattern recognition with wavelet-based temporal analysis. The most comprehensive hybrid approach, CNN_Tewa_Lin, integrates predictions from CNN_ens, TeWavelet, and Teleconnection models, creating a multi-method ensemble that leverages the strengths of each individual approach. Finally, TeWa_Lin provides a streamlined combination of TeWavelet and Teleconnection models, offering a balance between complexity and interpretability.

Pearson correlation and SMAE were the validating statistical methods selected to measure the performance of each of them.

Pearson Correlation

Pearson correlation is used to assess whether the forecasts reproduce the observed temporal variability and phase of anomalies, independently of any systematic bias or scaling errors. It therefore provides a robust, unitless measure of “pattern skill” that is comparable across stations, lead times and models.

Mean Temperature (TMean)

For the average temperature, the boxplots show that all models achieve modest but generally positive Pearson correlations for near-term leads, with median values typically between about 0.1 and 0.3 for lags 1-3. The best performance is obtained by the TeWa-based blends, particularly CNN_TeWa and TeWa_Lin, which consistently rank among the top models and maintain positive skill even at longer lead times (lags 4-6). The pure CNN ensemble (CNN_Ens) also provides robust, though slightly smaller, correlations across most lags. In contrast, the purely linear approaches (CNN_Lin, MultiLinear) and the standalone teleconnection model tend to degrade more rapidly with lead time, with medians hovering around zero or becoming negative beyond about 3-4 months. Overall, incorporating TeWa information clearly enhances temporal persistence of correlation skill in the seasonal temperature forecasts.

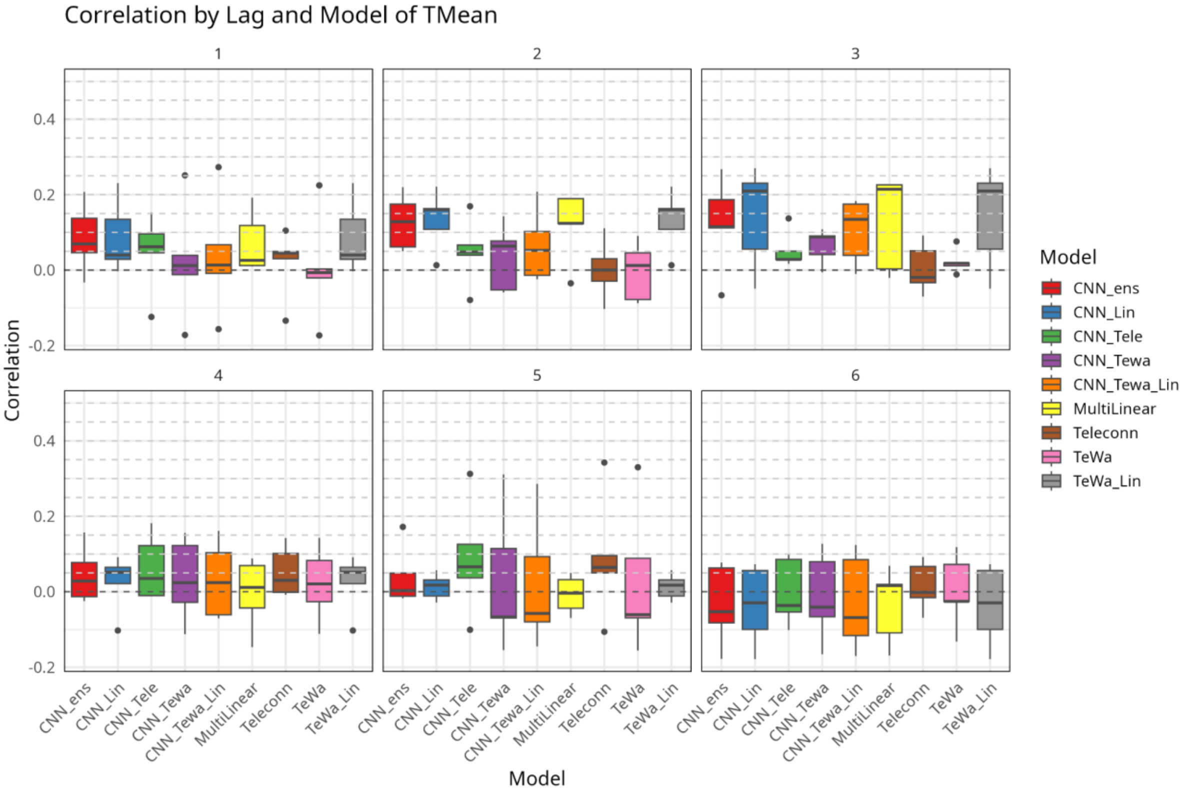


Figure 15. Combined Pearson Correlation for different Statistical based models for each lag and for Mean Temperature.

Among the observatories, the highest and most persistent correlations are found at the Mediterranean and continental sites in Greece (167160), Italy (160840) and Romania (151200), which show clearly positive median correlations at most lead months. The Czech Republic (114480) also exhibits modest but robust positive values, whereas the northern stations in Finland (028970) and Denmark (060700) often hover around zero or slightly negative correlations. Western sites in Belgium (064580), Spain (082210) and Estonia (260380) tend to display intermediate behaviour, with small but generally positive correlations that decay with lead time.

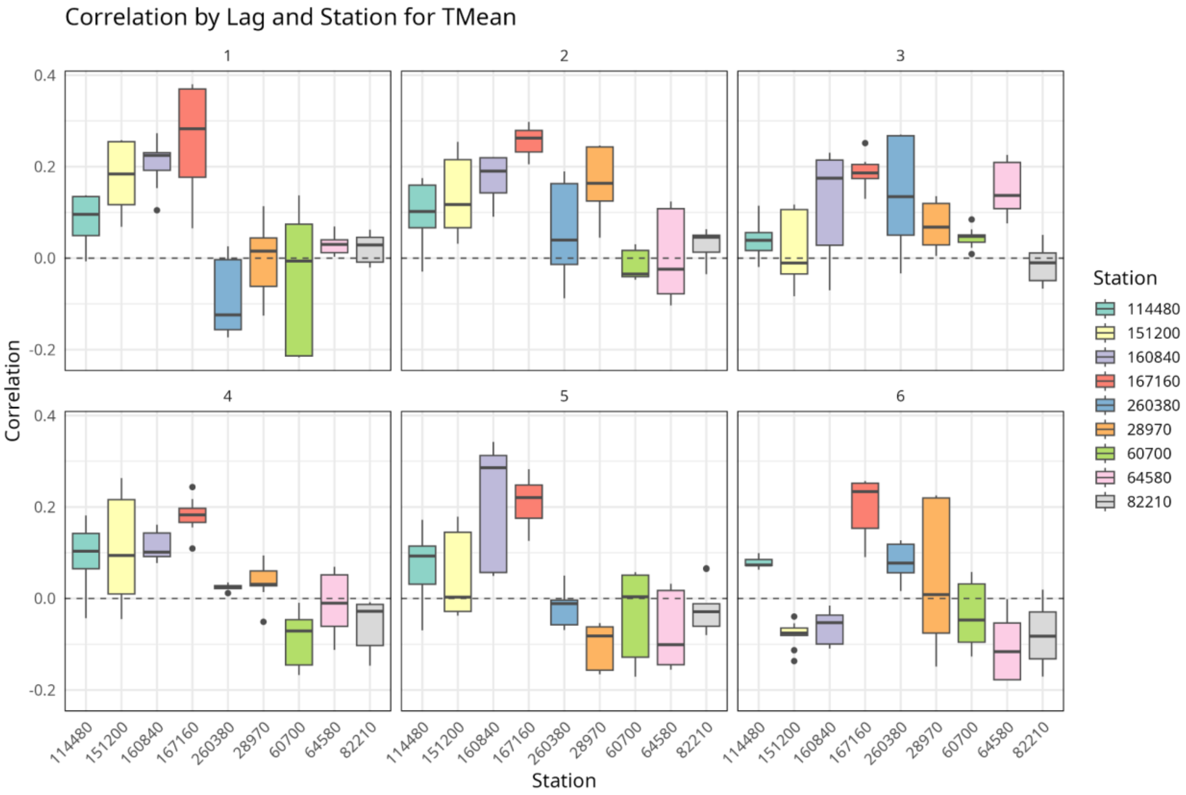


Figure 16. Combined Pearson Correlation for the observatories used for each lag (Mean Temperature).

Standardized Mean Absolute Error

SMAE was chosen as a quantify method to measure the magnitude of forecast errors in a scale-independent way, normalizing by the observed variability or climatology. This allows us to compare error levels across variables and locations with very different units and ranges, and to identify models that are not only well correlated but also quantitatively accurate.

Mean Temperature (TMean)

SMAE values for temperature are relatively stable with lead time and lie mostly between about 1.0 and 1.4 for the majority of CNN and Teleconnection/TeWa-Lin configurations. CNN_Ens consistently delivers the lowest errors (medians close to 1), with TeWa_Lin and, to a lesser extent, CNN_Lin and Teleconn forming a compact group with slightly higher but comparable SMAE. The standalone TeWa model is clearly an outlier, with substantially larger SMAE (often around 1.8–2.1), indicating that its raw forecasts are less accurate than the blends and CNN-based approaches, even when correlations are similar.

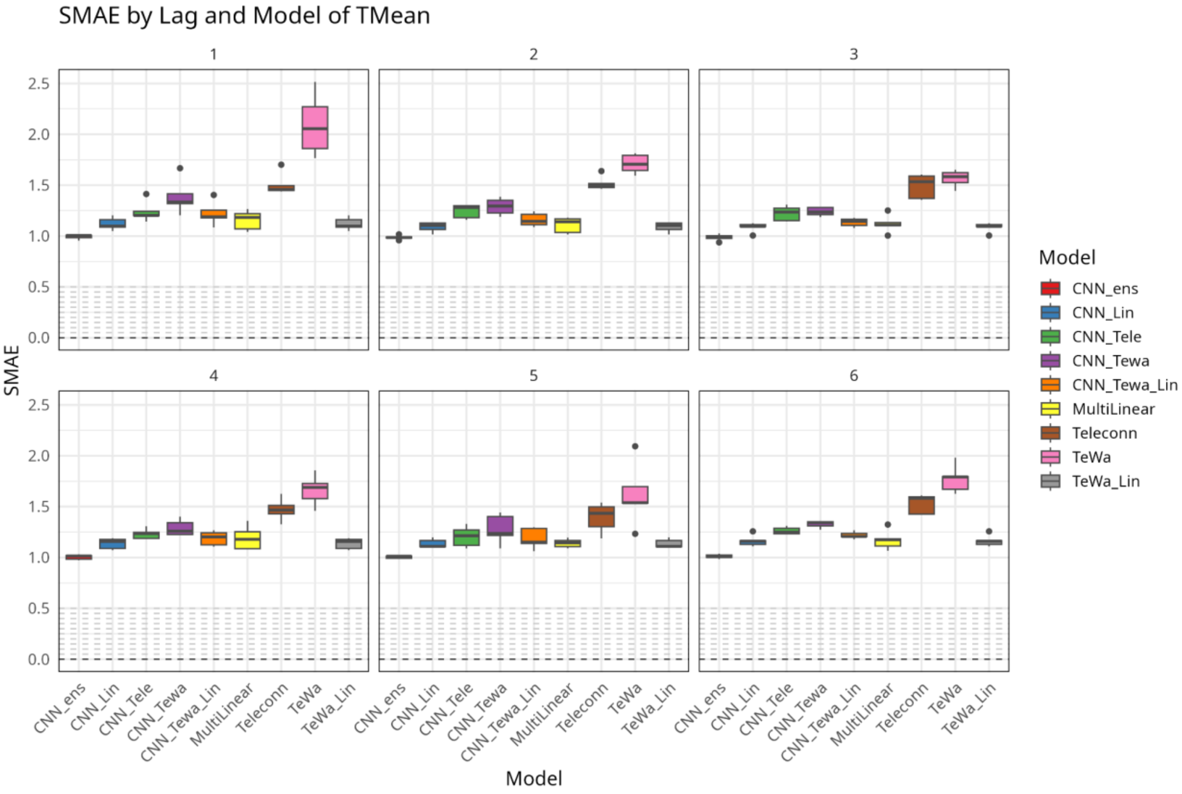


Figure 17. Combined SMAE for different Statistical based models for each lag and for Mean Temperature.

Temperature SMAE is fairly homogeneous across stations, but systematic differences are visible: Italy (Piacenza)[1], Czech Republic (Pilsen) and Spain (Marid-Barajas) usually appear among the lowest-error sites, with median SMAE close to 1–1.2 across lead months. Greece (Athens) and Romania (Cluj-Napoca) tend to show slightly larger errors, and the northern and western stations in Estonia, Belgium and Denmark also feature somewhat higher SMAE, suggesting that local variability and continental effects make temperature slightly difficult to predict there.

---

[1] City with the closest weather station accomplishing the required historical observed data length for Parma.

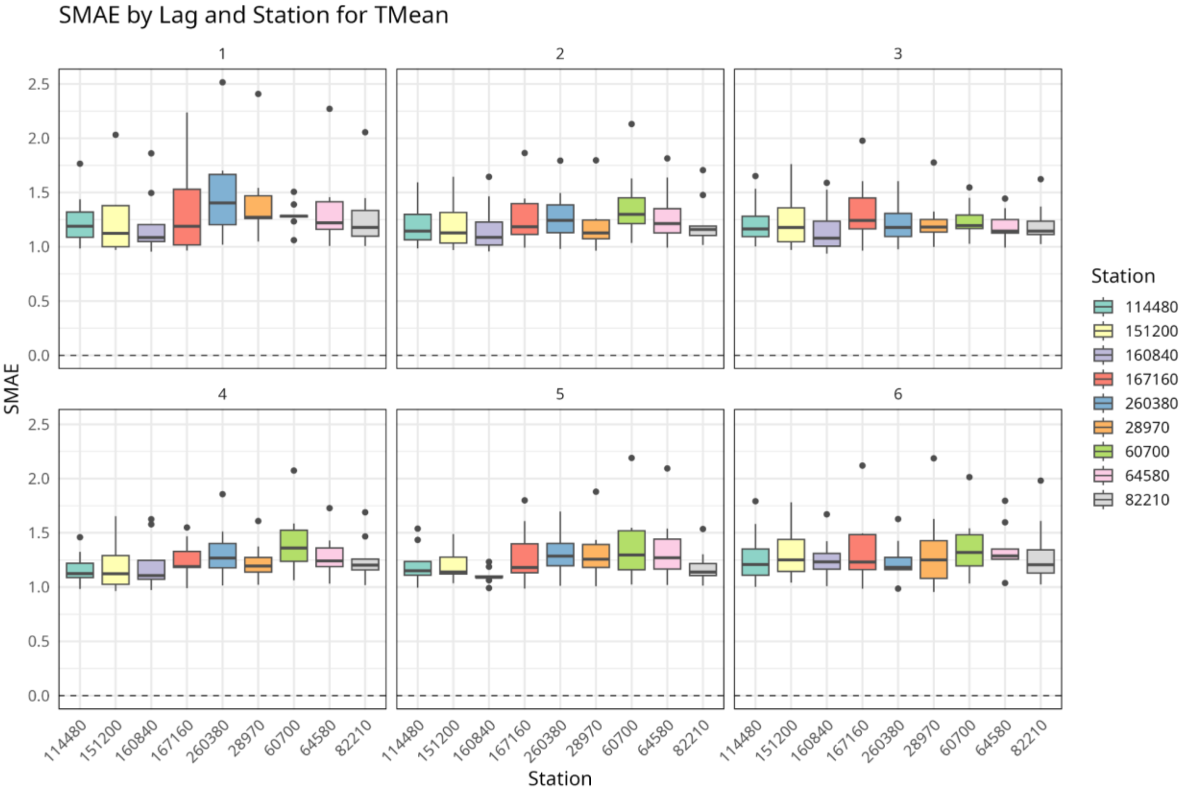


Figure 18. Combined SMAE for the observatories used for each lag using all models (Mean Temperature).

Key findings summary

Overall, the benchmarking confirms that the CNN_Ens, CNN_TeWa and TeWa_Lin configurations form the best-performing group across variables, lead times and metrics. For temperature and precipitation, they systematically provide the highest, or near-highest, Pearson correlations while keeping SMAE close to or below those of the remaining baselines. In particular, CNN_Ens tends to minimise SMAE for all variables, whereas TeWa_Lin and CNN_TeWa often show a slight advantage in correlation, especially at short and intermediate leads. Teleconnection- and TeWa-based blending therefore adds value with respect to purely linear or standalone TeWa approaches, and clearly outperforms the simple MultiLinear baseline.

At the same time, the results also highlight the intrinsic limitations of seasonal forecasting on a local scale. Correlation values are generally modest (often below 0.25) and decay with lead time, with mean wind speed showing the weakest and most variable skill. SMAE values around or above 1 indicate that large individual errors are still common, particularly for precipitation and wind. Skill is heterogeneous among stations and lead months, and some models occasionally produce near-zero or even slightly negative correlations. These features underline that, despite the relative gains brought by the CNN–TeWa ensembles, the predictive capacity remains limited and forecasts should be interpreted as probabilistic, low-signal guidance rather than as deterministic, high-accuracy predictions. Therefore, the statistical–dynamical approach based on CNN_Ens, CNN_TeWa and TeWa_Lin delivers station-scale skills that are broadly comparable to, and in some cases slightly better than, those of SEAS5. At lead month 1, Pearson correlations for temperature, precipitation and mean wind speed are of similar magnitude to the best SEAS5 members (typically up to ~0.2–0.3), while SMAE remains low and relatively stable with lead time, indicating a good control of local errors after bias correction and downscaling. In contrast, SEAS5 correlations and MAE improvements over climatology tend to collapse beyond months 2–3, whereas FICLIMA's post-processed models retain small but positive correlations and competitive error levels up to lead month 4–6. This suggests that the proposed methodology can extract at least as much usable signal as SEAS5 at short range and may provide added value for some variables and locations at longer seasonal leads, although forecast skill is still subject to substantial local variability in both systems.

4.1.2.3 Seasonal forecast - CFSv2 models results

Pearson Correlation

Temperature (Maximum Temperature)

For daily maximum temperature, CFSv2 shows strong short-range skill, with median correlations around 0.6-0.7 on day 1 and a gradual decline to about 0.2 by days 10-12. Beyond the second week, correlations drop to low but still positive values (≈0.1-0.15) and remain in that range for most of the 51-day horizon, with a small secondary peak around days 40-45. This pattern reflects clear added value at synoptic to sub-monthly scales, but only modest, residual predictability at longer leads.

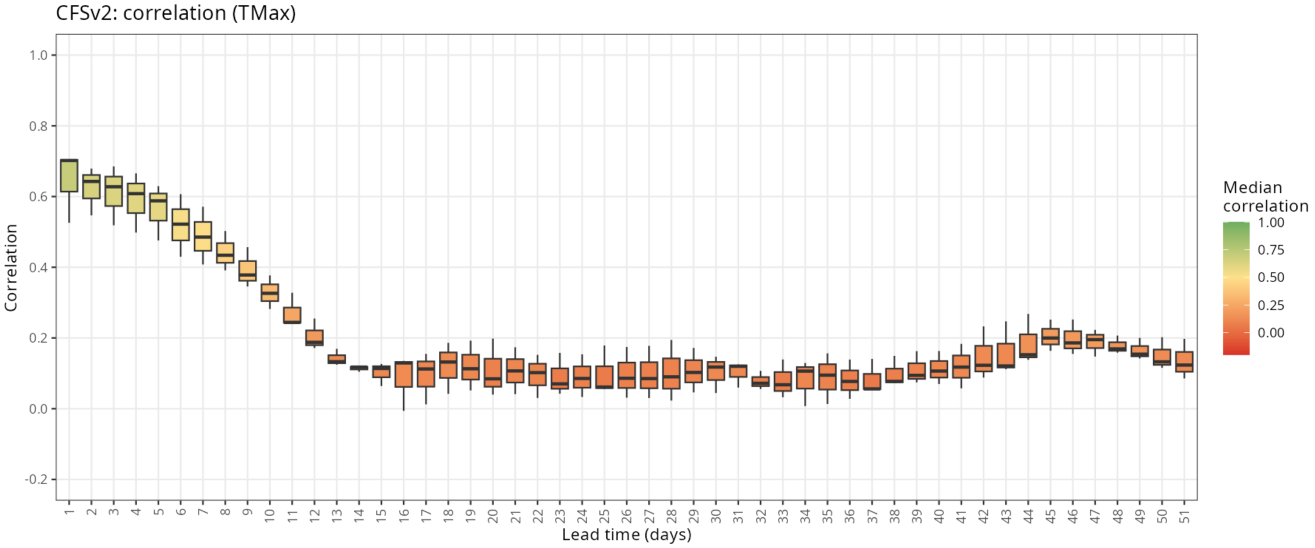


Figure 19. Pearson correlation for maximum temperature across the entire dataset of observations for each lead day.

Standardized Mean Absolute Error

Temperature (Maximum Temperature)

The standardized MAE for maximum temperature is lowest at very short leads (≈0.5 on day 1) and increases steadily during the first two to three weeks, stabilizing around SMAE ≈0.9-1.0 afterwards. This behaviour indicates that CFSv2 provides clear improvements over climatology at short lead times, while errors approach climatological levels as the forecast extends into the subseasonal window, with relatively small changes in SMAE beyond about day 20.

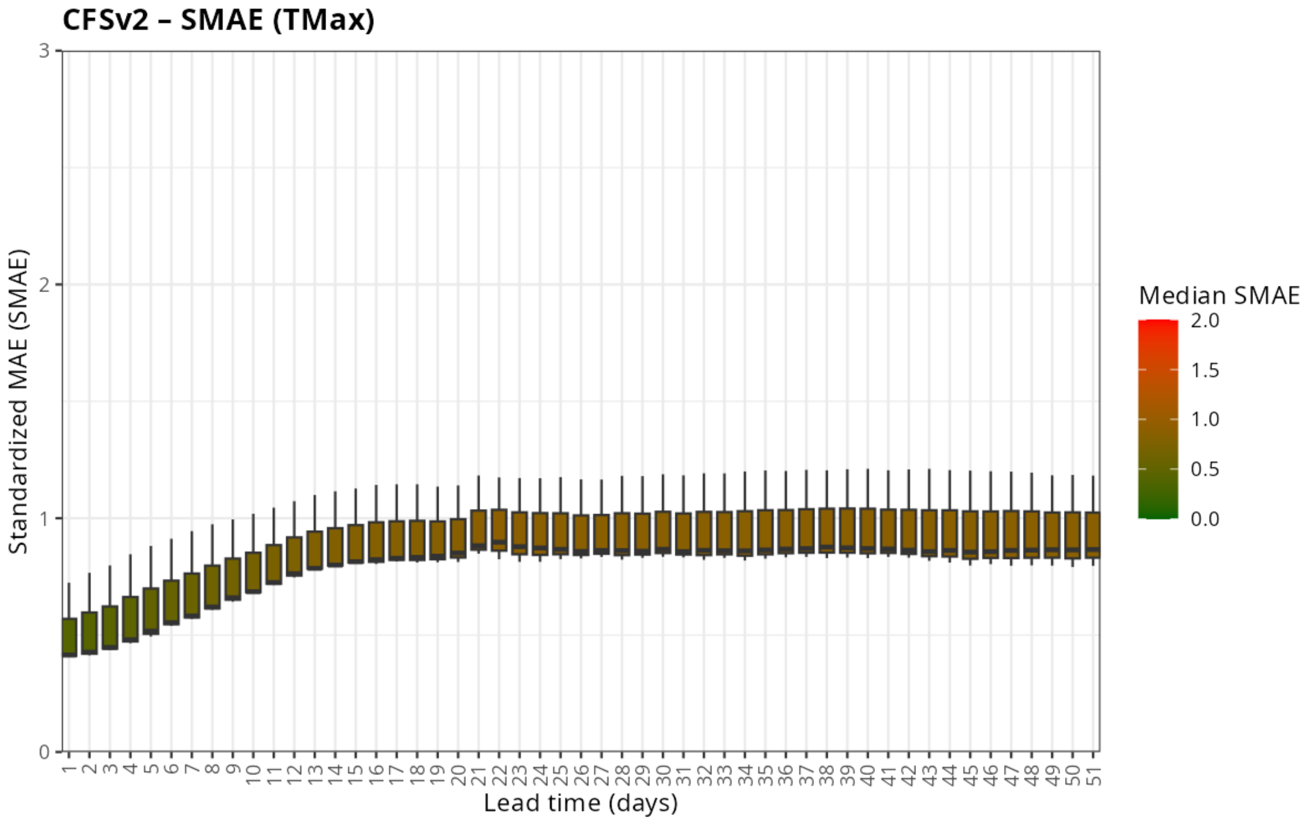


Figure 20. SMAE for maximum temperature across the entire dataset of observations for each lead day.

Key findings summary

Overall, the CFSv2 hindcast evaluation confirms that this dynamical system provides useful but strongly lead-time-dependent skill at the station scale. For temperature, CFSv2 achieves high short-range correlations and low SMAE during the first 1–2 weeks, indicating that it captures synoptic-scale variability reasonably well before its performance gradually degrades toward climatological levels. For precipitation, deterministic skill is much weaker and tends to vanish beyond ~10–15 days, with correlations close to zero and relatively large errors. For 10-m wind speed, the model shows moderate correlation at very short leads but systematically high SMAE, reflecting substantial local bias and amplitude errors that only decrease as the forecast damps toward climatology. Taken together, these results suggest that CFSv2 is most valuable as a subseasonal tool for the first 1–2 weeks, while its direct use for longer-range, point-scale forecasts of precipitation and wind should be treated with caution and ideally complemented by statistical post-processing.

4.1.2.4 Seasonal forecast summary and conclusions

Across the URBREATH stations and variables, the TeWA-based statistical Framework extended with CNN-derived teleconnection indices (CNN_Ens, CNN_TeWa, TeWa_Lin) consistently emerges as the best-performing group at the monthly scale. It delivers small but robust positive correlations (typically up to ~0.2–0.3 for temperature and precipitation at short leads) and SMAE values close to 1–1.4, and it maintains usable skill up to 4–6 months ahead for some stations and variables. This confirms that combining data-driven teleconnection indices with wavelet and ARIMA self-predictability effectively extracts a significant FRCaction of the available seasonal signal at station level.

The SEAS5 system shows clear added value at short lead times, especially for temperature, with median correlations around 0.2–0.3 and reduced MAE at lead month 1. However, its skill decays rapidly beyond 2–3 months, and at the local station scale its correlation and error metrics are generally comparable to or slightly weaker than those of the TeWA-based models. SEAS5 therefore provides useful large-scale dynamical guidance, but its direct monthly anomalies at station level remain limited in terms of deterministic skill for longer leads.

CFSv2, evaluated at the daily scale, exhibits strong short-range skill for temperature (correlations >0.6 on day 1 and still positive after 1–2 weeks) and modest positive correlations for precipitation during the first 10–15 days, with SMAE increasing toward climatological levels as lead time grows. For 10-m wind speed, CFSv2 captures part of the synoptic variability during the first week but shows relatively high standardized errors and rapidly damped anomalies, thereafter, reflecting difficulties in representing local wind at coarse resolution. Overall, CFSv2 offers valuable subseasonal information in the first 1–2 weeks, particularly for temperature, but its point-scale skill decreases markedly as the forecast horizon extends.

4.1.3 Climate projections

Daily projections have been generated for each of the four Frontrunner cities from 1981 to 2100. All possible combinations between the 10 selected CMIP6 models (see Table 3) and the four emission scenarios considered (SSP1-2.6, SSP2-4.5, SSP3-6.7, SSP5-8.5) have been considered. Therefore, a total of 40 possible climate projections is shown for each of the variables.

Results for Madrid

According to the climate change local scenarios generated, the main conclusions for each variable for Madrid are summarised below, together with an illustrative figure (figure A5). For reference, the remaining figures and extended information are included in Annex 5.

No significant changes can be expected in maximum relative humidity across all scenarios and time horizons (figure A5.1). Changes in values are expected to range from -0.3 to -0.4 per cent both by mid-century and by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 90.8%. No significant changes are also expected in minimum relative humidity values across all scenarios and time horizons (figure A5.2). Changes in these values are expected to range from -0.4 to -0.5 per cent by mid-century and from -0.2 to -0.5 per cent by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 67.4%. In the initial results, it has been observed an increase in temperature in both maximum and minimum values. As it warms, the atmosphere has a greater capacity to hold moisture. But at the same time, as global surface temperatures rise because of anthropogenic warming, the amount of water evaporated increases and, therefore, the water vapour content in the atmosphere increases. Therefore, this result is to be expected.

Regarding the radiation results, long wave (terrestrial) radiation and short wave (solar) radiation are studied. In terms of long wave radiation significant changes are expected as shown in figure A5.3. Changes in values are expected to range from 9.2 to 13.5 W/m^2 by mid-century and from 13 to 24.3 W/m^2 by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 294.7 W/m^2. In terms of short-wave radiation, significant changes are also expected; as shown in figure A5.4. Changes in values are expected to range from 6.9 to 6.3 W/m^2 by mid-century and from 8.5 to 10.5 W/m^2 by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 198.1 W/m^2. For perfect equilibrium, there must be a net balance at the top of the atmosphere, such that the amount of incoming energy from the sun (in the form of shortwave radiation) must be the same as the amount of energy emitted by the Earth (in the form of longwave radiation).

Regarding the wind speed results, some significant changes are expected as shown in figure A5.5. Changes in values are expected to be -0.2 km/h by mid-century and range from -0.3 to -0.4 km/h by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 7.5 km/h.

Regarding the Heat Wave results, the annual number of Heat Waves episodes and the average duration, mean intensity and maximum intensity of these episodes are studied. In terms of annual number of Heat Waves, significant changes are expected as shown in figure X1 (figure A5.6 in Annex 5). Changes in values are expected to range from 3.5 to 3.9 per year by mid-century and from 4.3 to 4.9 per year by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.3 episodes per year. This result is complemented by the characteristics analysed using the annual average for each episode: average duration, mean intensity and maximum intensity, analysed in figures A5.7-A5.9.

- Regarding average duration (figure A5.7), significant changes are expected. Its increases are expected to range from 2.2 to 3.5 days per year by mid-century and from 3.3 to 9.5 per year by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 4.9 days per year.
- Regarding average intensity (figure A5.8), significant changes are expected, especially by the end of the century. Its increases are expected to range from 0.5 (no significant) to 0.8 °C by mid-century and from 0.8 to 1.8 °C by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 39.4°C.
- Regarding maximum intensity (figure A5.9), significant changes are expected, especially by the end of the century. Its increases are expected to range from 1.2 to 1.6 °C by mid-century and from 1.8 to 3.4 °C by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 40.2°C.

Based on the results, not only an increase in the occurrence of heatwaves could be expected, but also in their average duration and intensity.

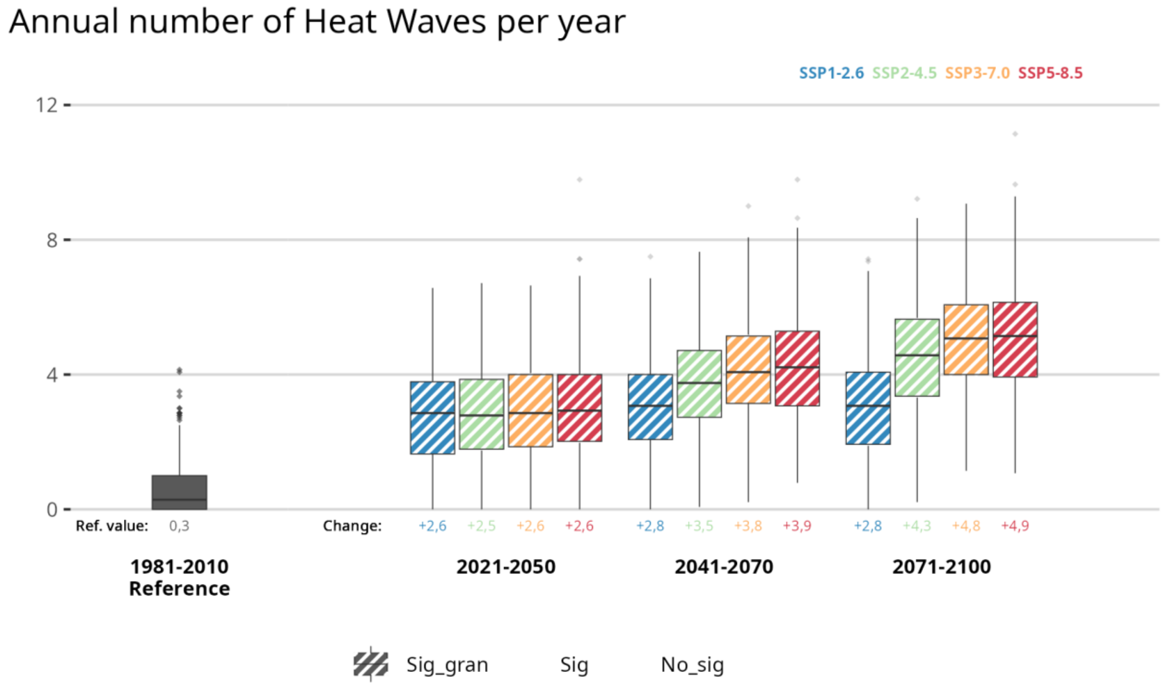


Figure 21. The graph shows the evolution of annual number of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of 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. [Figure A5.6 in Annex 5]

Regarding the SPI results, 3-month, 6-month, 12-month, 24-month and 60-month accumulation periods are studied to assess short-, medium- and long-term conditions (figures A5.10-A5.14). In summary, for all the accumulation periods analysed, it can be said that no significant changes are expected in any case, maintaining its value in the normal value category for each case during all periods and scenarios.

Regarding the SPEI results, 3-month, 6-month, 12-month, 24-month and 60-month accumulation periods are studied to assess short-, medium- and long-term conditions (as SPI) (figures A5.15-A5.19). In summary, for all the accumulation periods analysed, it can be said that significant decreases are expected in all cases.

- Regarding 3-month SPEI (figure A5.15), its decreases are expected to range from -0.7 to -0.8 by mid-century and from -0.9 to -1.5 (moderately dry) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0 (normal value).
- Regarding 6-month SPEI (figure A5.16), its decreases are expected to range from -0.7 to -1.1 by mid-century and from -1.1 to -1.9 (moderately-severely dry) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0 (normal value).
- Regarding 12-month SPEI (figure A5.17), its decreases are expected to range from -1 to -1.3 by mid-century and from -1.4 to -2.5 (moderately-extremely dry) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0 (normal value).
- Regarding 24-month SPEI (figure A5.18), its decreases are expected to range from -1.3 to -1.7 by mid-century and from -1.8 to -2.8 (severely-extremely dry) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0 (normal value).
- Regarding 60-month SPEI (figure A5.19), its decreases are expected to range from -1.7 to -2.1 by mid-century and from -2.3 to -3.5 (extremely dry) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0 (normal value).

Results for Leuven

According to the climate change local scenarios generated, the main conclusions for each variable for Leuven are summarised below, together with an illustrative figure. For reference, the remaining figures and extended information are included in Annex 5.

According to the climate change local scenarios generated for Leuven no significant changes can be expected in maximum relative humidity across all scenarios and time horizons (figure A5.20), except for the end of the century in the two most catastrophic scenarios. Changes in values are expected to range from -0.4 to -0.5 per cent both by mid-century and from -0.6 to -1.5 per cent by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 91.5%. No significant changes are also expected in minimum relative humidity values across all scenarios and time horizons (figure A5.21), except for the most catastrophic scenario by the end of the century. Changes in these values are expected to range around -0.6 per cent by mid-century and from -0.8 to -2 per cent by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 63.1%. In the initial results, an important increase in temperature has been observed in both maximum and minimum values. As it warms, the atmosphere has a greater capacity to hold moisture. But at the same time, as global surface temperatures rise because of anthropogenic warming, the amount of water evaporated increases and, therefore, the water vapour content in the atmosphere increases. The expected change is the result of the balance between this interaction.

Regarding the radiation results, long wave (terrestrial) radiation and short wave (solar) radiation are studied. In terms of long wave radiation significant changes are expected as shown in figure A5.22. Changes in values are expected to range from 8.1 to 11.5 W/m^2 by mid-century and from 10.4 to 23.4 W/m^2 by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 303.2 W/m^2. In terms of short-wave radiation, significant changes are also expected; as shown in figure X4. Changes in values are expected to range from 6.6 to 5.1 W/m^2 by mid-century and from 6 to 6.4 W/m^2 by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 124.7 W/m^2. For perfect equilibrium, there must be a net balance at the top of the atmosphere, such that the amount of incoming energy from the sun (in the form of shortwave radiation) must be the same as the amount of energy emitted by the Earth (in the form of longwave radiation).

Regarding the wind speed results, some significant changes are expected as shown in figure A5.24. Changes in values are expected to range from -0.3 to -0.4 km/h by mid-century and range from -0.2 to -0.4 km/h by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 13 km/h.

Regarding the Heat Wave results, the annual number of Heat Waves episodes and the average duration, mean intensity and maximum intensity of these episodes are studied. In terms of annual number of Heat Waves, significant changes are expected as shown in figure X2 (A5.25 in Annex 5). Changes in values are expected to range from 2 to 2.5 per year by mid-century and from 2.5 to 5 per year by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0 episodes per year. This result is complemented by the characteristics analysed using the annual average for each episode: average duration, mean intensity and maximum intensity, analysed in figures A5.26-A5.28.

- Regarding average duration (figure A5.26), significant changes are expected. Its increases are expected to range from 0.4 to 0.9 days per year by mid-century and from 0.7 to 1.7 per year by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 4.9 days per year.
- Regarding average intensity (figure A5.27), only significant changes are expected by the end of the century in the most catastrophic scenario. Its increases are expected to range from 0.7 to 0.8 °C by mid-century and from 0.8 to 1.3 °C by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 32.5°C.
- Regarding maximum intensity (figure A5.28), only significant changes are expected in the worst scenario by mid and end century. Its increases are expected to range from 1 to 1.4 °C by mid-century and from 1.3 to 2.6 °C by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 34°C.

Based on the results, a significant increase in the annual number of heat waves is expected, ranging from a median zero value in the reference period to a median of 2.5 or 5 heatwaves per year by mid-century and late-century, respectively, in the worst-case scenario. This is the most significant change expected in the analysed values. However, these heatwaves are also expected to intensify, particularly towards the end of the century, in terms of both average duration and average and maximum intensity, with significant increases in some cases.

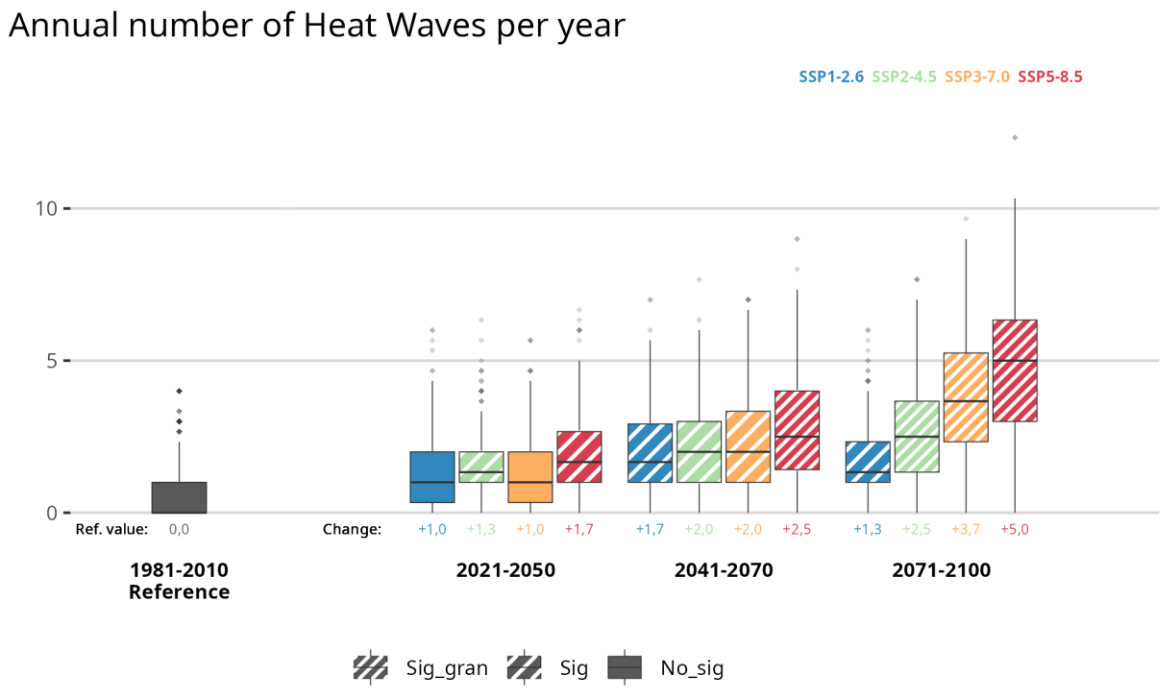


Figure 22. The graph shows the evolution of annual number of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of 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. [Figure A5.25 in Annex 5]

Regarding the SPI results, 3-month, 6-month, 12-month, 24-month and 60-month accumulation periods are studied to assess short-, medium- and long-term conditions (figures A5.29-A5.33). All indices start at a value of 0 in the reference period except for the 60-month accumulation index, which starts at a value of -0.1. Significant increases are expected by the end of the century in 3- and 6-month accumulations in all scenarios except the most sustainable one. For the 3-month SPI, a change from a value of 0 (normal value) in the reference period to a significant increase from 0.4 to 0.6 (normal-moderately wet) is expected by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively. For the 6-month SPI, a change from a value of 0 in the reference period to a significant increase from 0.5 to 0.8 (normal-moderately wet) is expected by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively. In 12- and 24-month accumulations significant increases are expected by mid-century in the worst scenario and by the end of the century in all scenarios except the most sustainable one. These changes for the 12-month accumulation period would range from 0.4 to 0.6 (normal-moderately wet) for mid-century and from 0.6 to 1.1 (moderately wet) for the end of the century in the SSP2-4.5 and SSP5-8.5 respectively. In the case of the 24-month accumulation period would range from 0.6 to 0.7 (moderately wet) for mid-century and from 0.9 to 1.7 (moderately-severely wet) for the end of the century in the SSP2-4.5 and SSP5-8.5. Finally, for the 60-month period, significant increases would be expected for all future periods in all scenarios except the most sustainable one, ranging from 1 to 1.3 (moderately wet) by mid-century and from 1.5 to 3 (moderately-extremely wet) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of -0.1.

Regarding the SPEI results, 3-month, 6-month, 12-month, 24-month and 60-month accumulation periods are studied to assess short-, medium- and long-term conditions (as SPI) (figures A5.34-A5.38). In summary, for 3-month, 6-month, 12-month accumulation periods, no significant decreases are expected from values of 0, 0 and 0.1 (normal values) in the reference period, respectively. For 24-month accumulation period, only a significant decrease is expected by the end of the century with a value of -0.9 from a 0.1 (moderately dry - normal value) value in the reference period. For 60-month accumulation period, almost all the expected changes are significant by mid and end century, ranging from -0.9 to -1.1 (moderately dry) by mid-century and from -0.7 to -1.5 (moderately dry) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.1 (normal values).

Results for Tallinn

According to the climate change local scenarios generated, the main conclusions for each variable for Tallinn are summarised below, together with an illustrative figure. For reference, the remaining figures and extended information are included in Annex 5.

According to the climate change local scenarios generated for Tallinn no significant changes can be expected in maximum relative humidity across all scenarios and time horizons (figure A5.39). Changes in values are expected to range from -0.3 to -0.4 per cent both by mid-century and by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 90.8%. No significant changes are also expected in minimum relative humidity values across all scenarios and time horizons (figure A5.40). Changes in these values are expected to range from -0.4 to -0.5 per cent by mid-century and from -0.2 to -0.5 per cent by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 67.4%. In the initial results, we observed an important increase in temperature in both maximum and minimum values. As it warms, the atmosphere has a greater capacity to hold moisture. But at the same time, as global surface temperatures rise as a result of anthropogenic warming, the amount of water evaporated increases and, therefore, the water vapour content in the atmosphere increases. The expected change in relative humidity values is the result of the balance between this interaction.

Regarding the radiation results, long wave radiation and short-wave radiation are studied. In terms of long wave radiation significant changes are expected as shown in figure A5.41. Changes in values are expected to range from 10.5 to 14.8 W/m^2 by mid-century and from 14.2 to 28.1 W/m^2 by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 282.9 W/m^2.Instead, in terms of short wave radiation only significant changes are expected at the less catastrophic scenario by mid and end century. As shown in figure A5.42, changes in values are expected to range from 3.1 to 2.9 W/m^2 by mid-century and from 3.5 to 1.7 W/m^2 by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 115.1 W/m^2. For perfect equilibrium, there must be a net balance at the top of the atmosphere, such that the amount of incoming energy from the sun (in the form of shortwave radiation) must be the same as the amount of energy emitted by the Earth (in the form of longwave radiation).

Regarding the wind speed results, no significant changes are expected as shown in figure A5.43. Changes in values are expected to range from 0 to -0.1 km/h by mid-century and range around -0.1 km/h by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 12.3 km/h.

Regarding the Heat Wave results, the annual number of Heat Waves episodes and the average duration, mean intensity and maximum intensity of these episodes are studied. In terms of annual number of Heat Waves, significant changes are expected as shown in figure X3 (figure A5.44 in Annex 5). Changes in values are expected to range from 1.5 to 2 per year by mid-century and from 2 to 3.8 per year by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.2 episodes per year. This result is complemented by the characteristics analysed using the annual average for each episode: average duration, mean intensity and maximum intensity, analysed in figures A5.45-A5.47.

- Regarding average duration (figure A5.45), some significant changes are expected. Its increase is expected to range from 0.6 (no significant value) to 1.5 days per year by mid-century and from 1.2 to 3.5 per year by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 4.5 days per year.
- Regarding average intensity (figure A5.46), there are no significant changes in any case. Its increases are expected to be around 0.5 °C by mid-century and from 0.6 to 0.9 °C by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 28.3°C.
- Regarding maximum intensity (figure A5.47), only significant changes are expected in the worst scenario by the end of the century. Its increases are expected to range from 0.9 to 1.2 °C by mid-

century and from 1 to 1.9 °C by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 29.7°C.

Based on the results, a significant increase in the annual number of heat waves is expected, ranging from a median of 0.2 events per year in the reference period to a median of 2.4 or 4 heatwaves per year by mid-century and late-century, respectively, in the worst-case scenario. This is the most significant change expected in the analysed values. However, these heatwaves are also expected to intensify in terms of average duration, especially by the end of the century. No significant changes are expected regarding average and maximum intensity.

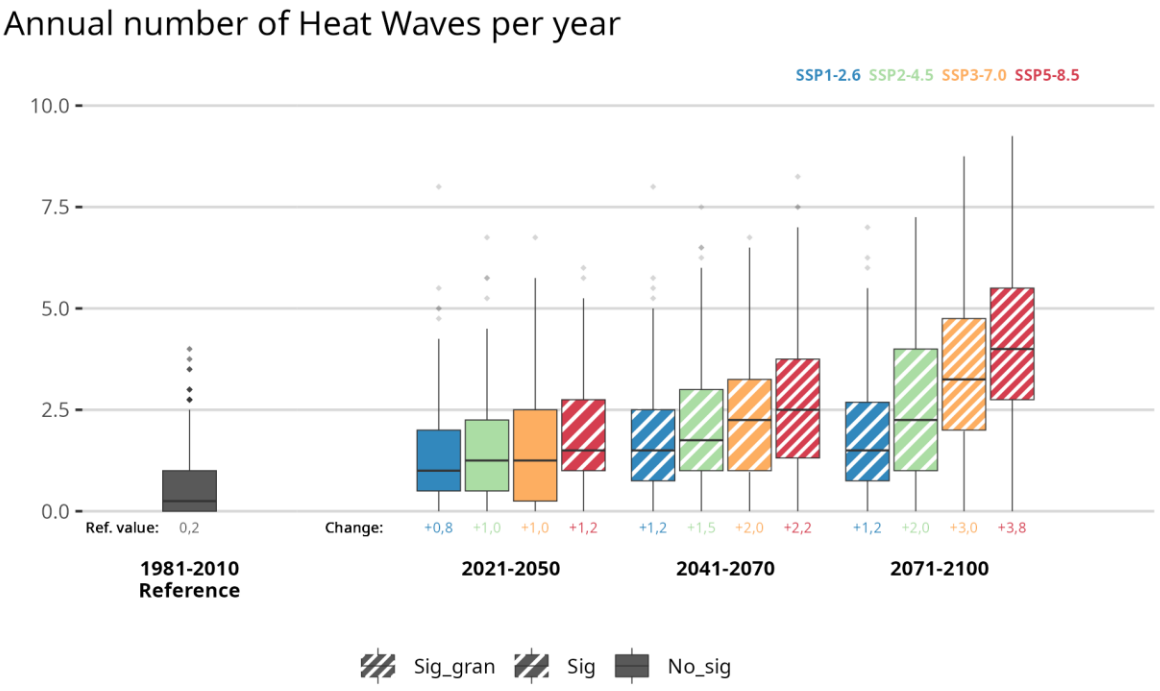


Figure 23. The graph shows the evolution of annual number of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of 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. [Figure A5.44 in Annex 5]

Regarding the SPI results, 3-month, 6-month, 12-month, 24-month and 60-month accumulation periods are studied to assess short-, medium- and long-term conditions (figures A5.48-A5.52). In summary, for all the accumulation periods analysed, it can be said that significant increases are expected in all cases.

- Regarding 3-month SPI (figure A5.48), its decreases are expected to range from 0.6 to 0.9 (moderately wet) by mid-century and from 0.9 to 1.7 (moderately-severely wet) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.
- Regarding 6-month SPI (figure A5.49), its decreases are expected to range from 0.9 to 1.2 (moderately wet) by mid-century and from 1.2 to 2.2 (moderately-severely wet) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.
- Regarding 12-month SPI (figure A5.50), its decreases are expected to range from 1.2 to 1.6 (moderately-severely wet) by mid-century and from 1.7 to 3 (severely-extremely wet) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.
- Regarding 24-month SPI (figure A5.51), its decreases are expected to range from 1.7 to 2.3 (severely-extremely wet) by mid-century and from 2.4 to 4.2 (extremely wet) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of -0.1.
- Regarding 60-month SPI (figure A5.52), its decreases are expected to range from 2.4 to 3.5 (extremely wet) by mid-century and from 3.6 to 5.5 (extremely wet) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.

Regarding the SPEI results, 3-month, 6-month, 12-month, 24-month and 60-month accumulation periods are studied to assess short-, medium- and long-term conditions (as SPI) (figures A5.53-A5.57). In summary, for all the accumulation periods analysed some significant increases are expected in all cases.

- Regarding 3-month SPEI (figure A5.53), its decreases are expected to range from 0.4 (no significant) to 0.6 (normal - moderately wet) by mid-century and from 0.5 to 1.2 (moderately-severely wet) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.
- Regarding 6-month SPEI (figure A5.54), its decreases are expected to range from 0.5 to 0.7 (moderately wet) by mid-century and from 0.7 to 1.4 (moderately wet) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.
- Regarding 12-month SPEI (figure A5.55), its decreases are expected to range from 0.8 to 1.1 (moderately wet) by mid-century and from 1.1 to 1.9 (moderately-severely wet) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.

- Regarding 24-month SPEI (figure A5.56), its decreases are expected to range from 1.1 to 1.4 (moderately wet) by mid-century and from 1.4 to 2.5 (moderately-extremely wet) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.
- Regarding 60-month SPEI (figure A5.57), its decreases are expected to range from 1.7 to 2.1 (severely-extremely wet) by mid-century and from 2.1 to 3.4 (extremely wet) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.

Results for Cluj-Napoca

According to the climate change local scenarios generated, the main conclusions for each variable for Cluj-Napoca are summarised below, together with an illustrative figure. For reference, the remaining figures and extended information are included in Annex 5.

According to the climate change local scenarios generated for Cluj-Napoca no significant changes can be expected in maximum relative humidity across all scenarios and time horizons (figure A5.58). Changes in values are expected to range from -1 to -1.5 per cent both by mid-century and from -0.9 to -3.1 by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 90.4%. Some significant changes are also expected in minimum relative (figure A5.59). Changes in these values are expected to range from -1.9 to -2.2 per cent by mid-century and from -1.6 to -4 per cent by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 59.2%. In the initial results, we observed an important increase in temperature in both maximum and minimum values. As it warms, the atmosphere has a greater capacity to hold moisture. But at the same time, as global surface temperatures rise because of anthropogenic warming, the amount of water evaporated increases and, therefore, the water vapour content in the atmosphere increases. The expected change in relative humidity values is the result of the balance between these interactions.

Regarding the radiation results, long wave radiation and short-wave radiation are studied. In terms of long wave radiation significant increases are expected in all cases as shown in figure A5.60. Changes in values are expected to range from 10.4 to 15 W/m^2 by mid-century and from 13.6 to 26.8 W/m^2 by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 285.6 W/m^2. In terms of short-wave radiation, significant increases are expected at all future periods and scenarios too. As shown in figure A5.61, changes in values are expected to range from 7.3 to 5.1 W/m^2 by mid-century and from 6.4 to 7.1 W/m^2 by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 150.3 W/m^2. For perfect equilibrium, there must be a net balance at the top of the atmosphere, such that the amount of incoming energy from the sun (in the form of shortwave radiation) must be the same as the amount of energy emitted by the Earth (in the form of longwave radiation).

Attending the wind speed results, some significant changes are expected as shown in figure A5.62 but not in all cases. Changes in values are expected to range from -0.2 to -0.1 km/h by mid-century and range around -0.3 km/h by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 6.5 km/h.

Regarding the Heat Wave results, the annual number of Heat Waves episodes and the average duration, mean intensity and maximum intensity of these episodes are studied. In terms of annual number of Heat Waves, significant changes are expected in all cases, as shown in figure X4 (figure A5.63 in Annex 5). Changes in values are expected to range from 2.5 to 3.2 per year by mid-century and from 3.5 to 5.1 per year by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.5 episodes per year. This result is complemented by the characteristics analysed using the annual average for each episode: average duration, mean intensity and maximum intensity, analysed in figures A5.64-A5.66.

- Regarding average duration (figure A5.64), significant increases are expected in almost all cases. Its increases are expected to range from 1.4 to 2.1 days per year by mid-century and from 1.9 to 4.9 per year by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 4.2 days per year.
- Regarding average intensity (figure A5.65), only significant changes are expected in the two worst scenarios by the end of the century. Its increases are expected to range from 0.2 to 0.7 °C by mid-century and from 0.7 to 1.6 °C by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 34.4°C.
- Regarding maximum intensity (figure A5.66), only significant changes are expected in the two worst scenarios by the end of the century. Its increases are expected to range from 0.3 to 0.9 °C by mid-century and from 0.9 to 2.5 °C by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 36°C.

Based on the results, a significant increase in the annual number of heat waves is expected, ranging from a median of 0.5 events per year in the reference period to a median of 3.7 or 5.6 heatwaves per year by mid-century and late-century, respectively, in the worst-case scenario. This is the most significant change expected in the analysed values along with the expected increase in the average duration of these events. However, these heatwaves are also expected to intensify in terms of both average and maximum intensity, but only in a significant way at the two worst scenarios by the end of the century.

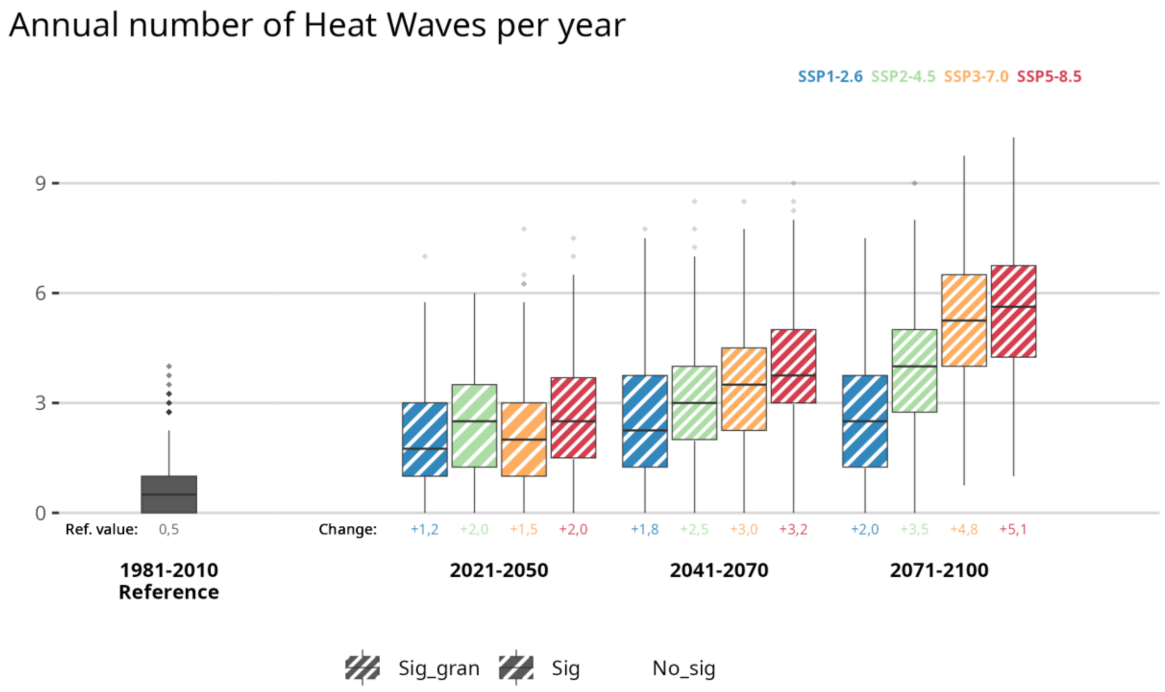


*Figure 24. The graph shows the evolution of annual number of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of 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. [**Figure A5.63**. in Annex 5]*

Regarding the SPI results, 3-month, 6-month, 12-month, 24-month and 60-month accumulation periods are studied to assess short-, medium- and long-term conditions (figures A5.67-A5.71). In summary, for all the accumulation periods analysed, it can be said that significant increases are expected in almost all cases.

- Regarding 3-month SPI (figure A5.67), its decreases are expected to range from 0.4 to 0.6 (normal-moderately wet) by mid-century and from 0.7 to 0.9 (moderately wet) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.
- Regarding 6-month SPI (figure A5.68), its decreases are expected to range from 0.5 to 0.7 (moderately wet) by mid-century and from 0.9 to 1.1 (moderately wet) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.

- Regarding 12-month SPI (figure A5.69), its decreases are expected to range from 0.6 to 0.9 (moderately wet) by mid-century and from 1.1 to 1.4 (moderately wet) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of -0.1.
- Regarding 24-month SPI (figure A5.70), its decreases are expected to range from 0.8 to 1.3 (moderately wet) by mid-century and from 1.6 to 2 (severely-extremely wet) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of -0.1.
- Regarding 60-month SPI (figure A5.71), its decreases are expected to range from 1.3 to 2 (moderately-extremely wet) by mid-century and from 2.4 to 2.6 (extremely wet) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of -0.1.

Regarding the SPEI results, 3-month, 6-month, 12-month, 24-month and 60-month accumulation periods are studied to assess short-, medium- and long-term conditions (as SPI) (figures A5.72-A5.76). In summary, for all the accumulation periods analysed some decreases are expected in all cases but only some are significant.

- Regarding 3-month SPEI (figure A5.72), its decreases are expected to range around -0.3 by mid-century and from -0.1 to -0.6 (significant) (normal-moderately dry) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.
- Regarding 6-month SPEI (figure A5.73), its decreases are expected to range around -0.4 by mid-century and from -0.3 to -0.9 (significant) (normal-moderately dry) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.
- Regarding 12-month SPEI (figure A5.74), its decreases are expected to range around -0.6 (moderately dry) by mid-century and from -0.4 to -1.3 (significant) (normal-moderately dry) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.
- Regarding 24-month SPEI (figure A5.75), its decreases are expected to range around -0.9 by mid-century (both significant) and from -0.5 to -1.8 (significant) (moderately-severely dry) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.
- Regarding 60-month SPEI (figure A5.76), its decreases are expected to range from -1.6 to -1.4 (both significant) (moderately-severely dry) by mid-century and from -1.1 to -2.3 (significant) (moderately-severely dry) by the end of the century in the SSP2-4.5 and SSP5-8.5 respectively with respect to the historical value of 0.

5. Displaying results in the final tool

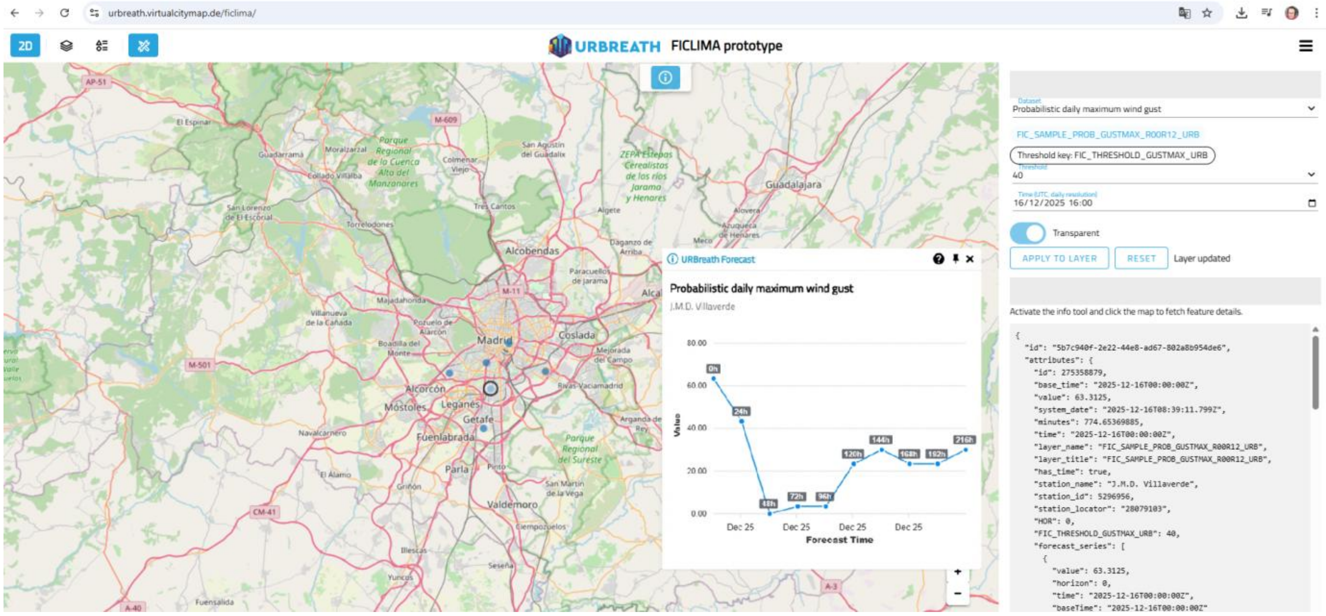
The increasing intensity and frequency of extreme weather events underscore the critical need for effective visualization tools to support urban decision-making and climate adaptation. A robust display mechanism is essential to translate complex model outputs, from the long-term to the immediate future, into actionable insights for city planners and stakeholders.

Effective climate risk management relies on exploiting the synergistic feedback that exists across all timescales, from climate projections and seasonal forecasting to short-term weather forecasting. Local downscaled climate change projections enable cities to identify long-term warning signals and risks to specific urban areas, which are typically represented by co-designed indices. Once these thresholds have been established, they can be translated and adapted to shorter timescales, such as seasonal and weather forecasts. This enables early warning indexes to anticipate impact risks and allow for the deployment of proactive protection measures for the short term. Furthermore, by comparing daily meteorological data with long-term climate simulations, cities can assess which climate change scenario is becoming reality and adjust their strategic planning accordingly. This approach establishes a seamless, bidirectional workflow of data and indices across temporal scales.

To ensure the efficient and reliable delivery of these multi-temporal results to the URBREATH decision-making tools, a dual data management strategy has been employed. The WMS (Web Map Service) protocol is utilized for all dynamic information that requires frequent updates, specifically the daily weather and monthly seasonal forecasts. For climate projections maps and historical benchmark data, a single batch delivery was done due to this is static information, MinIO is the repository and visualizer where this information can be consulted. MinIO is a tool provided by the project partner Engineering. Currently, the partners VCS and Municipia/Engineering are actively working to display final visualizer versions for the outputs that FICLIMA is providing via WMS (seasonal and weather forecast). Cities can access the currently available prototype web visualization here: https://URBREATH.virtualcitymap.de/ficlima/.

5.1 Weather forecast

At weather forecast Frontrunner cities can consult probabilistic forecasts for short term (1-10 days) for these variables: “minimum temperature”, “maximum temperature”, “wind gust” and “accumulated 24h precipitation”, placed at the dataset selector. They can select at the place where they are requiring the forecast from the available points and then select the threshold of interest. Next step is the selection of the starting point for the forecast at the “Time” section, this is displayed for the current day by default. Once the entire selection process is completed, pressing the “apply to layer button” the results are shown on screen clicking on the forecast point. See examples below (Figures 25-29):



Dataset: Probabilistic daily maximum wind gust

FIC_SAMPLE_PROB_GUSTMAX_R00R12_URB

Threshold key: FIC_THRESHOLD_GUSTMAX_URB

Threshold: 40

Time (UTC, daily resolution): 16/12/2025 16:00

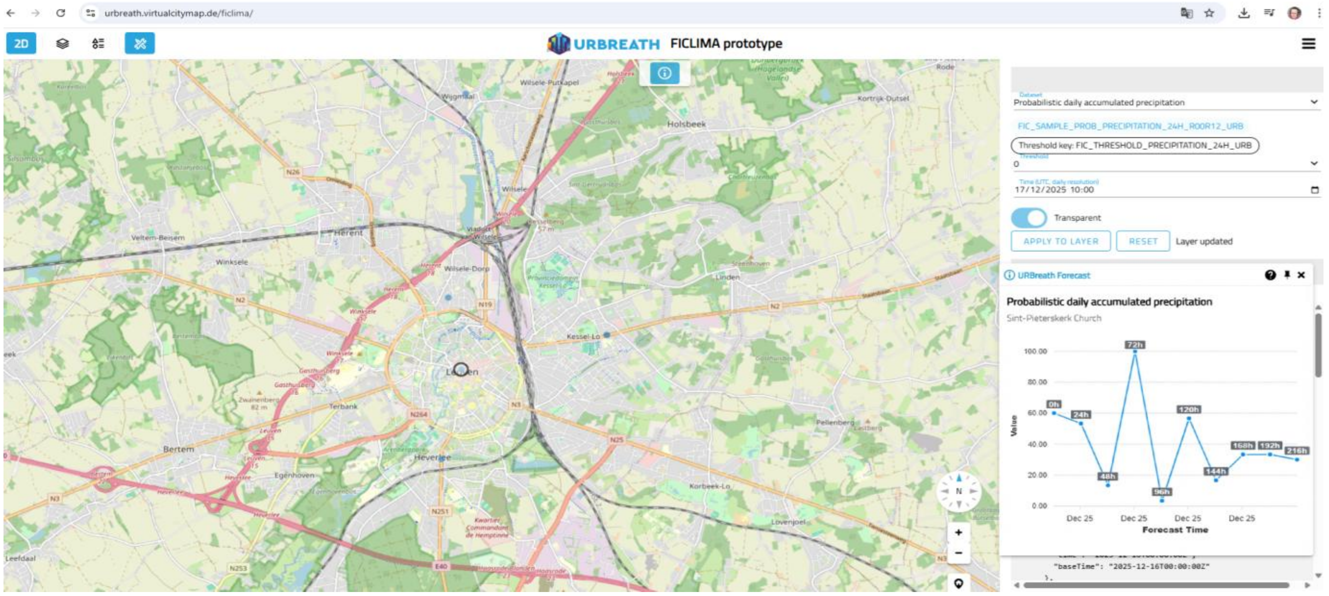
Transparent

Layer updated

Activate the info tool and click the map to fetch feature details.

```
{
  "id": "5b7c940f-2e22-44e8-ad67-802a8b954de6",
  "attributes": {
    "id": 275358879,
    "base_time": "2025-12-16T00:00:00Z",
    "value": 63.3125,
    "system_date": "2025-12-16T08:39:11.799Z",
    "minutes": 774.65369885,
    "time": "2025-12-16T00:00:00Z",
    "layer_name": "FIC_SAMPLE_PROB_GUSTMAX_R00R12_URB",
    "layer_title": "FIC_SAMPLE_PROB_GUSTMAX_R00R12_URB",
    "has_time": true,
    "station_name": "J.M.D. Villaverde",
    "station_id": 5296956,
    "station_locator": "28079103",
    "HOR": 0,
    "FIC_THRESHOLD_GUSTMAX_URB": 40,
    "forecast_series": [
      {
        "value": 63.3125,
        "horizon": 0,
        "time": "2025-12-16T00:00:00Z",
        "baseTime": "2025-12-16T00:00:00Z"
```

Figure 25. Screenshot of the FICLIMA prototype webpage for weather forecast by VCS. Are displayed results for Madrid city for maximum wind gust probability of overpassing 40kmk/h.



Dataset: Probabilistic daily accumulated precipitation

FIC_SAMPLE_PROB_PRECIPITATION_24H_R00R12_URB

Threshold key: FIC_THRESHOLD_PRECIPITATION_24H_URB

Threshold: 0

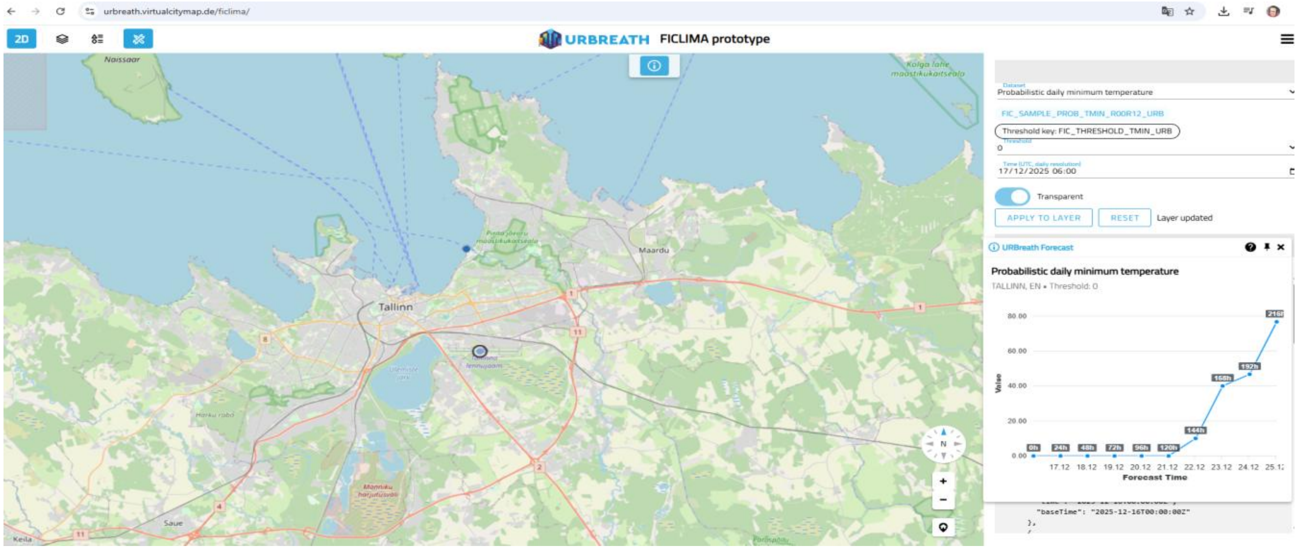
Time (UTC, daily resolution): 17/12/2025 10:00

Transparent

Layer updated

```
    "baseTime": "2025-12-16T00:00:00Z"
  },
```

Figure 26. Screenshot of the FICLIMA prototype webpage for weather forecast by VCS. Are displayed results for Leuven city for probabilistic daily accumulated precipitation at the threshold of appreciable precipitation (>0mm).



Dataset: Probabilistic daily minimum temperature

FIC_SAMPLE_PROB_TMIN_R00R12_URB

Threshold key: FIC_THRESHOLD_TMIN_URB

Threshold: 0

Time (UTC, daily resolution): 17/12/2025 06:00

Transparent

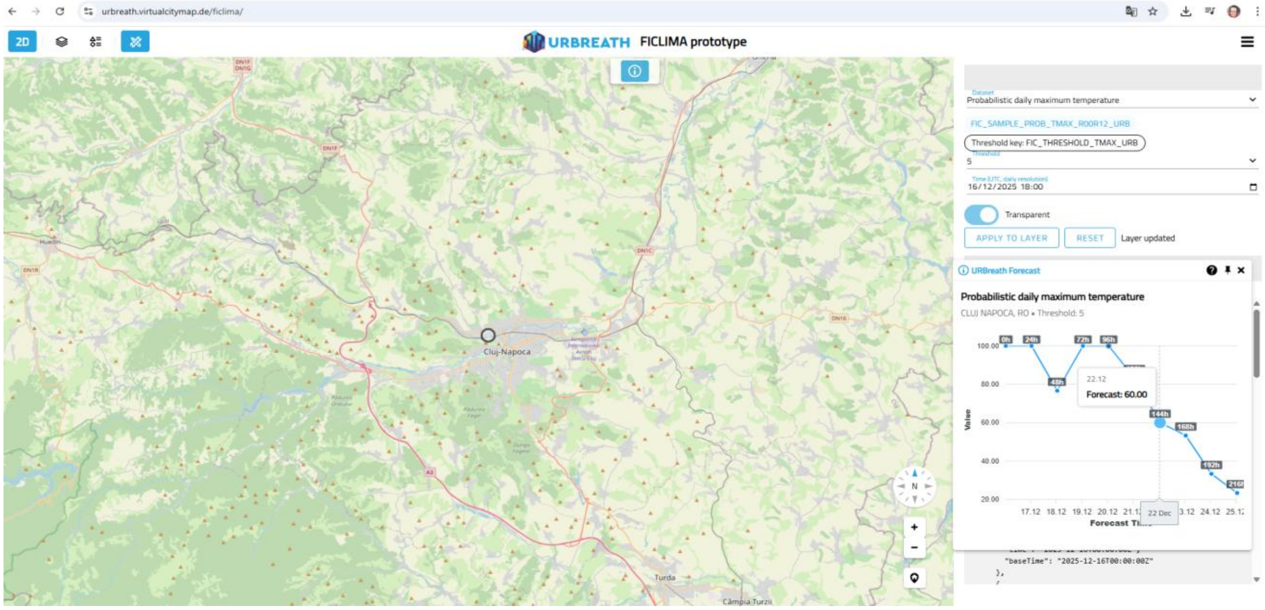
APPLY TO LAYER | RESET | Layer updated

Probabilistic daily minimum temperature

TALLINN, EN • Threshold: 0

```
    "baseTime": "2025-12-16T00:00:00Z"
  },
```

Figure 27. Screenshot of the FICLIMA prototype webpage for weather forecast by VCS. Are displayed results for Tallinn city for probabilistic for minimum temperature probability of overpassing below 0ºC.



Dataset: Probabilistic daily maximum temperature

FIC_SAMPLE_PROB_TMAX_R00R12_URB

Threshold key: FIC_THRESHOLD_TMAX_URB

Threshold: 5

Time (UTC, daily resolution): 16/12/2025 18:00

Transparent

APPLY TO LAYER | RESET | Layer updated

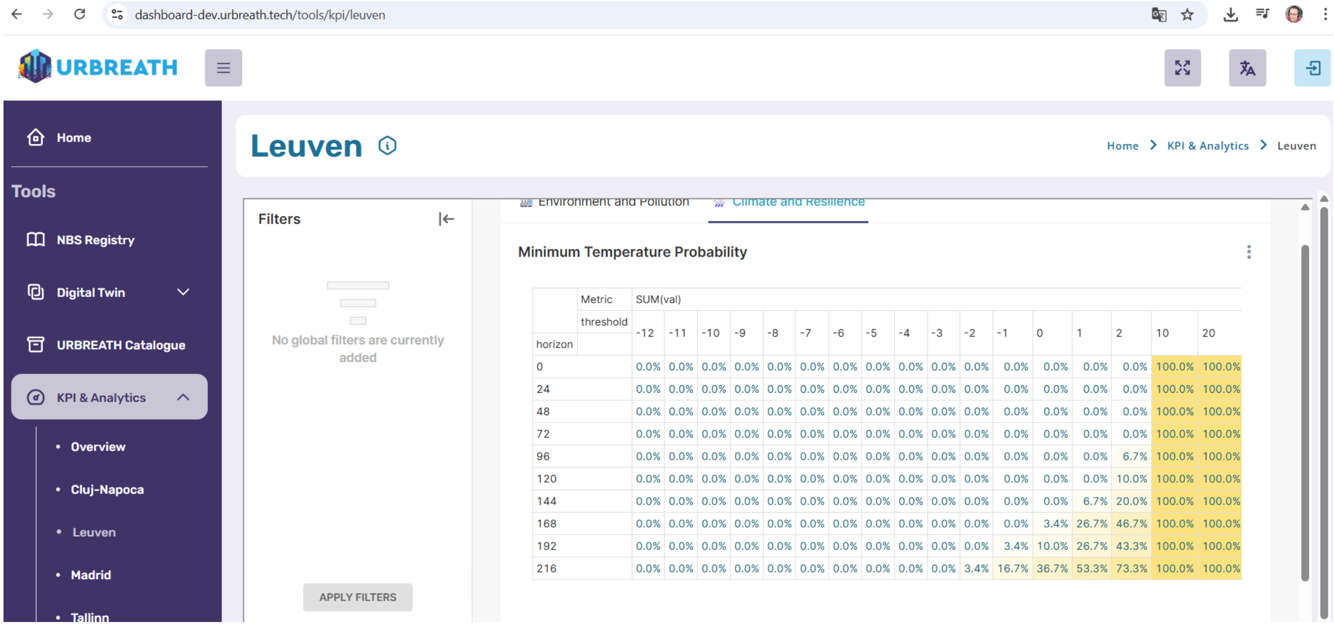
Probabilistic daily maximum temperature

CLUJ NAPOCA, RO • Threshold: 5

```
    "baseTime": "2025-12-16T00:00:00Z"
  },
```

Figure 28. Screenshot of the FICLIMA prototype webpage for weather forecast by VCS. Are displayed results for Cluj-Napoca city for probabilistic for maximum temperature probability of overpassing 5ºC

There is another URBREATH platform, prepared by Municipia, that is currently ingesting FICLIMA wms data in order to display weather and seasonal forecast outputs into its different tools, here is the link: https://dashboard-dev.URBREATH.tech/tools/kpi/leuven. The IT teams are working just in the moment of preparing this deliverable in order to optimize the user experience visualization, consequently the available sections and their dashboards are in a draft stage. Here we share an example as screenshot (Figure 29).



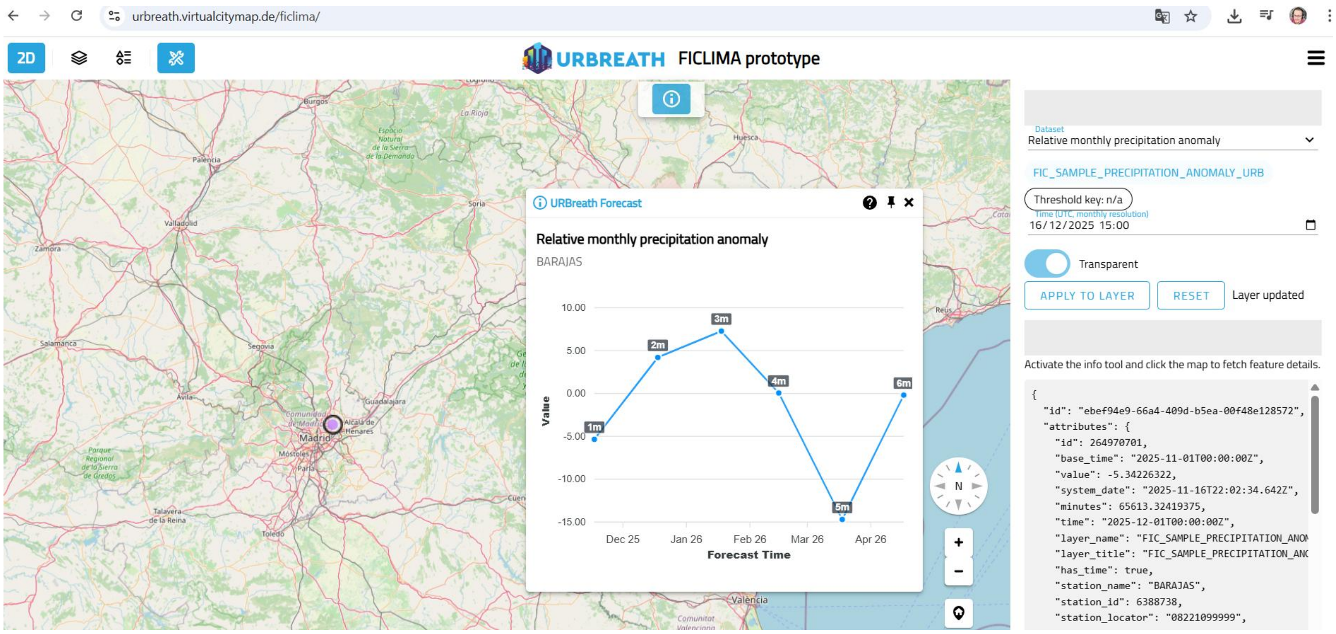
Minimum Temperature Probability

Metric	SUM(val)																
threshold	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	10	20
horizon																	
0	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	100.0%
24	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	100.0%
48	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	100.0%
72	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	100.0%
96	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	6.7%	100.0%	100.0%
120	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	10.0%	100.0%	100.0%
144	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	6.7%	20.0%	100.0%	100.0%
168	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	3.4%	26.7%	46.7%	100.0%	100.0%
192	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	3.4%	10.0%	26.7%	43.3%	100.0%	100.0%
216	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	3.4%	16.7%	36.7%	53.3%	73.3%	100.0%	100.0%

Figure 29. Screenshot of the FICLIMA prototype webpage for weather forecast by Municipia. In the picture can see displayed results for Leuven city for probabilistic for minimum temperature probability of overpassing every available threshold.

5.2. Seasonal forecast

For seasonal forecast, Frontrunner cities can consult anomalies forecasts for medium term (1-6 months) for these variables: “mean temperature”, “mean win” and “monthly precipitation”, placed at the dataset selector. The unique point of forecast displayed for each city is representative of the entire city, next step is the selection of the starting point for the forecast at the “Time” section, this is displayed for the current month by default. Once the entire selection process is completed, pressing the “apply to layer button” the results are shown on screen clicking on the forecast point. See examples below (Figures 30-32):



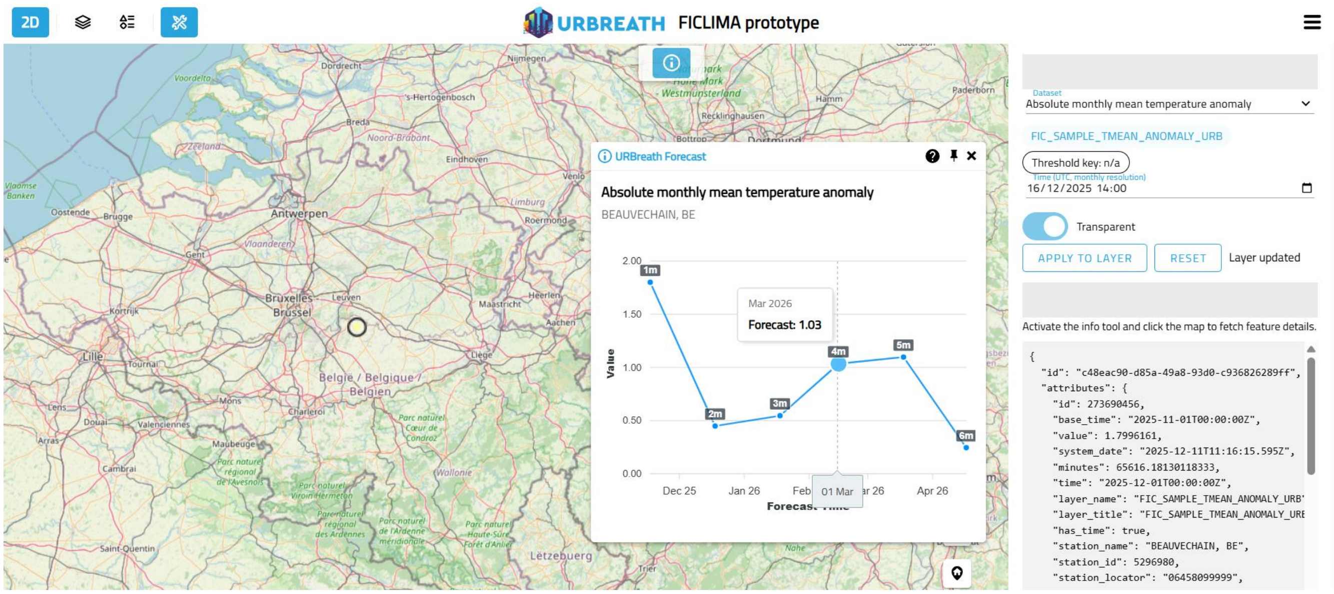
Relative monthly precipitation anomaly
BARAJAS

Dataset: Relative monthly precipitation anomaly
FIC_SAMPLE_PRECIPITATION_ANOMALY_URB
Threshold key: n/a
Time (UTC, monthly resolution): 16/12/2025 15:00
Transparent
Layer updated

Activate the info tool and click the map to fetch feature details.

```
{
  "id": "ebef94e9-66a4-409d-b5ea-00f48e128572",
  "attributes": {
    "id": 264970701,
    "base_time": "2025-11-01T00:00:00Z",
    "value": -5.34226322,
    "system_date": "2025-11-16T22:02:34.642Z",
    "minutes": 65613.32419375,
    "time": "2025-12-01T00:00:00Z",
    "layer_name": "FIC_SAMPLE_PRECIPITATION_ANOM
    "layer_title": "FIC_SAMPLE_PRECIPITATION_ANC
    "has_time": true,
    "station_name": "BARAJAS",
    "station_id": 6388738,
    "station_locator": "08221099999",
```

Figure 30. Screenshot of the FICLIMA prototype webpage for seasonal forecast by VCS. Are displayed results for Madrid city for relative monthly precipitation anomaly.



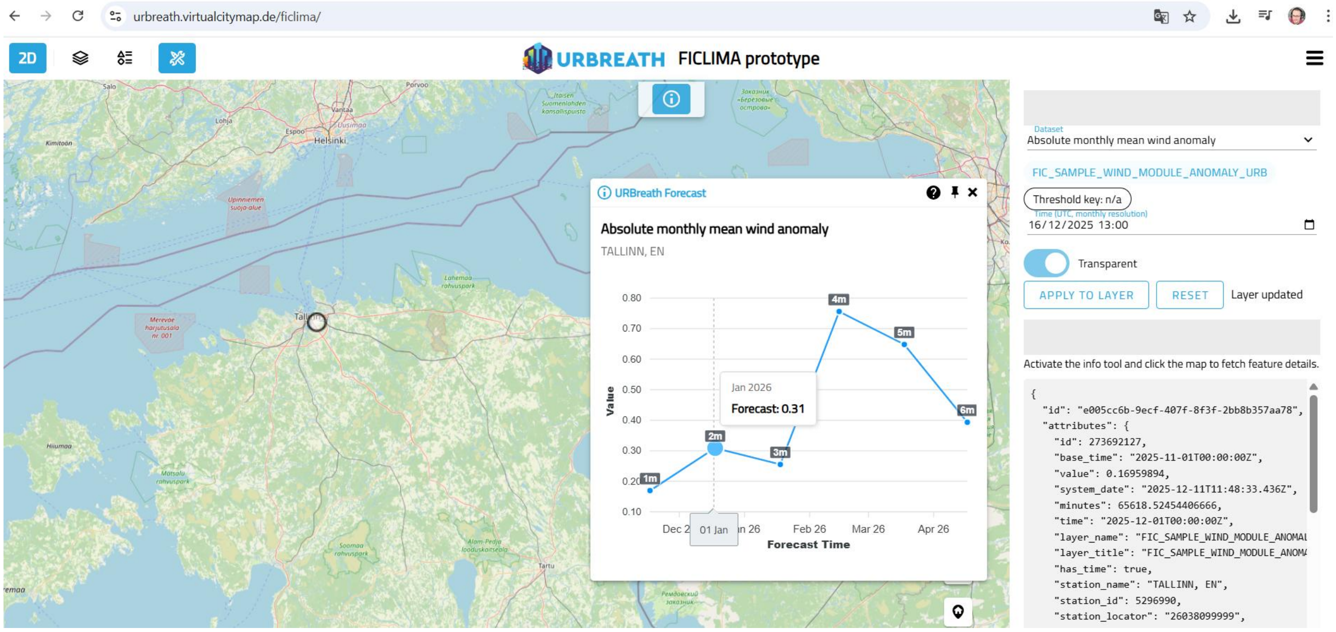
Absolute monthly mean temperature anomaly
BEAUVECHAIN, BE

Dataset: Absolute monthly mean temperature anomaly
FIC_SAMPLE_TMEAN_ANOMALY_URB
Threshold key: n/a
Time (UTC, monthly resolution): 16/12/2025 14:00
Transparent
Layer updated

Activate the info tool and click the map to fetch feature details.

```
{
  "id": "c48eac90-d85a-49a8-93d0-c936826289ff",
  "attributes": {
    "id": 273690456,
    "base_time": "2025-11-01T00:00:00Z",
    "value": 1.7996161,
    "system_date": "2025-12-11T11:16:15.595Z",
    "minutes": 65616.18130118333,
    "time": "2025-12-01T00:00:00Z",
    "layer_name": "FIC_SAMPLE_TMEAN_ANOMALY_URB'
    "layer_title": "FIC_SAMPLE_TMEAN_ANOMALY_URE
    "has_time": true,
    "station_name": "BEAUVECHAIN, BE",
    "station_id": 5296980,
    "station_locator": "06458099999",
```

Figure 31. Screenshot of the FICLIMA prototype webpage for seasonal forecast by VCS. Are displayed results for Madrid city for relative monthly precipitation anomaly.



Dataset: Absolute monthly mean wind anomaly

FIC_SAMPLE_WIND_MODULE_ANOMALY_URB

Threshold key: n/a

Time (UTC, monthly resolution): 16/12/2025 13:00

Layer updated

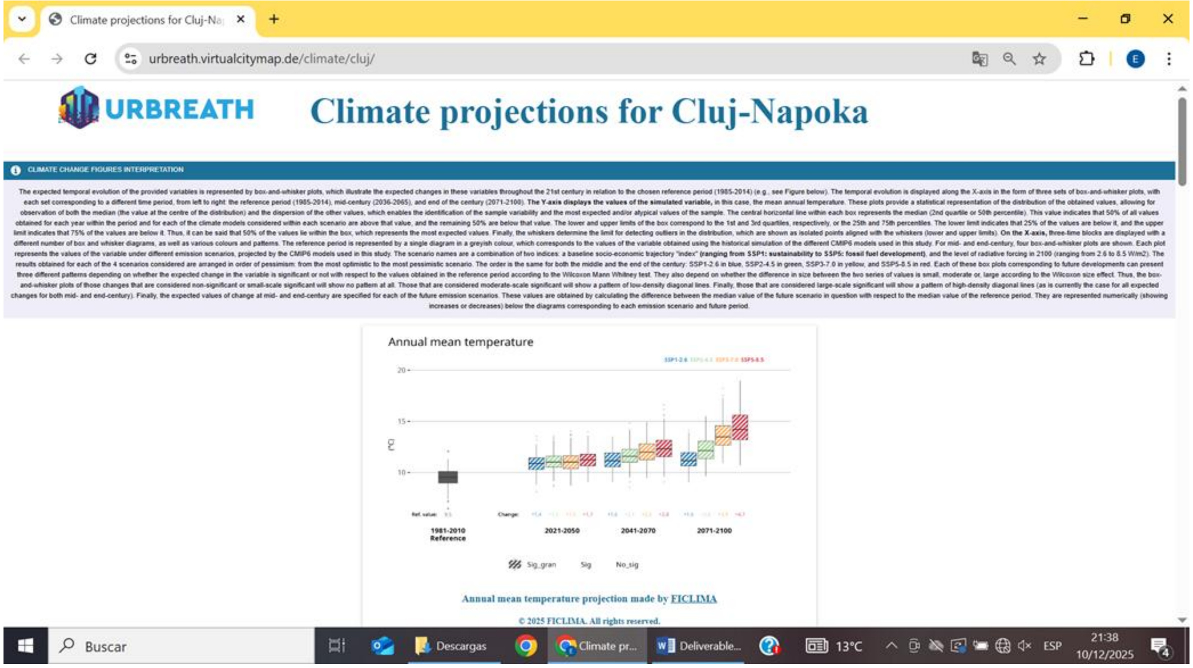
Activate the info tool and click the map to fetch feature details.

```
{
  "id": "e005cc6b-9ecf-407f-8f3f-2bb8b357aa78",
  "attributes": {
    "id": 273692127,
    "base_time": "2025-11-01T00:00:00Z",
    "value": 0.16959894,
    "system_date": "2025-12-11T11:48:33.436Z",
    "minutes": 65618.52454406666,
    "time": "2025-12-01T00:00:00Z",
    "layer_name": "FIC_SAMPLE_WIND_MODULE_ANOMAL
    "layer_title": "FIC_SAMPLE_WIND_MODULE_ANOMA
    "has_time": true,
    "station_name": "TALLINN, EN",
    "station_id": 5296990,
    "station_locator": "26038099999",
```

Figure 32. Screenshot of the FICLIMA prototype webpage for seasonal forecast by VCS. Are displayed results for Madrid city for relative monthly precipitation anomaly.

5.3. Climate projections

Within the project platform "Virtual city map", it is possible to view future climate scenarios for each of the simulated variables in graphical form, as well as a description of how to interpret the attached information. In this section is shown as example the FRC city of Cluj-Napoca, additionally it is shared here the link to the visualization tool: https://dashboard-dev.URBREATH.tech/tools/kpi/cluj-napoca. See examples below (Figures 33-34):

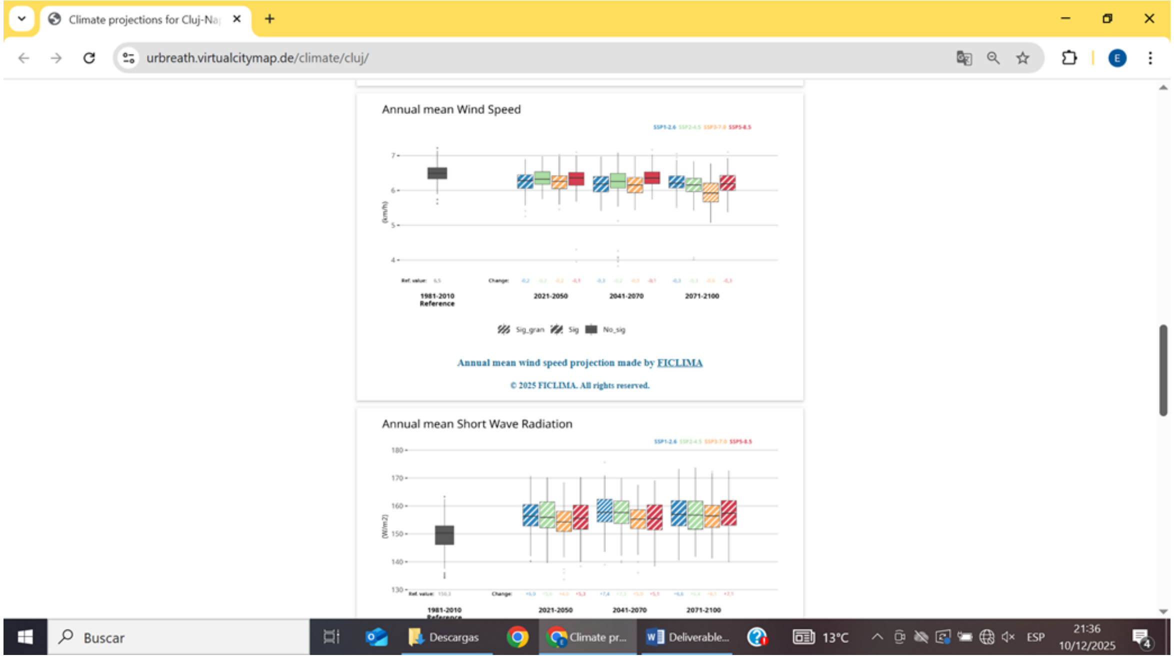


The expected temporal evolution of the provided variables is represented by box-and-whisker plots, which illustrate the expected changes in these variables throughout the 21st century in relation to the chosen reference period (1985-2014) (e.g., see Figure below). The temporal evolution is displayed along the X-axis in the form of three sets of box-and-whisker plots, with each set corresponding to a different time period, from left to right: the reference period (1985-2014), mid-century (2036-2065), and end of the century (2071-2100). **The Y-axis displays the values of the simulated variable,** in this case, the mean annual temperature. These plots provide a statistical representation of the distribution of the obtained values, allowing for observation of both the median (the value at the centre of the distribution) and the dispersion of the other values, which enables the identification of the sample variability and the most expected and/or atypical values of the sample. The central horizontal line within each box represents the median (2nd quartile or 50th percentile). This value indicates that 50% of all values obtained for each year within the period and for each of the climate models considered within each scenario are above that value, and the remaining 50% are below that value. The lower and upper limits of the box correspond to the 1st and 3rd quartiles, respectively, or the 25th and 75th percentiles. The lower limit indicates that 25% of the values are below it, and the upper limit indicates that 75% of the values are below it. Thus, it can be said that 50% of the values lie within the box, which represents the most expected values. Finally, the whiskers determine the limit for detecting outliers in the distribution, which are shown as isolated points aligned with the whiskers (lower and upper limits). **On the X-axis,** three-time blocks are displayed with a different number of box and whisker diagrams, as well as various colours and patterns. The reference period is represented by a single diagram in a greyish colour, which corresponds to the values of the variable obtained using the historical simulation of the different CMIP6 models used in this study. For mid- and end-century, four box-and-whisker plots are shown. Each plot represents the values of the variable under different emission scenarios, projected by the CMIP6 models used in this study. The scenario names are a combination of two indices: a baseline socio-economic trajectory "index" **(ranging from SSP1: sustainability to SSP5: fossil fuel development)**, and the level of radiative forcing in 2100 (ranging from 2.6 to 8.5 W/m2). The results obtained for each of the 4 scenarios considered are arranged in order of pessimism: from the most optimistic to the most pessimistic scenario. The order is the same for both the middle and the end of the century: SSP1-2.6 in blue, SSP2-4.5 in green, SSP3-7.0 in yellow, and SSP5-8.5 in red. Each of these box plots corresponding to future developments can present three different patterns depending on whether the expected change in the variable is significant or not with respect to the values obtained in the reference period according to the Wilcoxon Mann Whitney test. They also depend on whether the difference in size between the two series of values is small, moderate or, large according to the Wilcoxon size effect. Thus, the box-and-whisker plots of those changes that are considered non-significant or small-scale significant will show no pattern at all. Those that are considered moderate-scale significant will show a pattern of low-density diagonal lines. Finally, those that are considered large-scale significant will show a pattern of high-density diagonal lines (as is currently the case for all expected changes for both mid- and end-century). Finally, the expected values of change at mid- and end-century are specified for each of the future emission scenarios. These values are obtained by calculating the difference between the median value of the future scenario in question with respect to the median value of the reference period. They are represented numerically (showing increases or decreases) below the diagrams corresponding to each emission scenario and future period.

Annual mean temperature projection made by FICLIMA

© 2025 FICLIMA. All rights reserved.

Figure 33. Example of a visualisation of climate projections for average annual temperature within the URBREATH visualization tool, along with an explanation of how to interpret the results shown in the graph for the city of Cluj-Napoca.



Annual mean wind speed projection made by FICLIMA

© 2025 FICLIMA. All rights reserved.

Figure 34. Example of the visualisation of climate projections for annual mean wind speed and annual mean shortwave radiation in the city of Cluj-Napoca within the URBREATH visualization tool

6. Conclusions

This Deliverable successfully demonstrates the implementation and validation of advanced methodologies that significantly improve upon the current State-of-the-Art (SOTA) for urban climate services. The core contribution is the refining of predictions at the local scale through statistical post-processing, the integration of high-resolution local weather station data, and the systematic application of Artificial Intelligence (AI) models to enhance accuracy, correct systematic biases, and provide the necessary probabilistic uncertainty quantification for informed decision-making in the cities. These robust, local-scale climate and weather assessments reflect the urgent need for adaptation strategies through both short-term forecasts and long-term projections, in light of the profound implications of climate change as highlighted by the latest IPCC report. The current findings provide a robust operational Framework and starting point to mitigate the impacts of climate change for each FRC city in the coming decades.

This report's evaluation of ensemble weather prediction models has provided valuable information on the systematic biases and predictive capabilities of different models for urban-scale forecasting. As described in section 2.3.1, a historical comparison of two ensemble models (GEFS and ECMWF) has been successfully carried out for the desired period (2023–2024). While both models exhibit comparable limitations, particularly in scenarios involving convection, thermal inversions, or local wind phenomena, the ECMWF consistently demonstrated superior predictive capabilities. This is evident from the fact that the ECMWF's mean absolute errors for temperature are between 0.3°C and 0.5°C lower than those of the GEFS, and it also shows a higher correlation between observations and predictions. Regarding precipitation, both models demonstrate significant predictive capacity, as indicated by the AUC parameter. In all study areas except Cluj-Napoca, the AUC value ranges from 0.7 to 0.95 (with 1 representing an optimal prediction). Analysing the ROC curves shows that, in all study areas, the ECMWF improves on the GEFS prediction for all prediction horizons studied. The ECMWF also presents better RPS values, which analyse a given model's accuracy in terms of precipitation amount intervals. According to this analysis FICLIMA AI based model is based on ECMWF model.

In addition to the comparative evaluation of models, this report assessed the potential for correcting systematic bias in ensemble forecasts using an AI model based on the random forest algorithm. This model was trained using the full 2023–2024 ECMWF dataset and validated with data from 2025. Applying this AI correction model improved the accuracy of temperature forecasts, reducing the error by half in the first prediction horizons in areas where the model had the greatest systematic error. This model provides a foundation for future reports, in which the auto-training of the model will be refined to further reduce prediction error. These improvements aim to optimise probabilistic forecasts on an urban scale, providing more reliable predictions to support climate resilience efforts in all project's cities.

Regarding seasonal forecasts, the benchmarking confirms that the CNN_Ens, CNN_TeWa and TeWa_Lin configurations form the best-performing group, systematically outperforming the baseline models for temperature and precipitation, especially in terms of Pearson correlation and reduced Scaled Mean Absolute Error (SMAE). The integration of Teleconnection and Wavelet–ARIMA (self-predictability) within the TeWA Framework adds significant predictive value, and these statistical–dynamical post-processing models, fine-tuned by FIClima, deliver skills that are broadly comparable to, and in many cases slightly better than, the reference SEAS5 system, particularly by retaining positive correlations at longer seasonal lead times.

Overall, the climate study concludes that a meaningful climate risk assessment must be performed at local scale, translating global climate information into city-relevant variables and indicators through a robust, fit-for-purpose downscaling Framework. It combines a statistical downscaling method for temperature and precipitation (FICLIMA) with a parametric quantile–quantile mapping method for wind, relative humidity and radiation, anchored to reanalysis datasets (ERA5/ERA5-Land), and strengthens credibility through explicit verification and validation steps that test the ability to reproduce past conditions, mean climate and variability before producing future scenarios. Uncertainty is handled transparently via an ensemble of models and multiple SSP pathways, enabling results to be communicated as ranges rather than a single deterministic future.

It is necessary to emphasize operational relevance by moving beyond mean changes to analyze high-impact extremes, particularly heatwaves and meteorological drought, providing decision-ready metrics that better reflect real-world vulnerability and planning needs. Annex 5 reinforces this decision focus by compiling, for each FRC city and time horizon, ensemble-based evidence on key variables and indicators (humidity, radiation, wind, heatwave metrics, and SPI/SPEI at multiple accumulation scales), including inter-model spread, changes relative to the historical baseline and statistical significance, which supports prioritizing adaptation actions across short-, mid- and long-term horizons.

The successful validation of the methodologies across all time scales—from operational weather forecasting to long-term climate projections—provides a crucial foundation for more precise and impactful planning. Adjusting both weather prediction methodologies and climate projection techniques allows for a coherent approach, helping city stakeholders manage immediate risks, such as short-term heavy rain, while also guiding long-term decisions on crucial matters such as Nature based Solutions development and green space planning.

Looking ahead, the achievements validated in this Deliverable serve as the robust operational base for the project. This proven Framework will now be scaled up to encompass all Follower Cities, ensuring that all project participants benefit from the enhanced predictive developments. Furthermore, a fundamental next step is to re-direct the focus on the co-creation process with interested cities to develop ad-hoc indicators tailored to their specific and fine-tuned local needs, which are planned to be finalized by December 2026. To initiate this engagement and train cities in the use of the new tools and FICLIMA results, a dedicated webinar event is planned for January 20. This event will also fully explain the co-creation process, inviting cities that have not yet shown interest to participate actively in the development of these customized, decision-ready indicators.

7. References

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

Benestad, R.E. Downscaling precipitation extremes. Theor Appl Climatol 100, 1–21 (2010). https://doi.org/10.1007/s00704-009-0158-1

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.

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.

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

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.

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.

Monjo R., Caselles V., Chust G. (2014). Probabilistic correction of RCM precipitation in the Basque Country (Northern Spain). Theoretical and Applied Climatology, 117: 317-329. doi: 10.1007/s00704-013-1008-8.

Monjo, R.; Gaitán, E.; Pórtoles, J.; Ribalaygua, J.; Torres, L. (2016): Changes in extreme precipitation over Spain using statistical downscaling of CMIP5 projections. International Journal of Climatology, 36: 757-769. DOI: 10.1002/joc.4380.

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.

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

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.

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.

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. Verification metrics for weather and seasonal forecast

Weather forecast

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.

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

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).

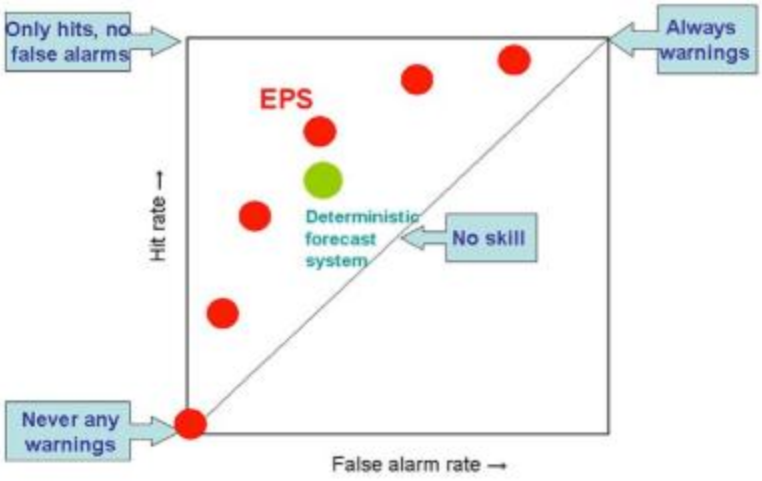


***Figure A1.1** Example ROC diagram: EPS vs. deterministic forecast. Source: ECMWF*

- 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.

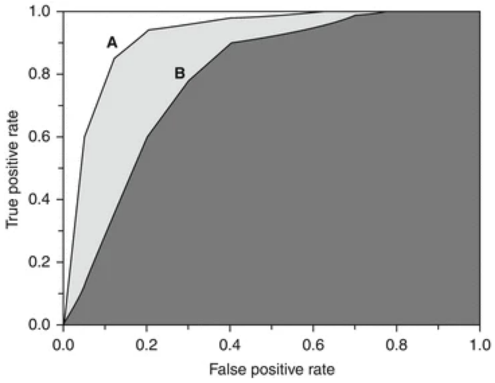
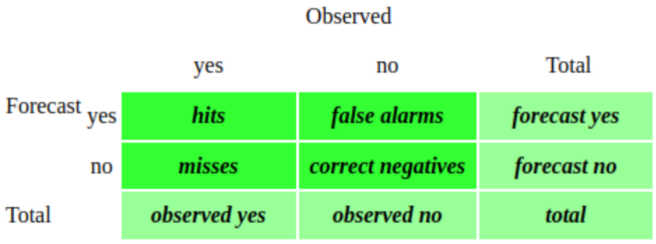


Figure A1.2 *Area under the ROC curves for two score classifiers, A and B. Source: Springer Nature website*

- **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 A1.3 *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:

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

- Root Mean Square Error (RMSE): measures the average difference between a statistical model's predicted values and the observed values. Mathematically, it is the standard deviation of the residuals. Residuals represent the distance between the regression line and the data points. It takes into account both random and systematic errors, making it useful for comparing models and monitoring model performance over time

$$RMSE \; = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \blacksquare (F_i - O_i)^2}$$

- Correlation: In mid-latitudes, this correlation measures the linear relationship between predictions and observations without taking possible bias (systematic error) into account. It is useful for determining the predictability of temperature increases and decreases (variability), as it is a practically Gaussian variable with linear changes.

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 A1.1*: Compilation of existing raw EPS models*

EPS Model	Organization	Temporal Resolution	Forecast Range	Ensemble Members	Finally available
ECMWF Ensemble Prediction System (ENS)	European Centre for Medium-Range Weather Forecasts (ECMWF)	6 hours	15 days	51	Yes
Global Ensemble Forecast System (GEFS)	NOAA/NCEP (USA)	6 hours	16 days	31	Yes
Met Office Global and UK Ensemble (MOGREPS-G & MOGREPS-UK)	UK Met Office	6 hours	7 days	18	No
Canadian Ensemble Prediction System (CEPS)	Environment Canada (CMC)	6 hours	16 days	20	No
Météo-FRCance ARPEGE Ensemble Prediction System	Météo-FRCance	6 hours	4 days	35	No
ICON Ensemble Prediction System (ICON-EPS)	Deutscher Wetterdienst (DWD, Germany)	3 hours	7 days	40	Yes

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 inFRCastructure 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: $\sigma = \sqrt{\frac{\sum \quad (F_i - F)^2}{10}}$

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

Seasonal Forecast

For each variable (temperature, precipitation and mean wind speed), model performance is assessed using a common set of complementary validation metrics. Pearson correlation is used to evaluate whether forecasts reproduce the temporal evolution and phase of observed anomalies, independently of any systematic bias. The Standardised Mean Absolute Error (SMAE) provides a dimensionless measure of the typical error magnitude, allowing us to compare skill across variables, stations and models with very different units and climatological ranges. Mean Absolute Error (MAE), expressed in the physical units of each variable, offers a simple measure of the average absolute error at station scale. Finally, the Root Mean Square Error (RMSE) also in physical units, highlights large errors more strongly and facilitates comparison with previous studies and operational benchmarks.

- 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.

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

Where: h is the lead time (e.g: month 1, 2, … 6), “i” is the index of each forecast-observation, Ph is the model prediction (for lead h and case i), and Oh is the observed value (for lead h and case i). SMAE is unitless, so a value of 1 means the average error is about one climatological standard deviation of the local variability.

- Mean Absolute Error (MAE): Provides a straightforward deterministic measure for prediction errors.

$$MAE_H = \frac{1}{N_h} \sum_{i=1}^{N_h} \left| P_{h,i} - O_{h,i} \right|$$

All the parameters are defined above. MAE provides a simple, deterministic measure of forecast accuracy in the original physical units of each variable.

- As an additional deterministic metric, it has been computed as the Root Mean Square Error (RMSE) to validate SEAS5 and each of the statistical models. For a time series of monthly (or daily) anomalies and observations over time steps, RMSE is defined as

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^{N_h} \left(P_{h,i} - O_{h,i}\right)^2}$$

Where Nh is the total number of forecast-observation pairs available for lead h. The rest of the parameters are defined above. RMSE is expressed in the same units as the variable as MAE does. It penalises larger errors more strongly because of the squared term.

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

Annex 2. Simulation of Heat weaves and Drought episodes

1. Heat weaves episodes

There are many varied terms used to define a heat wave. For example, the AEMET (Spanish Meteorological Agency) defines a heat wave as an "episode of at least three consecutive days, in which at least 10% of the stations considered record maximums above the 95% percentile of their series of maxima daily temperatures of the months of July and August of the period 1971–2000." The WMO (World Meteorological Organization) defines a heat wave as an extreme event with marked warming of the air or the invasion of very warm air over a large area, it usually lasts from a few days to a few weeks." The IPCC defines a heat wave as "a period [of] abnormally and uncomfortably hot weather". Those definitions are not the only ones accepted in scientific literature. Some heat-wave definitions have been used to identify heat waves in a time series of temperature data (Smith et al., 2013), and the choice of the heat wave definition can influence both projected heat-wave trends (Smith et al., 2013) and estimates of health risks during events (Anderson and Bell, 2009; Chen, et al. 2015, Kent, et al. 2014).

According to the recommendations of the WMO (WMO, 2010), a practical and qualitative definition of a heat wave must consider marked and unusual hot weather over a region during at least two consecutive days in the hot period of the year, based on local climate conditions, with thermal conditions recorded above given thresholds. On the basis of these recommendations, it has been defined as follows: at least three consecutive days with a maximum temperature above the 95th percentile of the maximum temperature series and calculated between the months of June to September during the period 1980–2000.

To better identify a heat wave, average duration, maximum intensity and average intensity have been evaluated as recommended by the WMO in the Guidelines on the definition and monitoring of extreme weather and Climate Events. The average duration refers to the average number of days the heat wave lasted. The average intensity represents the average value of the temperature during the heat wave. The maximum intensity is equal to the maximum value that the temperature reached during the heat wave.

2. Meteorological drought episodes

Drought is one of the most impactful extreme climate phenomena, affecting water availability, agriculture, tourism, human health and ecosystems worldwide. Its complexity, the influence of multiple atmospheric drivers and limited historical records make it difficult to characterize. Although precipitation is the primary driver, temperature, wind, humidity and evaporative demand also modulate drought intensity. For this reason, various institutions maintain comprehensive drought monitoring systems, and numerous standardized indices are used to quantify drought severity and enable comparison across regions and timescales. Among these, the Standardized Precipitation Index (SPI) and the Palmer Drought Severity Index (PDSI) are widely applied, yet SPI does not account for temperature-driven increases in atmospheric evaporative demand. To address this limitation, the Standardized Precipitation Evapotranspiration Index (SPEI) was developed, incorporating both precipitation and evaporative demand, thereby providing a more robust indicator under climate-change conditions.

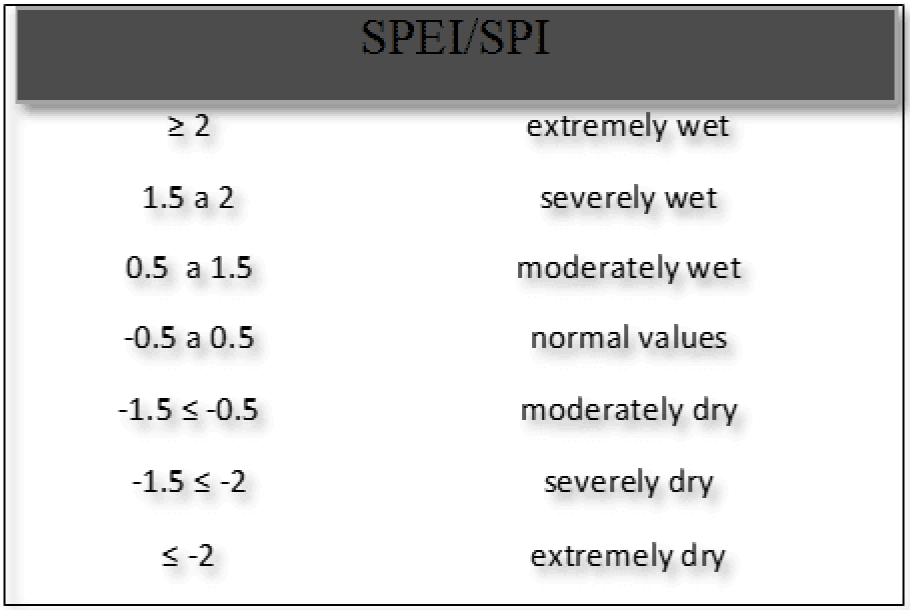
Future projections suggest that drought episodes may intensify in many regions, particularly in Europe and the Mediterranean, which are considered climate-change hotspots. However, drought projections remain uncertain because global climate models simulate temperature trends more reliably than precipitation patterns, especially in regions with high rainfall variability such as the Mediterranean basin. Additional complexities arise from factors, such as soil moisture, runoff behaviour and the physiological effects of rising CO_2 concentrations on plant water use. In areas where climate variability and complex terrain amplify drought risk, combining SPI and SPEI provides a more comprehensive assessment of meteorological drought. This is particularly relevant given that rising temperatures are expected to be one of the most significant climate signals throughout the 21st century.

SPI was developed by McKee et al. (1993) and is based on two assumptions: 1) that the variability of precipitation is greater than that of temperature and ETP, and 2) that the rest of the variables are stationary over time. The SPI value is defined as a numerical value that represents the number of standard deviations of precipitation, over the accumulation period in question, with respect to the average, once the original distribution of precipitation has been transformed into a normal distribution (i.e., mean of zero and standard deviation of 1). The SPI values can be interpreted as the number of standard deviations by which the observed anomaly deviates from the long-term mean.

SPEI, developed by (Vicente-Serrano et al., 2010) and revisited by (Begueria et al., 2014) is a variant of the widespread SPI; it has greater potential as a drought index since it considers the climate balance (through the difference between monthly precipitation and ETP). SPEI values can be interpreted in the same way as SPI values (number of standard deviations by which the observed anomaly deviates from the long-term mean).

The period 1976-2005 was used as a reference period, which represents the last 30 years of the Historical period. Based on this reference period, both the SPI and the SPEI were calculated for the period 2006-2100. The choice of the reference period was made to evaluate the future hydroclimatic conditions of the region with respect to the average conditions of the last 30 years of the Historical experiment.

Both indices were calculated using the R package 'SPEI' (Version 1.7). The SPI was calculated using Gamma distribution to fit the original precipitation series (WMO, 2012) and the SPEI was calculated using log-logistic distribution (Vicente-Serrano et al., 2015; Vicente-Serrano and Beguería, 2016). The parameters of these distributions were obtained by the method of unbiased probabilistic weighted moments (Vicente-Serrano and Beguería, 2016). The scale of SPI and SPEI values used in the study can be seen in Figure A2.1.



SPEI/SPI	
≥ 2	extremely wet
1.5 a 2	severely wet
0.5 a 1.5	moderately wet
-0.5 a 0.5	normal values
-1.5 ≤ -0.5	moderately dry
-1.5 ≤ -2	severely dry
≤ -2	extremely dry

Figure A2.1. Intensity scale for the SPeI/SPI indicators

The way in which the indexes have been analysed follows the guidelines of the WMO (WMO, 2017) which recommends analysis of a drought episode from three main aspects – magnitude (index values), duration (alternation between positive and negative values) and spatial extent – and all these aspects configure the severity of the episode.

Annex 3. Weather forecast results

LEUVEN

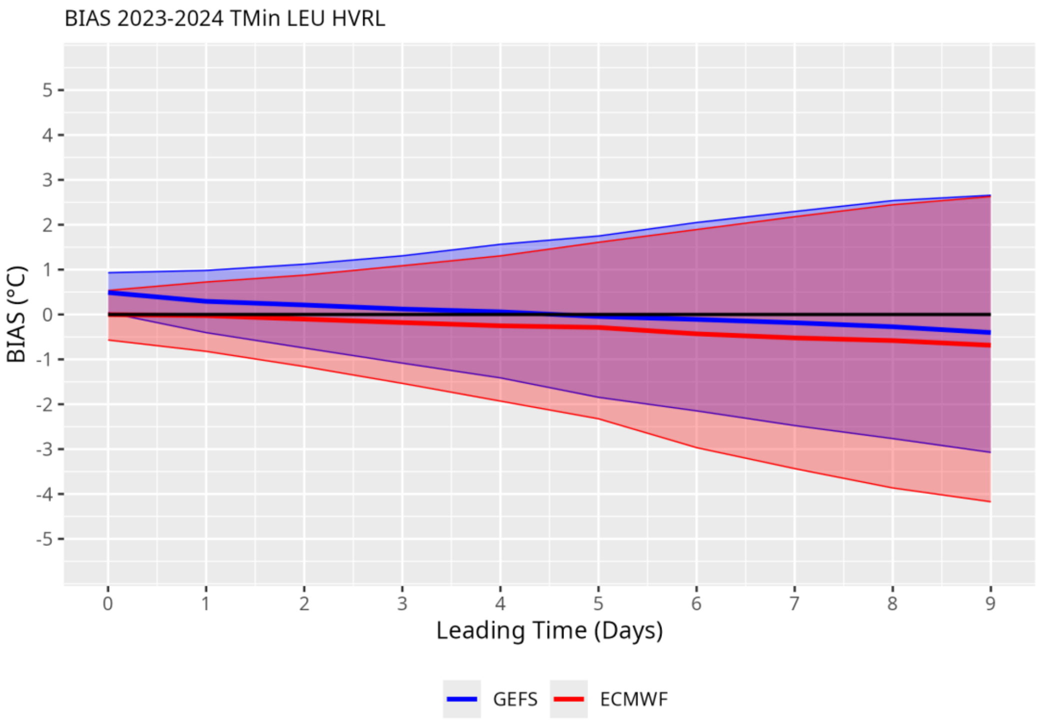


Figure A3.1. *BIAS for minimum temperature in Leuven.*

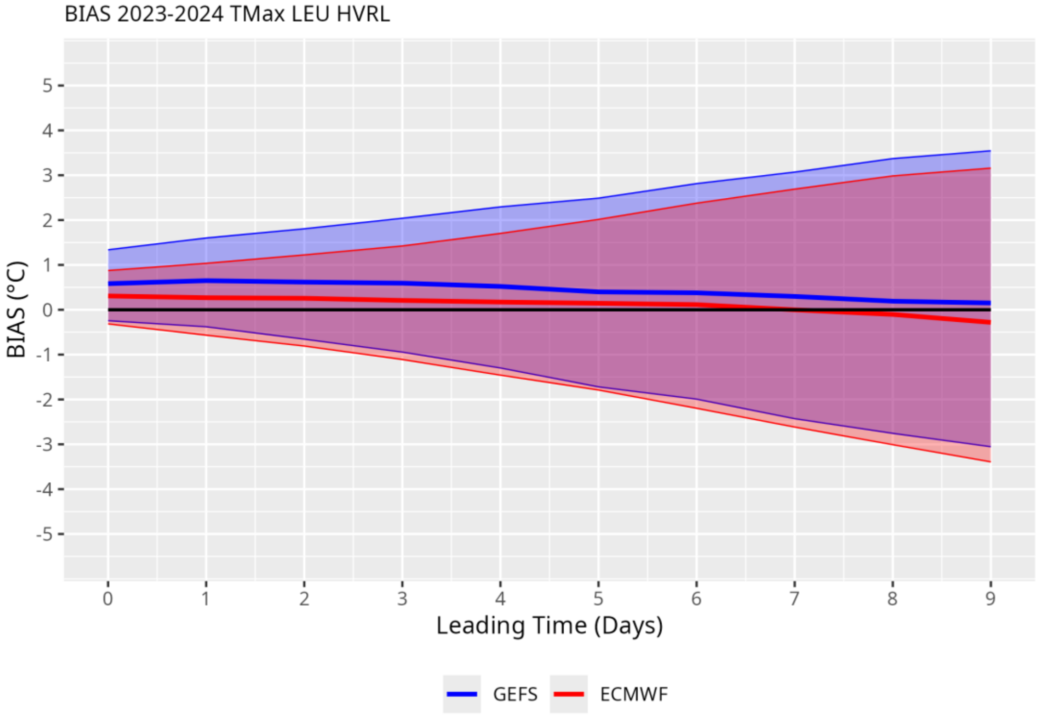


Figure A3.2. *BIAS for maximum temperature in Leuven.*

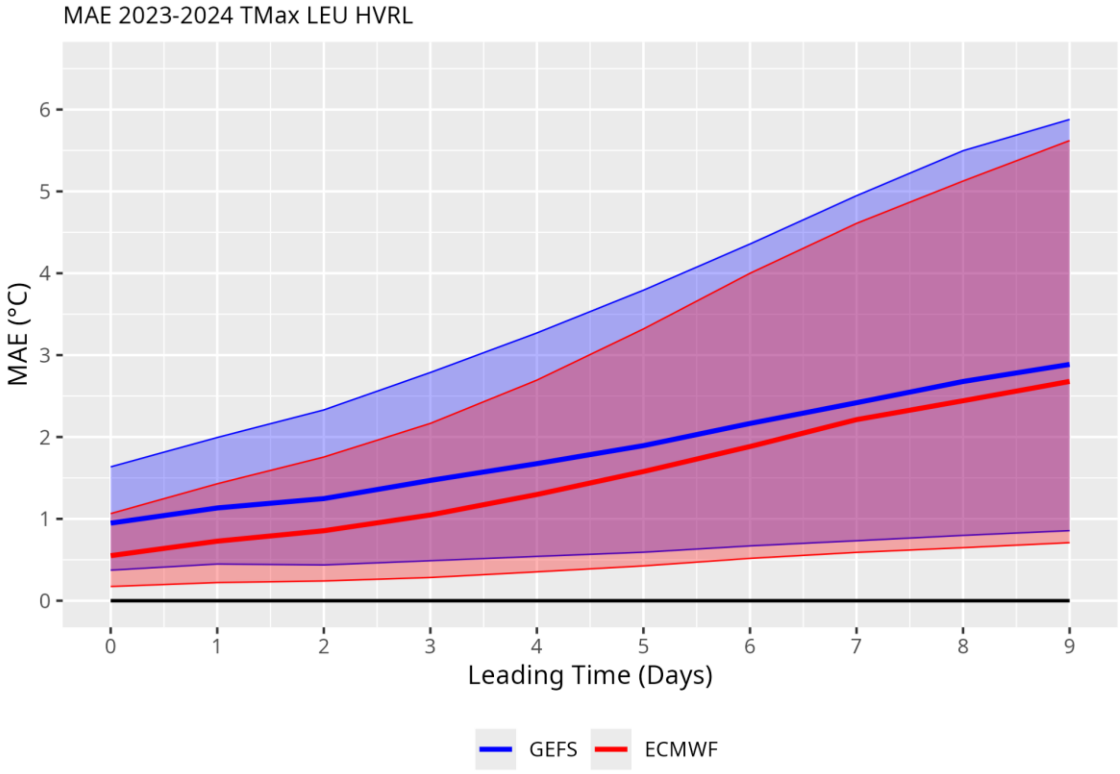


Figure A3.3. *MAE for maximum temperature in Leuven.*

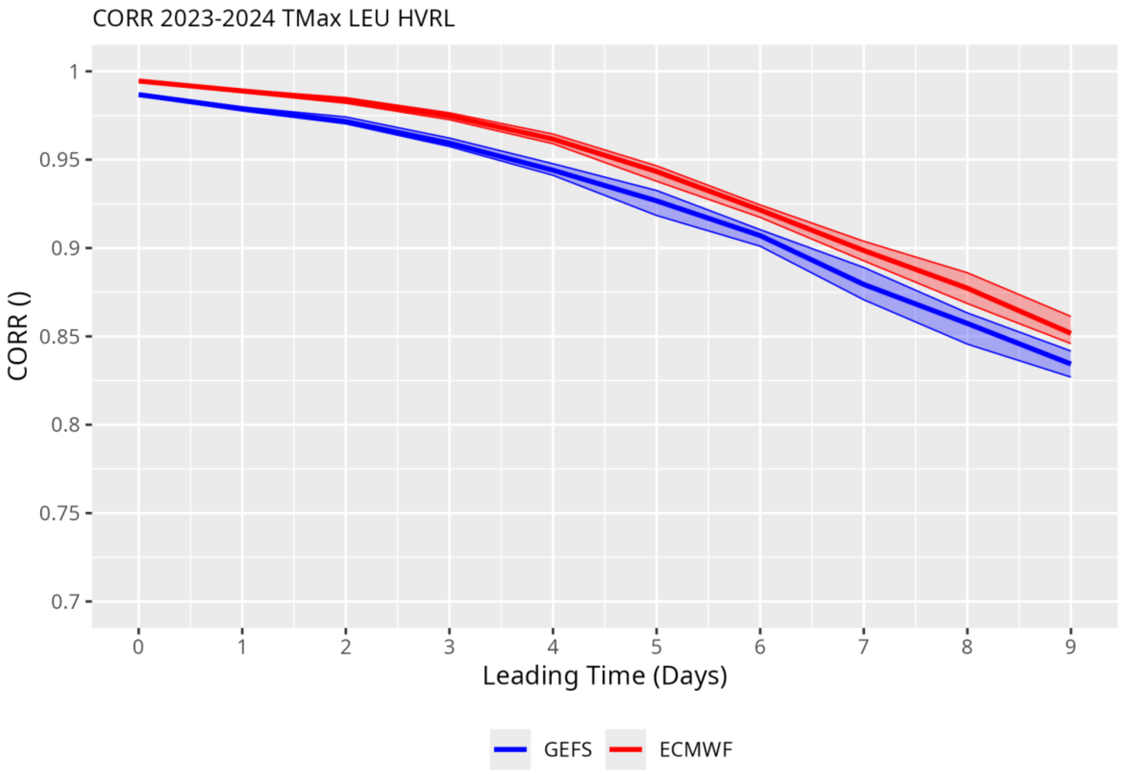


Figure A3.4. *Correlation between observations and predictions for maximum temperature in Leuven.*

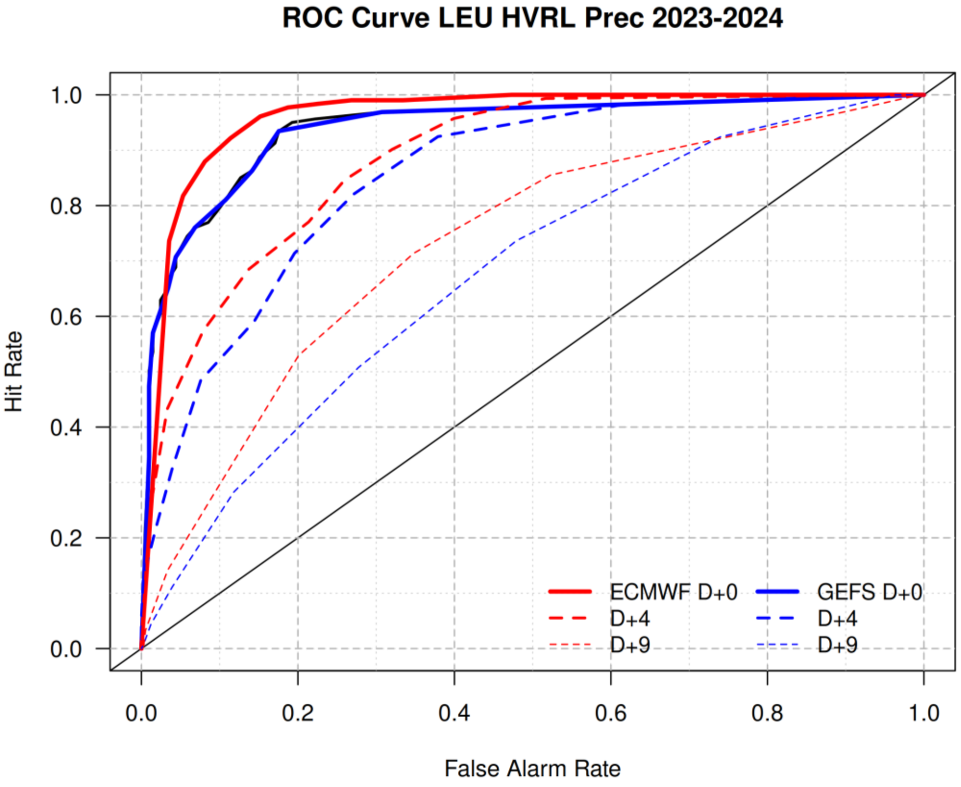


Figure A3.5. *ROC curve for precipitation forecast in Leuven.*

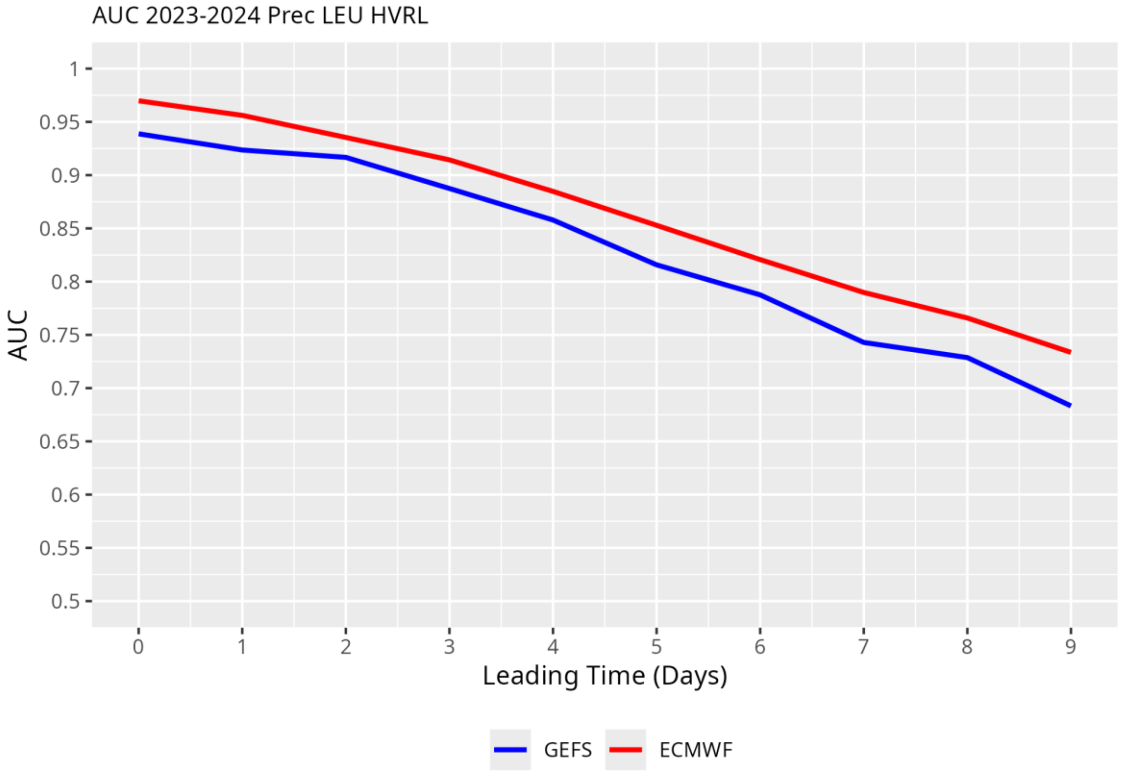


Figure A3.6. *Area under the ROC curve (AUC) for precipitation forecast in Madrid.*

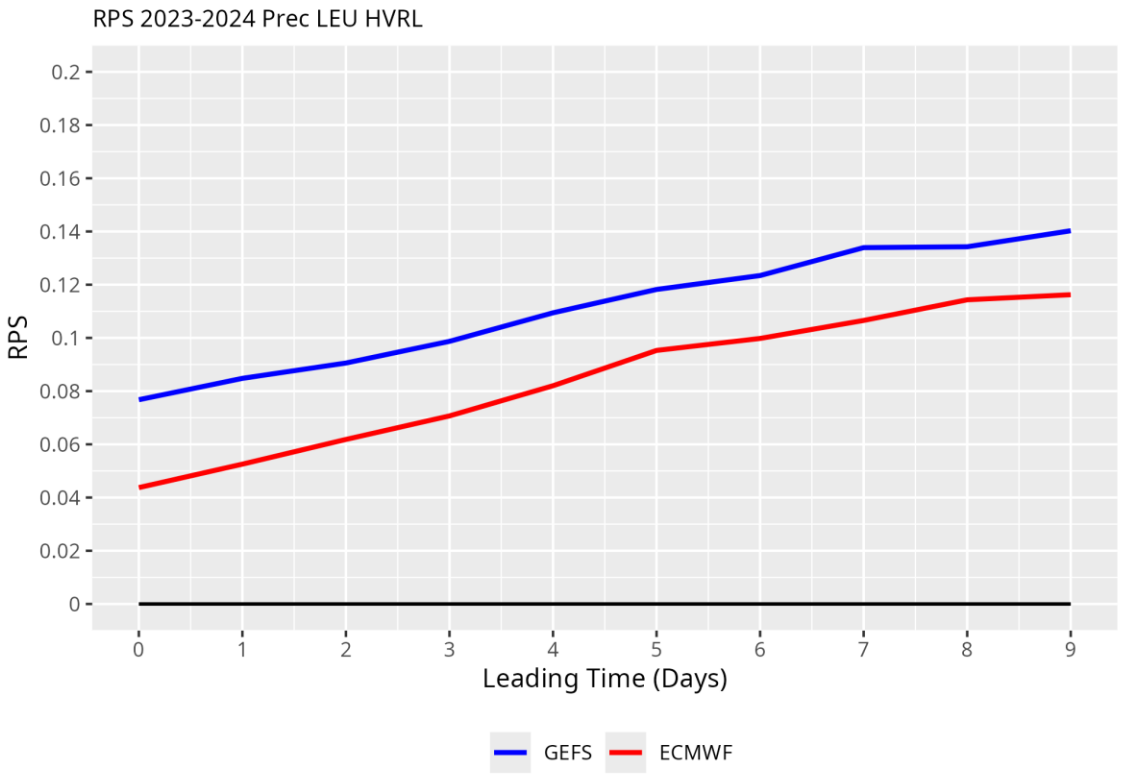


Figure A3.7. *RPS for precipitation forecast in Leuven.*

TALLINN

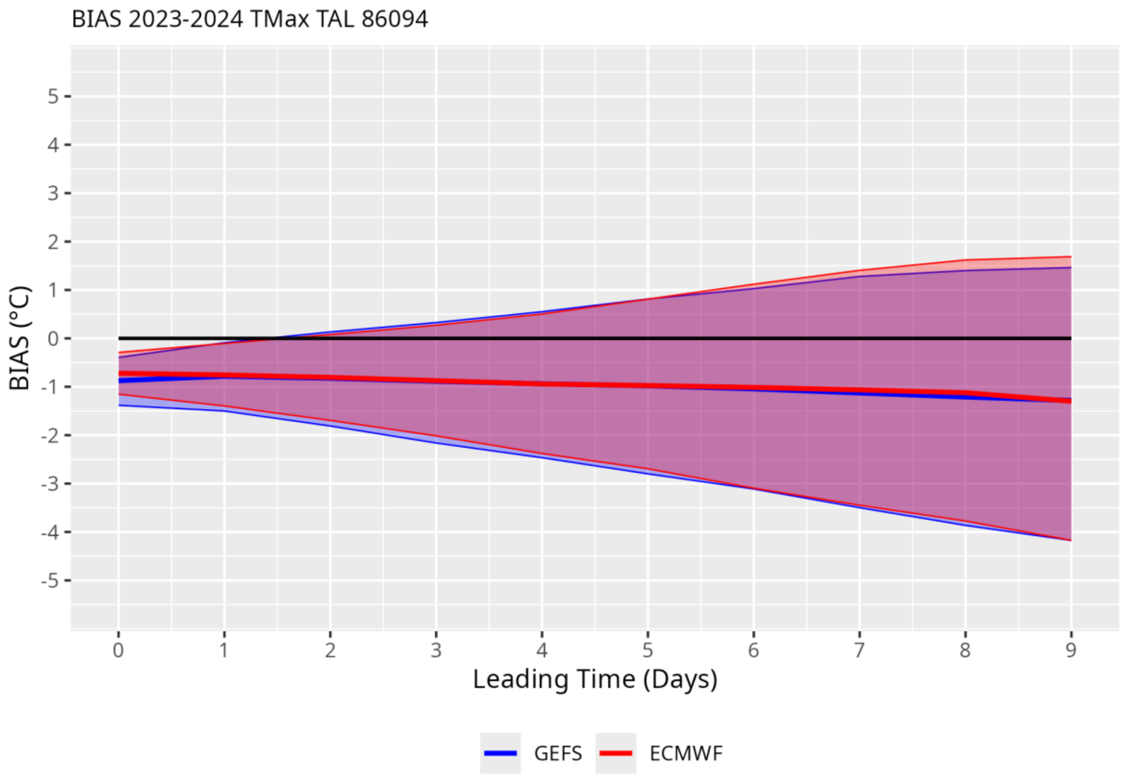


Figure A3.8. *BIAS for maximum temperature in Tallinn.*

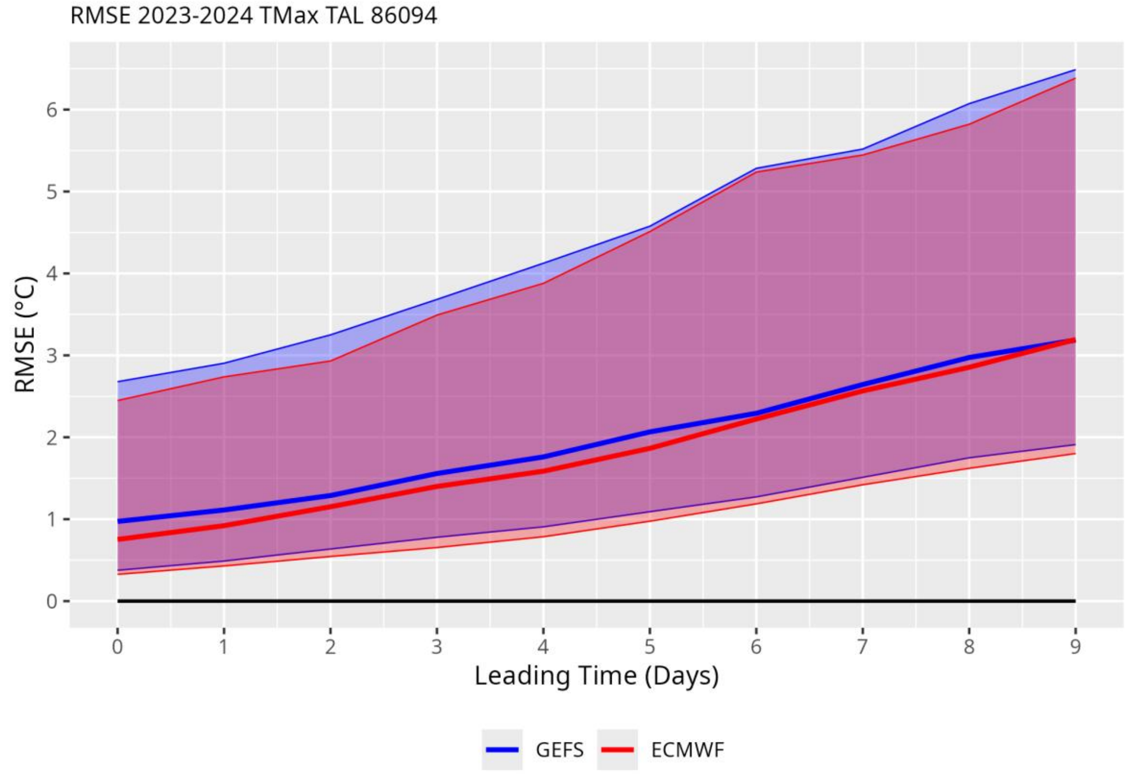


Figure A3.9. *Root Mean Square Error (RMSE) for maximum temperature in Tallinn.*

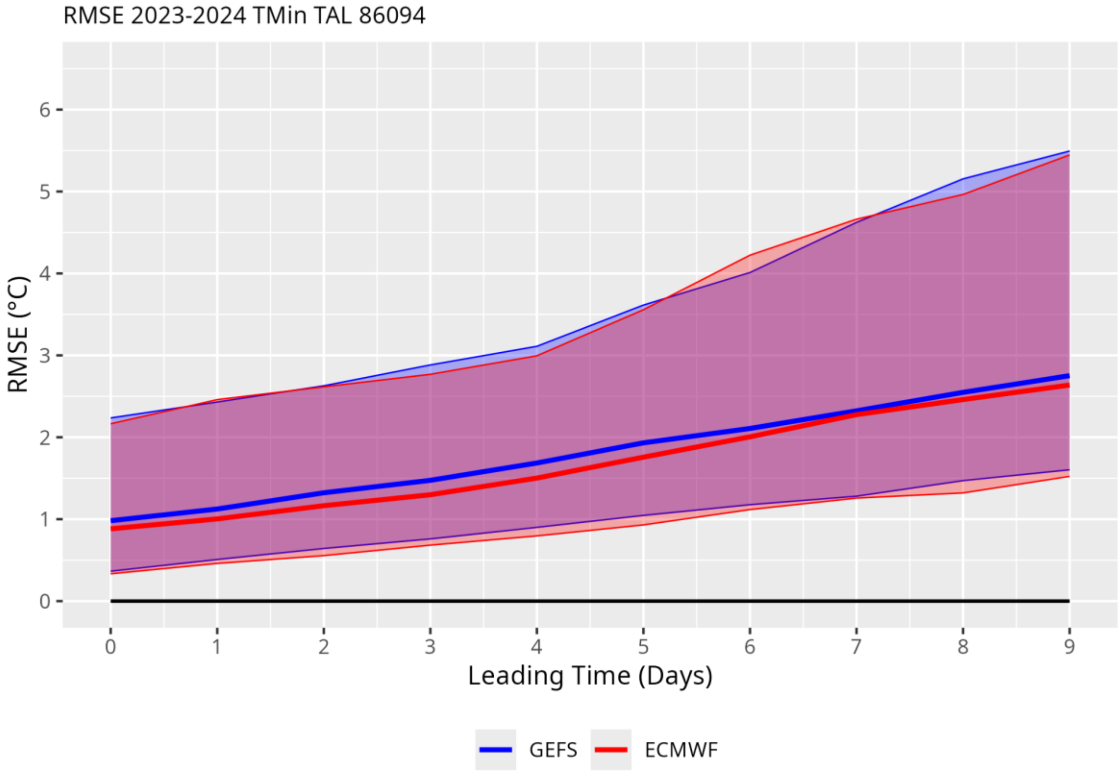


Figure A3.10. *Root Mean Square Error (RMSE) for minimum temperature in Tallinn.*

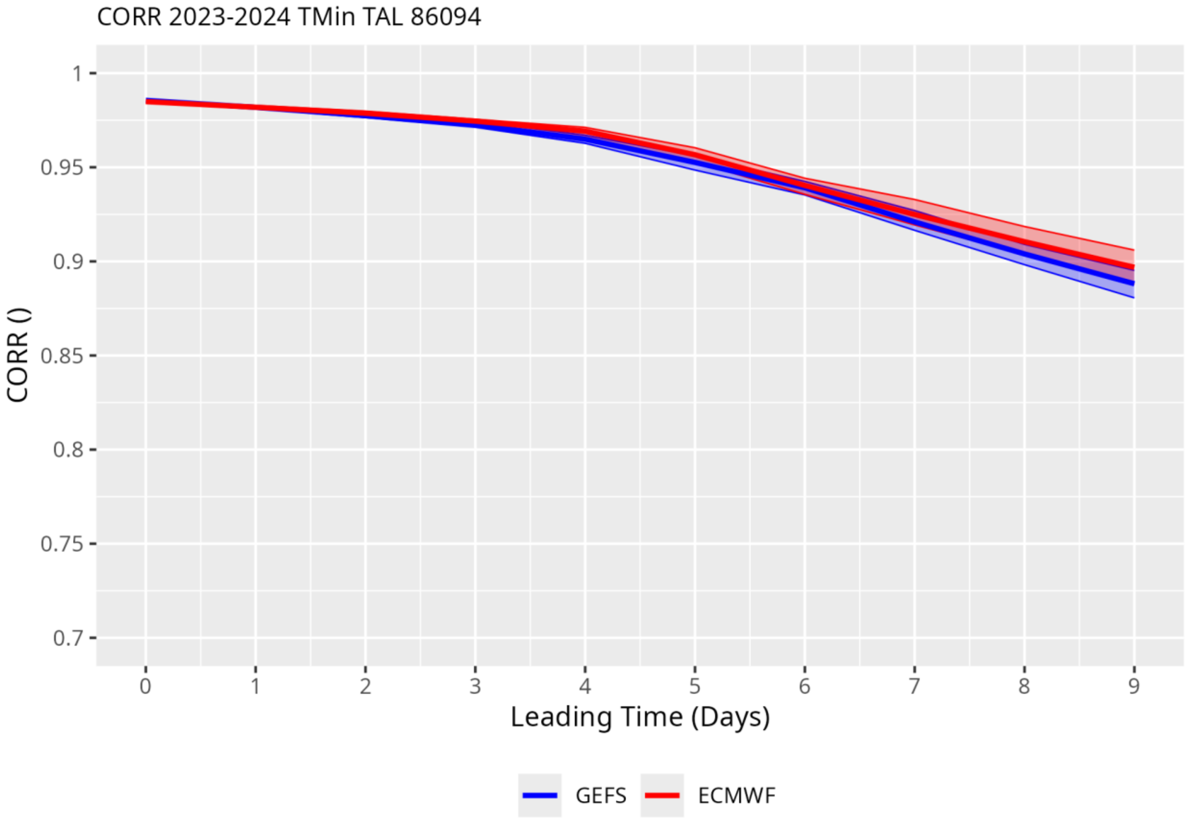


Figure A3.11. *Correlation between observations and predictions for minimum temperature in Tallinn.*

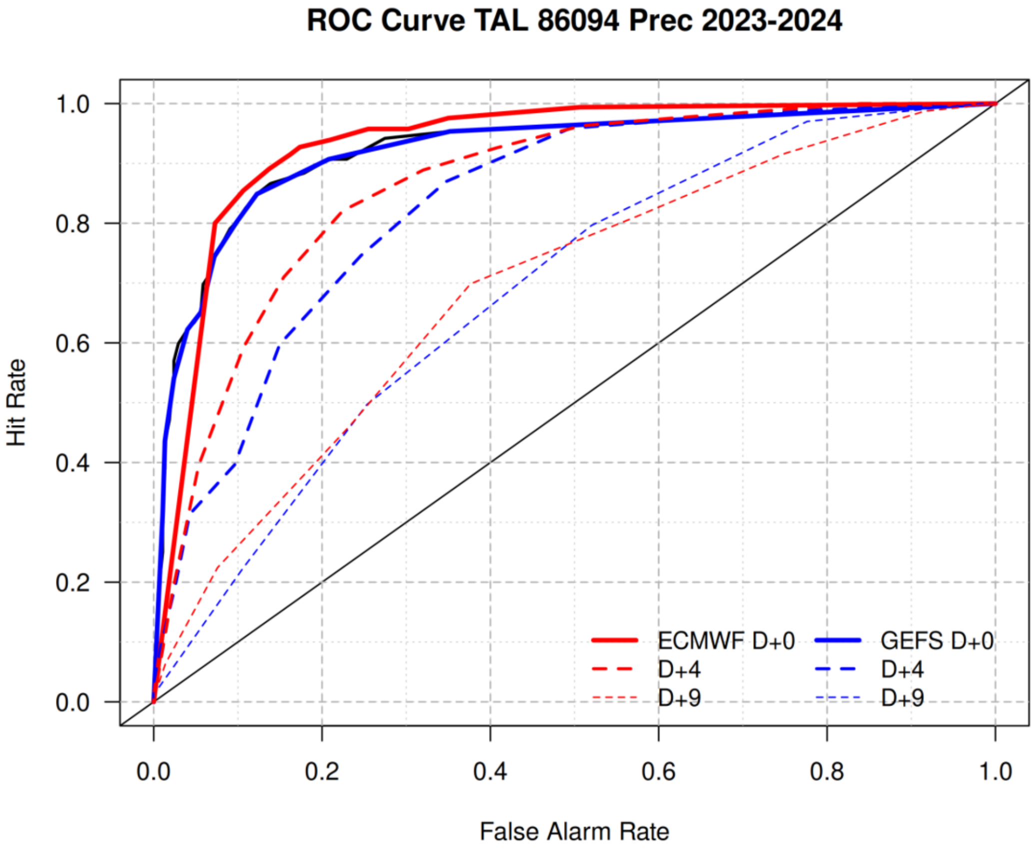


Figure A3.12. *ROC curve for precipitation forecast in Tallinn.*

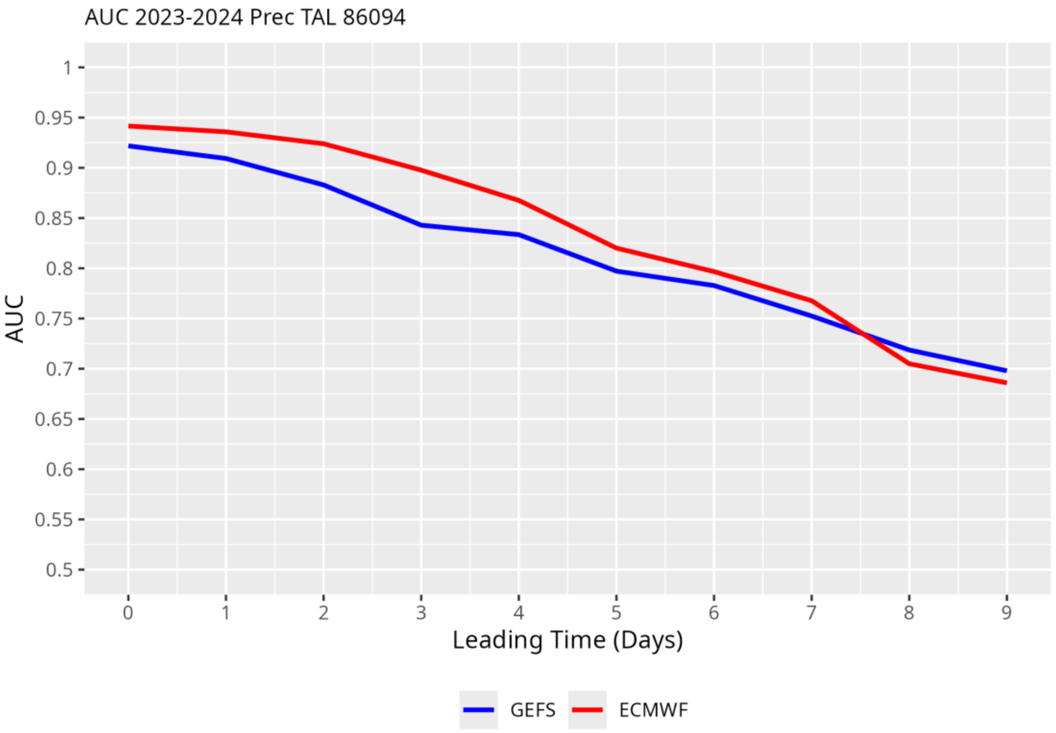


Figure A3.13. *Area under the ROC curve (AUC) for precipitation forecast in Tallinn.*

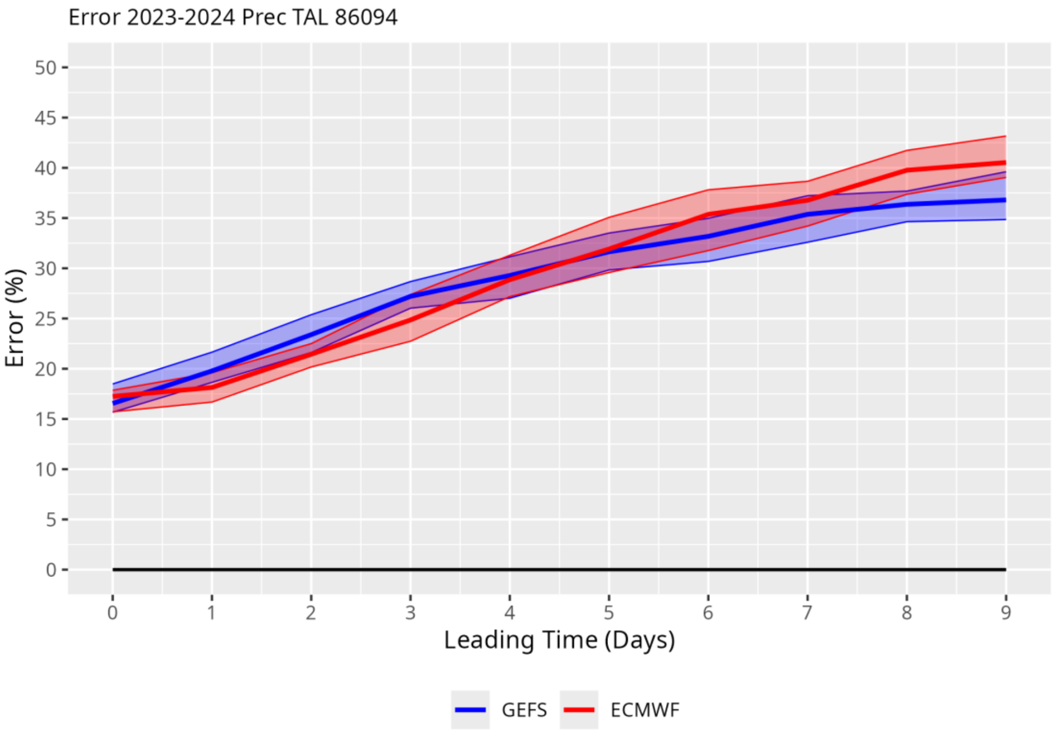


Figure A3.14. *Error rate (false positives + false negatives) for precipitation forecast in Tallinn.*

CLUJ-NAPOCA

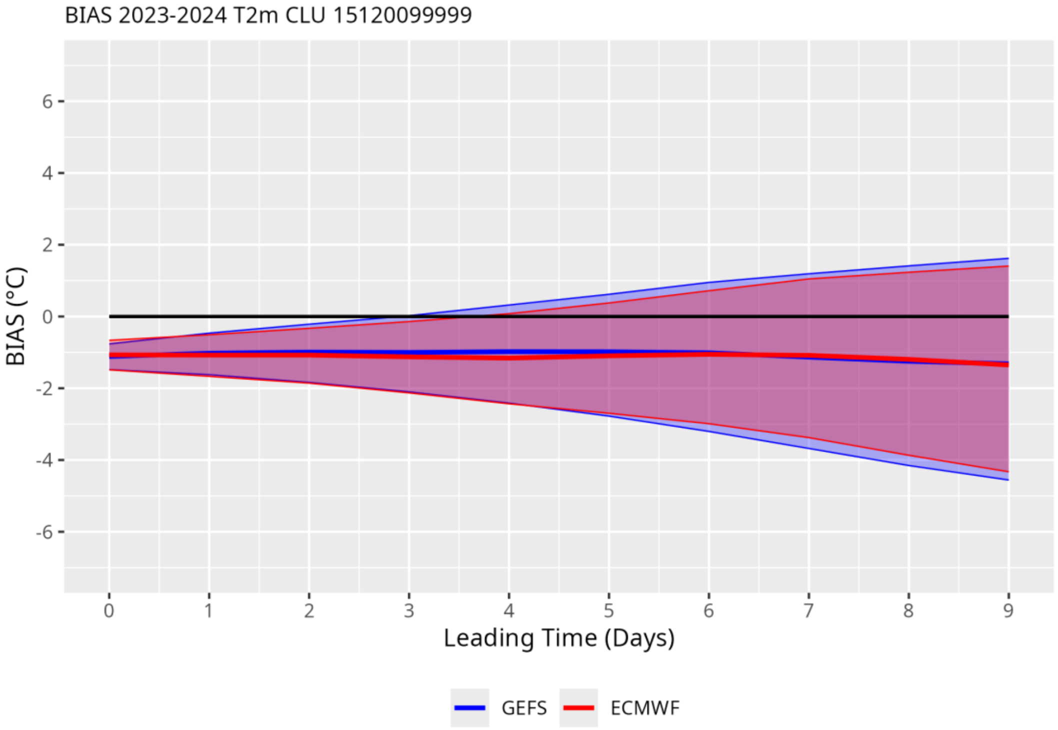


Figure A3.15. *BIAS for daily average temperature in Cluj-Napoca.*

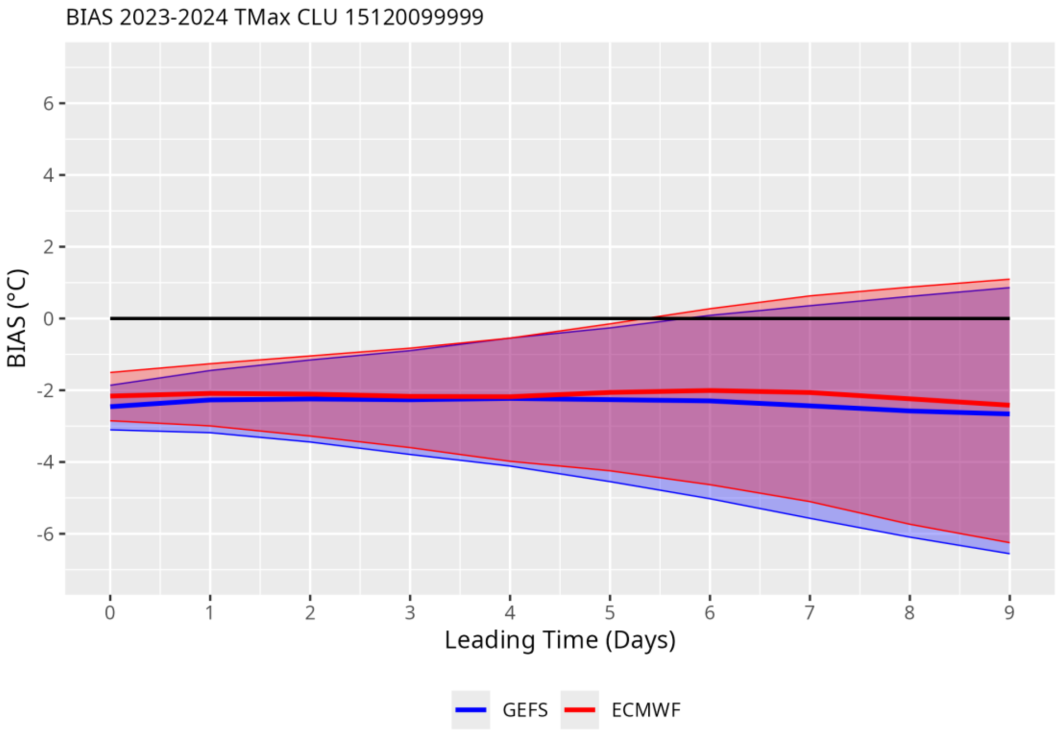


Figure A3.16. *BIAS for maximum temperature in Cluj-Napoca.*

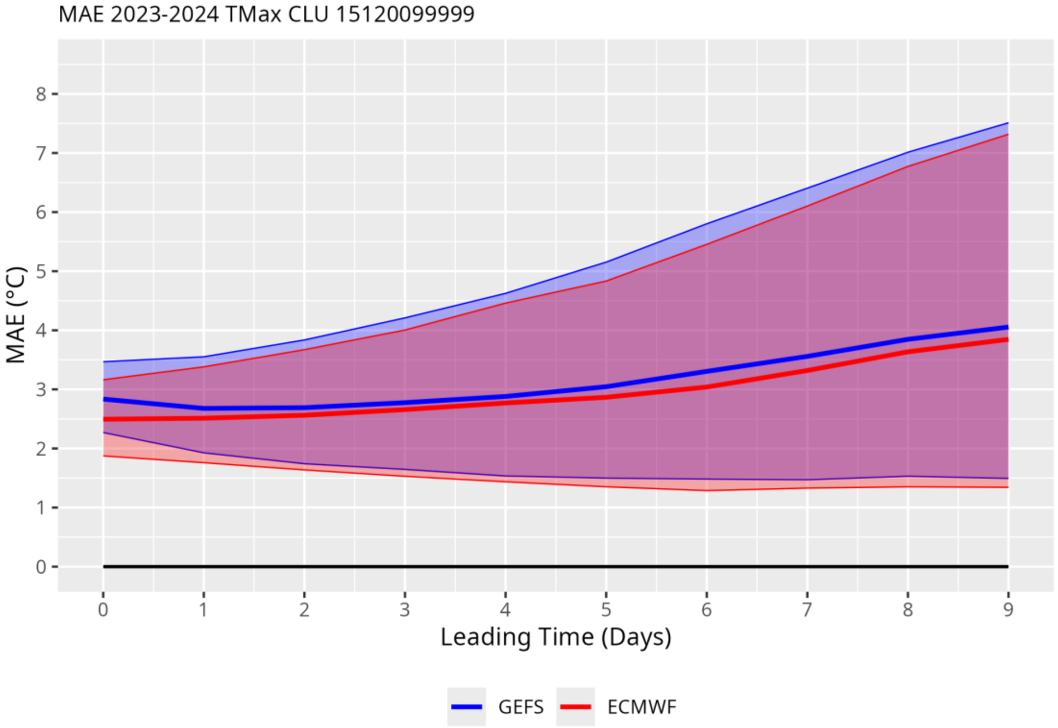


Figure A3.17. *MAE for maximum temperature in Cluj-Napoca.*

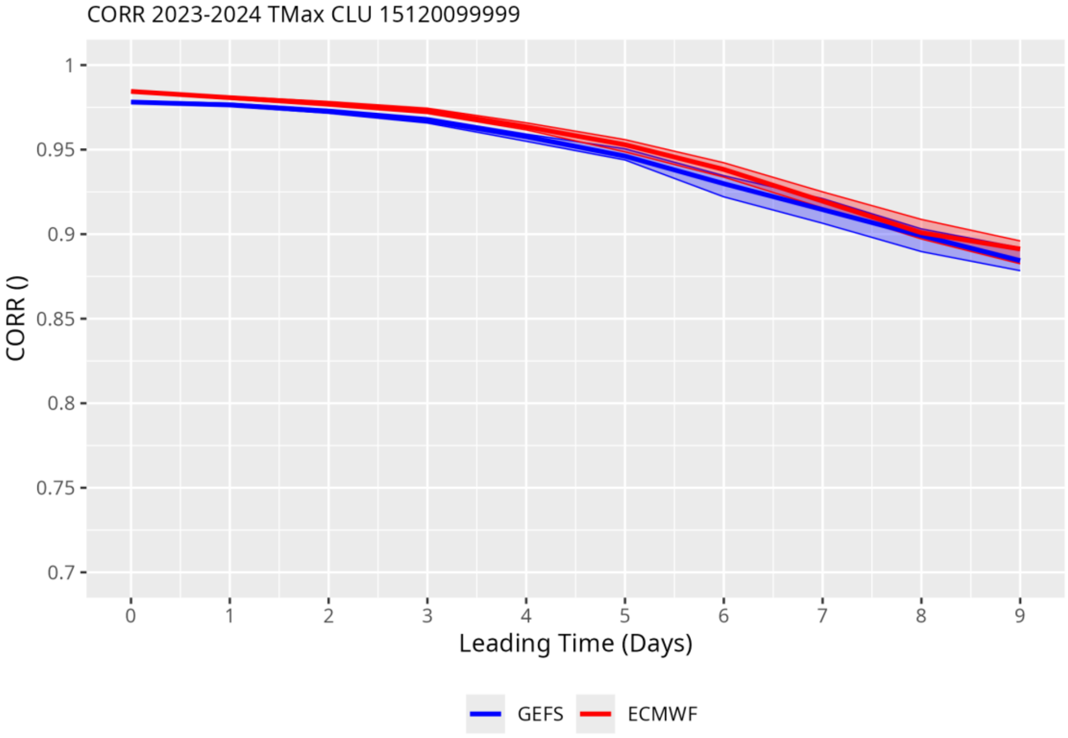


Figure A3.18. *Correlation for maximum temperature in Cluj-Napoca.*

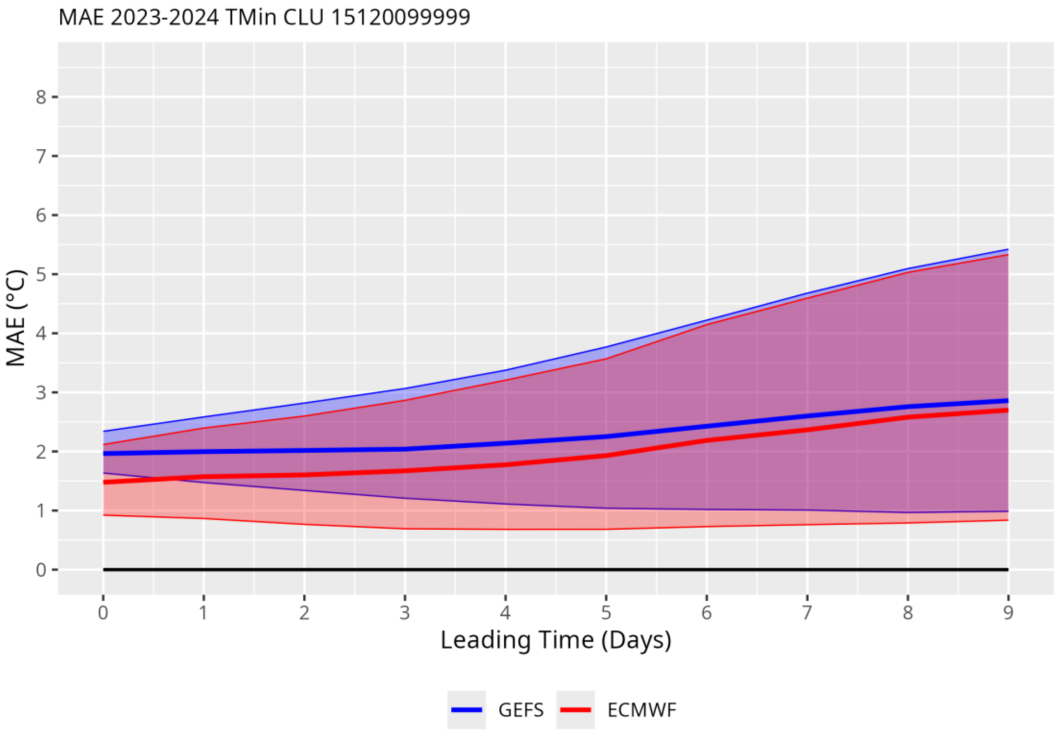


Figure A3.19. *MAE for minimum temperature in Cluj-Napoca.*

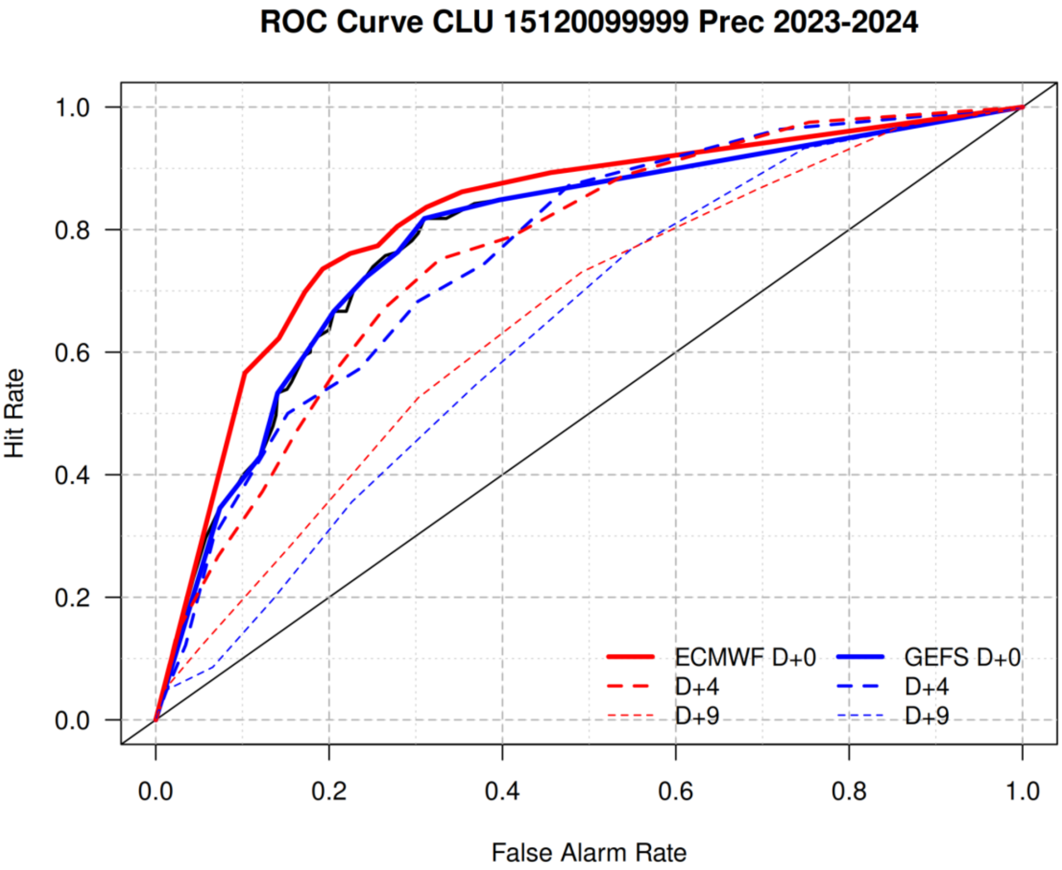


__Figure A3.20.__ ROC curve for precipitation forecast in Cluj-Napoca.

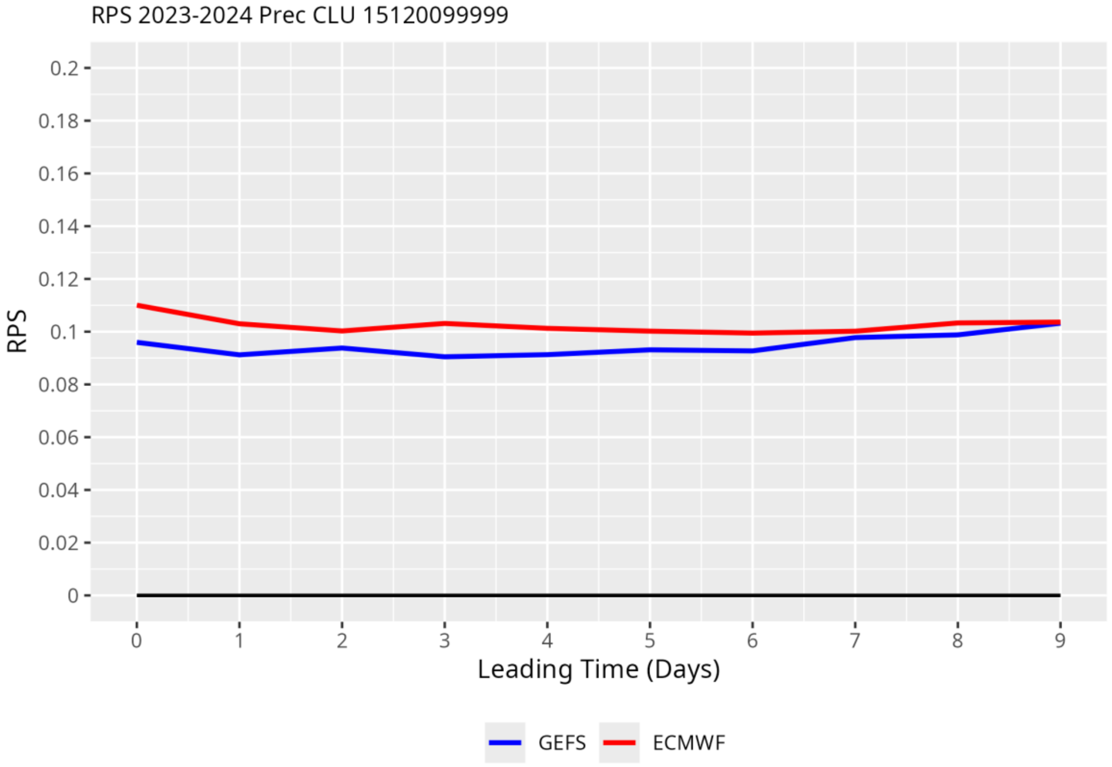


Figure A3.21. *RPS for precipitation forecast in Cluj-Napoca.*

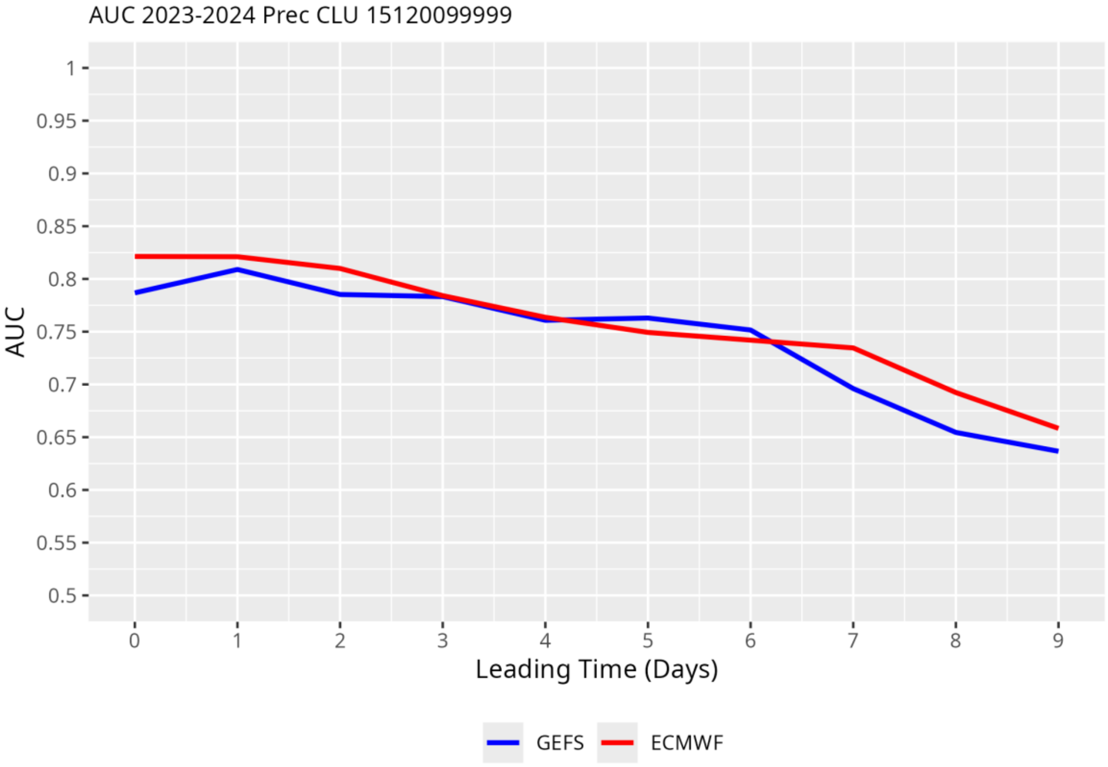


Figure A3.22. *Area under the ROC Curve (AUC) for precipitation forecast in Cluj-Napoca.*

Annex 4. Seasonal Forecast Full Results

Pearson Correlation (Ensemble Mean)

Mean Temperature (TMean)

The correlation analysis for mean temperature shows that forecast skill is strongest at the one-month lead, with positive correlations across most models and observation sites. Median values around 0.2 to 0.3 and upper extremes above 0.4 indicate meaningful predictability for short lead times. Beyond two months, correlations decline sharply toward zero, reflecting the challenges of longer-range forecasting. Models such as ECCC and ECMWF stand out with higher median correlations and greater consistency across stations, underscoring their relative strength for temperature prediction.

As an example, the validation results for the average temperature of the models that constitute the Copernicus seasonal prediction system are shown.

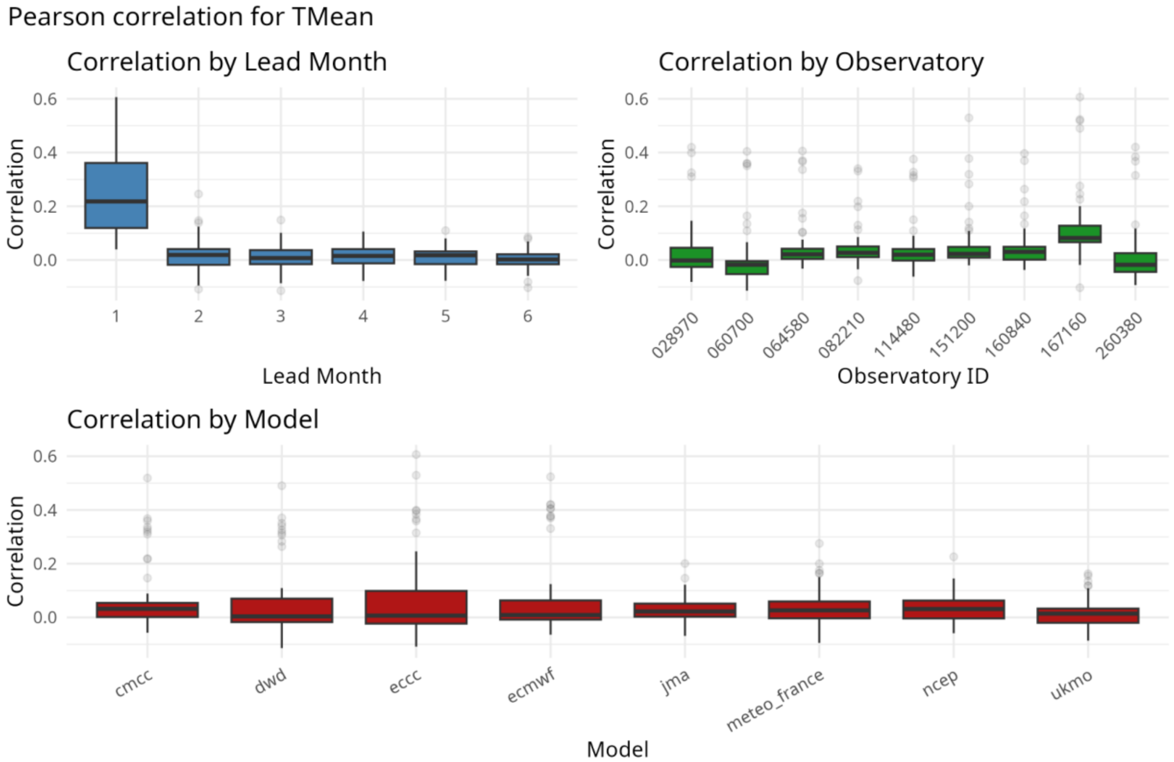


Figure A4.1. Pearson correlation for TMean, classified for lead month, observatory and model.

Precipitation (PRCP)

For precipitation, ensemble mean correlations also show short-term predictive value, with positive median correlations at lead month one, typically near 0.1 and upper ranges approaching 0.3 to 0.4. Skill decays quickly with lead time, with correlations close to or below zero by lead months three to six. Variability across observatories suggests localized differences in predictability, while model comparison highlights ECCC as generally outperforming others at short lead times, though all models struggle to maintain skill beyond two months.

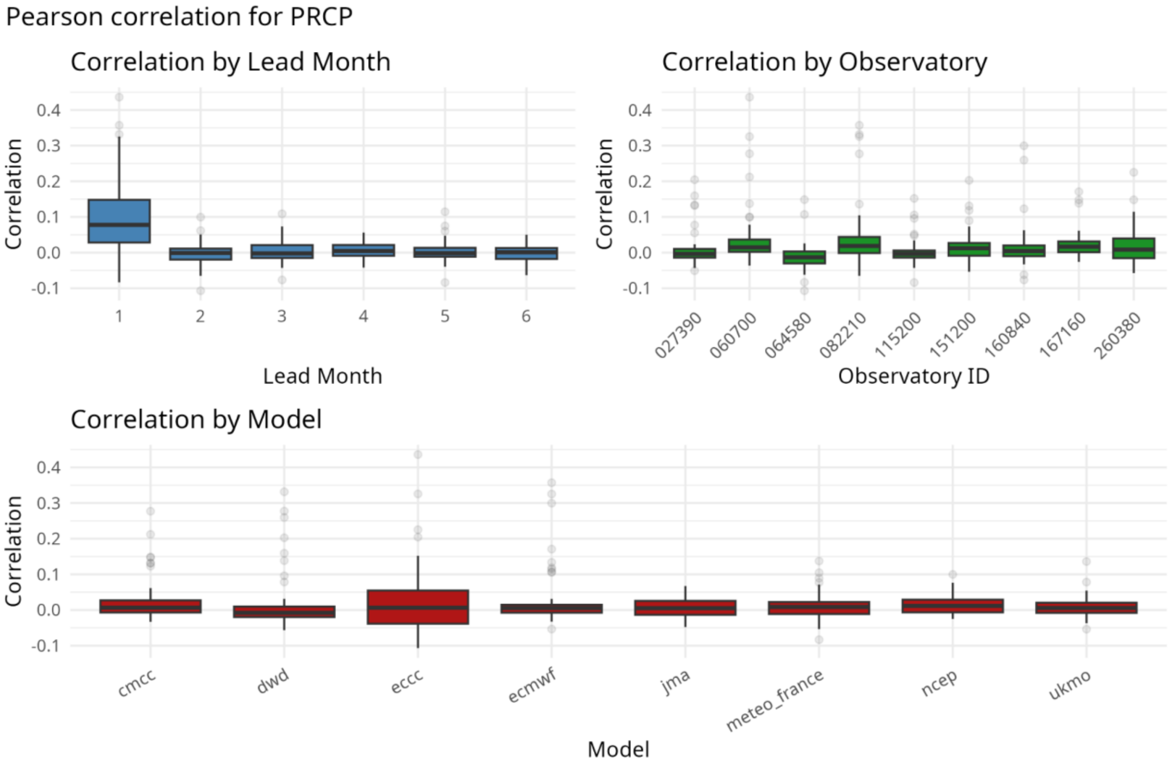


Figure A4.2 Pearson correlation for PRCP, classified by lead month, observatory, and model.

Mean Wind Speed

The correlation for wind speed follows a similar pattern, with modest but positive values at one-month lead, median correlations around 0.1 to 0.2, and some outliers reaching 0.4. Correlation diminishes markedly beyond lead month two. Spatial variability is evident, with certain observatories maintaining slightly better skill. Among models, ECCC consistently achieves higher median correlations at short leads, confirming its relative advantage in capturing wind anomalies, while other models show limited skill at longer horizons.

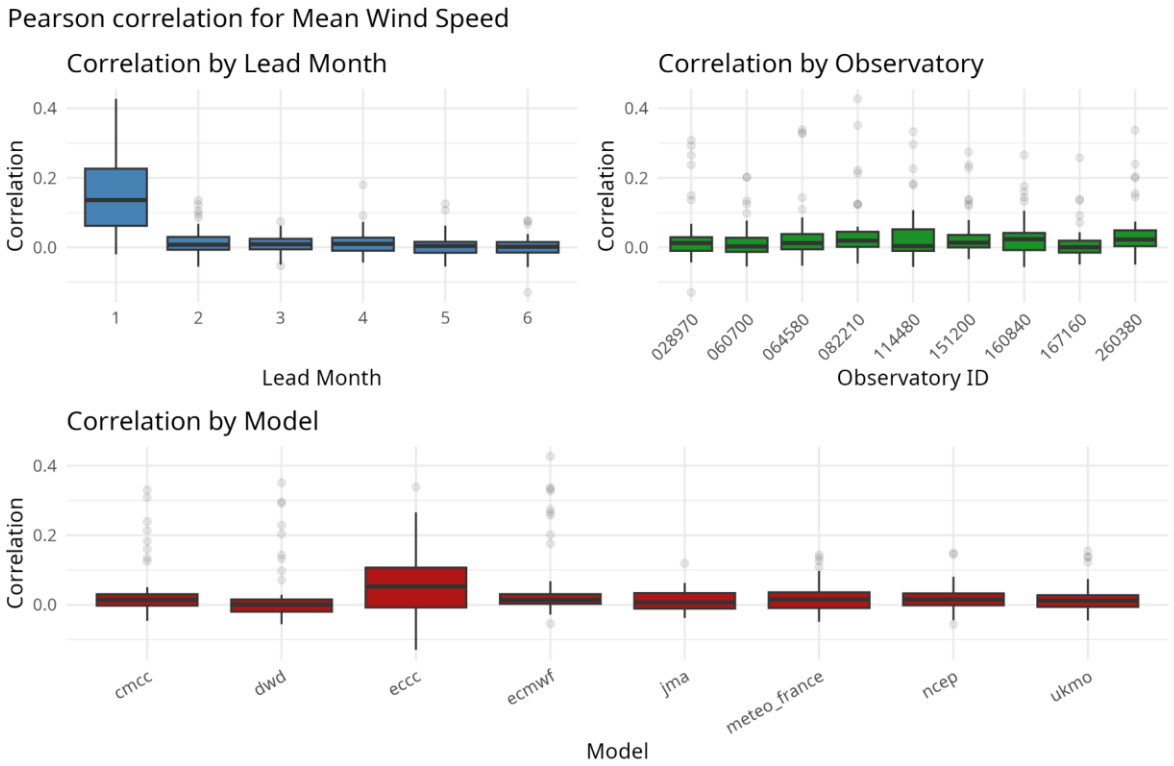


Figure A4.3. Pearson correlation for Mean Wind Speed, classified by lead month, observatory, and model.

Mean Absolute Error (MAE)

Mean Temperature (TMean)

The MAE analysis for temperature forecasts shows the lowest errors at one-month lead, with steady increases as the forecast month progresses, reflecting diminishing predictability. Improvement over climatology is most pronounced at short lead times, reinforcing the added value of dynamical forecasts early in the season. In order to avoid saturating the graphs with redundant information, only statistics up to the second month are shown, since the difference between the results obtained in the second and sixth months was negligible. Among the models, ECCC and ECMWF consistently achieve lower MAE values, highlighting their better performance, while differences among models narrow beyond lead month two.

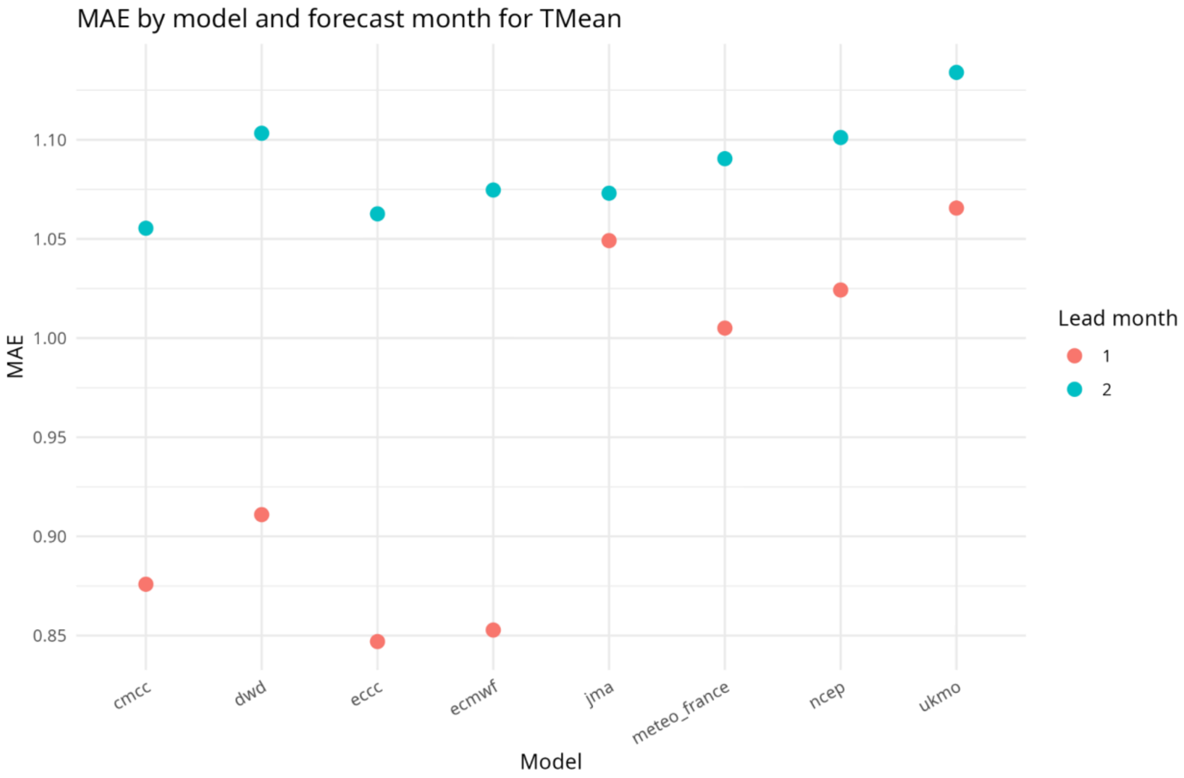


Figure A4.4. Mean Absolute Error for TMean, classified by lead month and model.

Precipitation (PRCP)

Precipitation forecasts display similar trends, with MAE lowest at one-month lead and increasing with forecast horizon. The short-term improvement over climatology is modest but significant, underscoring some ability to predict precipitation anomalies early in the forecast period. ECCC and ECMWF generally show lower MAE compared to other models at lead 1, while differences become negligible beyond two to three months, indicating the loss of forecast skill at longer ranges.

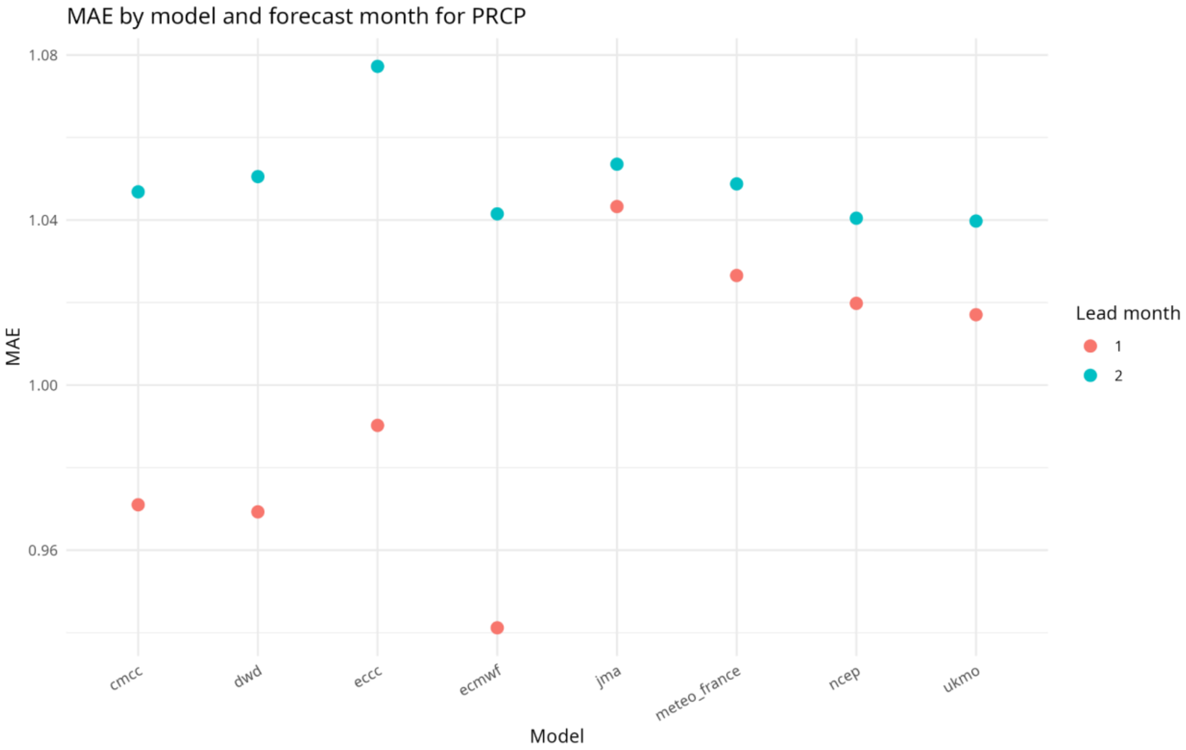


Figure A4.5. Mean Absolute Error for PRCP, classified by lead month and model.

Mean Wind Speed

Wind speed forecasts demonstrate their highest skill at short lead times, with minimum MAE at one month and steady increases thereafter. ECCC and ECMWF tend to achieve the lowest errors at lead 1, indicating their relative advantage. Beyond lead month two, model differences flatten and errors rise, consistent with declining predictability for seasonal wind anomalies.

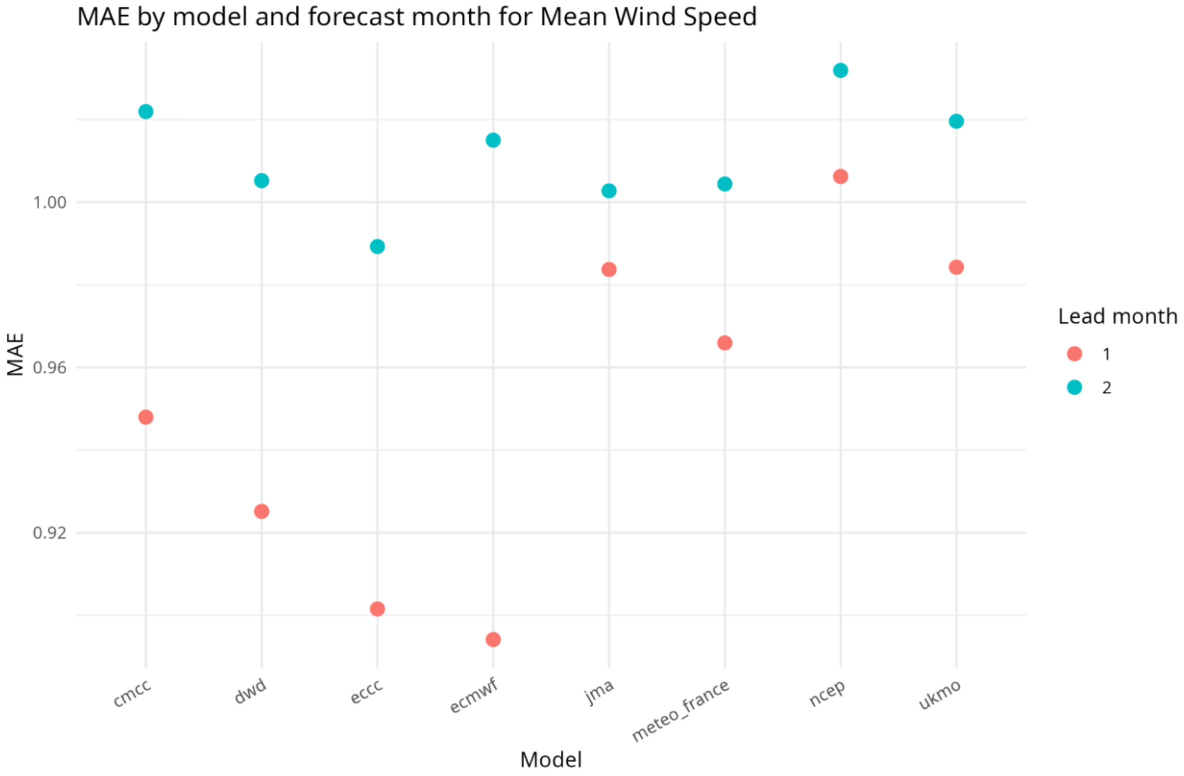


Figure A4.6. Mean Absolute Error for Mean Wind Speed, classified by lead month and model.

Root Mean Square Error (RMSE)

Mean Temperature (TMean)

RMSE patterns for temperature align closely with MAE results, showing the lowest errors at lead 1 and steady increases with lead time. This underscores the fading predictive skill beyond short leads. ECCC and ECMWF again emerge as the models with the lowest RMSE values at early leads, confirming their comparative strength in temperature forecasting.

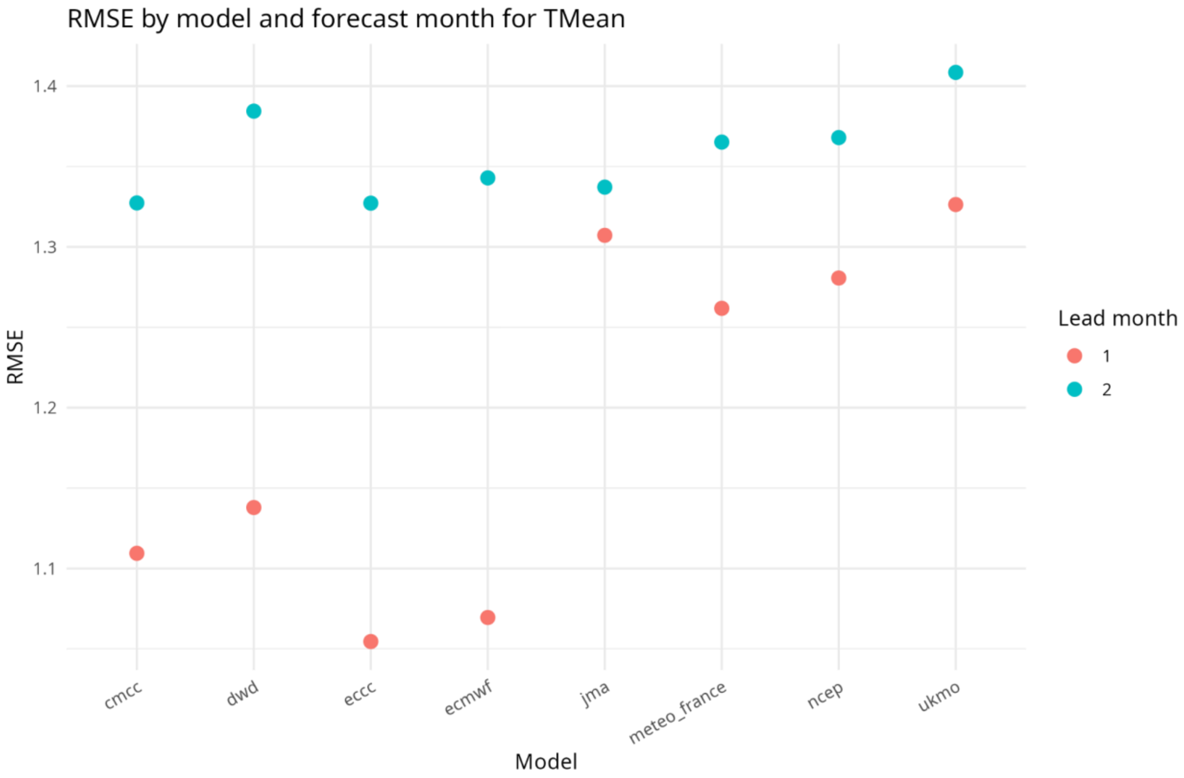


Figure A4.7. Root Mean Square Error for TMean, classified by lead month and model.

Precipitation (PRCP)

For precipitation, RMSE is lowest at lead 1 and grows steadily with longer forecasts, reflecting the limits of dynamical model skill. ECCC demonstrates better performance than peers at short leads, though differences flatten beyond lead 2 or 3, illustrating the difficulty in maintaining reliable forecasts for precipitation anomalies over longer time scales.

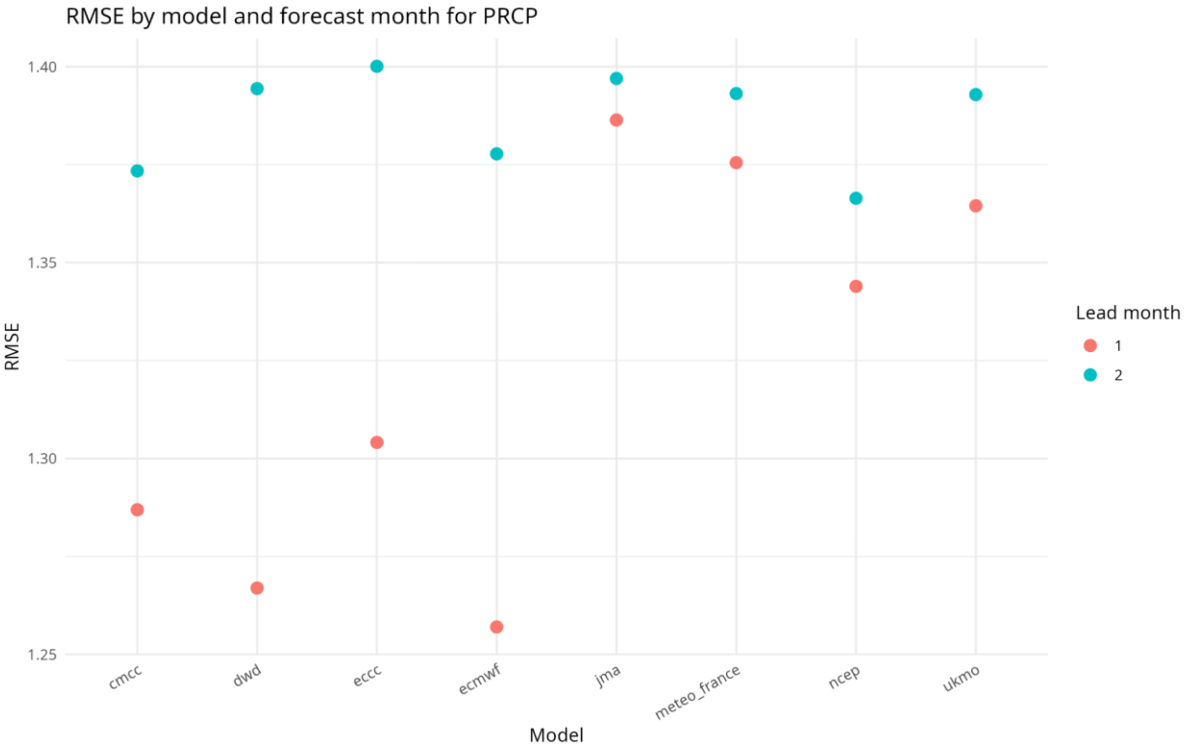


Figure 4.8. Root Mean Square Error for PRCP, classified by lead month and model.

Mean Wind Speed

Wind speed RMSE also shows clear lead-time dependence, with minimum errors at one-month lead and rising errors thereafter. ECCC and ECMWF generally maintain lower RMSE values at short lead times, highlighting their relative advantage in predicting wind anomalies. Model differences diminish at longer leads, reflecting the inherent limits of seasonal wind forecasts.

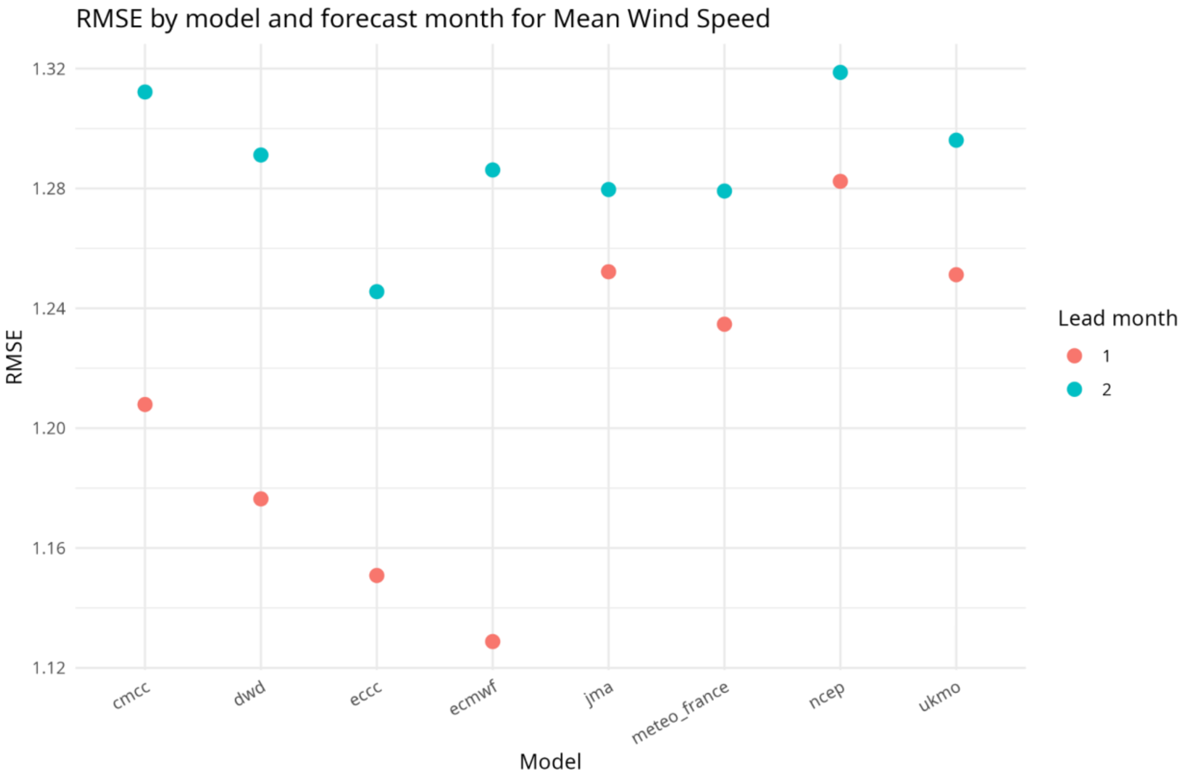


***Figure A4.9**. Root Mean Square Error for Mean Wind Speed, classified by lead month and model.*

Seasonal forecast - FICLIMA method models results

In order to determine which model fixes better, several statistical-based models were performed for each observatory, lag and variable.

Regarding the models used, the seasonal forecasting system originally employed nine proposed models, each combining different computational approaches to predict meteorological variables. These models represented a comprehensive ensemble of modern statistical and machine learning techniques tailored for long-range weather prediction.

The CNN_ens model represents the core deep learning approach, utilizing three different random seeds to create an ensemble of Convolutional Neural Networks trained on the gridded ERA5 predictor fields.

This ensemble approach helps capture spatial patterns in atmospheric data while providing robustness through model averaging. The CNN_Lin variant implements a simplified linear version within the Keras Framework, offering a more interpretable neural network architecture that maintains the convolutional structure but with linear activation functions.

For traditional statistical modeling, the MultiLinear model employs Ridge regression, a regularized linear regression technique that prevents overfitting while handling the high-dimensional predictor space. This serves as a baseline linear approach against which more complex models can be compared. The Teleconnection model implements a specialized R-based approach that identifies and utilizes atmospheric teleconnection patterns - large-scale atmospheric circulation patterns that influence weather across vast distances.

The TeWa model introduces wavelet-based time series analysis, combining temporal decomposition with atmospheric predictors to capture both frequency and time-domain relationships. This approach is particularly effective for identifying multi-scale temporal patterns in meteorological data. The hybrid models begin with CNN_Tele, which blends the CNN_ens predictions with Teleconnection-based forecasts, creating a synthesis of deep learning and atmospheric physics-based approaches.

Similarly, CNN_Tewa combines CNN_ens with TeWavelet predictions, merging spatial pattern recognition with wavelet-based temporal analysis. The most comprehensive hybrid approach, CNN_Tewa_Lin, integrates predictions from CNN_ens, TeWavelet, and Teleconnection models, creating a multi-method ensemble that leverages the strengths of each individual approach. Finally, TeWa_Lin provides a streamlined combination of TeWavelet and Teleconnection models, offering a balance between complexity and interpretability.

Pearson correlation and SMAE were the validating statistical methods selected to measure the performance of each of them.

Pearson Correlation

Pearson correlation is used to assess whether the forecasts reproduce the observed temporal variability and phase of anomalies, independently of any systematic bias or scaling errors. It therefore provides a robust, unitless measure of “pattern skill” that is comparable across stations, lead times and models.

Mean Temperature (TMean)

For the average temperature, the boxplots show that all models achieve modest but generally positive Pearson correlations for near-term leads, with median values typically between about 0.1 and 0.3 for lags 1–3. The best performance is obtained by the TeWa-based blends, particularly CNN_TeWa and TeWa_Lin, which consistently rank among the top models and maintain positive skill even at longer lead times (lags 4–6). The pure CNN ensemble (CNN_Ens) also provides robust, though slightly smaller, correlations across most lags. In contrast, the purely linear approaches (CNN_Lin, MultiLinear) and the standalone teleconnection model tend to degrade more rapidly with lead time, with medians hovering around zero or becoming negative beyond about 3–4 months. Overall, incorporating TeWa information clearly enhances temporal persistence of correlation skill in the seasonal temperature forecasts.

Figure A4.10. Combined Pearson Correlation for different Statistical based models for each lag and for Mean Temperature

Among the observatories, the highest and most persistent correlations are found at the Mediterranean and continental sites in Greece (167160), Italy (160840) and Romania (151200), which show clearly positive median correlations at most lead months. Czech Republic (114480) also exhibits modest but robust positive values, whereas the northern stations in Finland (028970) and Denmark (060700) often hover around zero or slightly negative correlations. Western sites in Belgium (064580), Spain (082210) and Estonia (260380) tend to display intermediate behaviour, with small but generally positive correlations that decay with lead time.

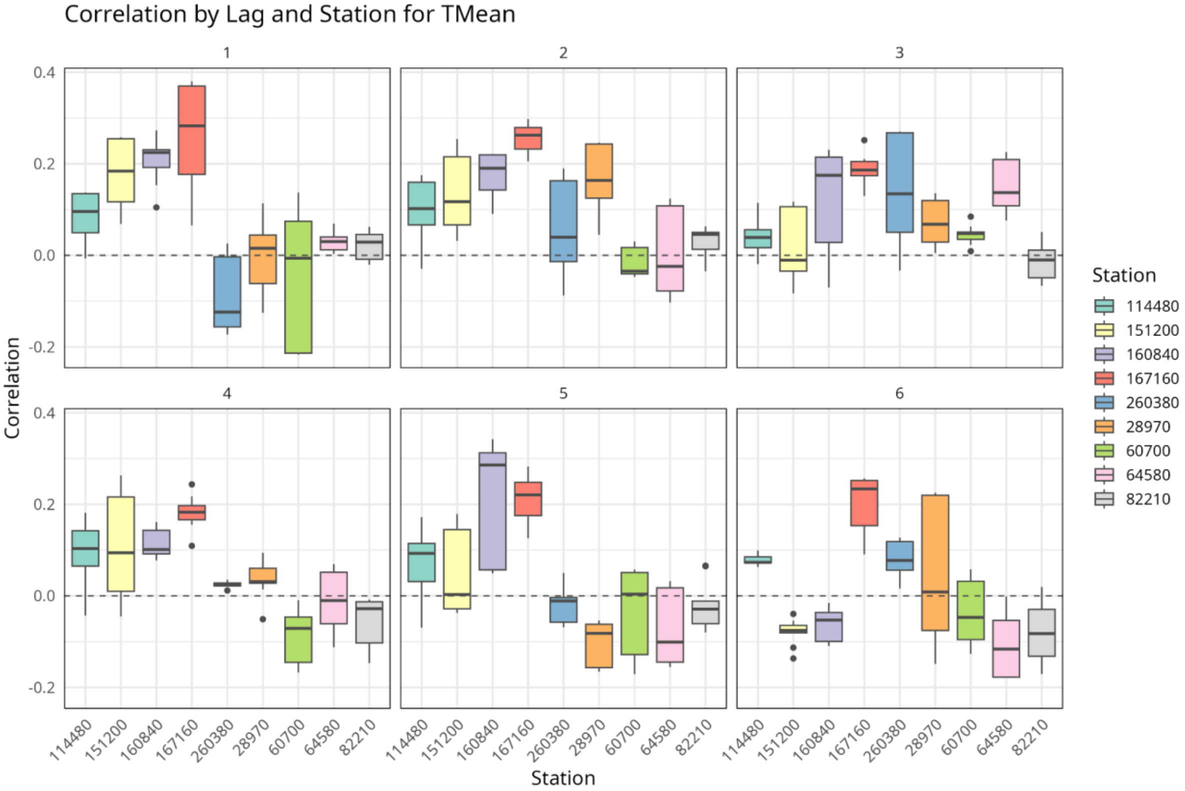


Figure A4.11. Combined Pearson Correlation for the observatories used for each lag (Mean Temperature)

Precipitation

For monthly precipitation, the Pearson correlation across the nine stations and leads 1–6 is generally positive but modest, with most medians between about 0 and 0.2. TeWa-based approaches (TeWa and especially TeWa_Lin) and the CNN_Ens model tend to occupy the upper end of this range at short and medium leads, while the purely linear MultiLinear model only becomes competitive at the longest lead. Skill decays with lead time, but TeWa_Lin and CNN–TeWa blends still retain a small positive advantage over the simpler CNN_Lin and MultiLinear baselines for most lags.

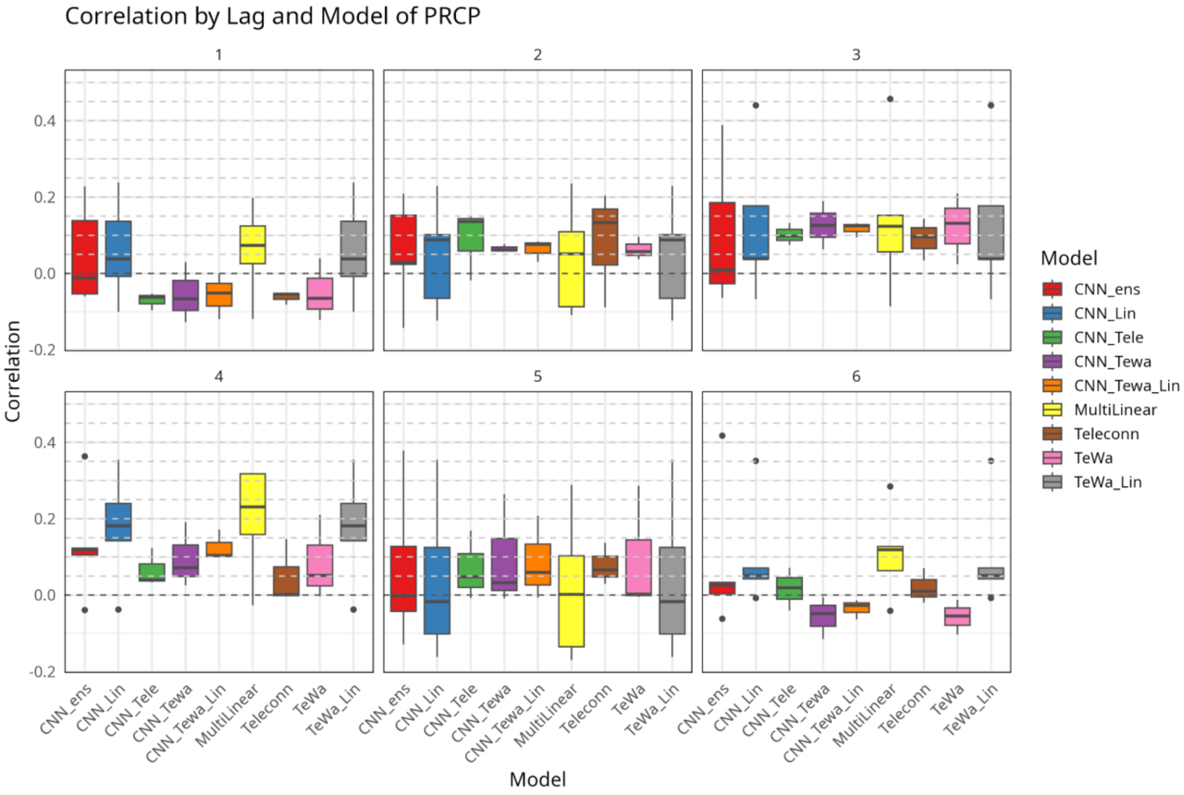


Figure A4.12. Combined Pearson Correlation for different Statistical based models for each lag and for Precipitation

Regarding the observatories, Madrid stands out with consistently positive correlations, especially at intermediate and longer leads (months 3–6), where it remains one of the few sites with non-negligible median correlations while most others collapse toward zero. Pilsen, Cluj-Napoca, Beauvechain[2] and, at some leads, Tallinn also show modest positive correlations, whereas Kajaani is predominantly negative and Piacenza and Aarhus fluctuate around zero. Athens generally exhibits weak or slightly negative correlations, indicating limited rainfall predictability there.

---

[2] City with the closest weather station accomplishing the required historical observed data length for Leuven

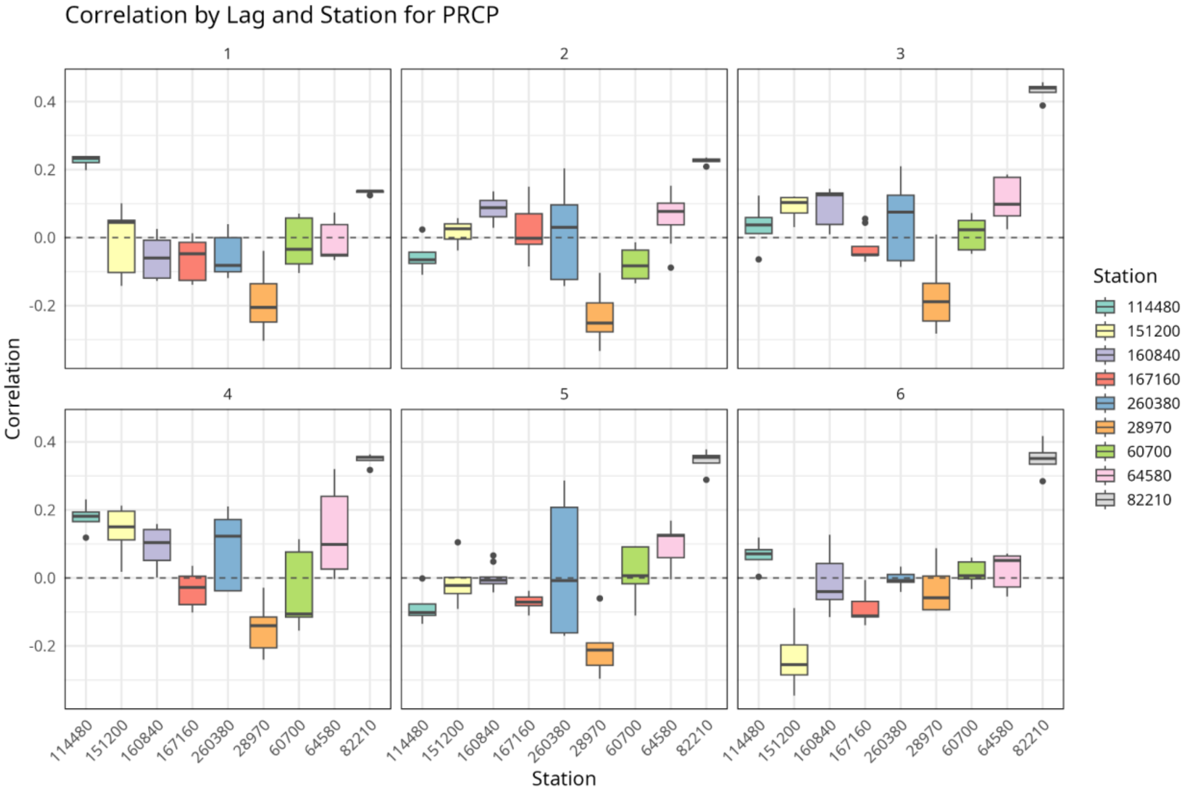


Figure A4.13. Combined Pearson Correlation for the observatories used for each lag (Precipitation)

Mean wind

For mean wind speed, correlations are weaker and more variable than for precipitation, with several models showing near-zero or slightly negative median correlations at short leads. Positive skill emerges mainly at intermediate and long leads, where TeWa and TeWa-based blends, together with some CNN variants (CNN_TeWa, CNN_Tele), reach median values up to around 0.2–0.25 in the best cases. Overall, correlation skill for wind remains modest, but the teleconnection/TeWa-augmented models consistently perform at least as well as the simpler CNN_Ens and linear baselines.

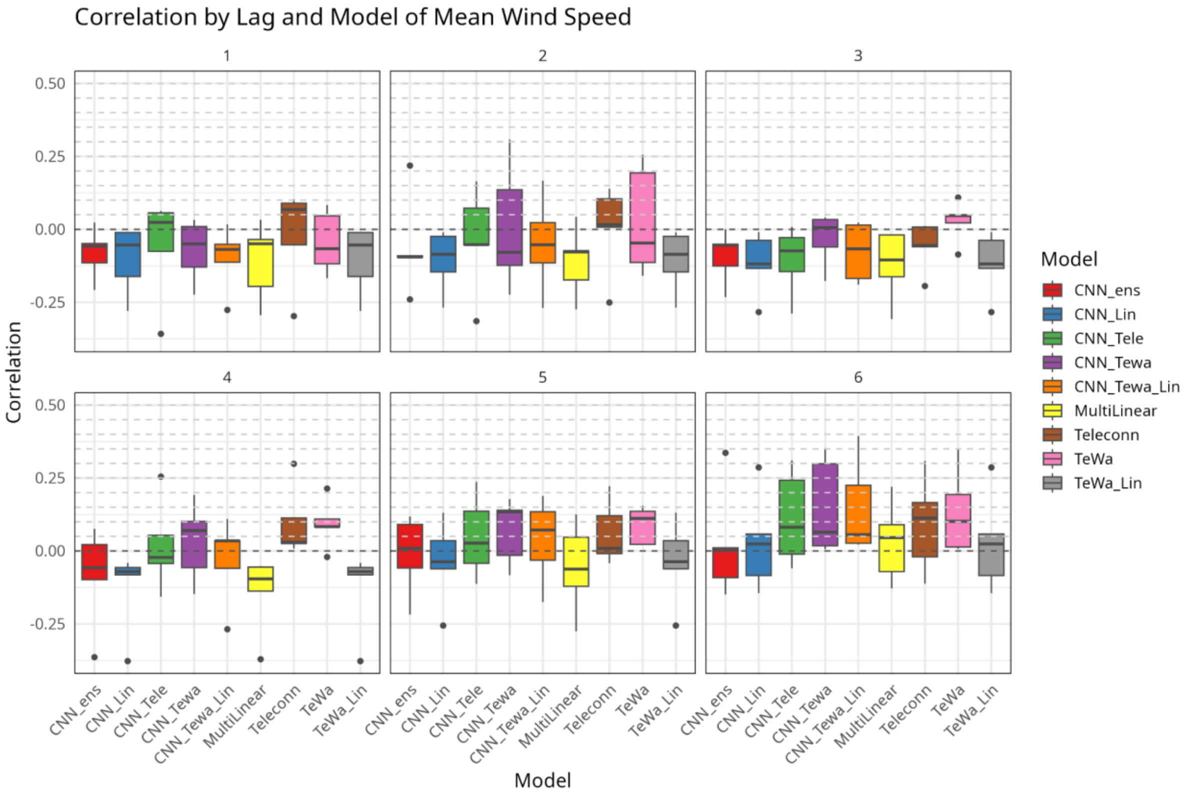


Figure A4.14. Combined Pearson Correlation for different Statistical based models for each lag and for Mean wind speed

For the observatories, the correlations are generally low and noisy, but some spatial patterns emerge. The Czech case FRCequently shows the highest positive correlations at longer leads (up to ~0.2–0.3), while Finnish and Danish achieve small positive values at several lead months. In contrast, Italian and Romanian observatories are often associated with clearly negative median correlations, and the Greek also tends to be negative at many leads. Belgish, Spanish and Estonish stations mostly stay close to zero, underlining the limited but spatially heterogeneous skill for wind anomalies.

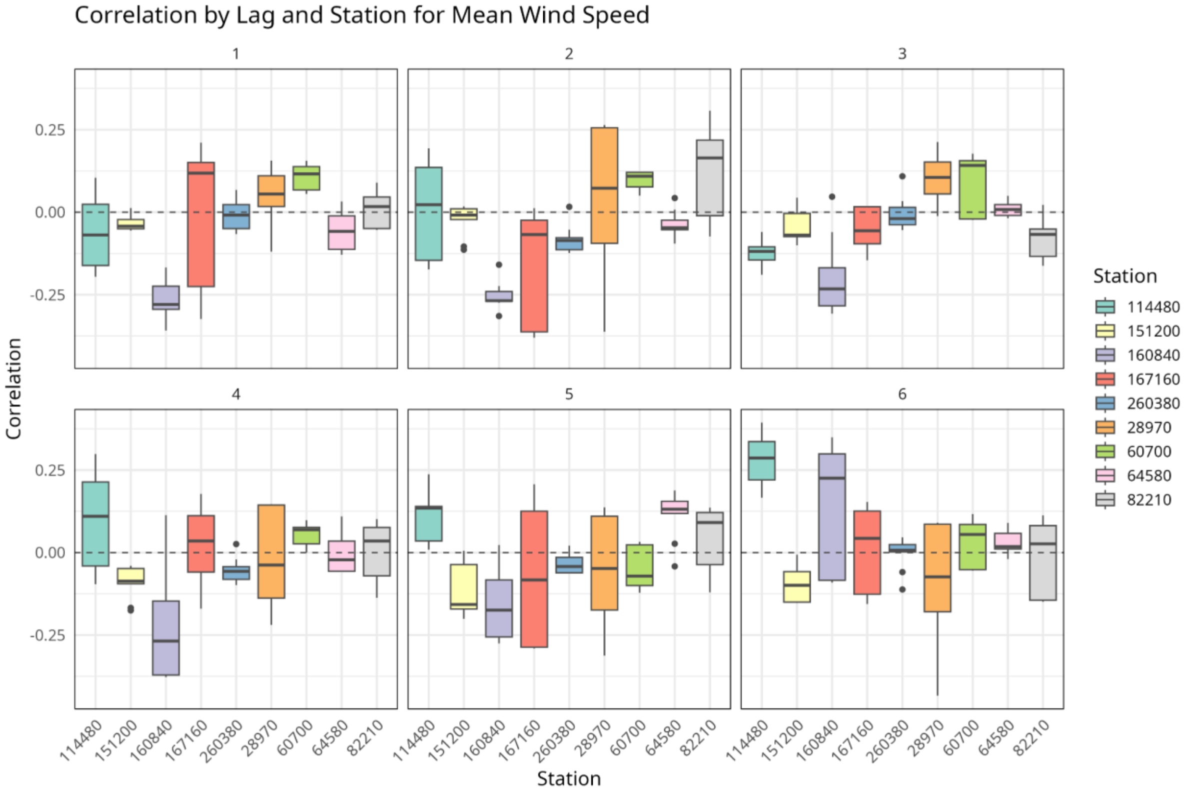


Figure A4.15. Combined Pearson Correlation for the observatories used for each lag (Mean Wind Speed)

Standardized Mean Absolute Error

SMAE was chosen as a quantify method to measure the magnitude of forecast errors in a scale-independent way, normalizing by the observed variability or climatology. This allows us to compare error levels across variables and locations with very different units and ranges, and to identify models that are not only well correlated but also quantitatively accurate.

Mean Temperature (TMean)

SMAE values for temperature are relatively stable with lead time and lie mostly between about 1.0 and 1.4 for the majority of CNN and Teleconnection/TeWa-Lin configurations. CNN_Ens consistently delivers the lowest errors (medians close to 1), with TeWa_Lin and, to a lesser extent, CNN_Lin and Teleconn forming a compact group with slightly higher but comparable SMAE. The standalone TeWa model is clearly an outlier, with substantially larger SMAE (often around 1.8–2.1), indicating that its raw forecasts are less accurate than the blends and CNN-based approaches, even when correlations are similar.

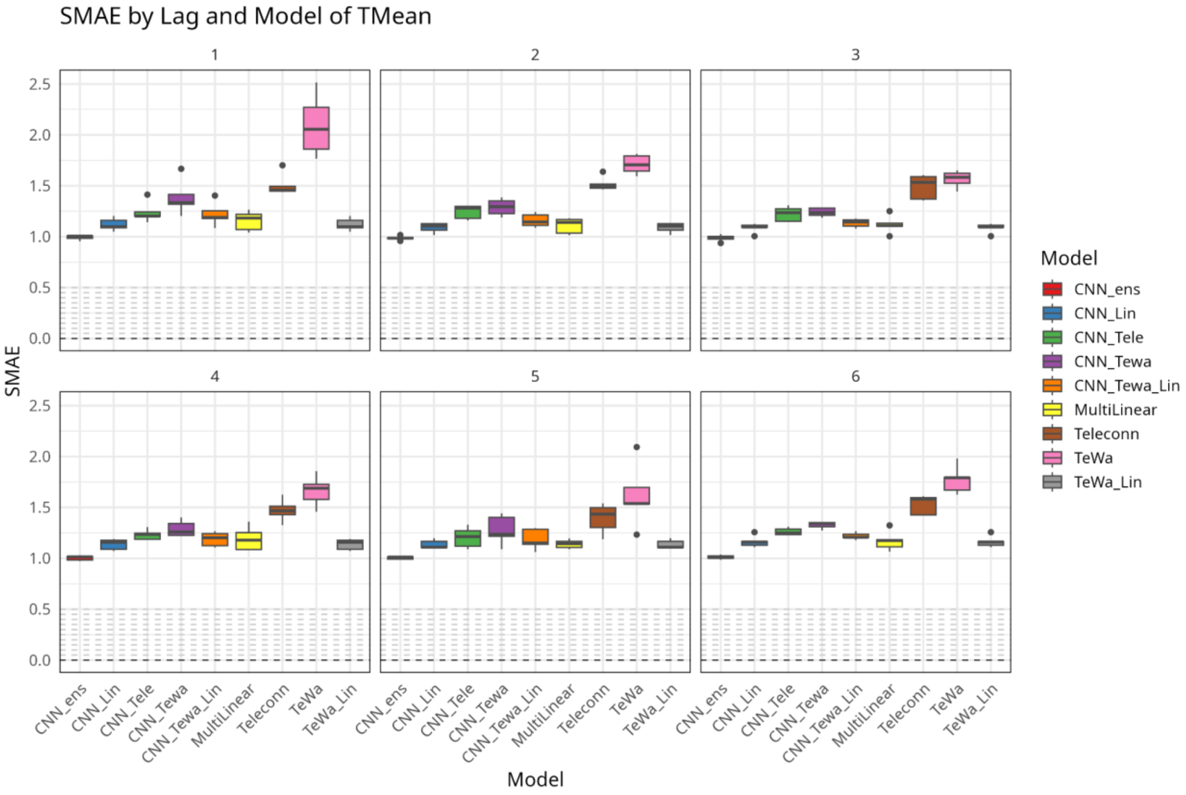


Figure A4.16. Combined SMAE for different Statistical based models for each lag and for Mean Temperature

Temperature SMAE is fairly homogeneous across stations, but systematic differences are visible: Italy (Piacenza), Czech Republic (Pilsen) and Spain (Madrid) usually appear among the lowest-error sites, with median SMAE close to 1–1.2 across lead months. Greece (Athens) and Romania (Cluj-Napoca) tend to show slightly larger errors, and the northern and western stations in Estonia, Belgium and Denmark also feature somewhat higher SMAE, suggesting that local variability and continental effects make temperature slightly difficult to predict there.

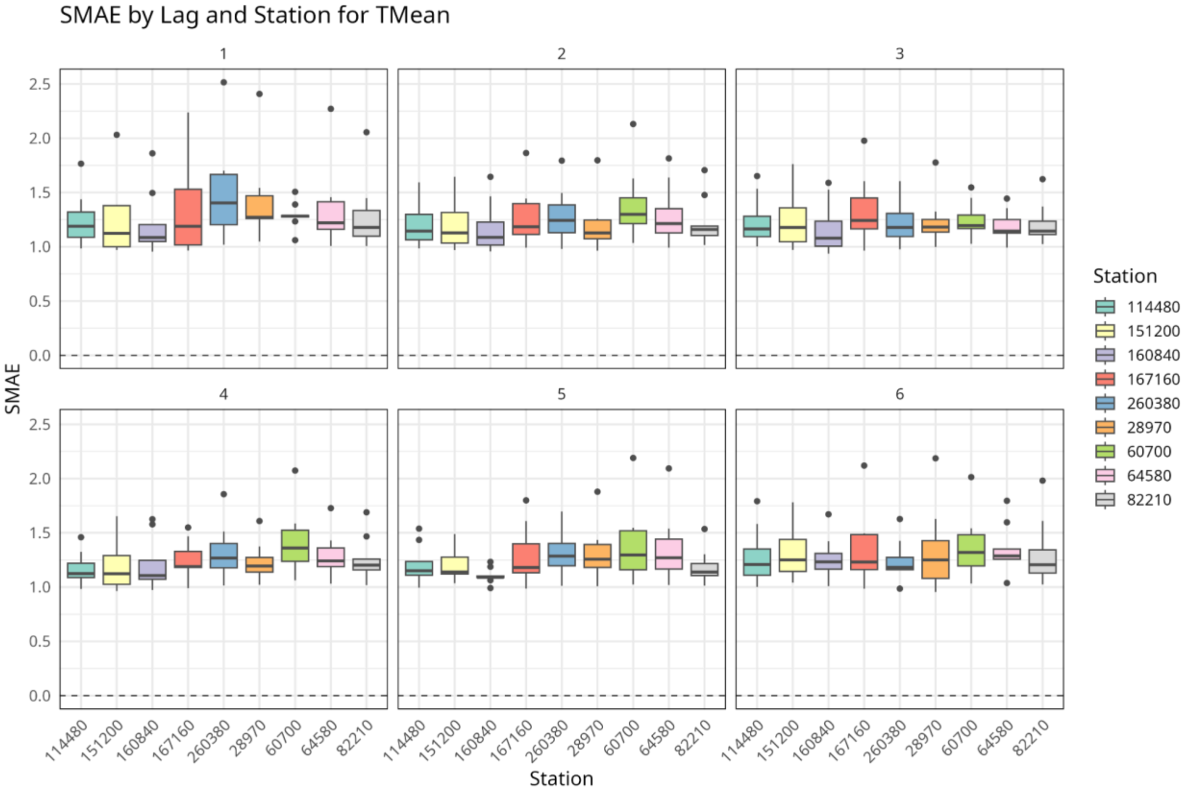


Figure A4.17. Combined SMAE for the observatories used for each lag using all models (Mean Temperature)

Precipitation

In terms of SMAE for precipitation, CNN_Ens clearly provides the lowest errors across all lead times, with median values typically close to 1–1.2, and only a very weak deterioration with increasing lead. Most other CNN and Teleconnection models (CNN_Lin, CNN_Tele, CNN_TeWa, CNN_TeWa_Lin, Teleconn, TeWa_Lin) cluster around slightly higher SMAE values (~1.2–1.6), while the standalone TeWa configuration shows noticeably larger errors (often above 2) at most lags. This means that the three models kept for operational testing (CNN_Ens, CNN_TeWa and TeWa_Lin) lie solidly within the best-performing group from an error perspective.

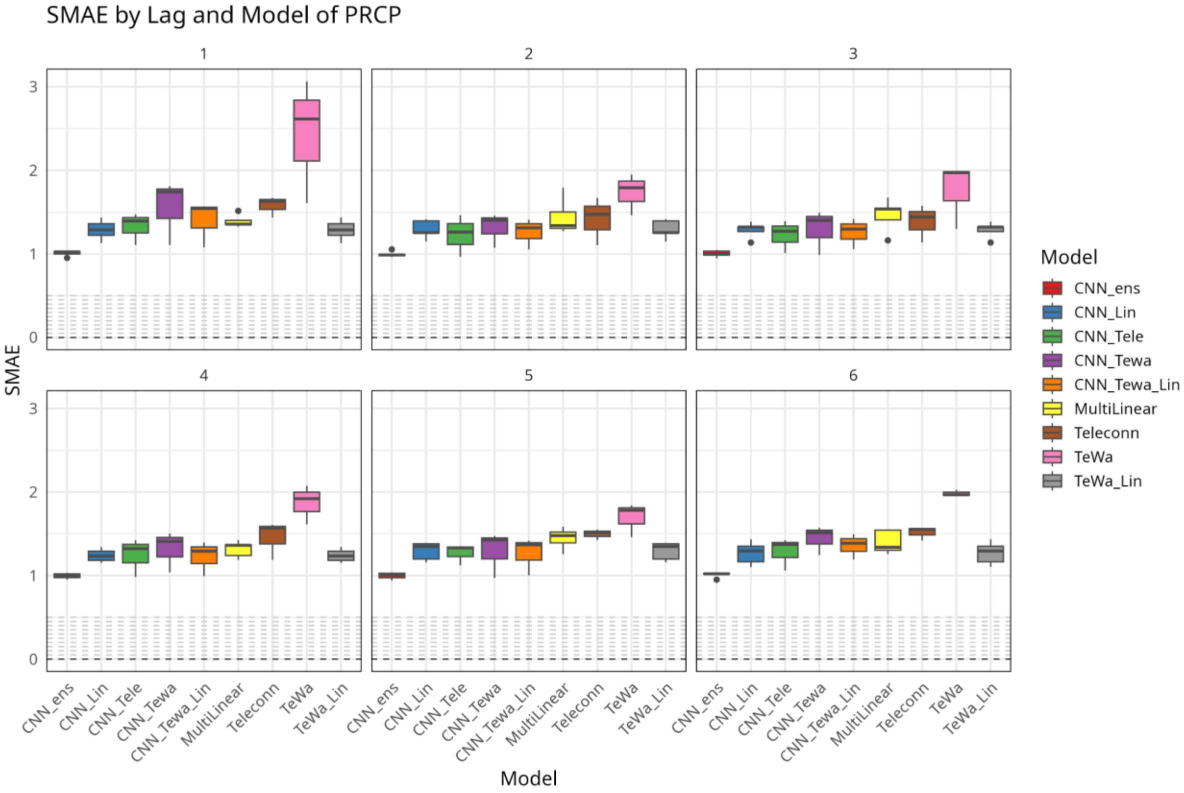


Figure A4.18. Combined SMAE for different Statistical based models for each lag and for Precipitation

Precipitation SMAE again reveals clear spatial contrasts. The lowest errors are typically obtained at Madrid, Piacenza and Beauvechain, where median SMAE stays close to 1–1.3 for most lead months. Pilsen, Aarhus and Tallinn occupy an intermediate range, while Cluj-Napoca and especially Athens exhibit larger errors (SMAE often above ~1.6–2), confirming that these eastern and southeastern sites are more challenging for precipitation forecasting.

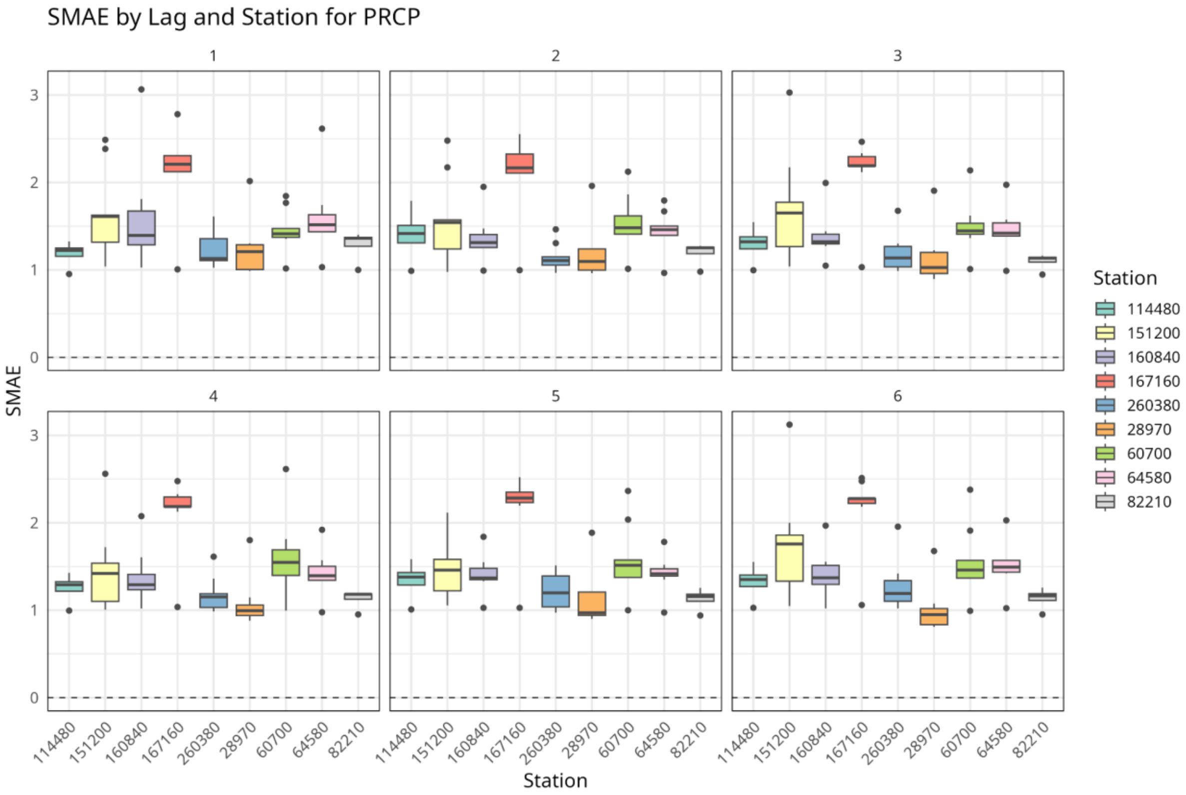


Figure A4.19. Combined SMAE for the observatories used for each lag using all models (Precipitation)

Mean wind

For mean wind speed, SMAE values are generally between about 1 and 2 for most models and lags, again with a limited degradation as lead time increases. CNN-based methods (particularly CNN_Ens and CNN_Tele) tend to yield the lowest errors, while TeWa and TeWa_Lin show slightly higher SMAE but remain within a comparable range. The clear outlier is the MultiLinear baseline, which systematically exhibits much larger errors (often above 2.5), confirming that the nonlinear CNN and TeWa-augmented approaches are better suited for capturing the variability of mean wind speed in this seasonal prediction setting.

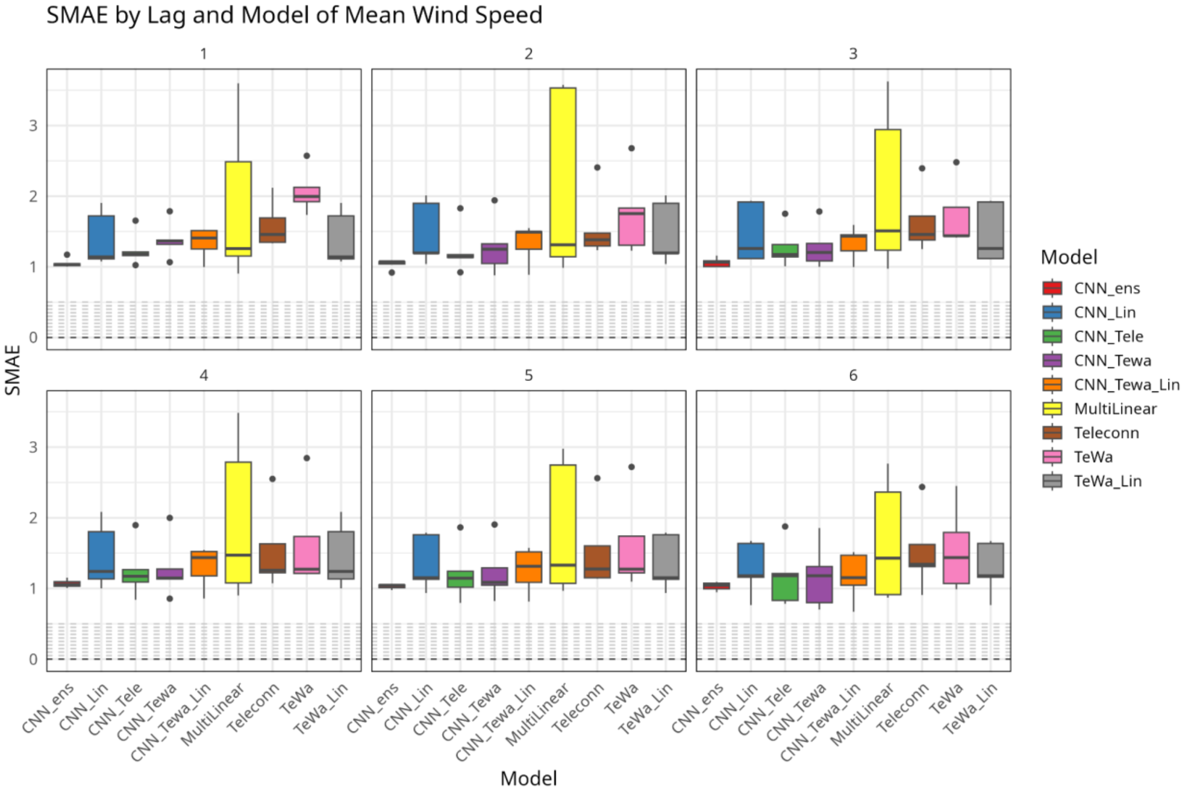


Figure A4.20. Combined SMAE for different Statistical based models for each lag and for Mean wind speed

Wind SMAE highlights those spatial differences in error amplitude. Czech Republic (Pilsen) and Denmark (Aarhus) usually record the smallest errors (SMAE around 1–1.3), indicating relatively well-behaved local wind statistics after post-processing. Romania (Cluj-Napoca) and Finland (Kajaani) occupy an intermediate range, whereas Italy (Piacenza), Estonia (Tallinn), Greece (Athens) and, to a lesser extent, Spain (Madrid-Barajas) and Belgium (Beauvechain) often show higher SMAE (up to ~2 or more at some leads). This confirms that, despite the overall modest skill, both correlation and error metrics display a clear geographic signal, with certain stations offering more favourable conditions for seasonal wind prediction than others.

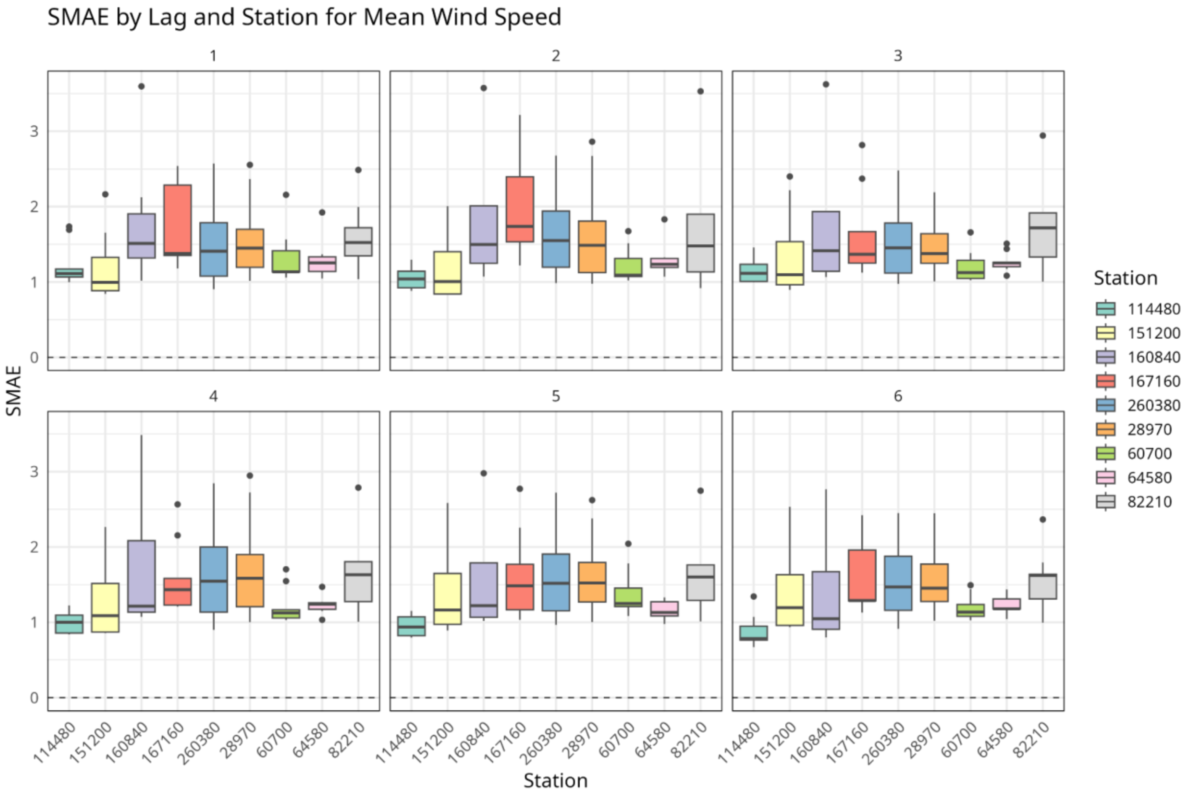


Figure A4.21. Combined SMAE for the observatories used for each lag using all models (Mean Wind Speed)

Seasonal forecast - CFSv2 models results

Pearson Correlation

Temperature (Maximum Temperature)

For daily maximum temperature, CFSv2 shows strong short-range skill, with median correlations around 0.6–0.7 on day 1 and a gradual decline to about 0.2 by days 10–12. Beyond the second week, correlations drop to low but still positive values (≈0.1–0.15) and remain in that range for most of the 51-day horizon, with a small secondary peak around days 40–45. This pattern reflects clear added value at synoptic to sub-monthly scales, but only modest, residual predictability at longer leads.

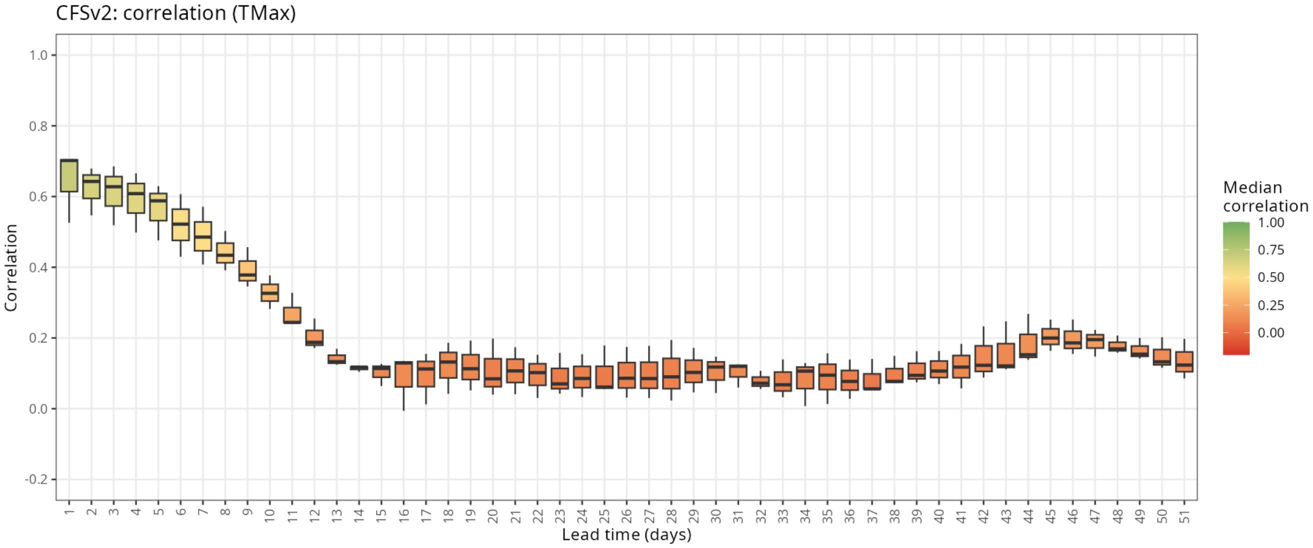


Figure A4.22. Pearson correlation for maximum temperature across the entire dataset of observations for each lead day.

Precipitation

For precipitation, correlations are much weaker and noisier. Median values start near 0.15–0.2 in the first few days but quickly decrease, reaching values close to zero around days 15–20 and hovering near or slightly below zero for most of the remaining lead times. Only a small and spatially variable positive signal re-emerges around days 40–50. Overall, this indicates very limited deterministic skill for daily rainfall anomalies beyond roughly two weeks.

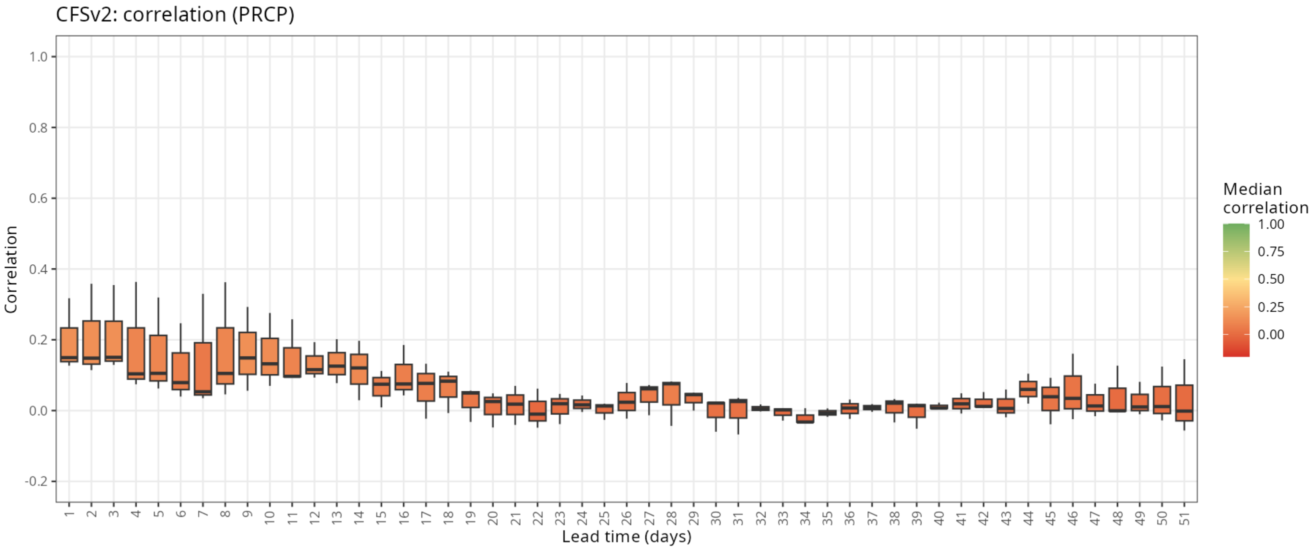


Figure A4.23. Pearson correlation for precipitation across the entire dataset of observations for each lead day.

Mean Wind Speed

Regarding wind speed, CFSv2 exhibits good correlation at short lead times (up to 5–7 days), with values between 0.3 and 0.6, indicating that the model captures reasonably well the temporal evolution of wind anomalies.

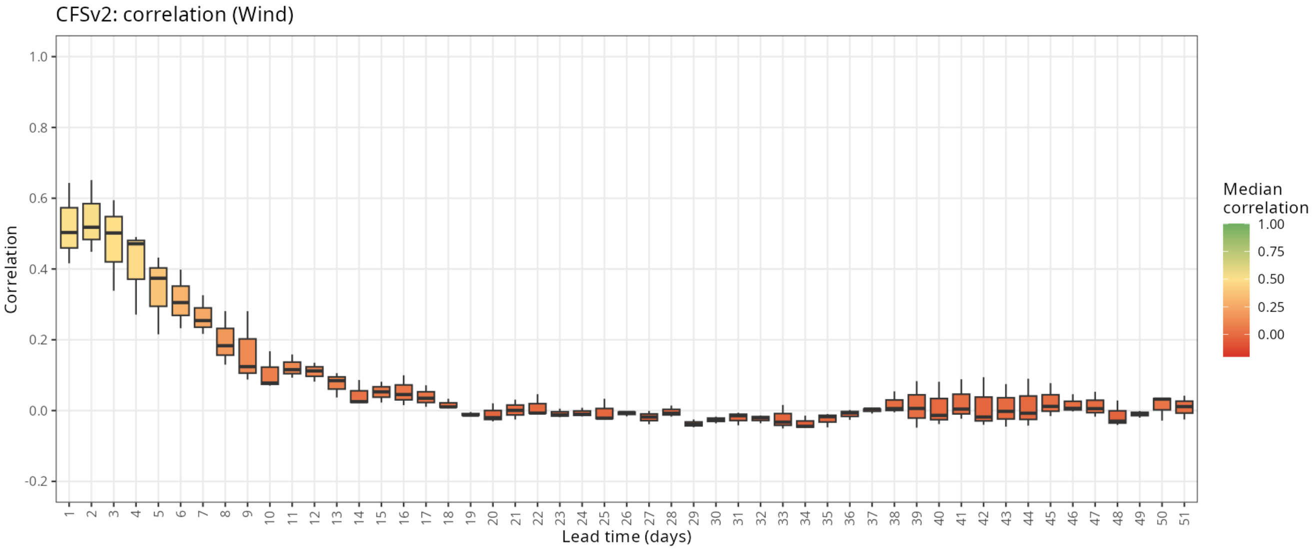


Figure A4.24. Pearson correlation for Mean Wind Speed across the entire dataset of observations for each lead day.

Standardized Mean Absolute Error

Temperature (Maximum Temperature)

The standardized MAE for maximum temperature is lowest at very short leads (≈0.5 on day 1) and increases steadily during the first two to three weeks, stabilizing around SMAE ≈0.9–1.0 afterwards. This behaviour indicates that CFSv2 provides clear improvements over climatology at short lead times, while errors approach climatological levels as the forecast extends into the subseasonal window, with relatively small changes in SMAE beyond about day 20.

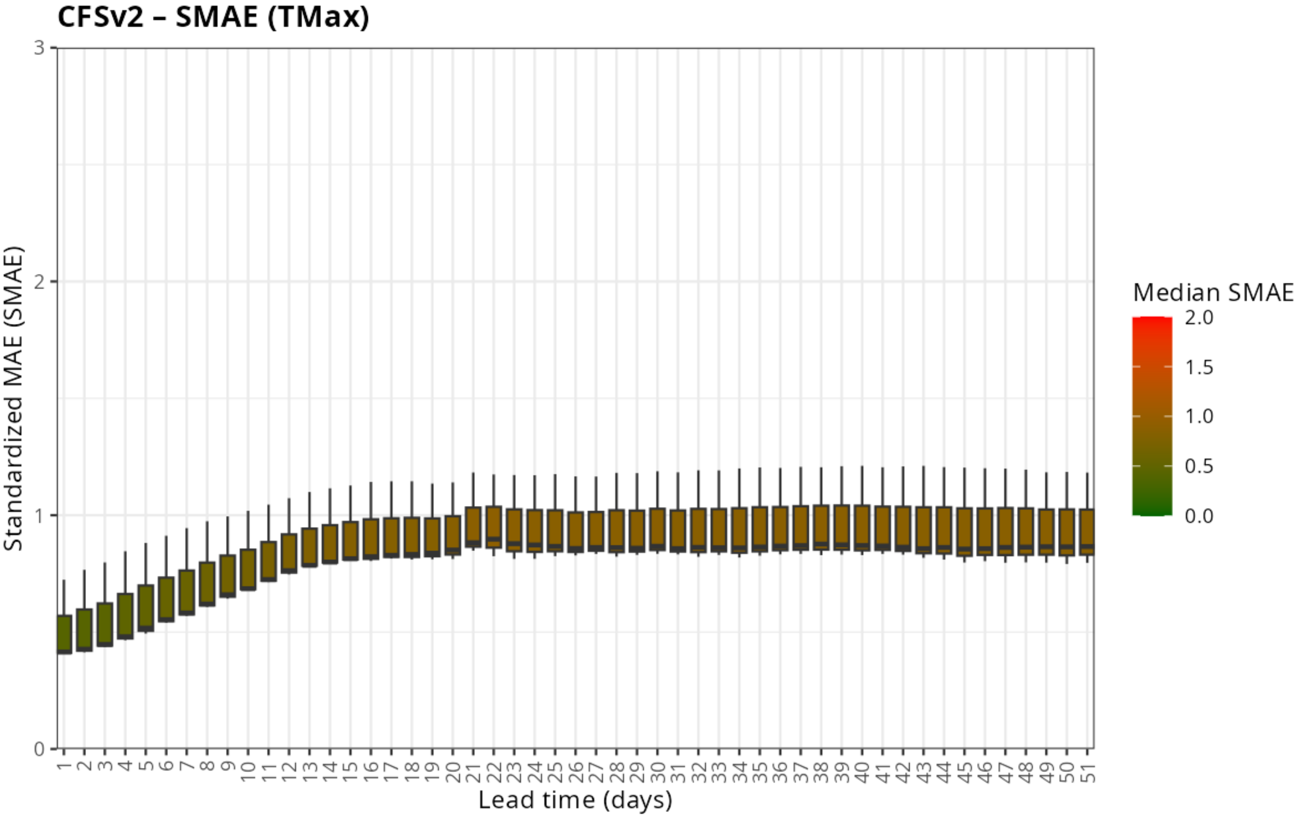


Figure A4.25. SMAE for maximum temperature across the entire dataset of observations for each lead day.

Precipitation (PRCP)

For precipitation, SMAE values are higher and more variable, with medians typically between about 0.8 and 1.3 across the 51-day horizon. Errors are comparatively lower during the first 10–15 days, then increase and peak around days 20–30, where some stations approach standardized errors near 1.4–1.5, before decreasing slightly again toward the end of the period. This confirms that daily rainfall forecasts from CFSv2 rapidly become noisy, with only modest and transient improvements over a climatological reference.

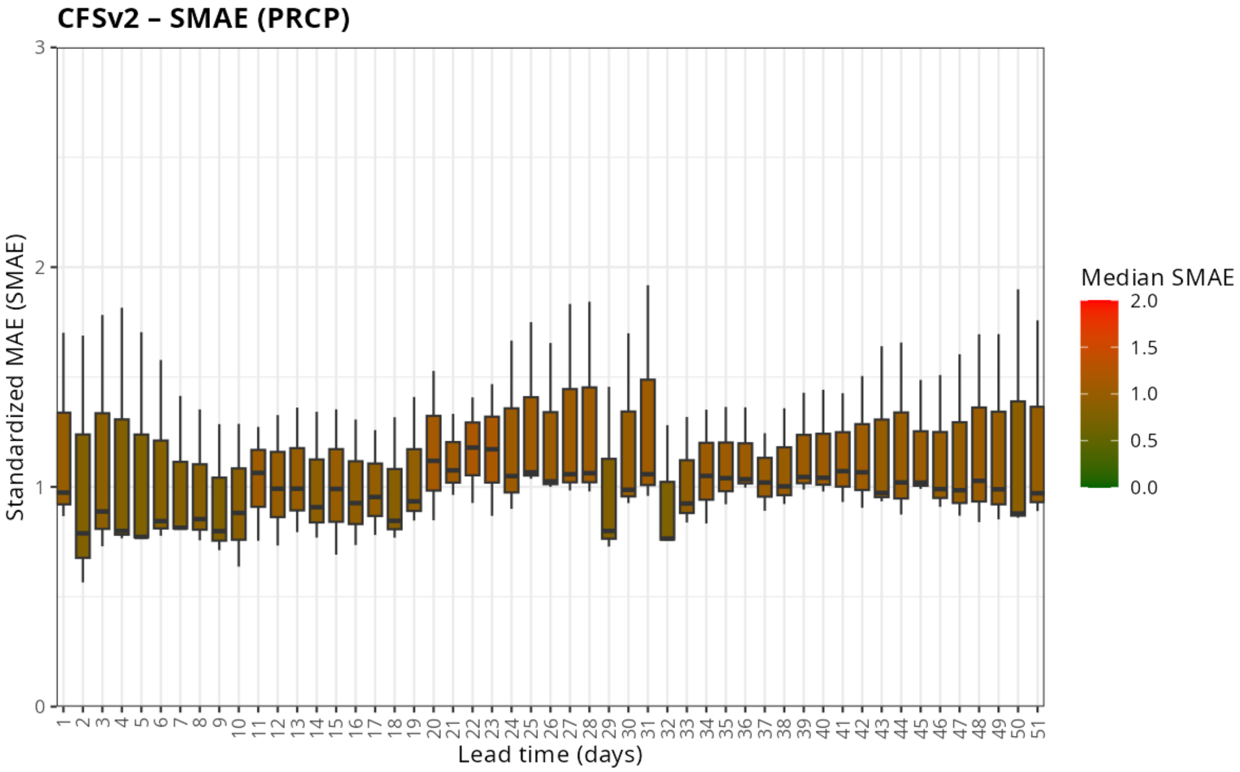


Figure A4.26. SMAE for precipitation across the entire dataset of observations for each lead day.

Mean Wind Speed

For 10-m wind speed, CFSv2 standardized MAE remains high at these ranges and only decreases towards values close to 1 after 10 days.

This behaviour is consistent with a forecast that reproduces the synoptic-scale variability but suffers from substantial local bias and amplitude errors at station scale: the model tends to over- or under-estimate wind speeds, so absolute errors are large even when anomalies co-vary with observations. As lead time increases, the forecast gradually damps towards climatology, which reduces the amplitude of forecast anomalies and hence the absolute error, leading to smaller SMAE values despite the loss of temporal correlation.

This pattern is typical for near-surface wind in global models at relatively coarse resolution, where unresolved local topography and surface effects limit the skill of point-scale wind predictions, especially beyond 1–2 weeks.

Annex 5. Climate projection results

Climate Change Figures Interpretation

The expected temporal evolution of the provided variables is represented by box-and-whisker plots, which illustrate the expected changes in these variables throughout the 21st century in relation to the chosen reference period (1985-2014) (e.g., see Figure A5.0). The temporal evolution is displayed along the X-axis in the form of three sets of box-and-whisker plots, with each set corresponding to a different time period, from left to right: the reference period (1985-2014), mid-century (2036-2065), and end of the century (2071-2100). The Y-axis displays the values of the simulated variable, in this case, the mean annual temperature.

These plots provide a statistical representation of the distribution of the obtained values, allowing for observation of both the median (the value at the centre of the distribution) and the dispersion of the other values, which enables the identification of the sample variability and the most expected and/or atypical values of the sample. The central horizontal line within each box represents the median (2nd quartile or 50th percentile). This value indicates that 50% of all values obtained for each year within the period and for each of the climate models considered within each scenario are above that value, and the remaining 50% are below that value. The lower and upper limits of the box correspond to the 1st and 3rd quartiles, respectively, or the 25th and 75th percentiles. The lower limit indicates that 25% of the values are below it, and the upper limit indicates that 75% of the values are below it. Thus, it can be said that 50% of the values lie within the box, which represents the most expected values. Finally, the whiskers determine the limit for detecting outliers in the distribution, which are shown as isolated points aligned with the whiskers (lower and upper limits).

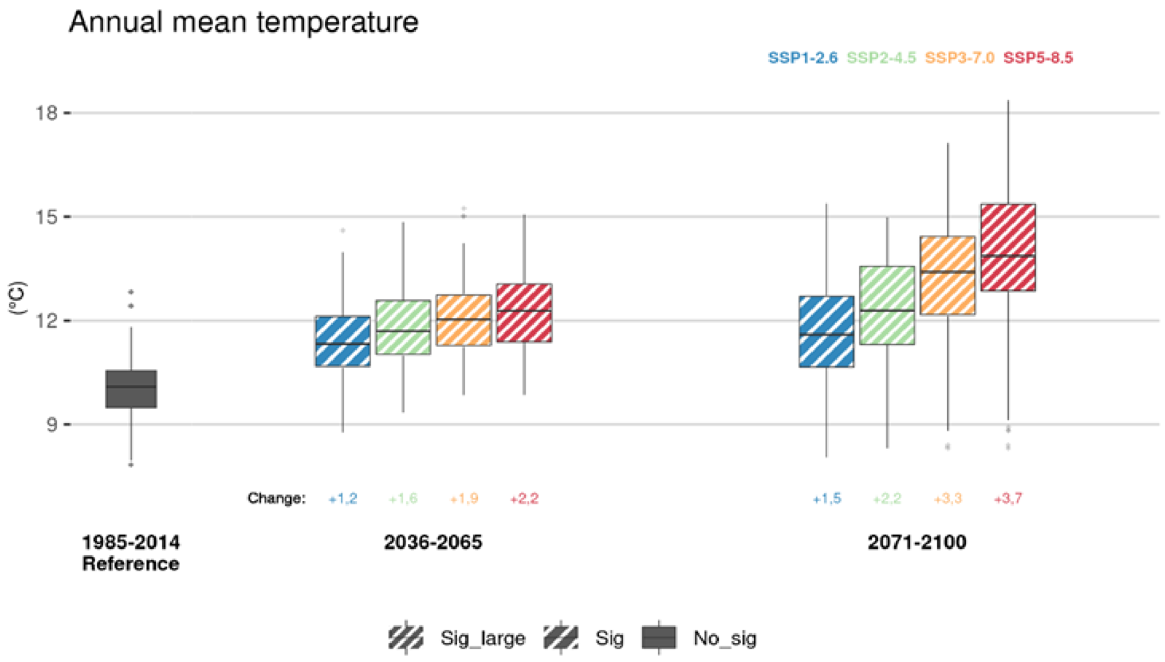


Figure A5.0. Graphic example: Expected evolution for the annual mean temperature variable.

On the X-axis, three-time blocks are displayed with a different number of box and whisker diagrams, as well as various colours and patterns. The reference period is represented by a single diagram in a greyish colour, which corresponds to the values of the variable obtained using the historical simulation of the different CMIP6 models used in this study. For mid- and end-century, four box-and-whisker plots are shown. Each plot represents the values of the variable under different emission scenarios, projected by the CMIP6 models used in this study. The scenario names are a combination of two indices: a baseline socio-economic trajectory "index" (ranging from SSP1: sustainability to SSP5: fossil fuel development), and the level of radiative forcing in 2100 (ranging from 2.6 to 8.5 W/m2) (O'Neill et al., 2016).

The results obtained for each of the 4 scenarios considered are arranged in order of pessimism: from the most optimistic to the most pessimistic scenario. The order is the same for both the middle and the end of the century: SSP1-2.6 in blue, SSP2-4.5 in green, SSP3-7.0 in yellow, and SSP5-8.5 in red. Each of these box plots corresponding to future developments can present three different patterns depending on whether the expected change in the variable is significant or not with respect to the values obtained in the reference period according to the Wilcoxon Mann Whitney test. They also depend on whether the difference in size between the two series of values is small, moderate or, large according to the Wilcoxon size effect. Thus, the box-and-whisker plots of those changes that are considered non-significant or small-scale significant will show no pattern at all. Those that are considered moderate-scale significant will show a pattern of low-density diagonal lines. Finally, those that are considered large-scale significant will show a pattern of high-density diagonal lines (as is currently the case for all expected changes for both mid- and end-century).

Finally, the expected values of change at mid- and end-century are specified for each of the future emission scenarios. These values are obtained by calculating the difference between the median value of the future scenario in question with respect to the median value of the reference period. They are represented numerically (showing increases or decreases) below the diagrams corresponding to each emission scenario and future period.

Results for Madrid

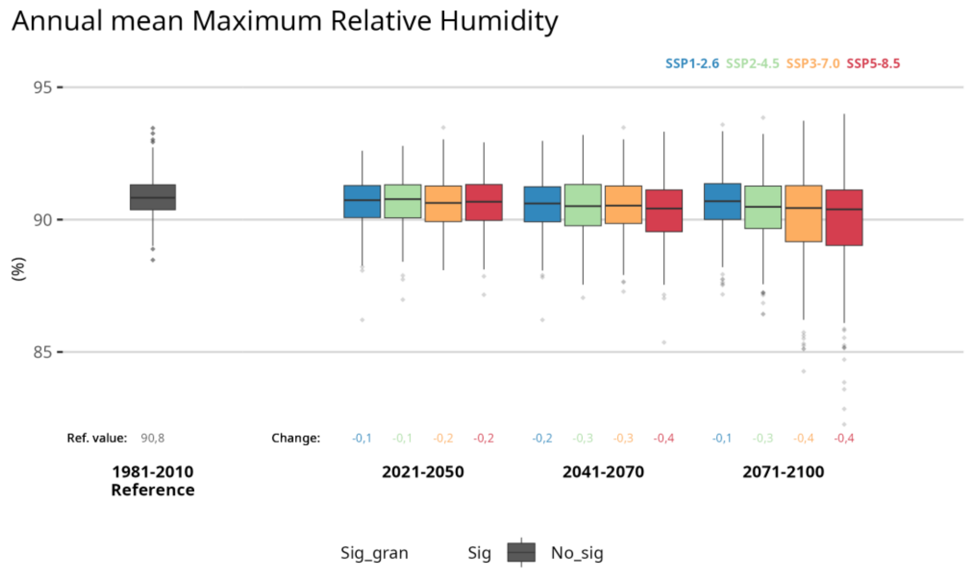


Figure A5.1. *The graph shows the evolution of maximum relative humidity throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the 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.*

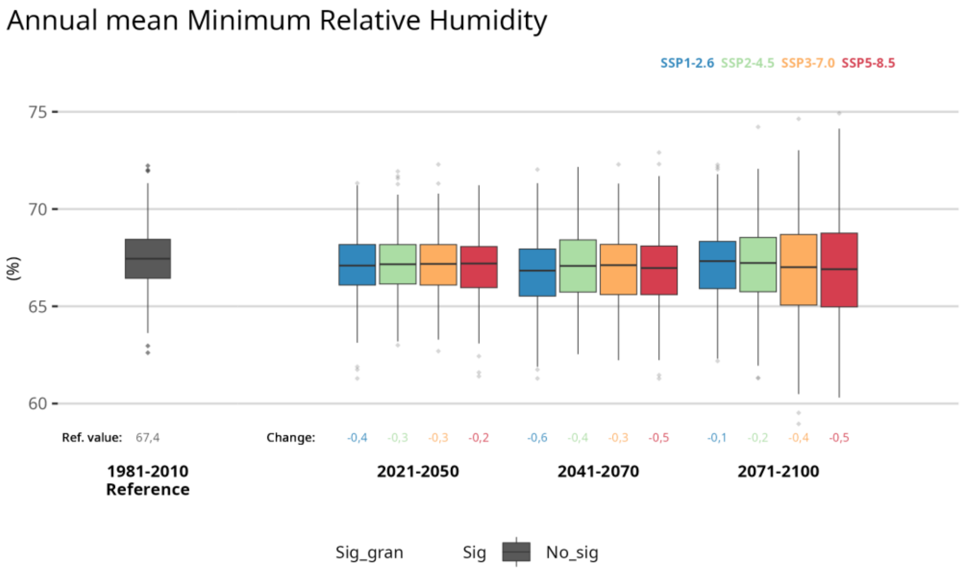


Figure A5.2. *The graph shows the evolution of minimum relative humidity throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the 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.*

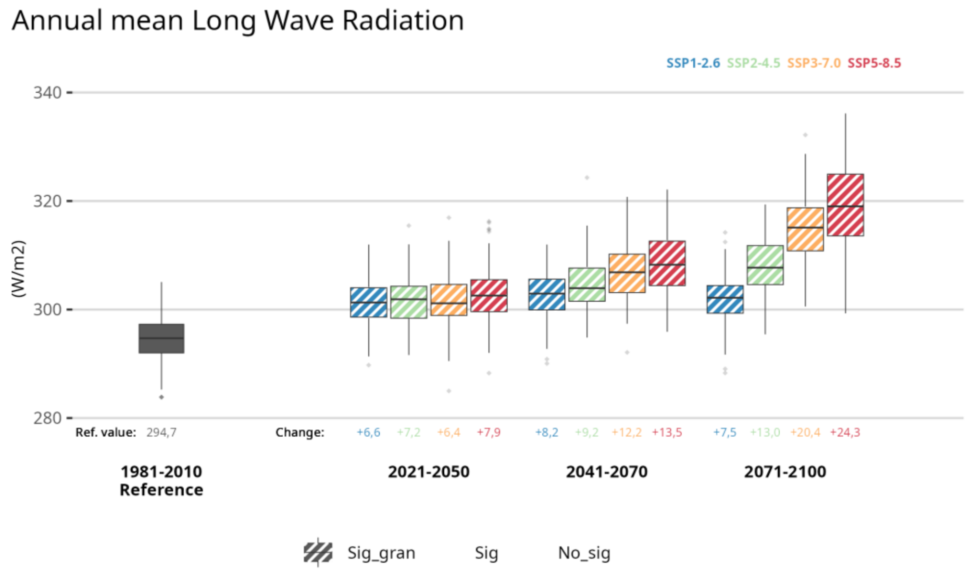


***Figure A5.3**. The graph shows the evolution of long wave radiation throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

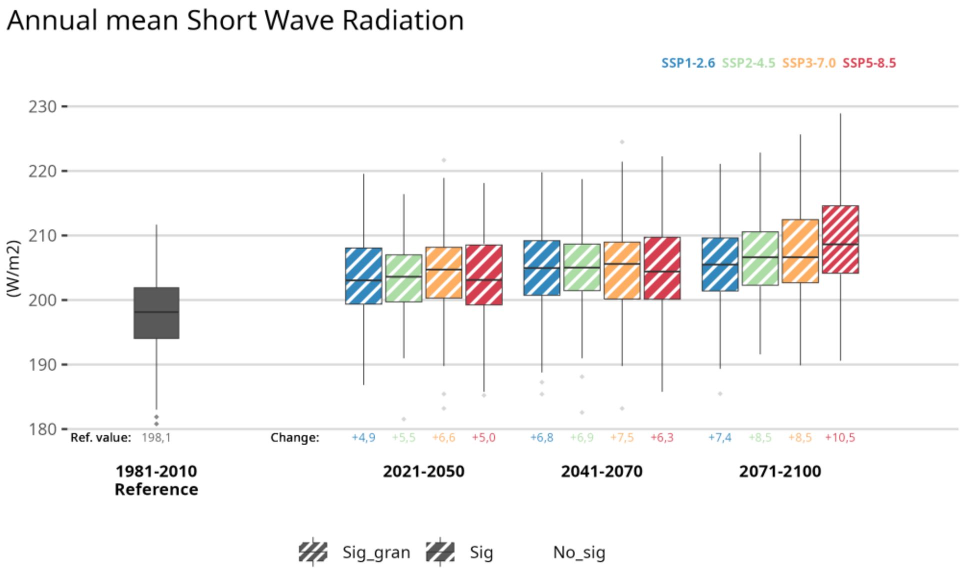


***Figure A5.4**. The graph shows the evolution of short wave radiation throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the 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.*

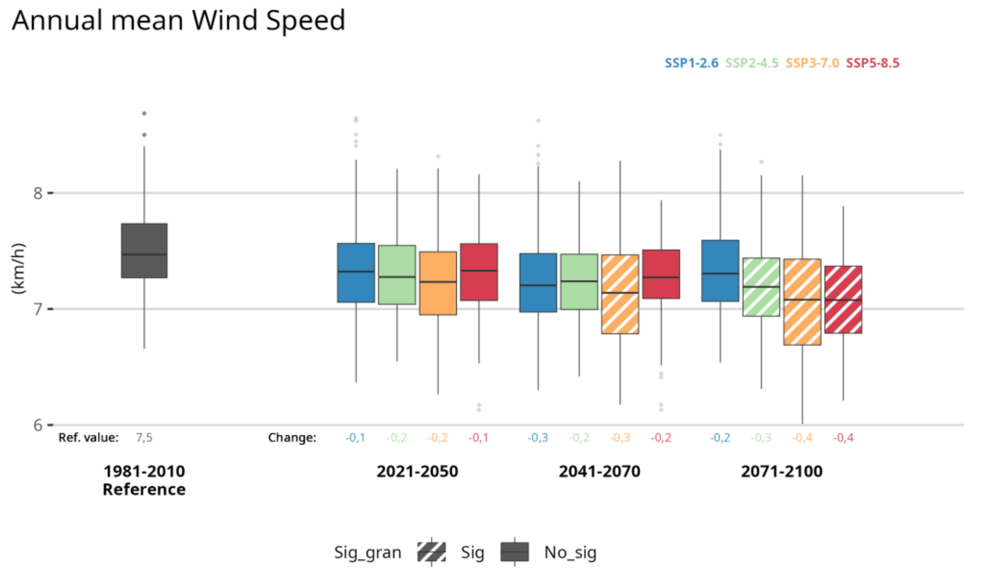


***Figure A5.5**. The graph shows the evolution of wind speed throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

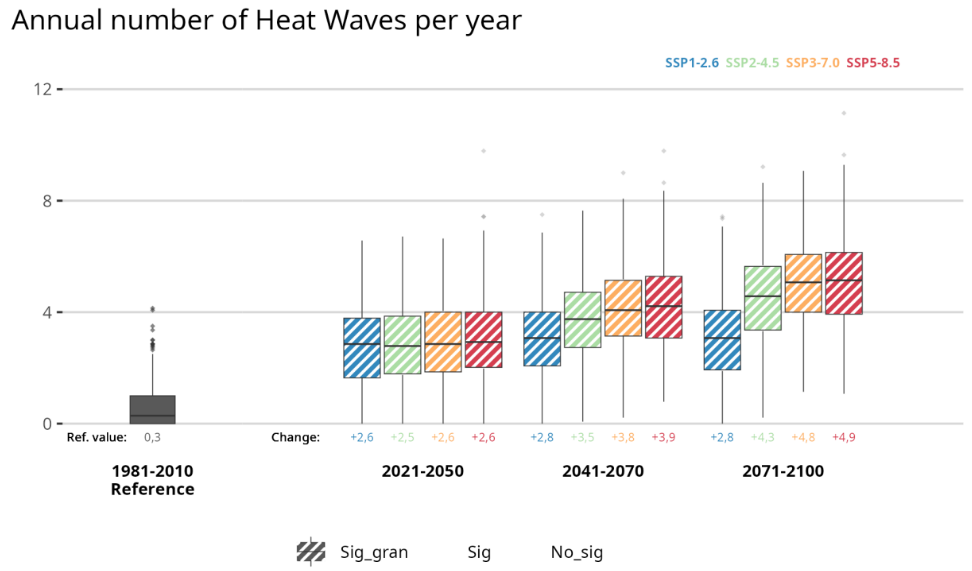


***Figure A5.6**. The graph shows the evolution of annual number of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the 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.*

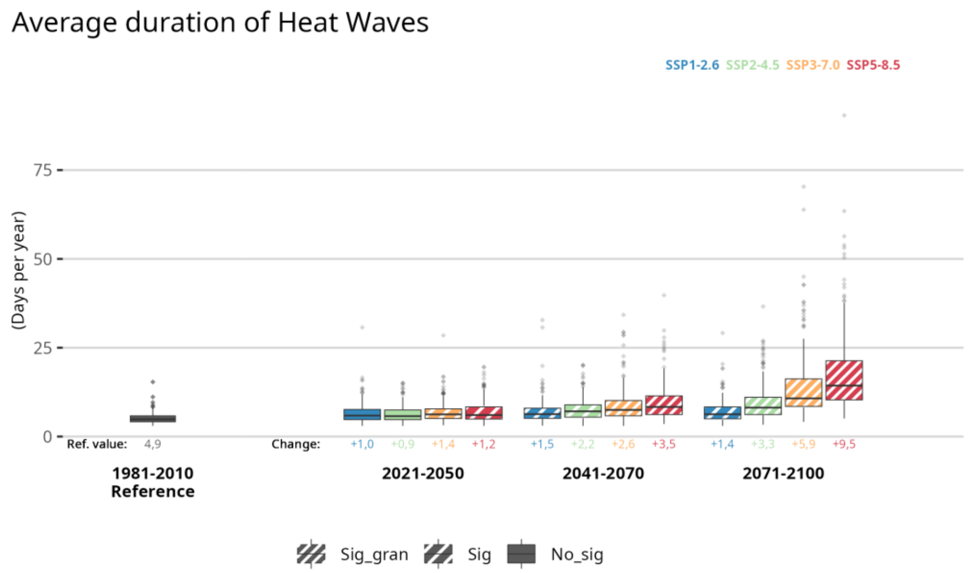


Figure A5.7. *The graph shows the evolution of average duration of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the 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.*

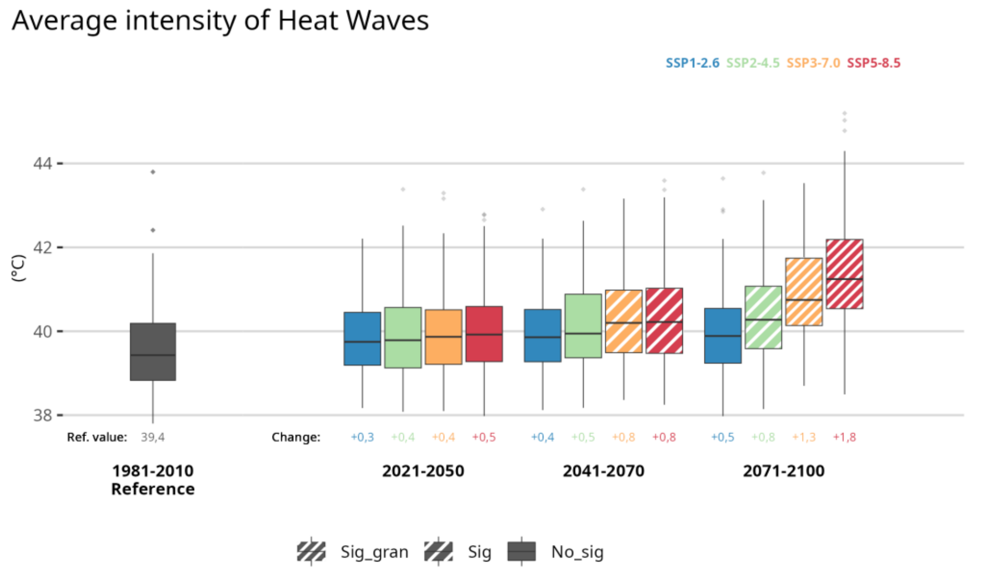


Figure A5.8. *The graph shows the evolution of average intensity of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the 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.*

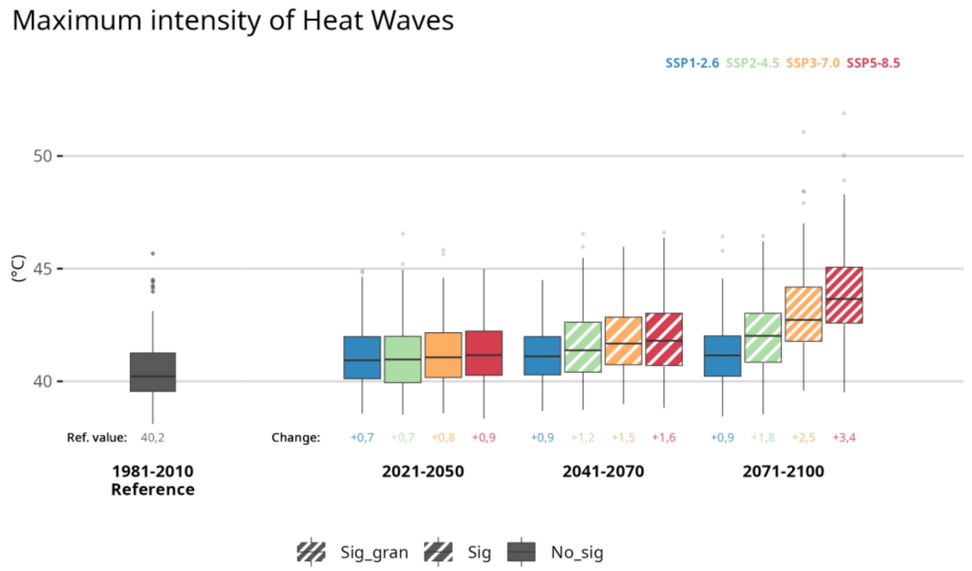
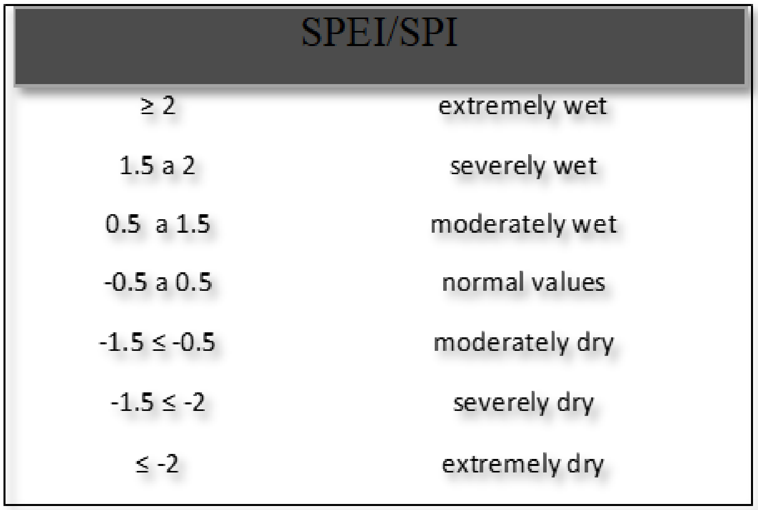


***Figure A5.9**. The graph shows the evolution of maximum intensity of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for 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.*

Table A5.1. *SPI and SPEI values classification table.*



SPEI/SPI	
≥ 2	extremely wet
1.5 a 2	severely wet
0.5 a 1.5	moderately wet
-0.5 a 0.5	normal values
-1.5 ≤ -0.5	moderately dry
-1.5 ≤ -2	severely dry
≤ -2	extremely dry

Figure A5.10*. The graph shows the evolution of 3-month SPI index 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.*

***Figure A5.11**. The graph shows the evolution of 6-month SPI index 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.*

Figure A5.12*. The graph shows the evolution of 12-month SPI index 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.*

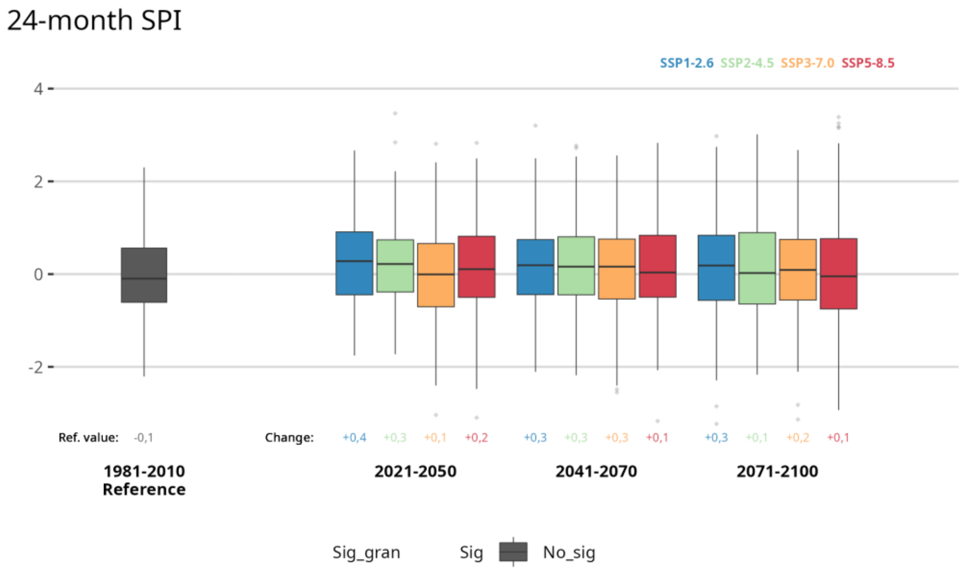


Figure A5.13. *The graph shows the evolution of 24-month SPI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

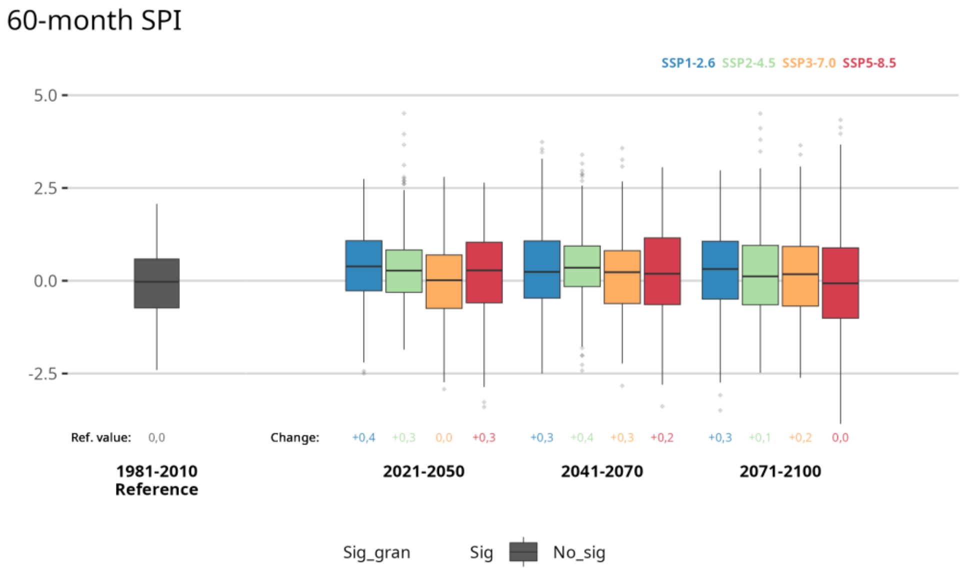


Figure A5.14. *The graph shows the evolution of 60-month SPI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

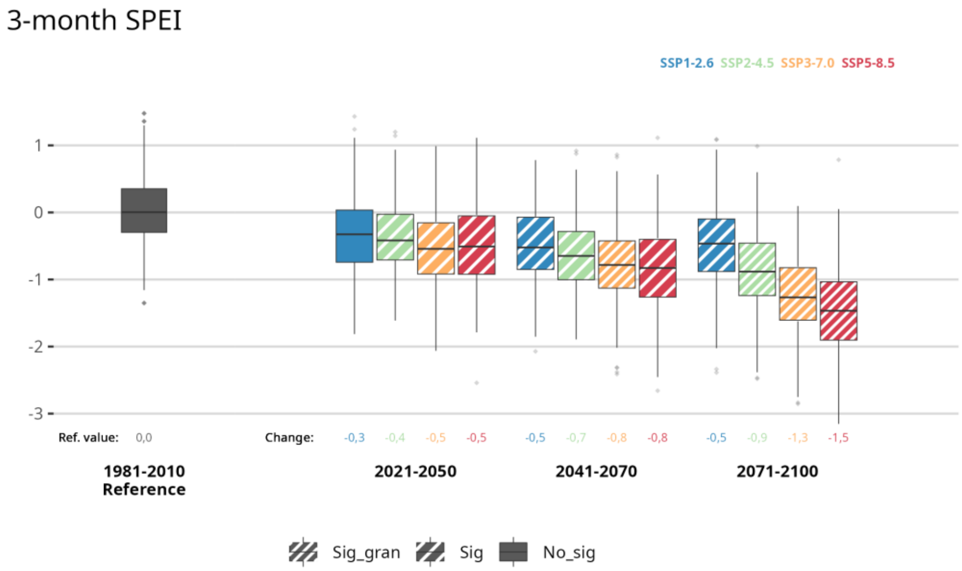


Figure A5.15. *The graph shows the evolution of 3-month SPEI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

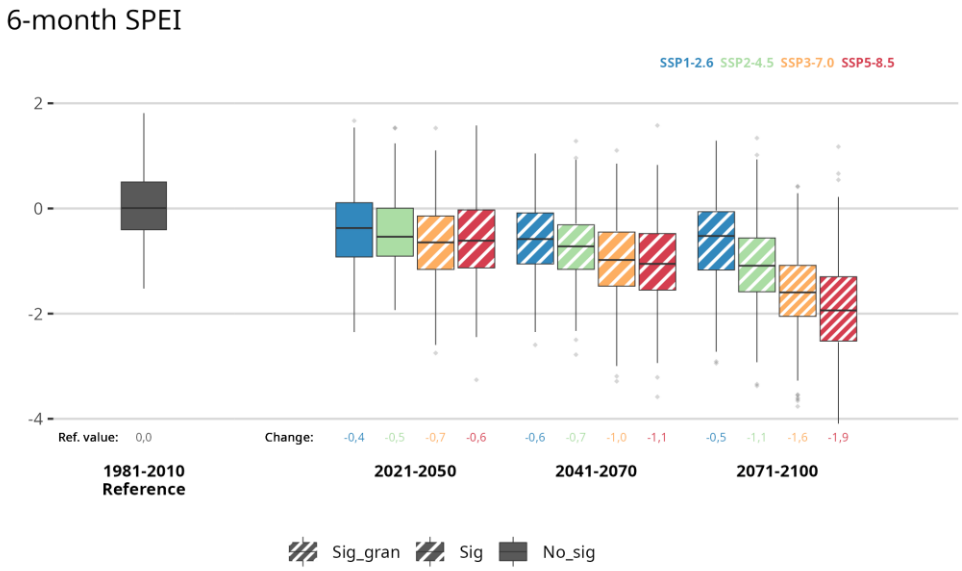


Figure A5.16. *The graph shows the evolution of 6-month SPEI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

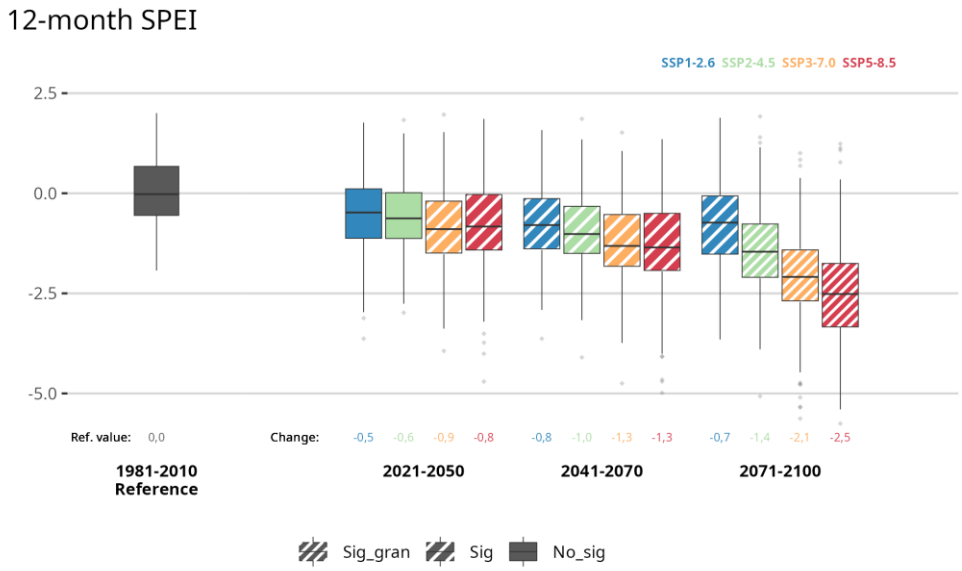


Figure A5.17. *The graph shows the evolution of 12-month SPEI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

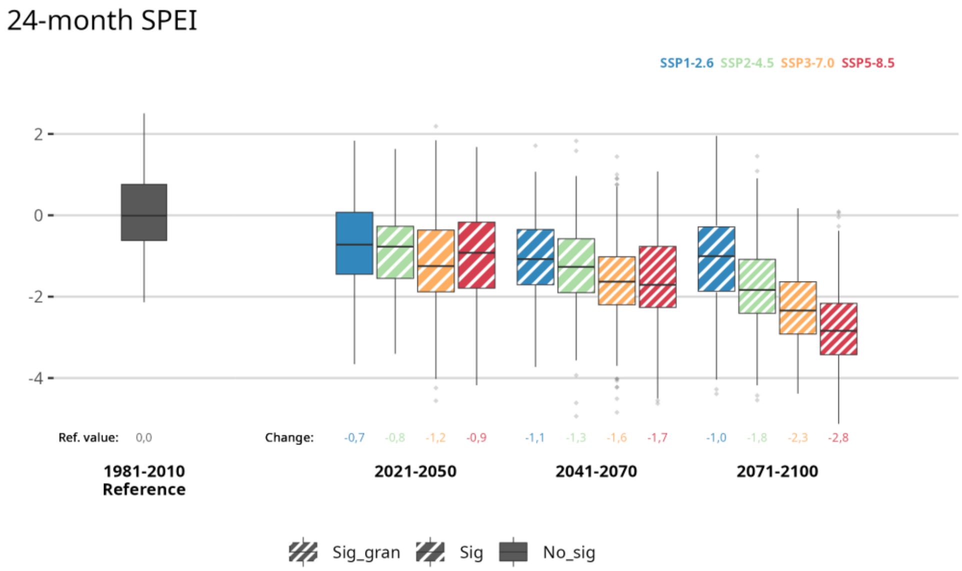


***Figure A5.18**. The graph shows the evolution of 24-month SPEI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the 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.*

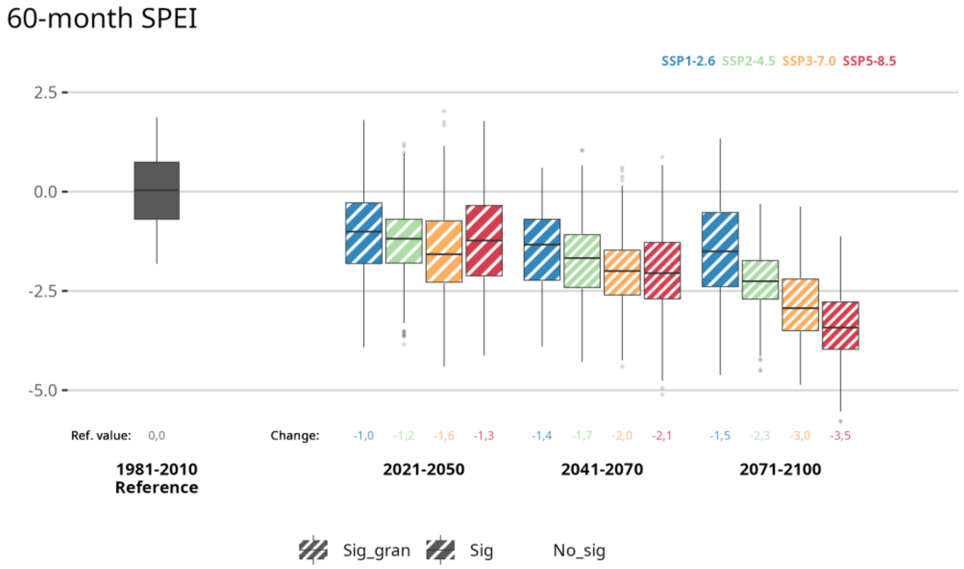


Figure A5.19. *The graph shows the evolution of 60-month SPEI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

Results for Leuven

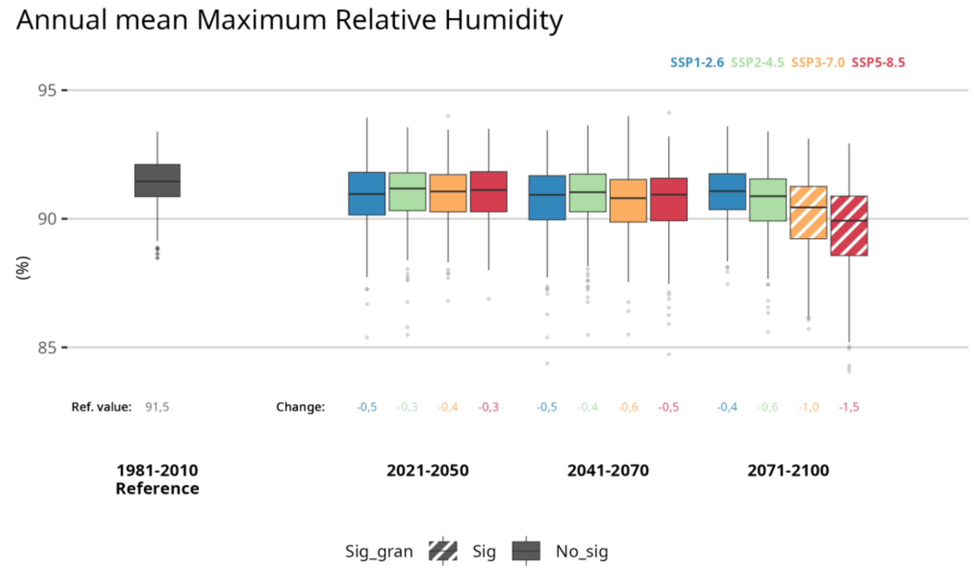


Figure A5.20*. The graph shows the evolution of maximum relative humidity throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the 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.*

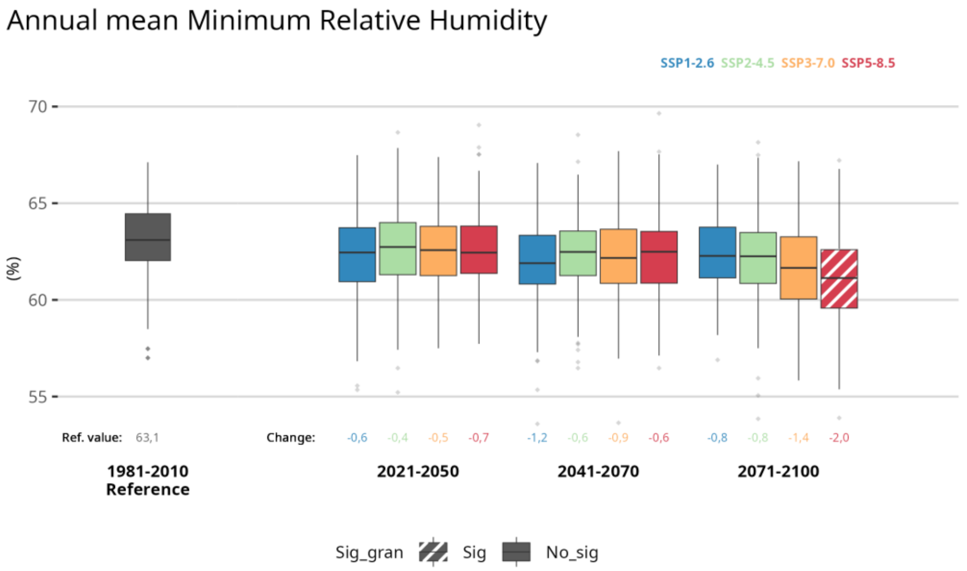


Figure A5.21*. The graph shows the evolution of minimum relative humidity throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the 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.*

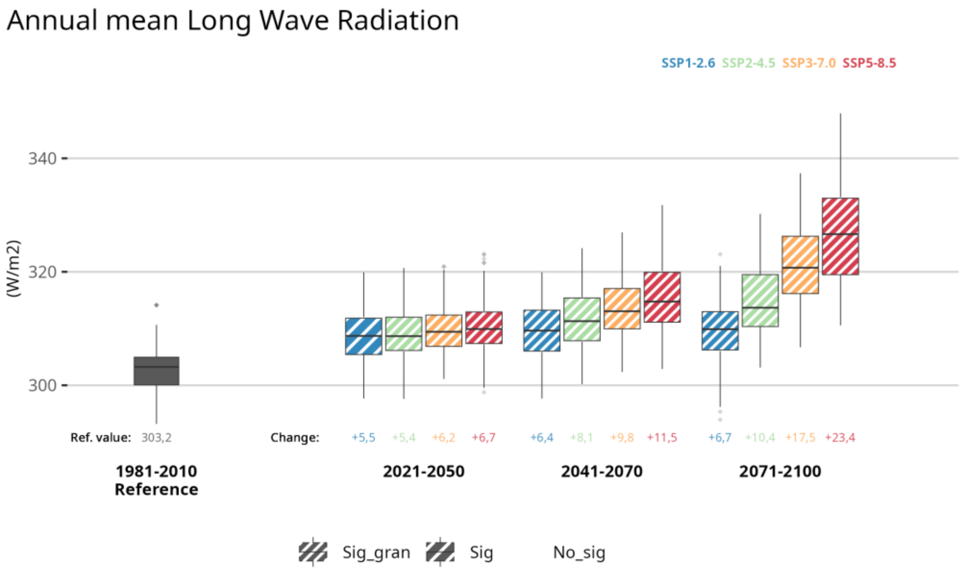


***Figure A5.22**. The graph shows the evolution of long wave radiation throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the 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.*

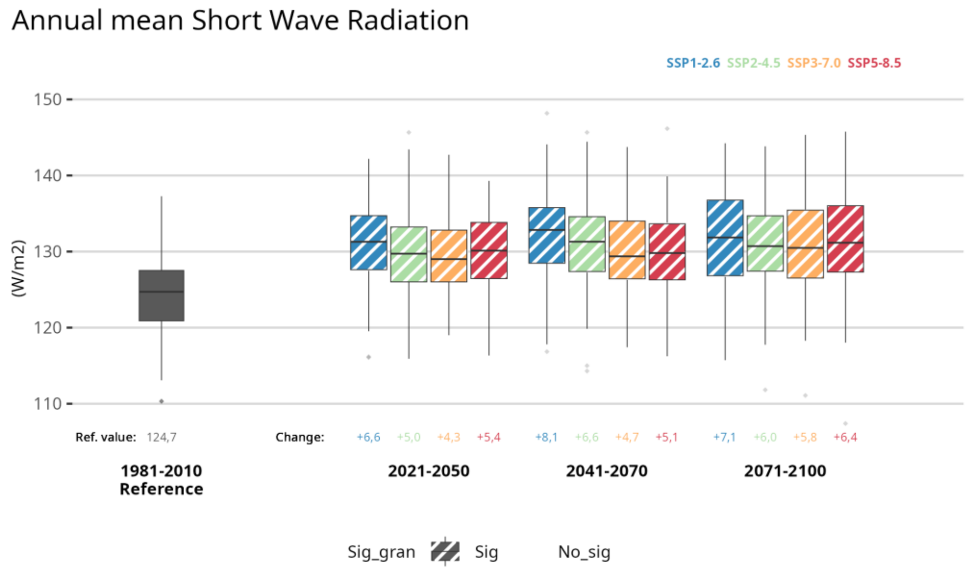


***Figure A5.23**. The graph shows the evolution of short wave radiation throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the 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.*

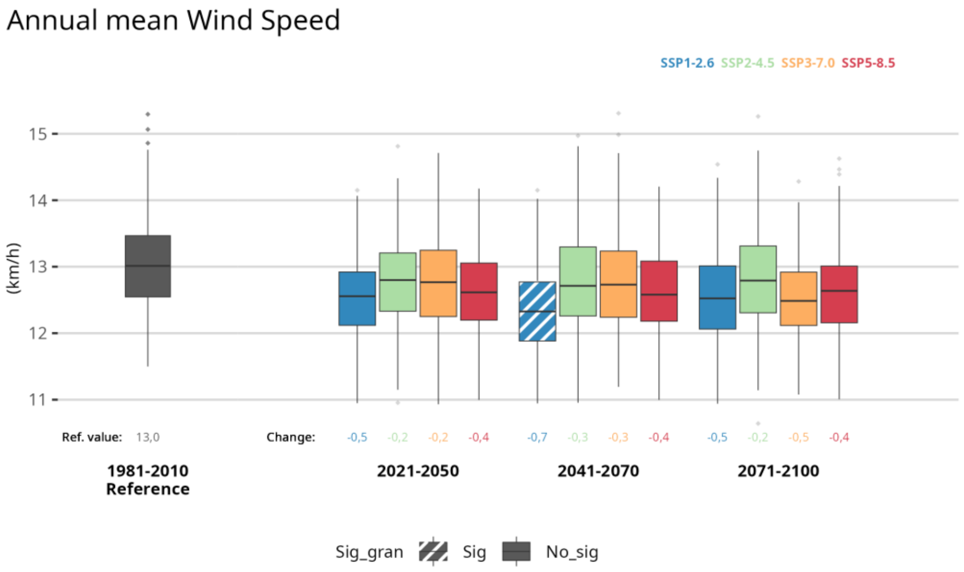


***Figure A5.24**. The graph shows the evolution of wind speed throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

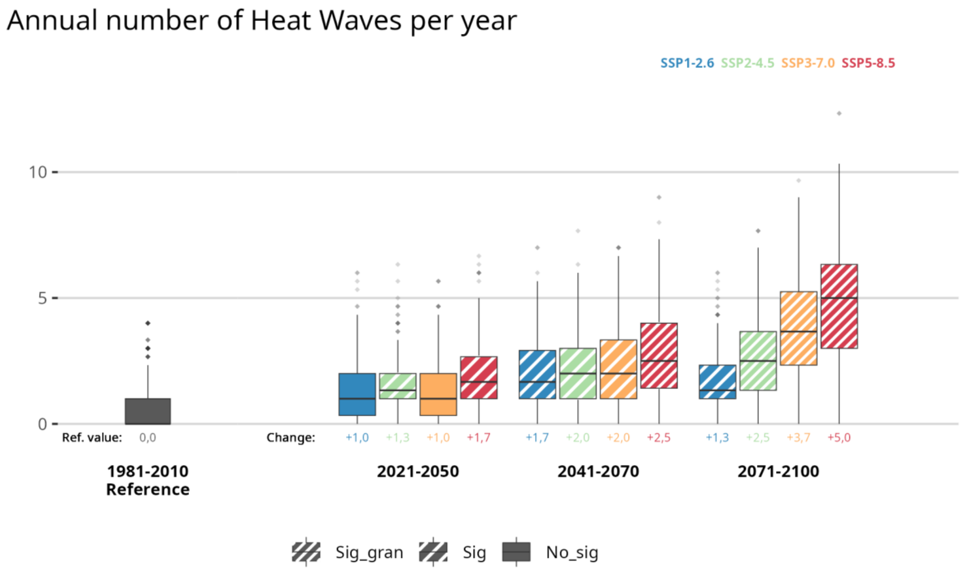


Figure A5.25. *The graph shows the evolution of annual number of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set 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.*

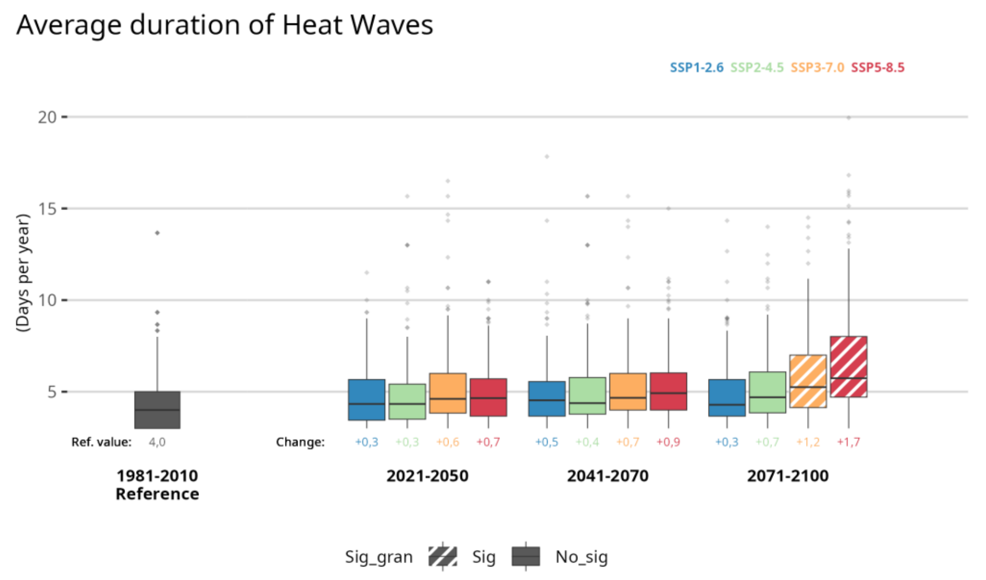


Figure A5.26. *The graph shows the evolution of average duration of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the 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.*

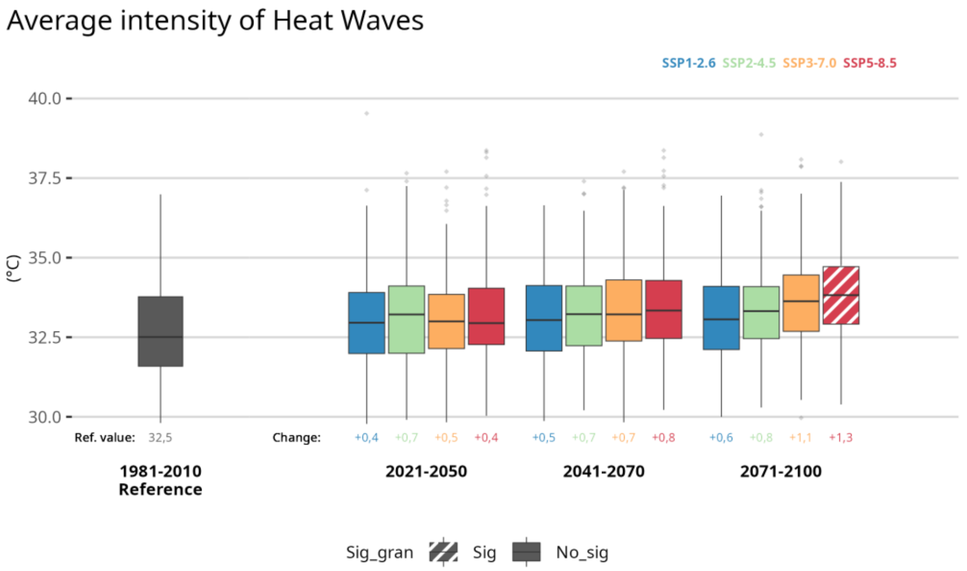


Figure A5.27. *The graph shows the evolution of average intensity of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the 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.*

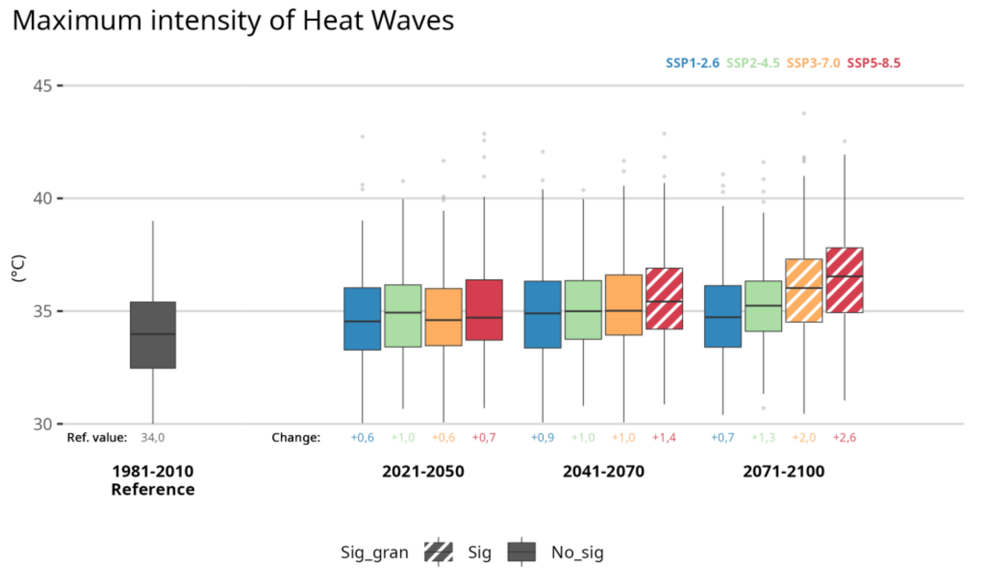


Figure A5.28. *The graph shows the evolution of maximum intensity of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the 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.*

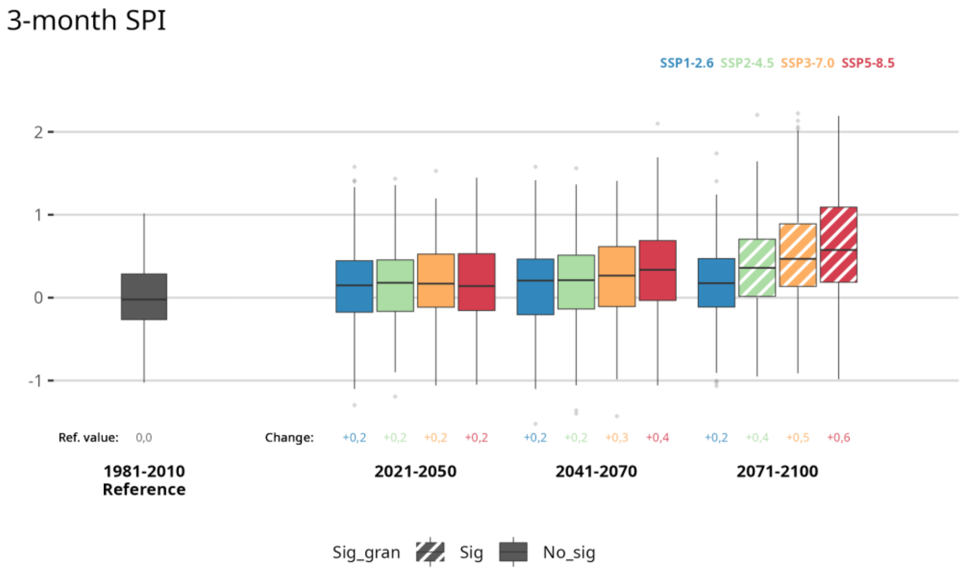


***Figure A5.29**. The graph shows the evolution of 3-month SPI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

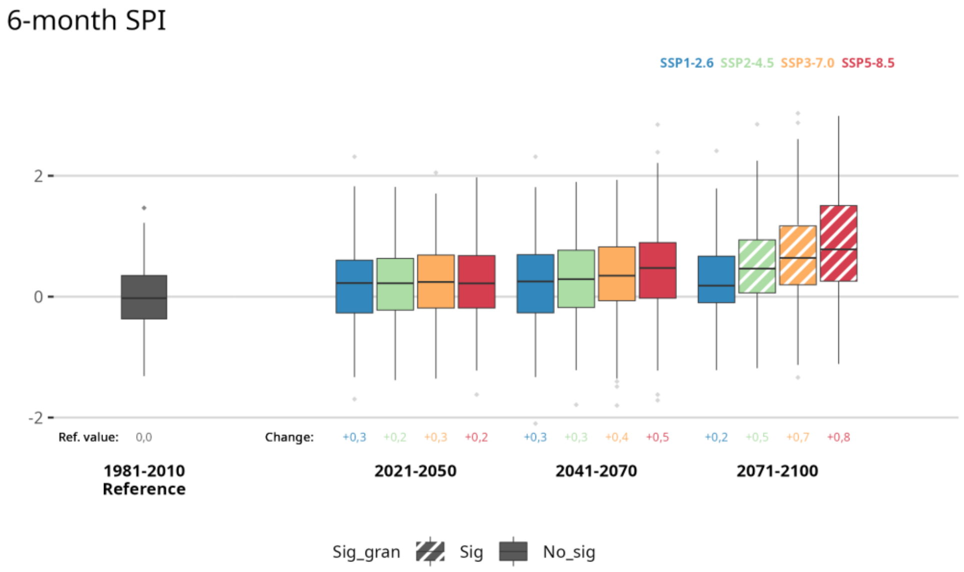


Figure A5.30. *The graph shows the evolution of 6-month SPI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

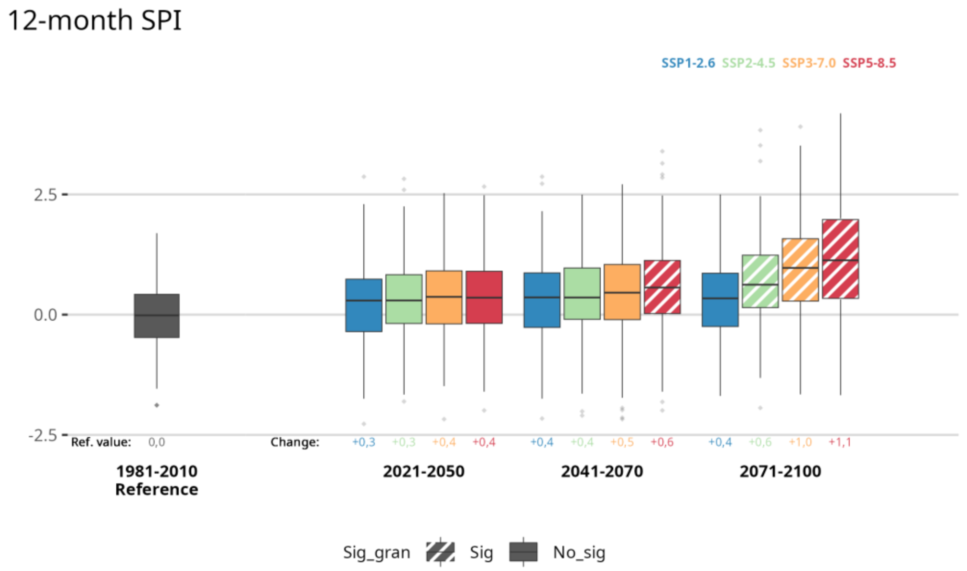


***Figure A5.31**. The graph shows the evolution of 12-month SPI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

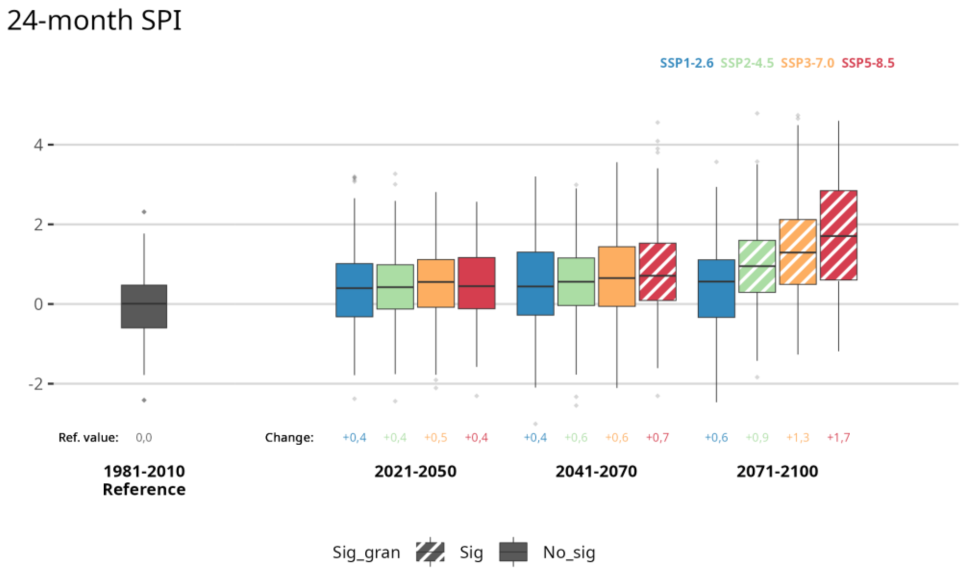


Figure A5.32. *The graph shows the evolution of 24-month SPI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

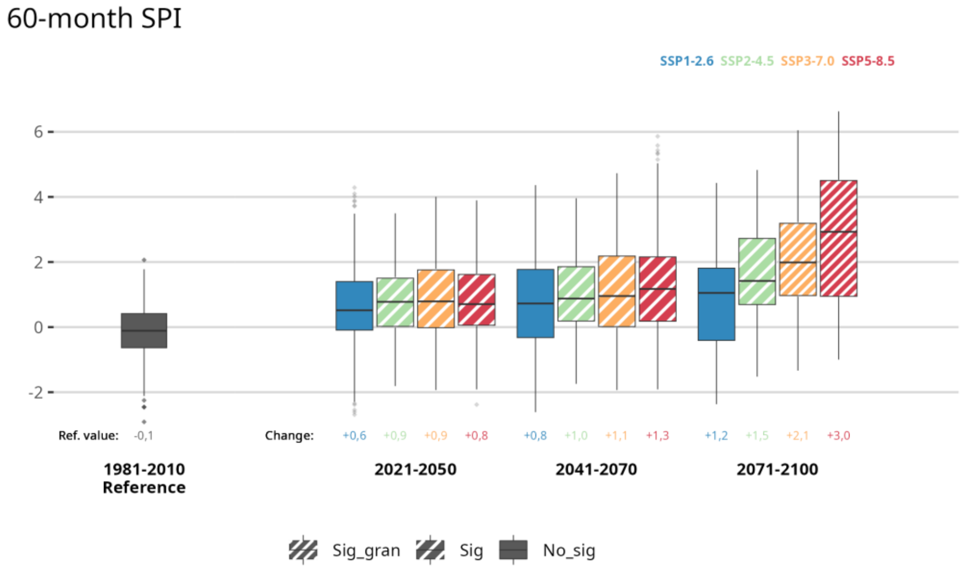


Figure A5.33. *The graph shows the evolution of 60-month SPI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

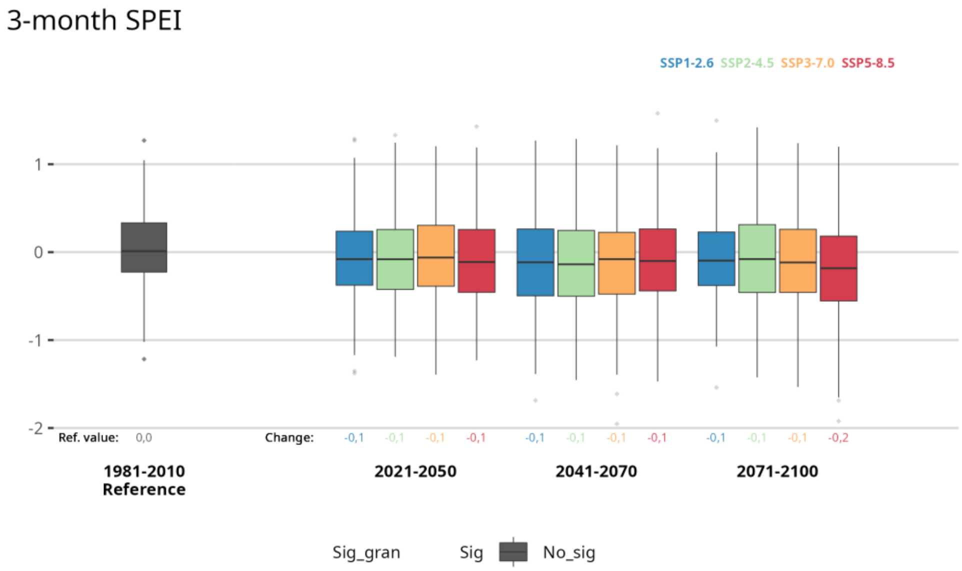


Figure A5.34. *The graph shows the evolution of 3-month SPEI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

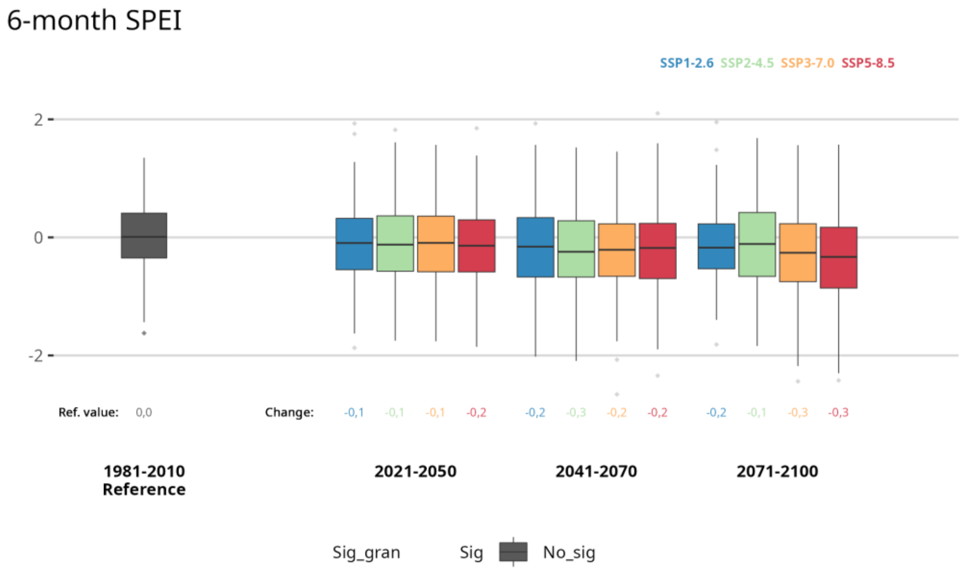


Figure A5.35. *The graph shows the evolution of 6-month SPEI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

Figure A5.36. *The graph shows the evolution of 12-month SPEI index 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.*

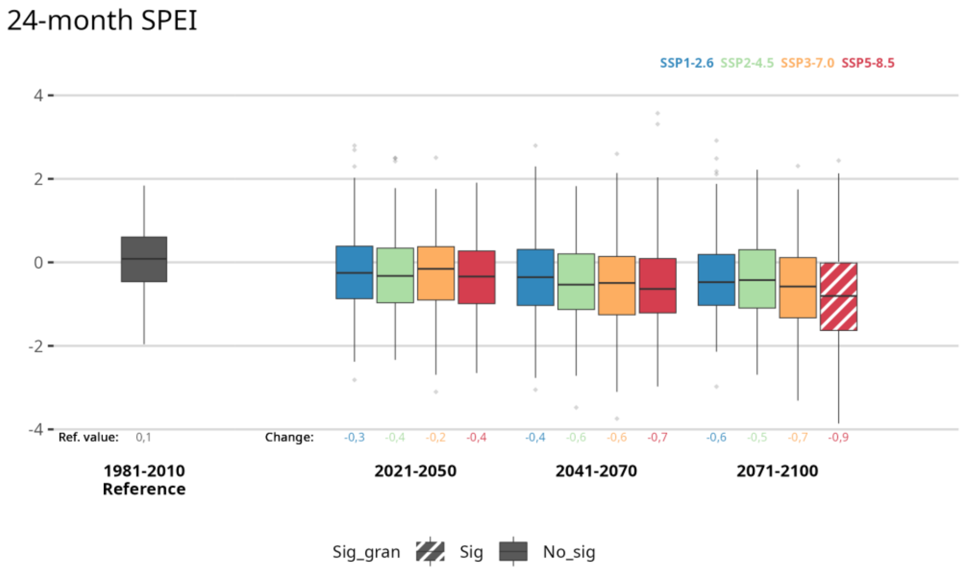

24-month SPEI

***Figure A5.37**. The graph shows the evolution of 24-month SPEI index 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.*

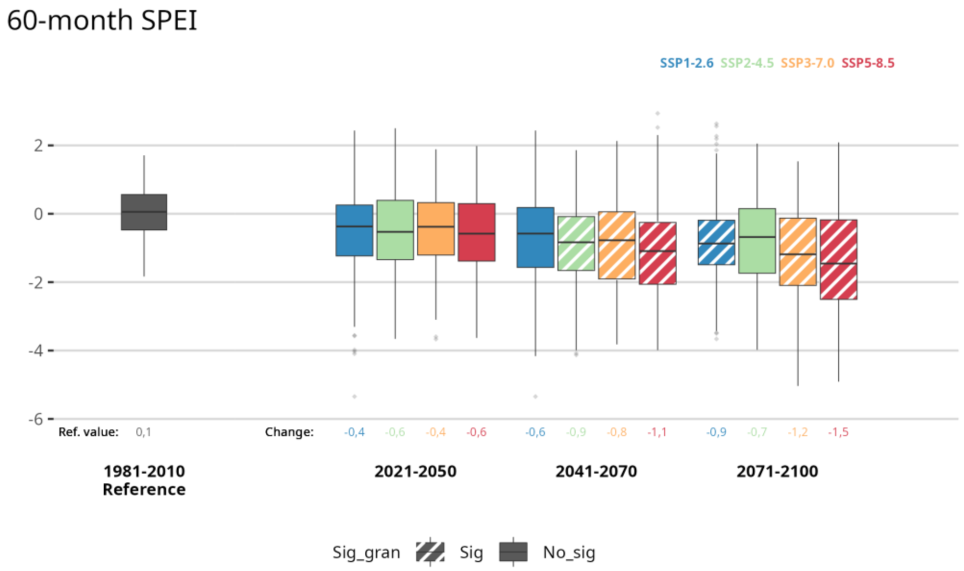


Figure A5.38. *The graph shows the evolution of 60-month SPEI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

Results for Tallinn

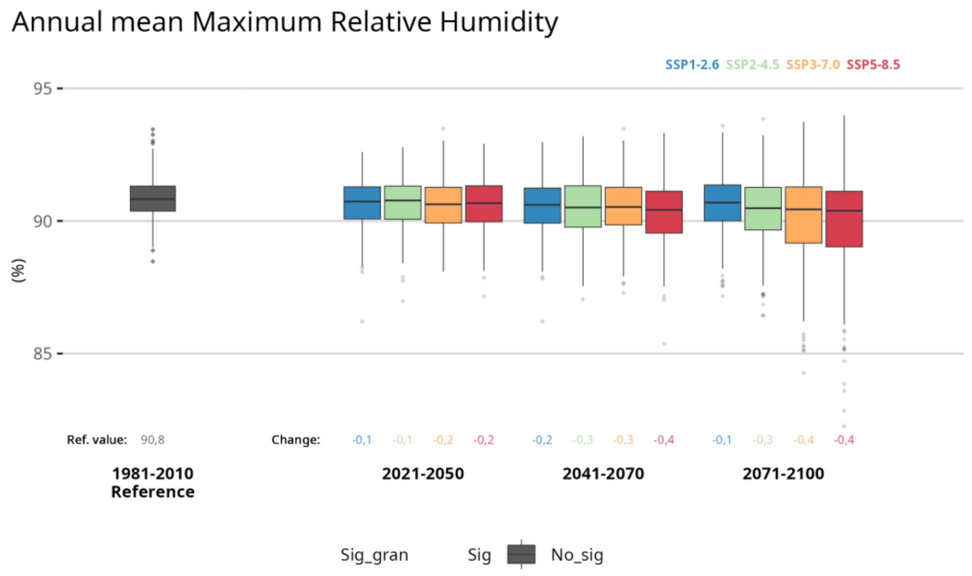


Figure A5.39. *The graph shows the evolution of maximum relative humidity throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the 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.*

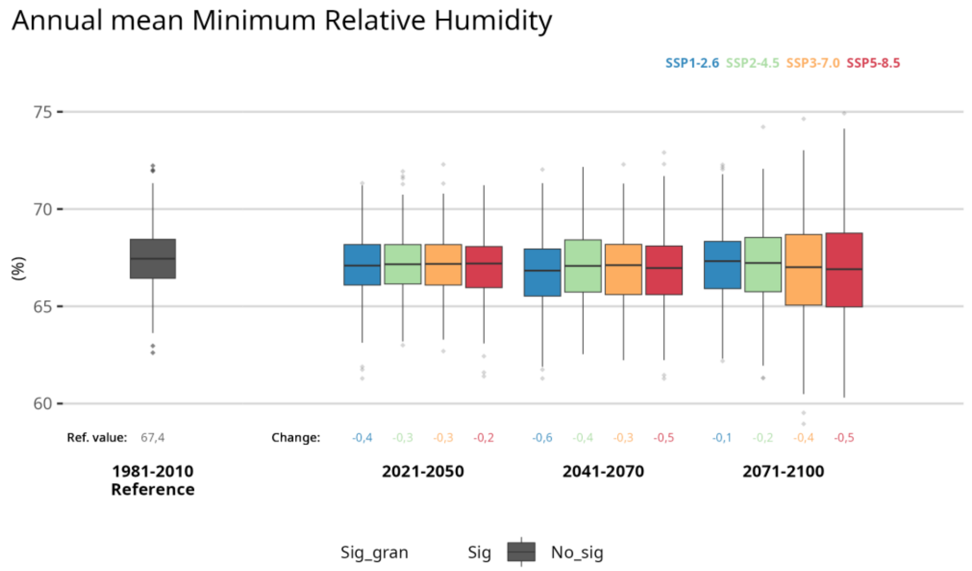


***Figure A5.40**. The graph shows the evolution of minimum relative humidity throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all 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.*

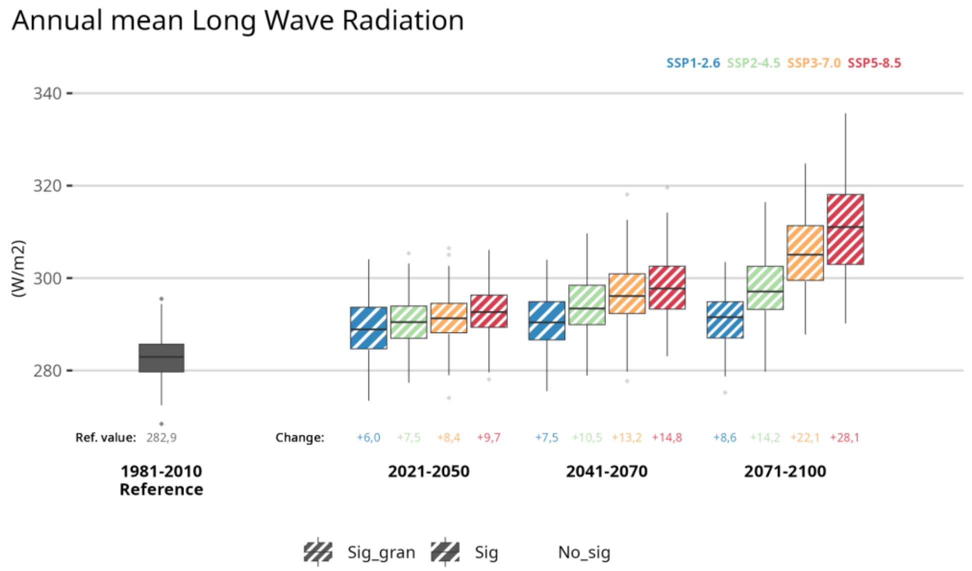


__Figure A5.41__. The graph shows the evolution of long wave radiation throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.

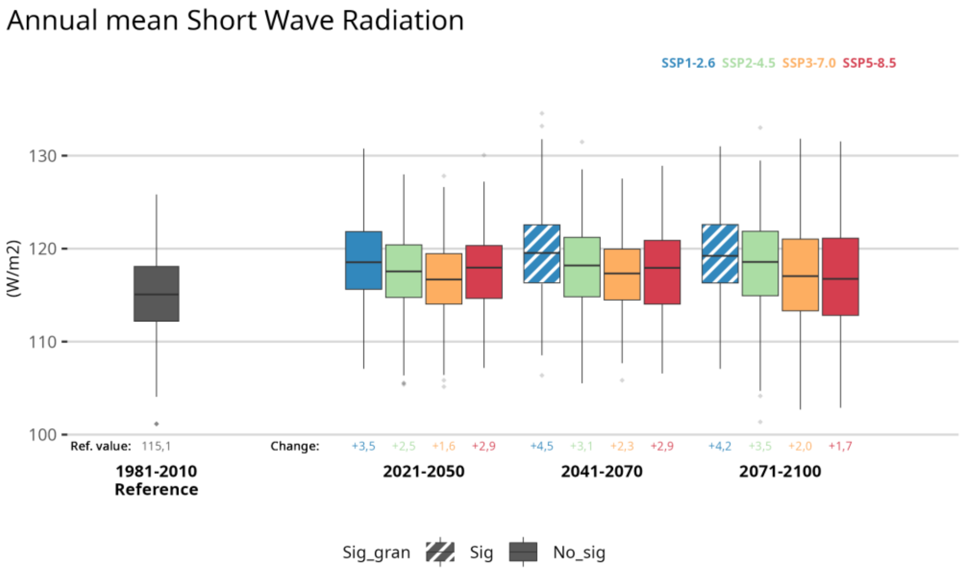


***Figure A5.42**. The graph shows the evolution of short wave radiation throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the 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.*

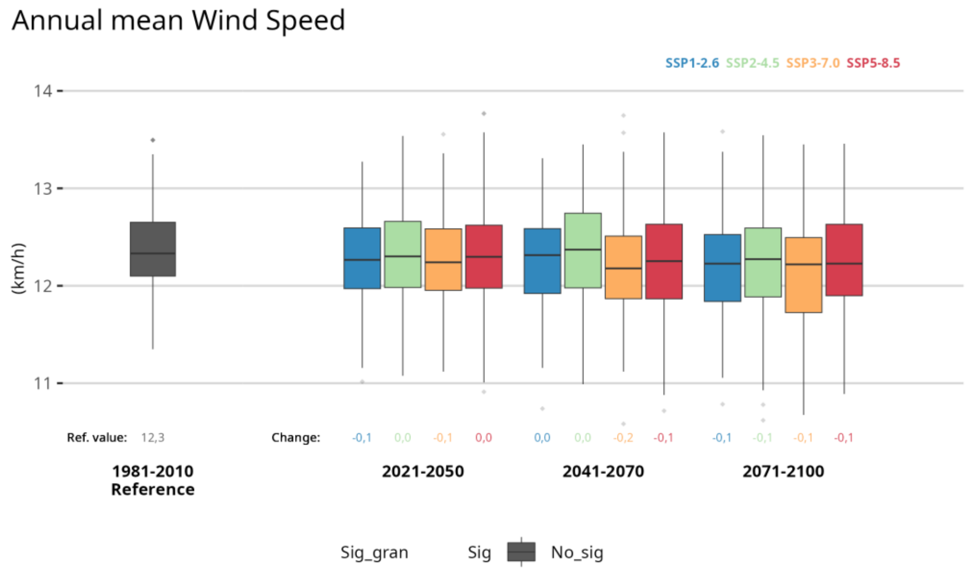


***Figure A5.43**. The graph shows the evolution of wind speed throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

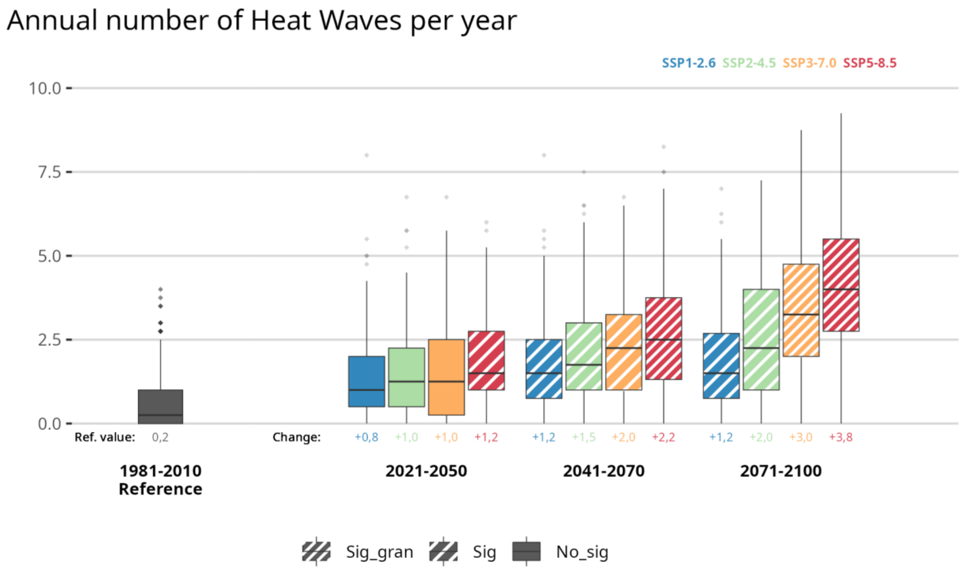


Figure A5.44*. The graph shows the evolution of annual number of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set 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.*

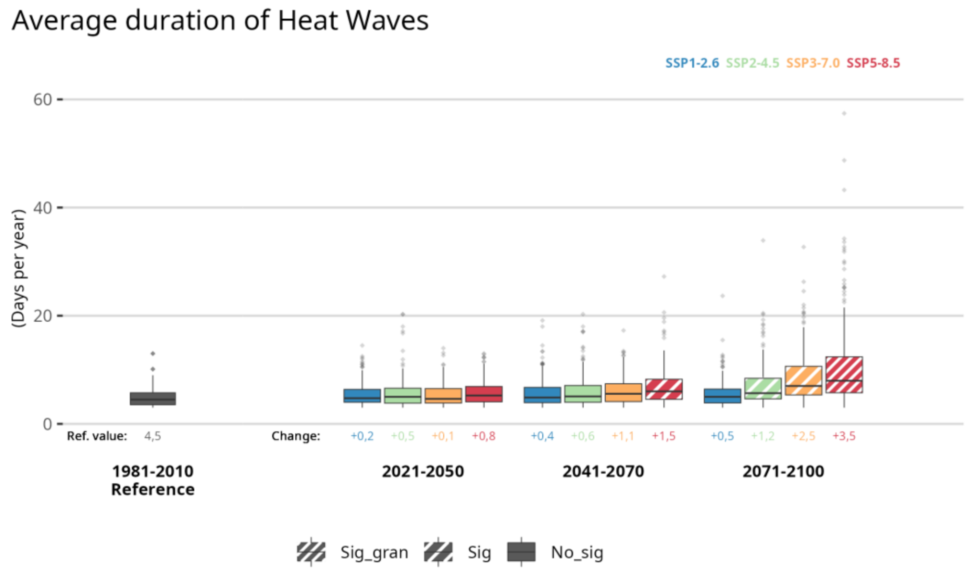


***Figure A5.45**. The graph shows the evolution of average duration of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for 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.*

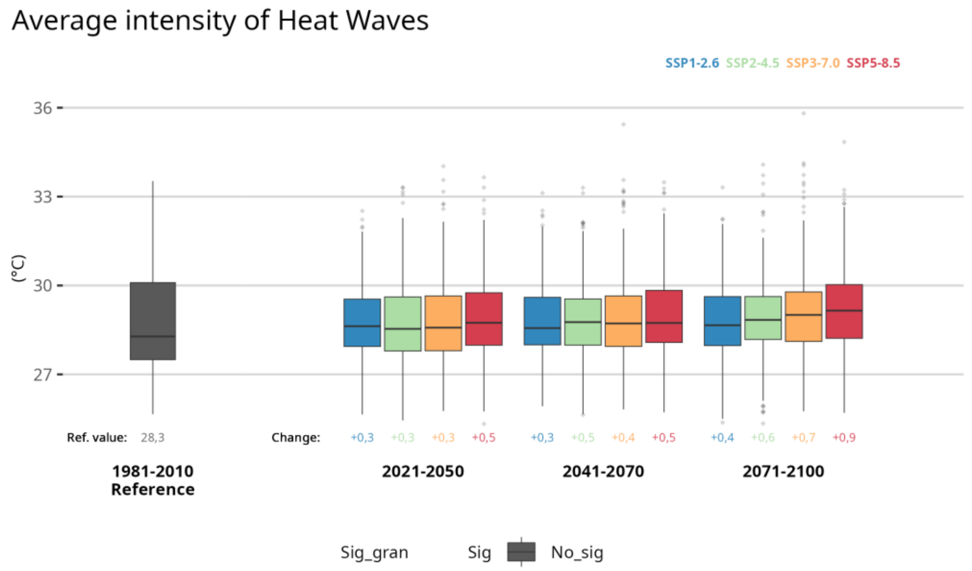


Figure A5.46. *The graph shows the evolution of average intensity of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the 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.*

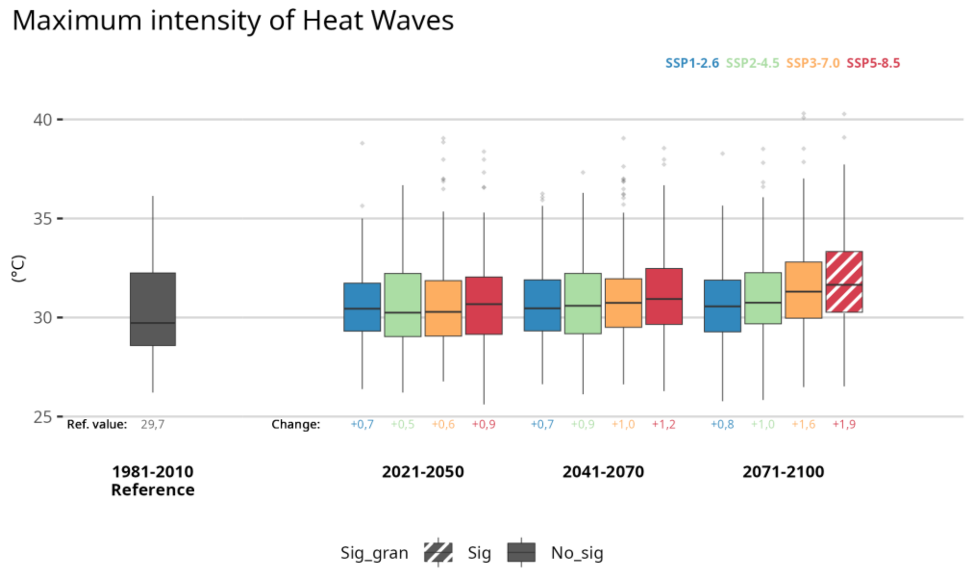


Figure A5.47. *The graph shows the evolution of maximum intensity of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the 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.*

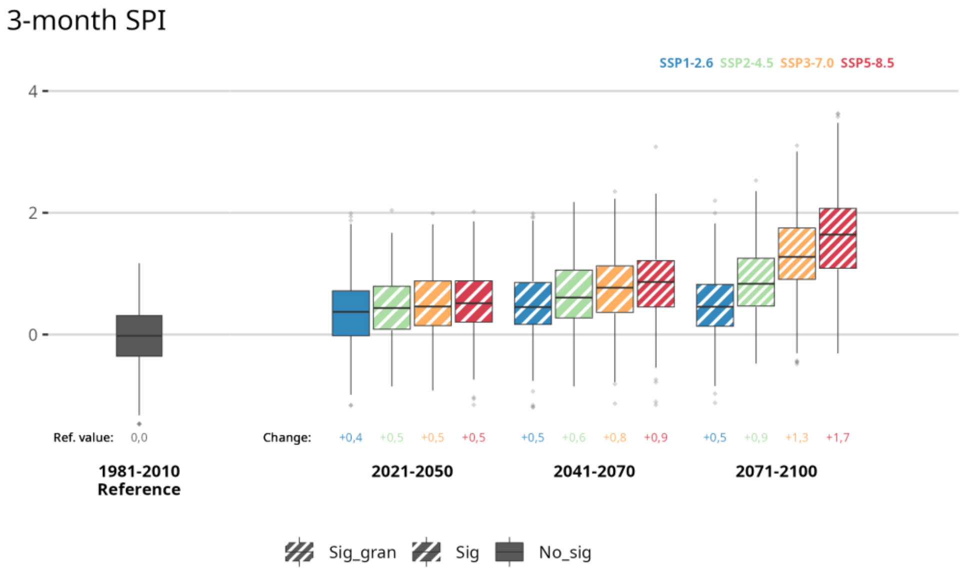


Figure A5.48*. The graph shows the evolution of 3-month SPI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

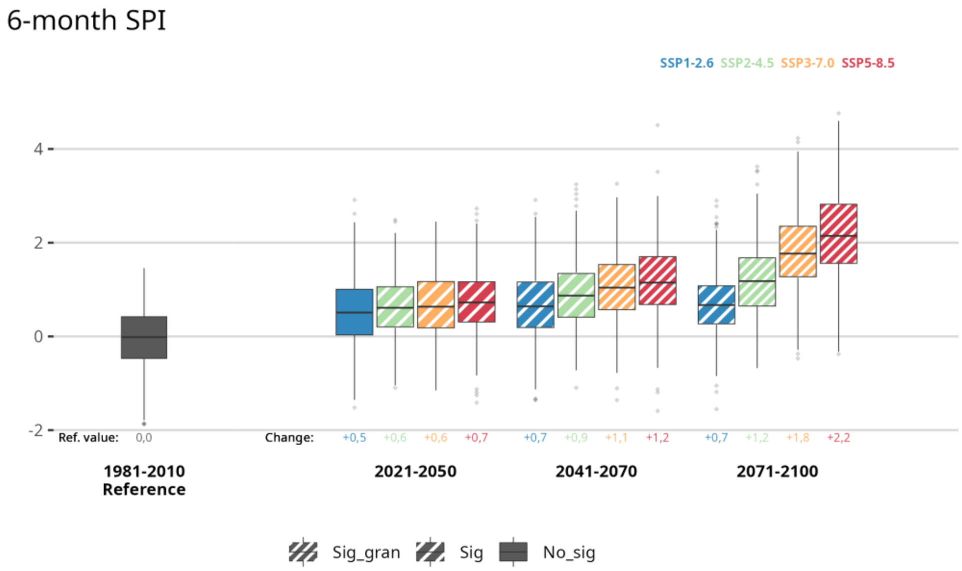


Figure A5.49. *The graph shows the evolution of 6-month SPI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

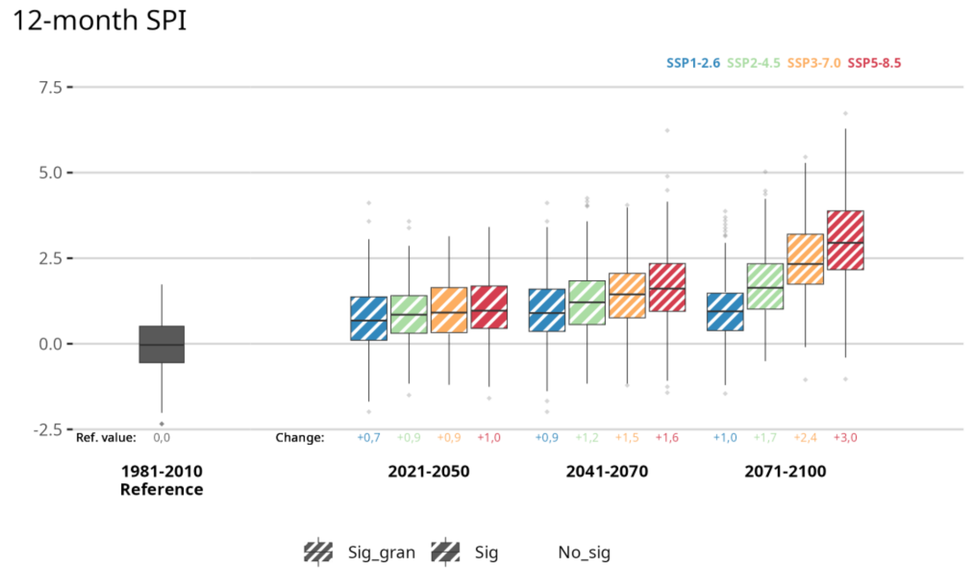


Figure A5.50. *The graph shows the evolution of 12-month SPI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

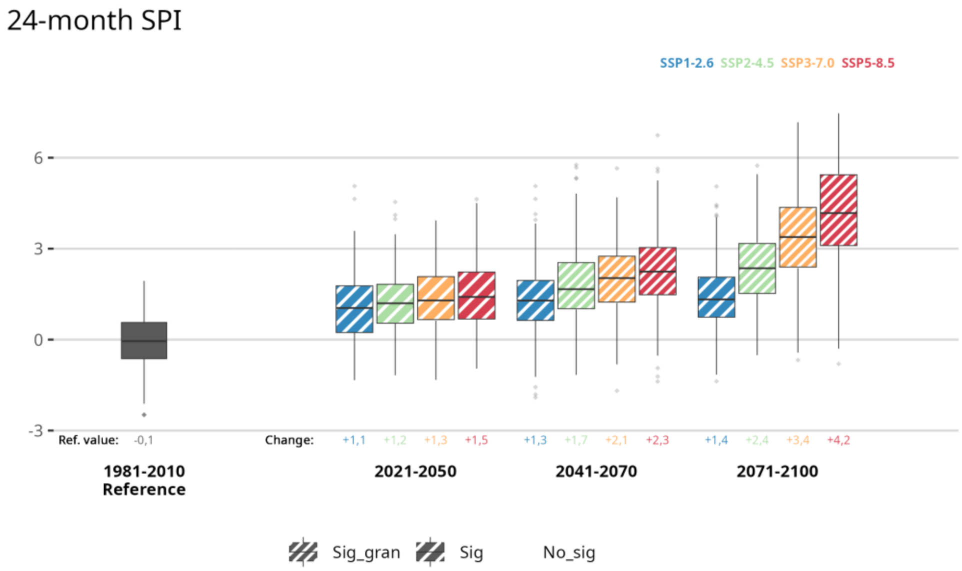


Figure A5.51. *The graph shows the evolution of 24-month SPI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

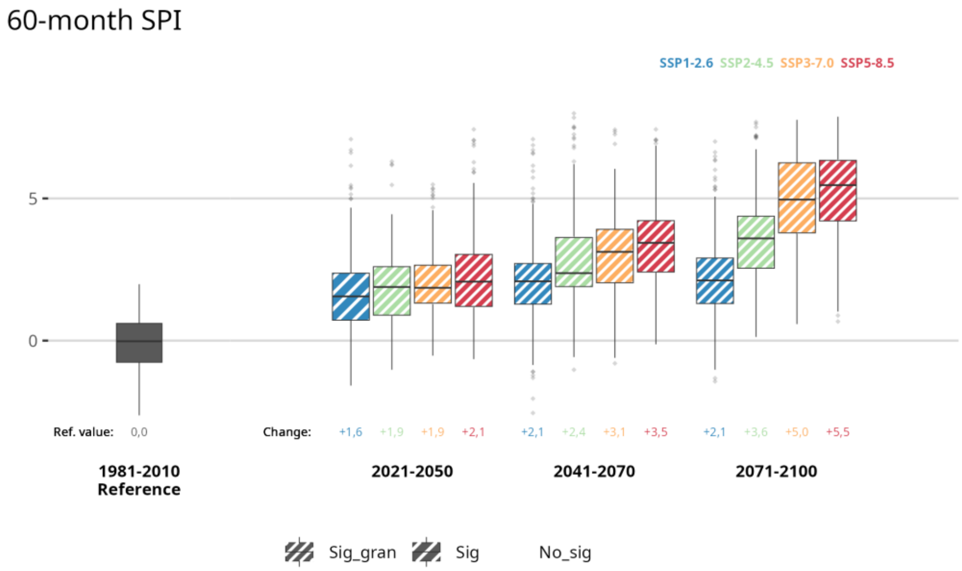


***Figure A5.52**. The graph shows the evolution of 60-month SPI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

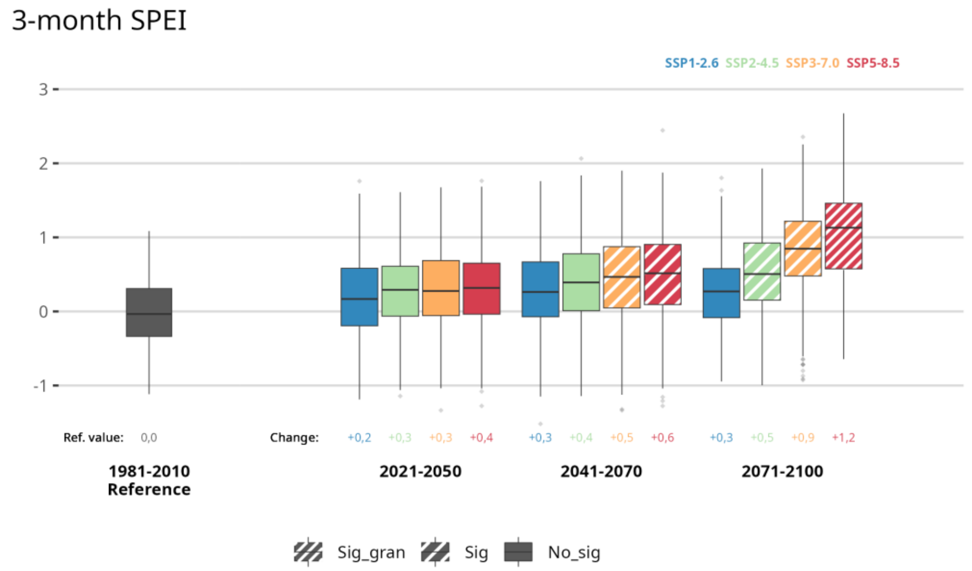


***Figure A5.53**. The graph shows the evolution of 3-month SPEI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

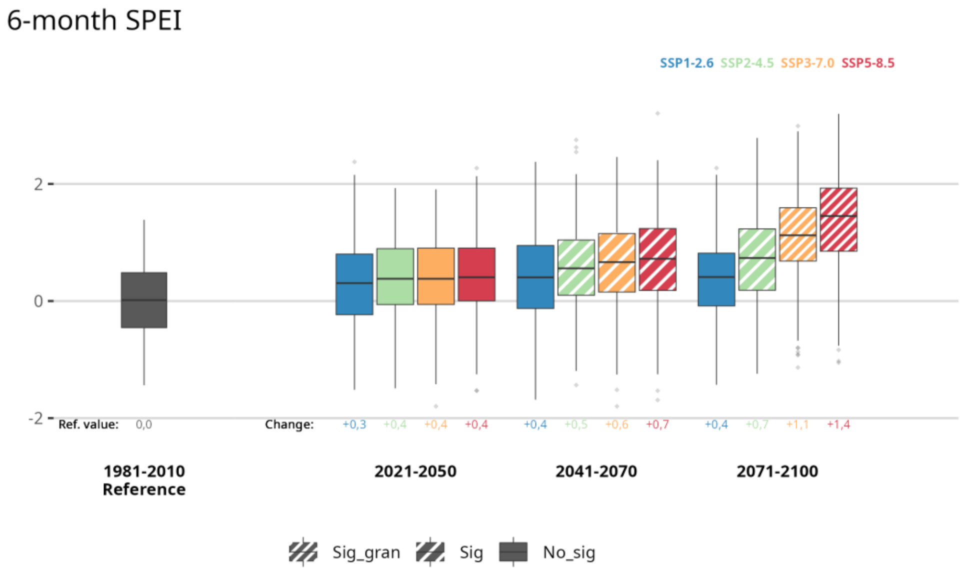


***Figure A5.54**. The graph shows the evolution of 6-month SPEI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

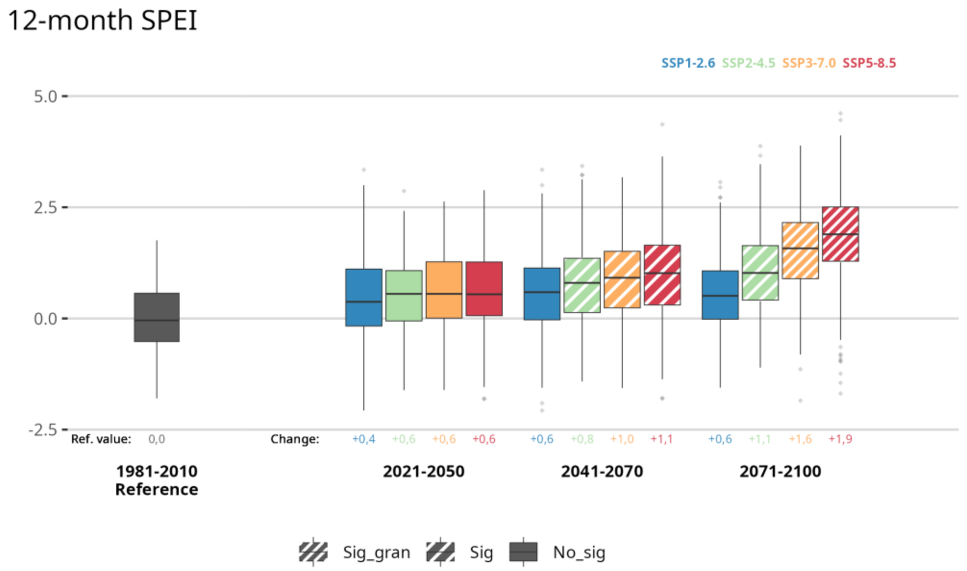


Figure A5.55. *The graph shows the evolution of 12-month SPEI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

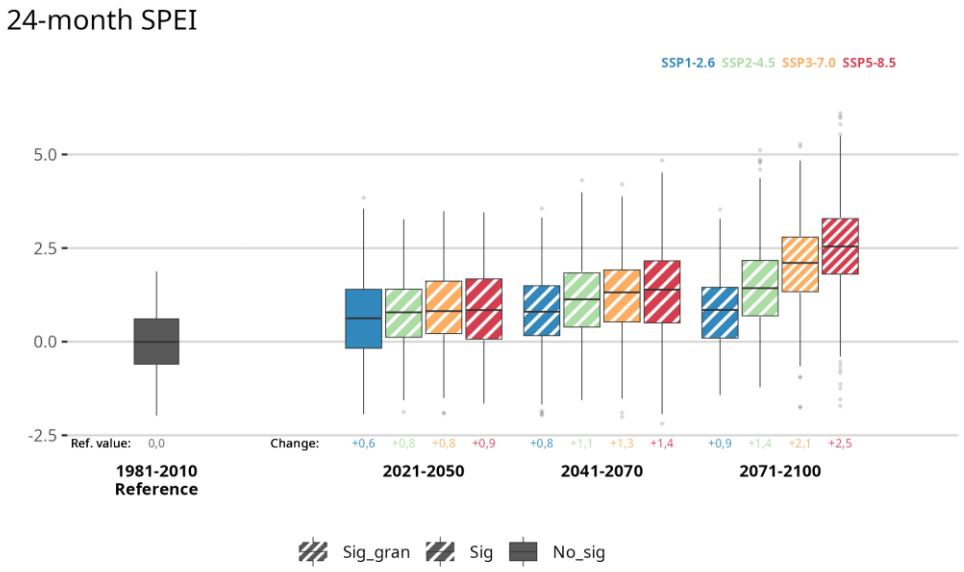


Figure A5.56. *The graph shows the evolution of 24-month SPEI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

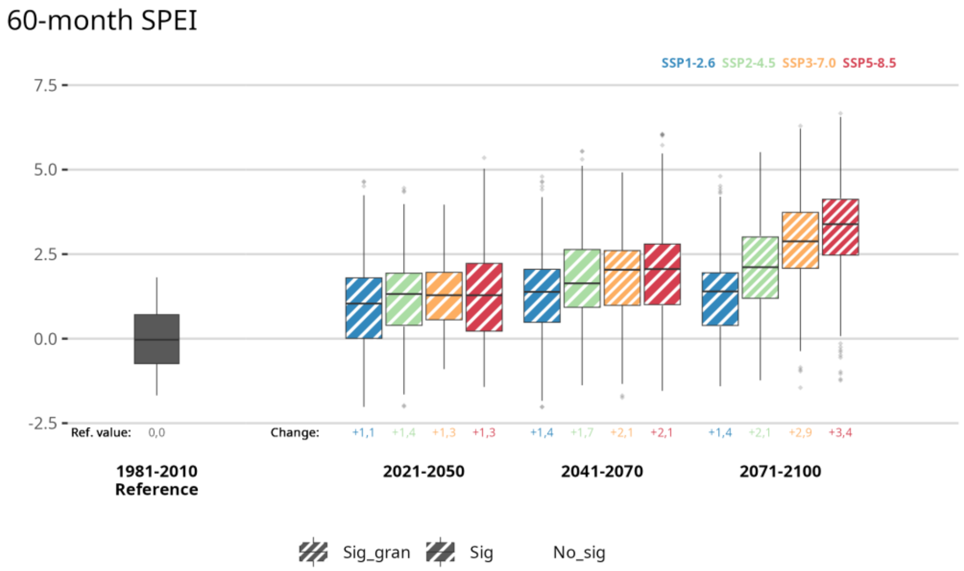


Figure A5.57. *The graph shows the evolution of 60-month SPEI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

Results for Cluj-Napoca

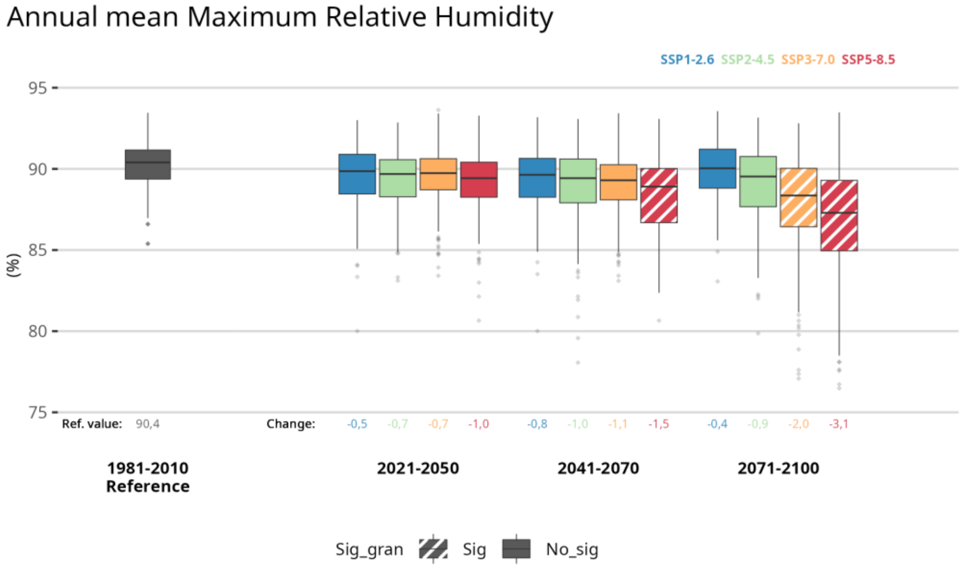


Figure A5.58. *The graph shows the evolution of maximum relative humidity throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the 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.*

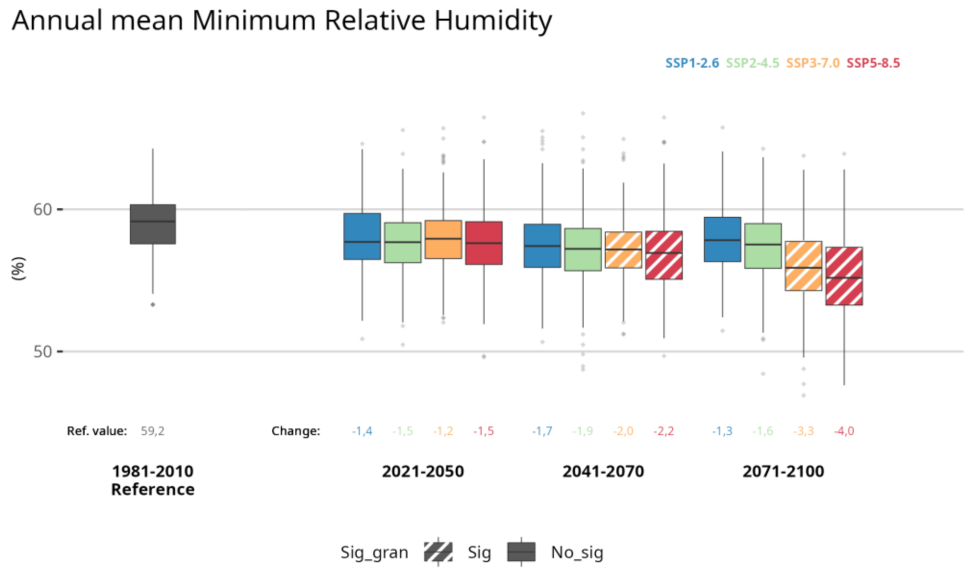


Figure A5.59. *The graph shows the evolution of minimum relative humidity throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the 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.*

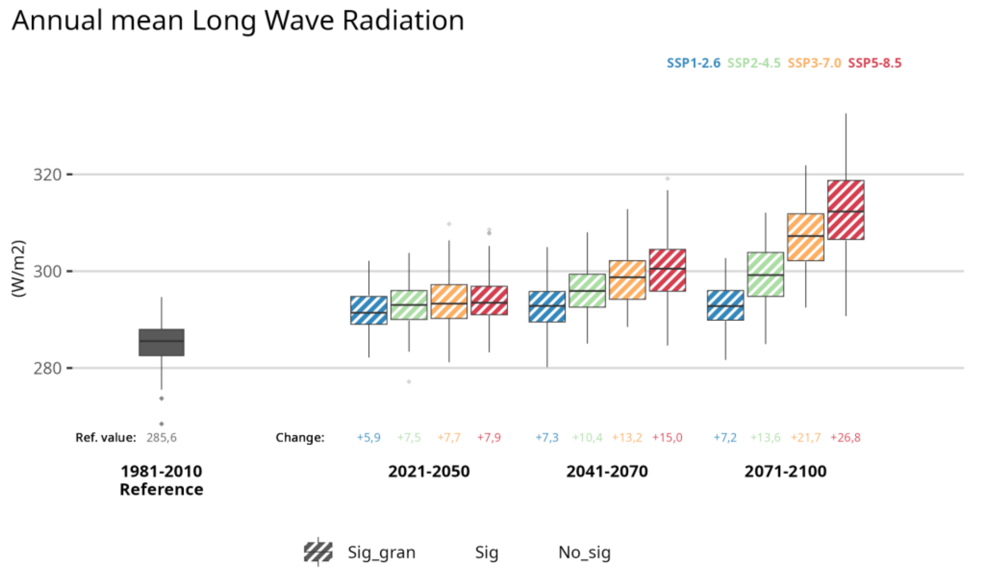


Figure A5.60. *The graph shows the evolution of long wave radiation throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

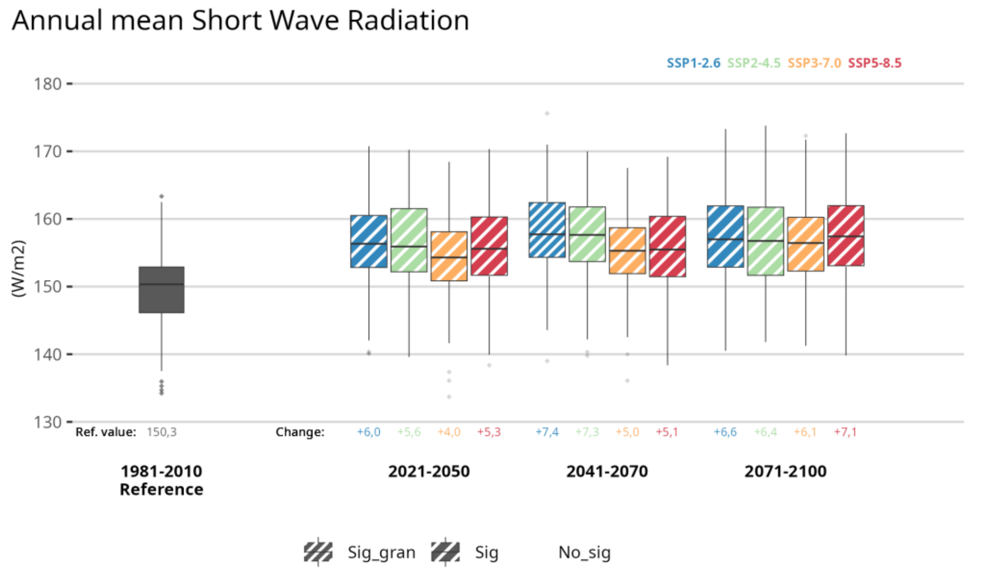


***Figure A5.61**. The graph shows the evolution of short wave radiation throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the 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.*

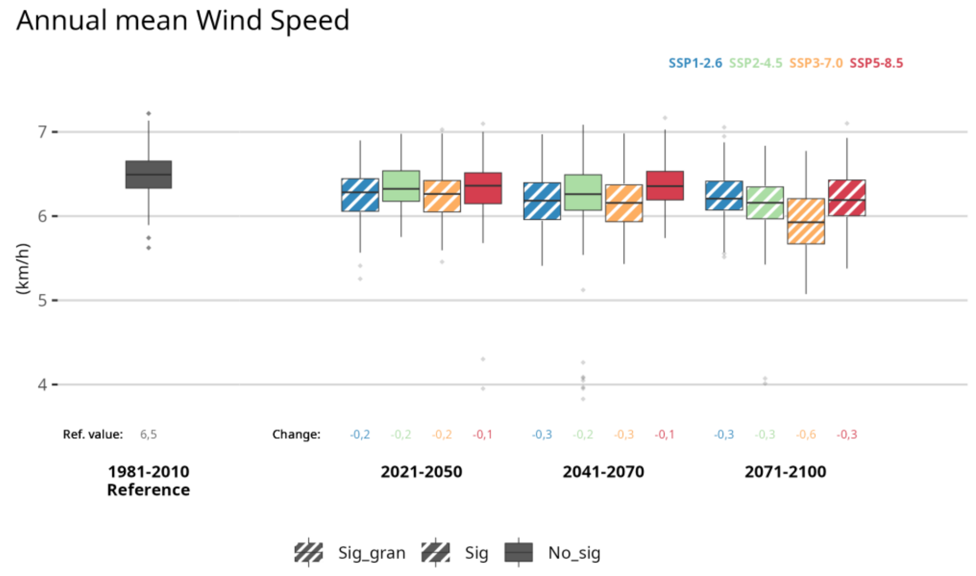


***Figure A5.62**. The graph shows the evolution of wind speed throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

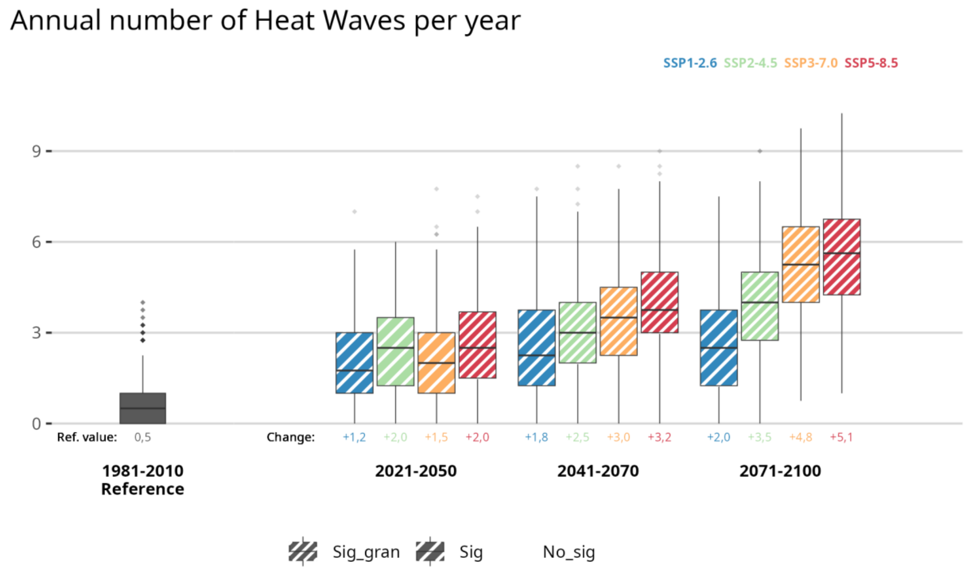


Figure A5.63. *The graph shows the evolution of annual number of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set 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.*

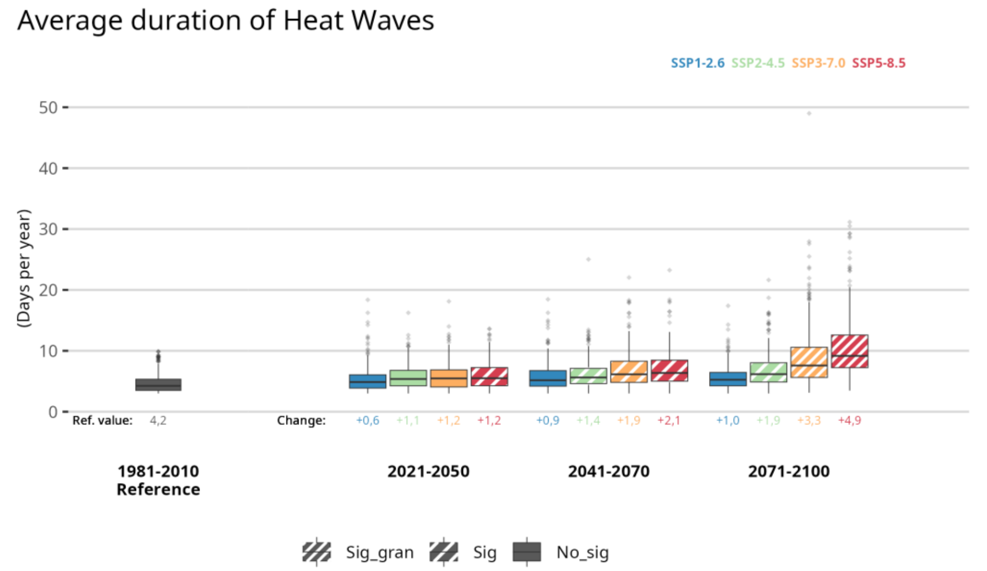


Figure A5.64. *The graph shows the evolution of average duration of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the 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.*

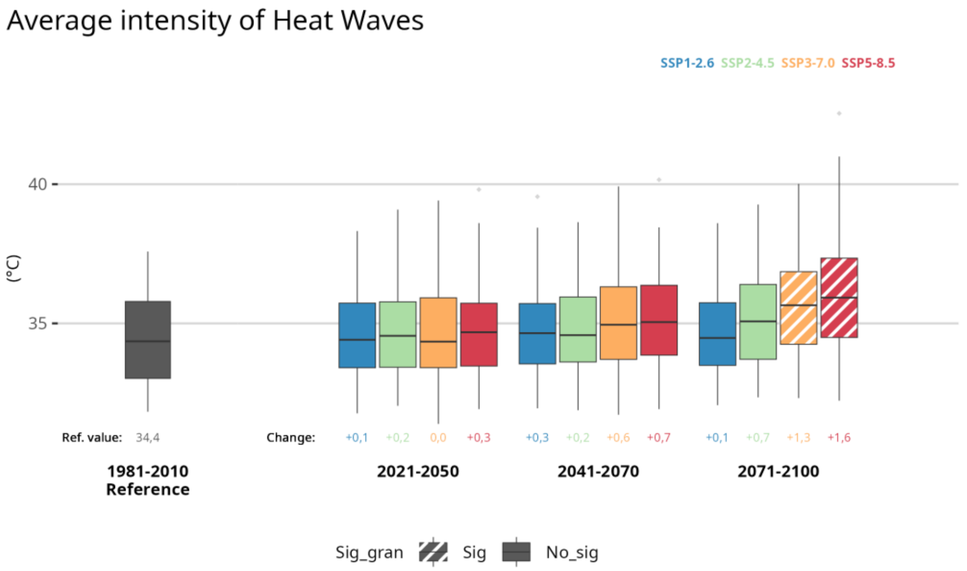


Figure A5.65. *The graph shows the evolution of average intensity of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the 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.*

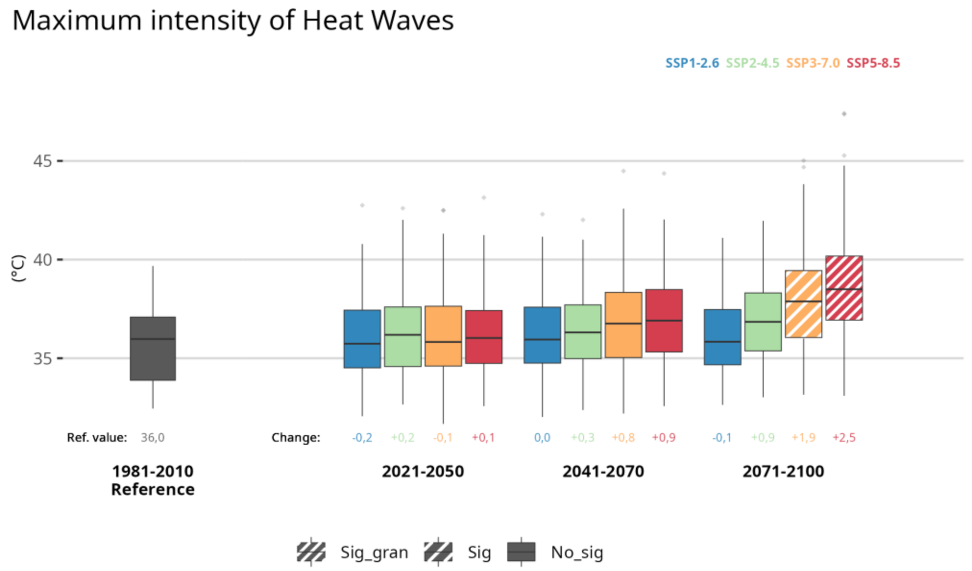


Figure A5.66. *The graph shows the evolution of maximum intensity of Heat Waves per year throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the 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.*

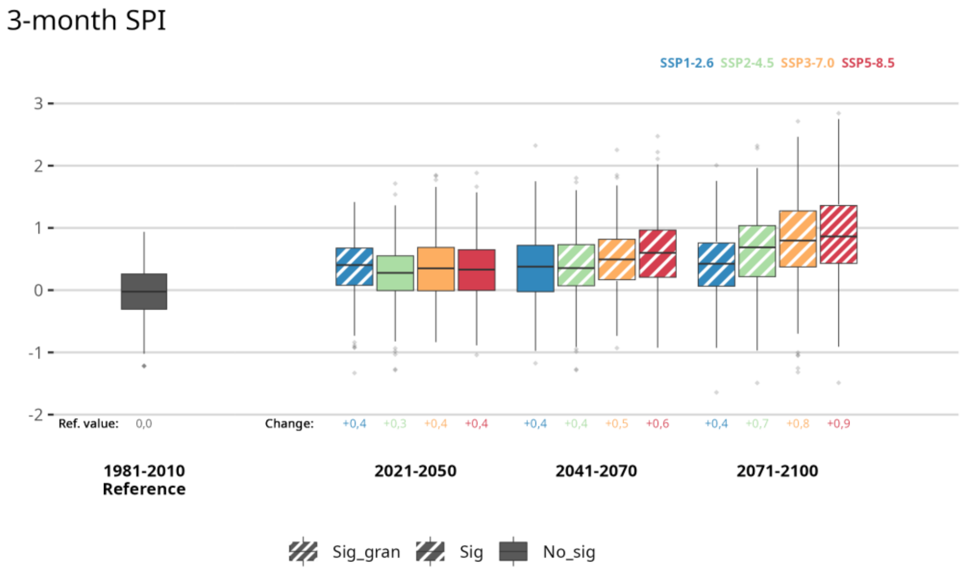


Figure A5.67. *The graph shows the evolution of 3-month SPI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

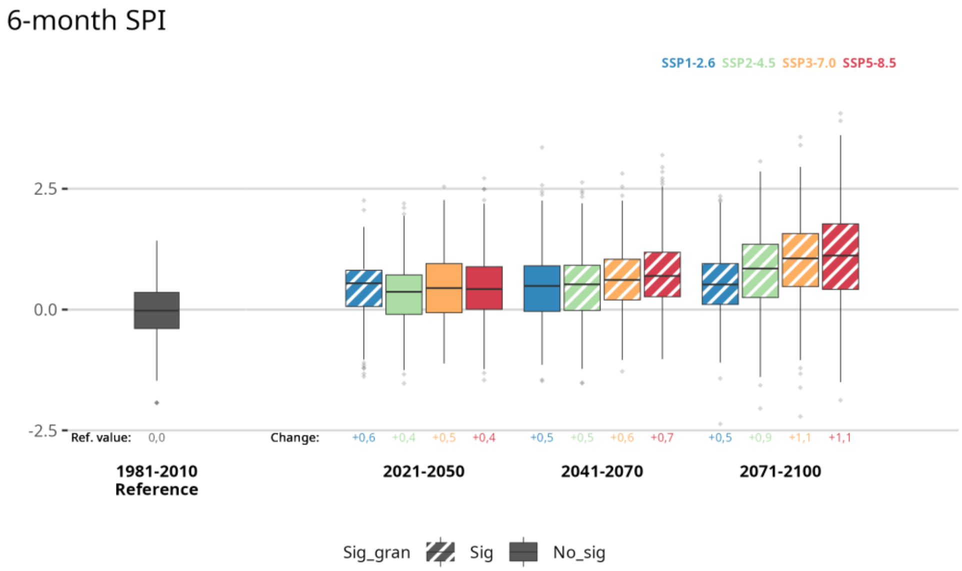


Figure A5.68. *The graph shows the evolution of 6-month SPI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

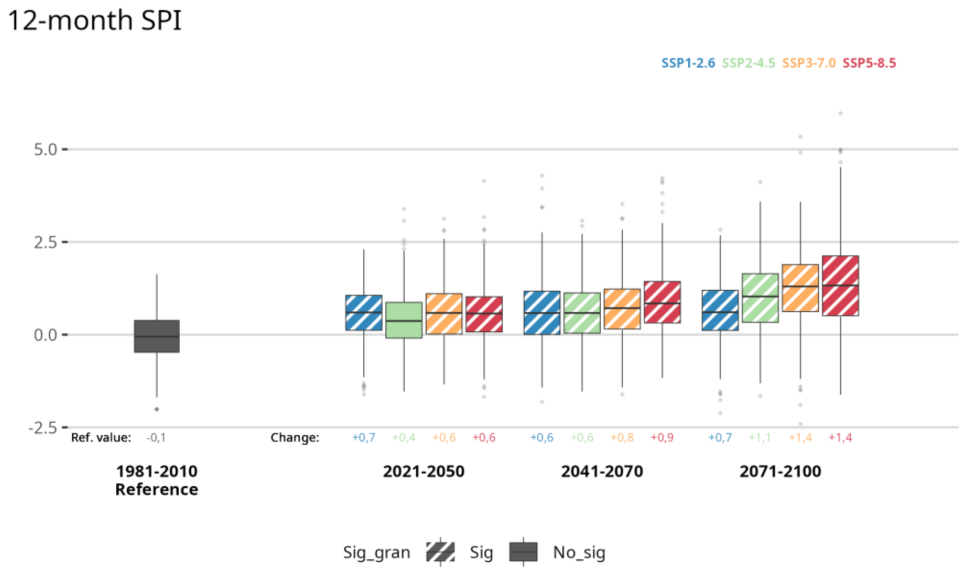


Figure A5.69. *The graph shows the evolution of 12-month SPI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

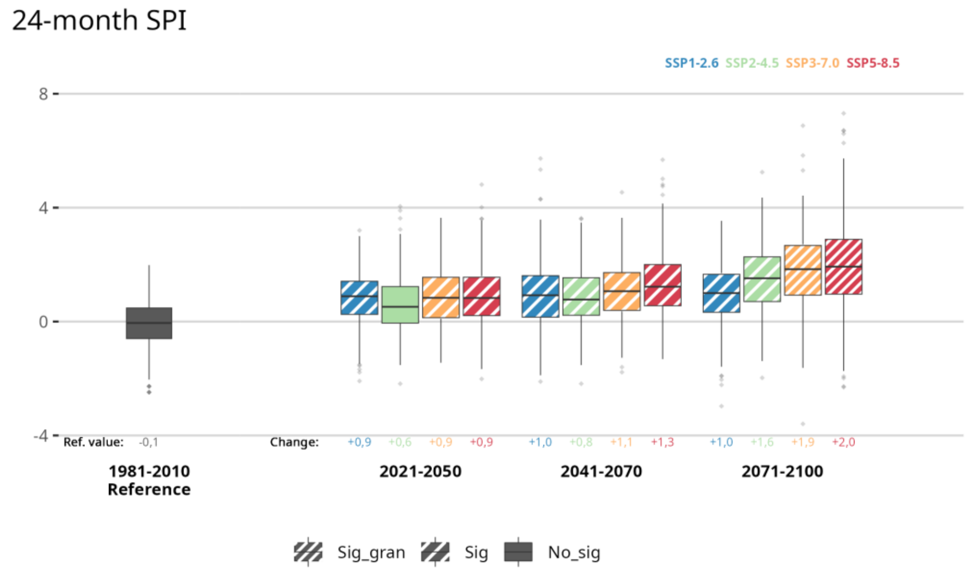


Figure A5.70. *The graph shows the evolution of 24-month SPI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

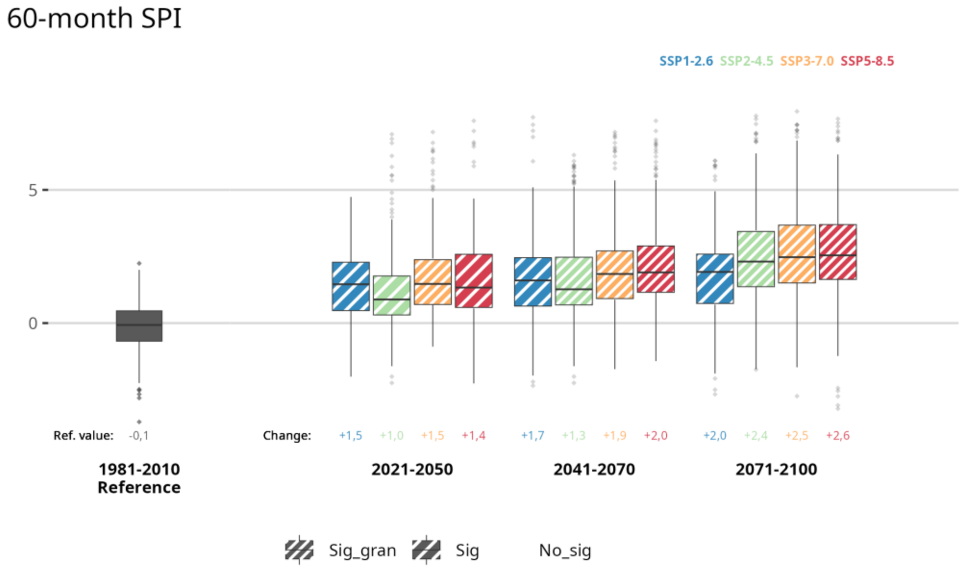


Figure A5.71. *The graph shows the evolution of 60-month SPI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

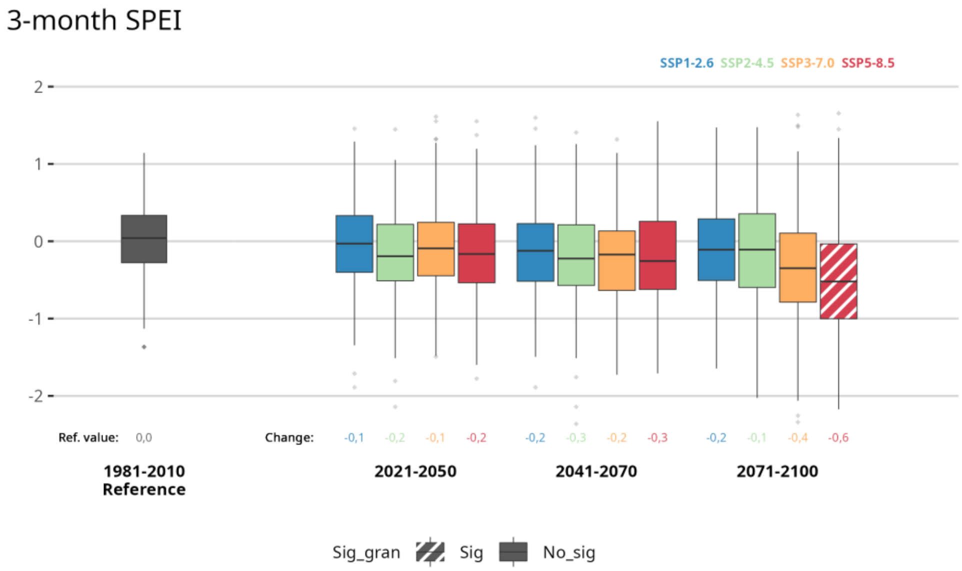


Figure A5.72. *The graph shows the evolution of 3-month SPEI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

Figure A5.73. *The graph shows the evolution of 6-month SPEI index 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.*

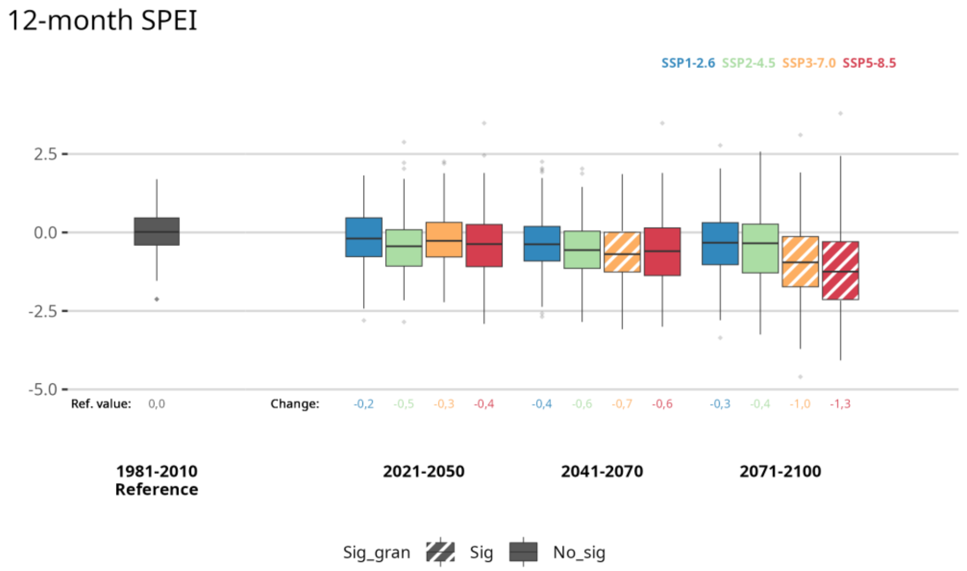


Figure A5.74. *The graph shows the evolution of 12-month SPEI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

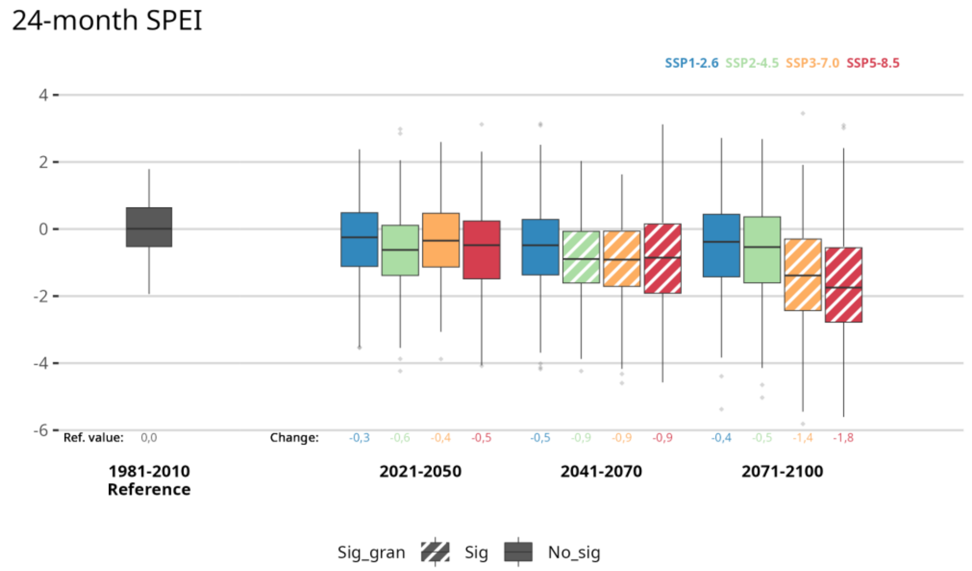


***Figure A5.75**. The graph shows the evolution of 24-month SPEI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the 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.*

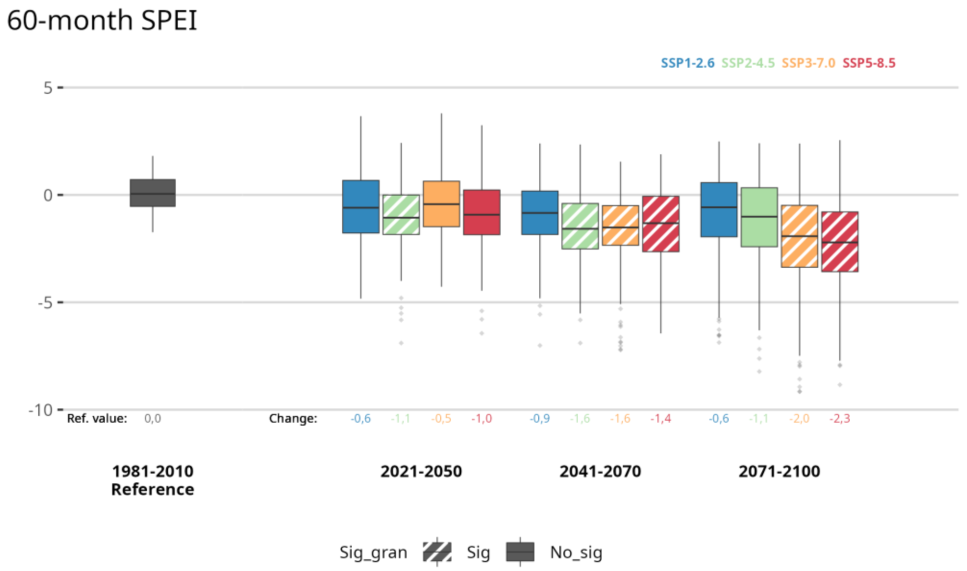


Figure A5.76. *The graph shows the evolution of 60-month SPEI index throughout the 21st century for the periods Historical (1981-2010), short (2021-2050), medium (2041-2070) and long-term (2071-2100) under four emission scenarios: SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5 for the set of all the observatories 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.*

Annex references

Anderson BG, Bell ML. Weather-Related Mortality How Heat, Cold, and Heat Waves Affect Mortality in the United States. Epidemiology 2009; 20: 205-213.

Begueria S, Vicente-Serrano SM, Reig F, Latorre B. Standardized precipitation evapotranspiration index (SPEI) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. International Journal of Climatology 2014; 34: 3001-3023.

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

Chen YD, Li JF, Zhang Q. Changes in site-scale temperature extremes over China during 2071-2100 in CMIP5 simulations. Journal of Geophysical Research-Atmospheres 2016; 121: 2732-2749.

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.

Kent ST, McClure LA, Zaitchik BF, Smith TT, Gohlke JM. Heat Waves and Health Outcomes in Alabama (USA): The Importance of Heat Wave Definition. Environmental Health Perspectives 2014; 122: 151-158.

Mckee T, Doesken N and Kleist J. The Relationship of Drought FRCequency and Duration Times Scales. American Meteorological Society. 8th Conference on Applied Climatology. 1993: January 17–22 Anaheim, California, pp. 179–184.

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.

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.

Smith TT, Zaitchik BF, Gohlke JM. Heat waves in the United States: definitions, patterns and trends. Climatic Change 2013; 118: 811-825.

Vicente-Serrano SM, Begueria S, Lopez-Moreno JI. A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. Journal of Climate 2010; 23: 1696-1718.

Vicente-Serrano SM, Lopez-Moreno JI. Hydrological response to different time scales of climatological drought: an evaluation of the Standardized Precipitation Index in a mountainous Mediterranean basin. Hydrology and Earth System Sciences 2005; 9: 523-533.

Vicente-Serrano, S.M.; Beguería, S. Comment on "Candidate Distributions for Climatological Drought Indices (SPI and SPEI)" by James H. Stagge et al. Int. J. Clim. 2016, 36, 2120–2131.

WMO (World Meteorological Organization), 2010. No. 1055. ISBN 978-92-63-11055-8. Switzerland www.wmo.int

WMO (World Meteorological Organization), 2012. In: Svoboda, M., Hayes, M. and Wood, D.(Eds.) Standardized Precipitation Index User Guide. Geneva: WMO

WMO (World Meteorological Organization), 2015. Guidelines on the Definition and Monitoring of Extreme Weather and Climate Events.

WMO (World Meteorological Organization), 2017. Statement on the State of the Global Climate in 2016.Nº 1189.2017. ISBN 978-92-63-11189-0