https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

Seasonal forecasting of health-related heat stress indices and its enhancement through teleconnection-based subsampling

Luca Famooss Paolini[1,2], Claudia Di Napoli[3], Fredrik Wetterhall[3], Paolo Ruggieri[1], Salvatore Pascale[1], Erika Brattich[1], and Silvana Di Sabatino[1]

[1] Department of Physics and Astronomy "Augusto Righi", University of Bologna, Bologna, Italy
[2] CMCC Foundation - Euro-Mediterranean Center on Climate Change, Bologna, Italy
[3] European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom

Correspondence: Luca Famooss Paolini (luca.famoosspaolini@unibo.it; luca.famooss@cmcc.it)

Abstract. Seasonal prediction systems (SPSs) have the potential to inform forecasts of heat-related risks, but their application to human heat stress remains limited by the availability of radiation data required to compute thermal stress indices and by the generally low predictive skill for European summer climate. Here, we develop a methodological framework to derive the Universal Thermal Climate Index (UTCI) and Wet Bulb Globe Temperature (WBGT) from SPS outputs by reconstructing the radiation components required for their computation. Validation against ERA5 shows that the proposed methodology provides accurate estimates of both indices, with root-mean-square error (RMSE) generally below 1.2 K for UTCI and 0.4 K for WBGT. We then investigate whether teleconnection-based subsampling can improve seasonal predictions of human heat stress by enhancing the representation of the summer North Atlantic Oscillation (NAO) in the European Centre for Medium-Range Weather Forecasts (ECMWF) SPS SEAS5. The subsampling is implemented within a "fair-sliding" framework designed to emulate real-time forecasting conditions over 2003–2024. Results show that subsampling enhances NAO prediction skill, increasing the correlation between the SEAS5 and ERA5 summer NAO indices from 0.40 to 0.63, while reducing the RMSE from 1.10 to 0.91. The enhanced NAO skill is accompanied by better predictions of summer-mean UTCI and WBGT, as well as the number of summer days exceeding heat stress thresholds, particularly over western, central, and northern Europe. The gains in thermal stress prediction are primarily driven by better skill for temperature- and radiation-related components, while improvements in humidity and wind are weaker or more regionally confined. Overall, these results demonstrate a pathway to extend SPS capabilities toward seasonal forecasting of heat stress conditions, supporting health-oriented climate services.

1 Introduction

Seasonal prediction systems (SPSs) have emerged as valuable tools for anticipating climate variability and hazards months in advance, bridging the gap between short-range weather forecasts (up to 10–15 days) and long-term climate projections (Hewitt et al., 2012; Bruno Soares et al., 2018). Their ability to provide actionable climate information on seasonal timescales offers important opportunities to support proactive decision-making across multiple sectors, including infrastructure management, agriculture, energy, and public health (De Felice et al., 2015; Colón-González et al., 2021; Gualdi et al., 2026). In particular,

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

reliable seasonal forecasts of heat stress conditions could strengthen climate services aimed at reducing the impacts of hot summer on human health, by informing preparedness and adaptation measures well before the onset of the warm season.

Despite this potential, two major limitations have hindered the use of SPSs to predict human heat stress. The first concerns the availability of meteorological variables required to compute thermal stress indices from current SPS outputs (Bröde et al., 2012; Budd, 2008), while the second relates to the limited skill of SPSs in predicting extratropical summer climate (Prodhomme et al., 2022). Regarding the first limitation, the assessment of human thermal stress requires indices that account not only for air temperature, but also for the combined effects of humidity, wind speed, and solar and thermal radiation (Jendritzky et al., 2012). Among these, the Universal Thermal Climate Index (UTCI) and Wet Bulb Globe Temperature (WBGT) are widely adopted to characterize thermal stress in outdoor, occupational, and sport settings (Budd, 2008; Krüger, 2021; Parajuli et al., 2025). However, these indices are not directly provided by current SPSs and their computation is hindered by the availability of radiation data. In particular, current SPSs outputs do not provide all the radiation variables required to estimate the mean radiant temperature (MRT), a key input for the computation of thermal stress indices such as the UTCI and WBGT (Brimicombe et al., 2022). Moreover, the available radiation variables are generally provided as daily aggregates, making their temporal resolution inconsistent with the instantaneous meteorological inputs required for these indices. Developing a robust framework to derive the UTCI and WBGT from the variables available in SPS outputs is therefore a necessary step toward extending seasonal climate prediction to heat stress applications.

Beyond the computation of thermal stress indices, another major challenge arises from the generally limited predictive skill of SPSs for extratropical summer climate, with Europe being a particularly challenging region (Famooss Paolini et al., 2025). For instance, predictive skill for extreme temperatures and heat waves is generally modest, if not absent, across large parts of the continent (Prodhomme et al., 2022). The same holds true for precipitation and extreme rainfall, posing significant challenges for the prediction of summer droughts and compound drought–heatwave events (Pepler et al., 2015; Mishra et al., 2019; Wu et al., 2026). A major factor contributing to these limitations is the misrepresentation of North Atlantic teleconnections, which prevents SPSs from fully exploiting important sources of seasonal predictability (Eade et al., 2014; Beverley et al., 2019; Dunstone et al., 2023). Consistently, many SPSs exhibit a low signal-to-noise ratio over the Euro-Atlantic sector, whereby the predictable component of climate variability is overwhelmed by chaotic internal variability within model ensembles (Scaife and Smith, 2018; Weisheimer et al., 2024). This deficiency has been linked to unresolved model errors, including biases in local atmosphere–ocean interactions (Roberts et al., 2021; Famooss Paolini et al., 2022; Patrizio et al., 2023), deficiencies in parameterized processes such as gravity waves (Achatz et al., 2024), weak eddy–mean flow interactions (Hardiman et al., 2022; Smith et al., 2022), and inaccuracies in remote teleconnections such as ENSO (Williams et al., 2023; O'Reilly, 2025).

Notably, a growing body of evidence indicates that the representation of North Atlantic teleconnections in SPSs can be improved through statistical post-processing techniques, leading to enhanced predictive skill for several climate processes (Neddermann et al., 2019; Düsterhus, 2020; Gargiulo et al., 2026). In this context, teleconnection-based subsampling has recently emerged as a promising methodology to address deficiencies in ensemble prediction systems. This approach selects a reduced subset of ensemble members that best represents the expected phase of a large-scale circulation pattern during the target season, based on predictors available one season in advance (Dalelane et al., 2020; Dobrynin et al., 2022). Using this approach,

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

Famooss Paolini et al. (2025) have recently demonstrated substantial improvements in the prediction skill of the summer North Atlantic Oscillation (NAO) in a multi-model SPS ensemble over the hindcast period 1993–2016. In turn, constraining the ensemble to members that realistically capture NAO teleconnections is reflected into enhanced skill for variables strongly influenced by this circulation pattern. In particular, skill improves for extreme temperature conditions over western, central, and northern Europe, where current SPSs generally exhibit the largest limitations in predicting summer climate variability (Prodhomme et al., 2022). These results suggest that improving the representation of large-scale circulation patterns may provide a pathway to enhance seasonal predictions of heat-related hazards. However, the potential of teleconnection-based subsampling to improve seasonal forecasts of human thermal stress has not yet been explored.

Here, we address both the methodological challenges associated with deriving thermal stress indices from SPS outputs, and the SPS limitations in predicting European summer climate. To this end, we use the European Centre for Medium-Range Weather Forecasts (ECMWF) SPS SEAS5, which is the longest-running and most consistent system contributing to the Copernicus Climate Change Service (C3S; Johnson et al., 2019). First, we develop and validate a methodological framework to derive the UTCI and WBGT from meteorological variables available in SEAS5, enabling the assessment of thermal stress predictability from operational seasonal forecasts. Second, we implement a teleconnection-based subsampling approach for the summer NAO to investigate whether an improved representation of large-scale atmospheric circulation can enhance seasonal prediction of human heat stress. The subsampling procedure is implemented within a "fair-sliding" framework, to ensure a realistic assessment of forecast performance by preventing the use of future information that could artificially inflate skill estimates (Risbey et al., 2021). We evaluate the resulting forecasts over the period 2003–2024, considering both summer (June–July–August, JJA) mean UTCI and WBGT and the number of summer heat stress days (HSDs) exceeding predefined temperature thresholds. By combining the derivation of thermal stress indices from SPS outputs with an operational teleconnection-based subsampling, this study provides a framework to advance seasonal predictions of human heat stress, potentially supporting more effective health-oriented climate services through earlier and more reliable information on heat-related risks.

The article is structured as follows. Section 2 describes the data, the derivation of the UTCI and WBGT from SPS outputs, the definition of HSDs, the implementation of the teleconnection-based subsampling in the "fair-sliding" operational set up, and the metrics used to assess prediction skill. Section 3 first evaluates the derived thermal stress indices against ERA5, then assesses the SEAS5 prediction skill for the summer NAO with and without subsampling, and finally examines the resulting impact on forecasts of thermal stress indices, HSDs, and their underlying climate variables. Section 4 summarises the main findings of this work, its limitations, and the implications for health-oriented climate services.

2 Methods

2.1 Data

All analyses conducted in this study are performed using the ECMWF SPS SEAS5 (Johnson et al., 2019), one of the systems contributing to the C3S. SEAS5 is here chosen because it provides the most comprehensive and long-running dataset available on the C3S, providing data from 1981 to the present date. Moreover, this dataset is produced using a single model version,

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

ensuring data consistency throughout the record. Differently, other SPSs contributing to the C3S are not used here, as they provide shorter time series (mainly starting from 1993), which are generated with different model versions. In this context, we use the simulations initialised on May 1 during the 1981–2024 time period. The years 1981–2016 correspond to the SEAS5 hindcast (or reforecast) period, for which 25 ensemble members are available, whereas the years from 2017 onward correspond to the real-time forecast period, for which 51 ensemble members are provided. To ensure consistency across the entire period of time, only the first 25 members are used in the analysis. Additionally, the ERA5 atmospheric reanalysis (Hersbach et al., 2020) is used as a proxy for observations over the same time period.

2.2 Computation of thermal stress indices

The UTCI and WBGT are both computed with the Python library *thermofeel* (Brimicombe et al., 2022), which requires instantaneous values of air temperature at 2 m height (T2m), intensity of wind speed at 10 m height (U10), dew point temperature (Tdew), and MRT for their calculations. While the first three variables can be directly downloaded from the C3S website, the MRT is not provided and is hence computed using *thermofeel* based on the solar and thermal radiation components listed in Table 1.

Table 1. Solar and thermal radiation components ($\mathrm{W\,m^{-2}}$) needed for the calculation of MRT (K) with the Python library *thermofeel*. The left column shows the full names of the solar and thermal radiation components, while the right column reports their corresponding acronyms as used in *thermofeel*.

Radiation component	**Acronym in *thermofeel***
Surface solar radiation downwards	SSRD
Surface net solar radiation	SSR
Direct solar irradiance	DSRP
Direct solar radiation at the surface	FDIR (function of DSRP)
Surface thermal radiation downwards	STRD
Surface net thermal radiation	STRR

However, two main limitations prevent the direct calculation of the UTCI and WBGT for SPSs. First, the SPSs do not provide the direct component of the surface solar radiation (FDIR in Table 1), which is needed for the MRT estimation. Second, the solar and thermal radiation components are only available as 24-hour aggregated values for SPSs, making them incompatible with the instantaneous climatic variables required for the computation of thermal stress indices. To address these limitations, the following approach is adopted.

The first limitation is addressed through the use of empirical decomposition models, which have been previously proposed to decompose the global horizontal irradiance (GHI, or SSRD in Table 1) into its direct and diffusive components when mea-

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

surements are unavailable (Perez et al., 1990). In general, these models decompose the GHI relying on statistical relationships between its components and the clearness index (the ratio between GHI and the horizontal extraterrestrial irradiance), derived from observations across various geographic regions worldwide. Specifically, the Erbs (Erbs et al., 1982) and DISC (Direct Insolation Simulation Code; Maxwell, 1987) models are tested here. The Erbs model estimates the direct component of GHI using empirical relationships between the diffusive fraction of GHI and the clearness index. On the other hand, the DISC model estimates the direct component of GHI by applying an empirical approach that relates the direct transmittance (the ratio between the direct component of GHI and the horizontal extraterrestrial irradiance) to both the clearness index and the air mass.

The second limitation in computing the thermal stress indices is addressed by estimating the hourly values of solar and thermal radiation components as a percentage of their 24-hour aggregates. Specifically, for each variable in Table 1, the ERA5 data are downloaded both as instantaneous values at 12:00 UTC and as 24-hour aggregated values. The 12:00 UTC value is then expressed as a percentage of the daily total. This calculation is repeated for all summer days over the 1981–2024 time period to derive a daily climatological percentage. Finally, these percentages are applied to the ECMWF 24-hour aggregated data to approximate the corresponding values at 12:00 UTC. The 12:00 UTC is here selected as time step of interest because, among the 6-hourly outputs available from the C3S SPSs (00:00, 06:00, 12:00, 18:00), it usually represents the period of maximum heat stress for the human body during summer days (Błażejczyk et al., 2022).

Before applying the approach described above to SEAS5 data, it is first tested with ERA5 data. Specifically, the method is used to estimate MRT, which is then employed to compute the UTCI and WBGT. The resulting values are compared against the indices' values computed using MRT obtained directly from ERA5 hourly solar and thermal radiation components. The root-mean-square error (RMSE) and bias (i.e., the mean difference between the estimated and ERA5 data) are used to assess the accuracy of both UTCI and WBGT estimates. In all cases, the T2m, U10 and Tdew at 12:00 UTC are directly retrieved from ERA5. This assessment serves to evaluate the impact of the proposed approach on the final representation of thermal stress conditions.

2.3 Heat stress conditions hazardous for human health

After computing the UTCI and WBGT, we first evaluate the ECMWF seasonal prediction skill of their averages during the summer season. We then assess the skill for heat stress conditions that pose risk to human health and affect working environments. In particular, we count the number of summer days when the UTCI exceeds 26°C, a threshold associated with moderate heat stress (Zare et al., 2018). Therefore, these days encompass a range of conditions, from moderate to strong, very strong, and extreme heat stress, all of which have been associated with adverse health impacts (Di Napoli et al., 2019). Similarly, we count the number of summer days when the WBGT is above 23°C, a level at which recommendations for work–rest cycles typically begin (Brimicombe et al., 2023). This allows us to identify days when preventive actions or guidelines for physical activity should be considered. To avoid confusion with the seasonal mean UTCI and WBGT, we refer to the counts of HSDs as $\mathrm{HSD_{UTCI}}$ and $\mathrm{HSD_{WBGT}}$. Although results are shown only for the temperature thresholds mentioned above, the methodology can be applied to any range of temperature values associated with specific thermal stress categories. We choose these

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

moderate thresholds to allow for a more comprehensive assessment of the prediction skill for heat stress conditions across a wider geographical extent in Europe. Focusing exclusively on the most extreme categories would exclude large regions from the analysis, especially in northern Europe, where extreme heat stress conditions as defined by fixed thresholds rarely or never occur (Mărmureanu et al., 2025).

2.4 Teleconnection-based subsampling adapted for an operational forecast environment

The teleconnection-based subsampling is a statistical post-processing technique developed to enhance prediction skill by refining multi-member model ensembles (e.g., Dobrynin et al., 2018; Düsterhus, 2020; Gargiulo et al., 2026). The methodology selects only those ensemble members that better represent the observed statistical relationship between a large-scale atmospheric circulation pattern in a given season and its predictors one season ahead. In this way, only members that properly capture the teleconnections affecting a specific region are retained, thereby improving the representation of large-scale low-frequency atmospheric variability over the region of interest. By combining statistical member selection criteria with dynamical model outputs, the teleconnection-based subsampling is recognized as a hybrid statistical–dynamical approach (Slater et al., 2023).

The teleconnection-based subsampling is here applied to improve seasonal prediction skill of both summer thermal stress indices (UTCI and WBGT) and heat stress conditions hazardous for human health ($\mathrm{HSD_{UTCI}}$ and $\mathrm{HSD_{WBGT}}$). Here, we focus on the summer NAO as the weather regime of interest. Indeed, results in Famooss Paolini et al. (2025) show that selecting the ensemble members consistent with summer NAO teleconnections improves seasonal predictions of summer European climate, including the occurrence of extreme temperatures. The improvement is most pronounced over western, central and northern Europe, where state-of-the-art SPSs exhibit the lowest prediction skill (Prodhomme et al., 2022). The approach is not applied to other large-scale atmospheric circulation patterns describing summer low-frequency variability over the North Atlantic sector, such as the Eastern Atlantic pattern. Model biases in representing the spatial structures of these modes may limit the potential improvements obtained with the teleconnection-based subsampling technique (Famooss Paolini et al., 2025).

The subsampling approach is here implemented to mimic a real-time, operational forecast environment. To do so, the NAO predictor indices are defined using only information that would be available at the forecast time, adopting the "fair-sliding" approach described in Risbey et al. (2021). For each forecast year, the geographical regions exhibiting positive and statistically significant correlation between the summer NAO and its predictors are identified using a 22-year sliding training window that ends one year prior to the forecast. We then compute predictor indices as detrended area-weighted anomalies over these regions. In this context, we define 2003–2024 as the real forecast period, thus balancing the length of the training and forecast samples (22 years each) while providing a sufficiently long record to properly assess forecast skill. For example, for the 2003 forecast, the areas of significant correlation between the summer NAO and its predictors are identified using data from 1981 to 2002, and the corresponding predictor indices are computed as detrended area-weighted anomalies over these regions, with the 2003 value retained as the statistical prediction of summer NAO. The same procedure applies to each forecast year by shifting the training window forward (e.g., 1982–2003 for 2004, 1983–2004 for 2005, etc.), thereby avoiding any leakage of future information into the predictor definitions. The fair-sliding approach is preferred over alternative "fair" methods, such as

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

the "fair-all" strategy, in which the training period is progressively extended from a fixed baseline (Risbey et al., 2021). This is because the sliding-window framework enables to estimate the NAO–predictors relationships over a recent period, which is particularly relevant given the non-stationarity of teleconnections over North Atlantic sector (Famooss Paolini et al., 2024; Outten and Davy, 2024). In this way, the identified relationships are more likely to reflect the evolving climate background rather than being influenced by longer-term historical variability.

As predictors of the summer NAO, we use the spring state of climate components previously identified as precursors of the North Atlantic summer circulation and as drivers of NAO-like anomalies (Screen, 2013; Matsumura et al., 2014; Dunstone et al., 2023). Specifically, we use Arctic sea-ice concentration (SIC), Northern Hemisphere snow cover (SNOWC), and lower-stratospheric zonal wind, zonally averaged over the North Atlantic sector (75°W–15°E) as in Wang and Ting (2022). Consistent with previous studies adopting subsampling approaches (Dobrynin et al., 2022; Famooss Paolini et al., 2025), SIC and SNOWC predictors are evaluated during the season preceding the one of interest, that is from late-winter to early spring (February, March, and April). Within the fair-sliding framework described above, the optimal reference month is selected independently for each forecast year as the one leading to the best NAO prediction skill over the corresponding 22-year training window. As a result, the selected month can vary from year to year. This strategy is adopted to minimize subjective choices in the prediction setup, and ensures that also in the reference month selection only information available at forecast time is used. At the same time, such approach allows for potential non-stationarity in the relationships between the summer NAO and its precursors. Since the subsampling framework relies on statistical NAO predictions, better NAO statistical predictions are expected to translate into more effective subsampling. Differently, the stratospheric predictor is evaluated only in May, consistent with the stratosphere–troposphere interaction timescale discussed in Dunstone et al. (2023), and no analogous month-by-month optimisation is applied to this variable.

Notably, the stratospheric predictor is defined directly using SEAS5 data, and not ERA5 data as for the other predictors. Since we analyse simulations on May 1, the May lower-stratospheric variability is not yet available in ERA5 at the time of initialization and thus cannot be used to predict the summer NAO. However, SPSs have been shown to skillfully represent stratospheric variability (Dunstone et al., 2023), making their predictions one month ahead a suitable proxy for ERA5 reanalysis. To test this approach, we cross-correlate the zonally averaged zonal wind during May in SEAS5 with the same variable in ERA5 from 1981 to 2024. Consistent with previous studies, Figure S1 shows that the ECMWF SPS initialized on May 1 is able to properly capture the average stratospheric variability in May, including in the region where zonal-mean zonal wind is mostly correlated with the summer NAO. Thus, we use SEAS5 zonal-mean zonal wind in May as a proxy of the corresponding ERA5 field and as a predictor of the summer NAO.

Following previous studies adopting subsampling approach (Dobrynin et al., 2022; Famooss Paolini et al., 2025), the relationships between the summer NAO and the SIC and SNOWC predictors are evaluated with a leave-one-out cross-validated correlation analysis. For each of these predictors and left-one-out year, we identify the regions where the correlation is significantly different from zero at the 90% confidence level through a two-tailed Student's t test. Then, the regions of positive correlation identified in each cross-validation steps are combined into a single area, and the predictor index is defined as the detrended, area-weighted anomalies over that region and normalized by its standard deviation. Differently, the stratospheric

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

predictor index is defined as the detrended, area-weighted anomalies within the extratropical lower stratosphere (green box in Figure S1; 30–100 hPa, 50°N–75°N), standardized by its respective standard deviation and without the cross-validated step. This choice is motivated by the fact that, in some years, the NAO–stratosphere correlation is not significant in any latitude–pressure levels points of interest, which would produce missing index values. This intermittency is consistent with Dunstone et al. (2023), who showed that the spring stratospheric variability influences the summer NAO only during specific windows of opportunity. Therefore, we opt to calculate this index as the zonal-mean zonal wind anomalies averaged over a fixed region exhibiting the strongest correlation with the summer NAO for the full 1981–2024 time period (green box in Figure S1), ensuring a continuous predictor time series.

We note that, unlike our previous teleconnection-based subsampling work (Famooss Paolini et al., 2025), North Atlantic sea surface temperature (SST) is not included here as an additional predictor of the summer NAO. This choice is motivated by its consistently weak predictive skill within the operational subsampling framework adopted in this study. Specifically, the correlation between the summer NAO in the SST-subsampled ensemble and the corresponding ERA5 index remains systematically lower than the skill of the full SEAS5 ensemble, regardless of the subsampled ensemble size (not shown). As a result, including SST as an additional predictor does not provide added value and does not lead to further improvements in the prediction of summer thermal stress indices or heat stress conditions over 2003–2024. The comprehension of this reduced performance, particularly in comparison with our results obtained over the hindcast period (Famooss Paolini et al., 2025), is beyond the scope of the present study and would require targeted investigation. However, it may reflect non-stationarity and/or non-linearity of NAO–predictor relationships (Dunstone et al., 2023), as well as a possible overestimation of their predictive strength when assessed outside a real-time operational prediction framework (Risbey et al., 2021).

The predictor indices defined as described above represent the statistical predictions of the summer NAO. These predictions are then used to subsample the SEAS5 ensemble, selecting 3 ensemble members whose summer NAO state is closest to each statistical prediction. The selected members are then merged into one ensemble, ensuring that members selected by multiple predictors are not counted twice. As a result, the subsampled ensemble size may range from a minimum of 3 members (complete overlap across predictors) to a maximum of 9 members (no overlap) depending on the forecast year. To assess how ensemble size influences forecast skill, we repeat the teleconnection-based subsampling while changing the number of members selected for each predictor from 1 up to 25 (i.e., the full SEAS5 ensemble).

2.5 Weather regime index

For ERA5, the summer NAO index is defined as the first Empirical Orthogonal Function (EOF) of summer geopotential height anomalies at 500 hPa (Z500) over the North Atlantic sector (25°–80°N, 90°W–40°E). Before computing the covariance matrix, the Z500 data are weighted by the square-root of the cosine of latitude, so that every cell contributes to the field variance in proportion to its surface area. For SEAS5, the EOF analysis follows the same procedure, but all members are concatenated along the time dimension to derive a single EOF for the full ensemble. Subsequently, we decompose the principal component of summer Z500 back to the number of model ensemble members. This approach is preferred compared with performing

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

EOF analysis on individual ensemble members separately, as it has been shown to produce more robust and stable modes of variability (Dobrynin et al., 2022).

To ensure a fully retrospective definition of the summer NAO index, consistent with the approach used for the predictors, the EOF analysis for any given forecast year is performed using only the preceding 22 years. The Z500 anomalies of the forecast year are then projected onto the corresponding leading EOF. The values of the summer NAO indices are finally merged into one timeseries, constructed without relying on future information.

2.6 The summer NAO role in explaining the thermal stress conditions in Europe

To assess the influence of the summer NAO on the European thermal stress conditions, we perform a linear regression analysis of the summer UTCI and WBGT, as well as summer $\mathrm{HSD_{UTCI}}$ and $\mathrm{HSD_{WBGT}}$, onto the summer NAO index using ERA5 data. The analysis is conducted over the period 2003–2024, corresponding to the real forecast period considered in this study. The statistical significance of regression coefficients is tested with a two-tailed Student's t test under the null hypothesis of coefficients equal to zero at the 90% confidence level. We then use R^2 coefficient to quantify the proportion of variance in the thermal stress indices explained by the summer NAO (DelSole and Tippett, 2022).

2.7 Seasonal prediction skill and statistical test

The performance of SEAS5 in predicting the summer NAO is evaluated through correlation analysis and RMSE. The correlation significance is assessed through a two-tailed Student's t test, providing the confidence level for which the null hypothesis of correlation equal to zero is rejected. Furthermore, we evaluate the prediction skill of both full and subsampled SEAS5 ensembles for summer thermal stress conditions by computing the anomaly correlation coefficients (ACCs) for summer UTCI and WBGT, as well as $\mathrm{HSD_{UTCI}}$ and $\mathrm{HSD_{WBGT}}$. To assess the drivers of any changes introduced by subsampling, we also compute ACCs for each individual variable used to derive the UTCI and WBGT, that is T2m, U10, Tdew and MRT. The ACC statistical significance is assessed through a two-tailed Student's t test against the null hypothesis of ACCs equal to zero at the 90% confidence level. To evaluate the effect of subsampling, we compute the differences between the ACCs computed from the subsampled ensemble and those from the full ensemble. The statistical significance of the ACC differences is determined using a z test on Fisher z-transformed correlation coefficients (Fisher, 1920; Hinkle et al., 2003), against the null hypothesis of zero differences at the 90% confidence level.

3 Results

As a first step, we assess the accuracy of the approach described in Methods to overcome the limitations in computing thermal stress indices from SPS outputs. Indeed, SPSs do not provide all solar and thermal radiation fluxes required for the UTCI and WBGT estimation (Table 1). In particular, FDIR is not available, while the remaining radiation fluxes are provided as 24-h aggregated fields. We address these limitations by (i) using empirical decomposition models to estimate FDIR, and (ii) estimating the 12:00 UTC values of radiation components from their corresponding 24-aggregated values (see Methods).

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

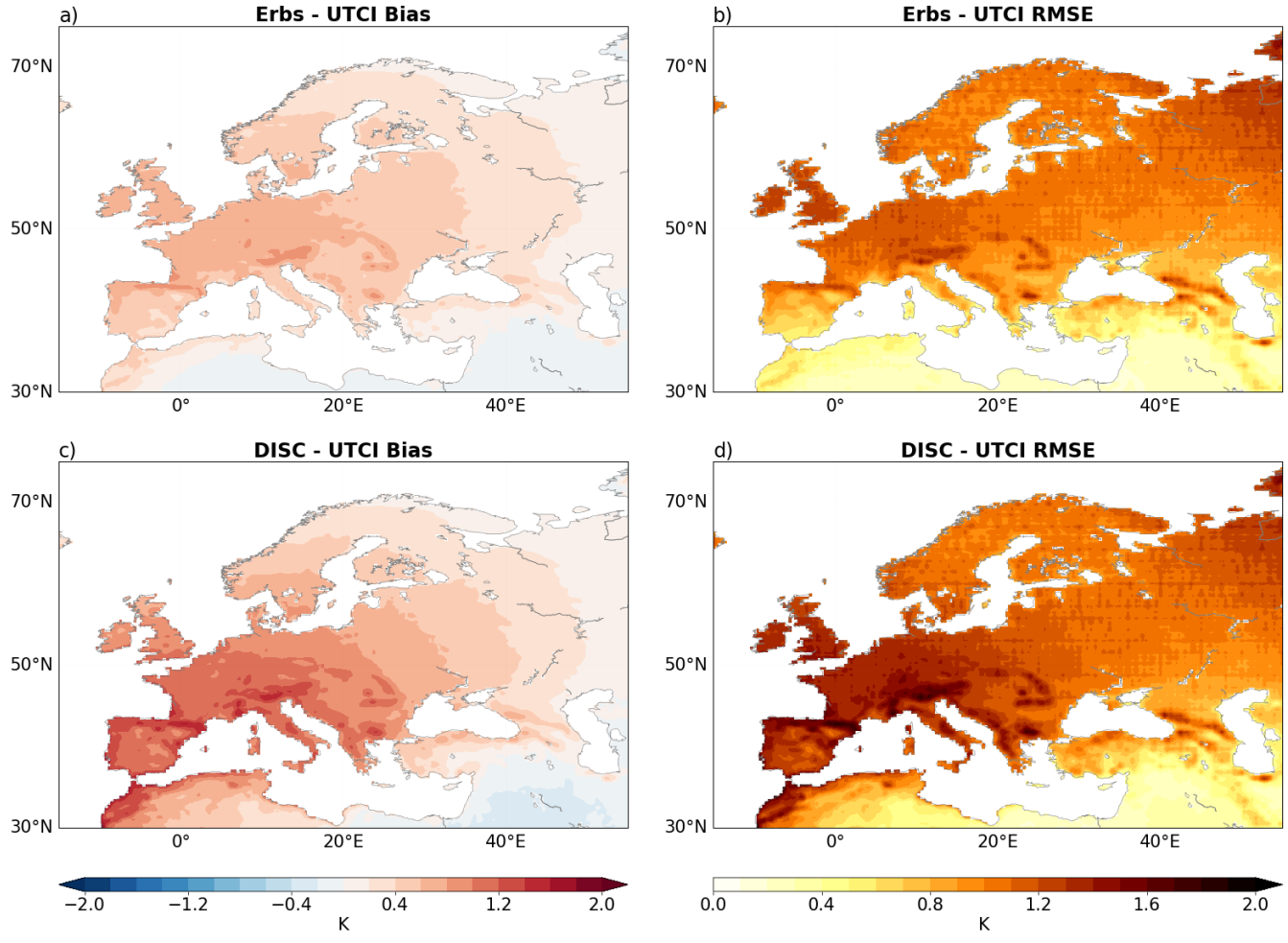

Figure 1. UTCI bias from using approximated solar and thermal radiation components. a) Bias of the UTCI estimation based on the Erbs empirical decomposition model and the 12:00 UTC values of radiation components in Table 1 derived from their 24-aggregated values. This estimation is compared to the UTCI calculated directly using hourly radiation components. ERA5 data during 1981–2024 are used for this comparison. b) Same as a) but for RMSE. c) and d) Same as a) and b) but using the DISC empirical decomposition model.

Figure 1 shows that using Erbs empirical decomposition model, together with hourly values of radiation components derived from their 24-aggregated values, induces a systematic positive bias in the UTCI over Europe. This bias is particularly evident over western and central Europe, where values reach about 0.6 K. The RMSE show a similar spatial pattern, with values generally below 1.2 K. Results obtained with the DISC empirical decomposition model are qualitatively similar but generally exhibit larger errors, with bias reaching about 1.2 K and RMSE approaching 2 K. Comparable conclusions hold for WBGT, although both bias and RMSE are more concentrated over Mediterranean Europe and remain lower than those found for the UTCI (Figure 2). For the Erbs model, WBGT bias is generally below about 0.2–0.3 K, while RMSE remains mostly below

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

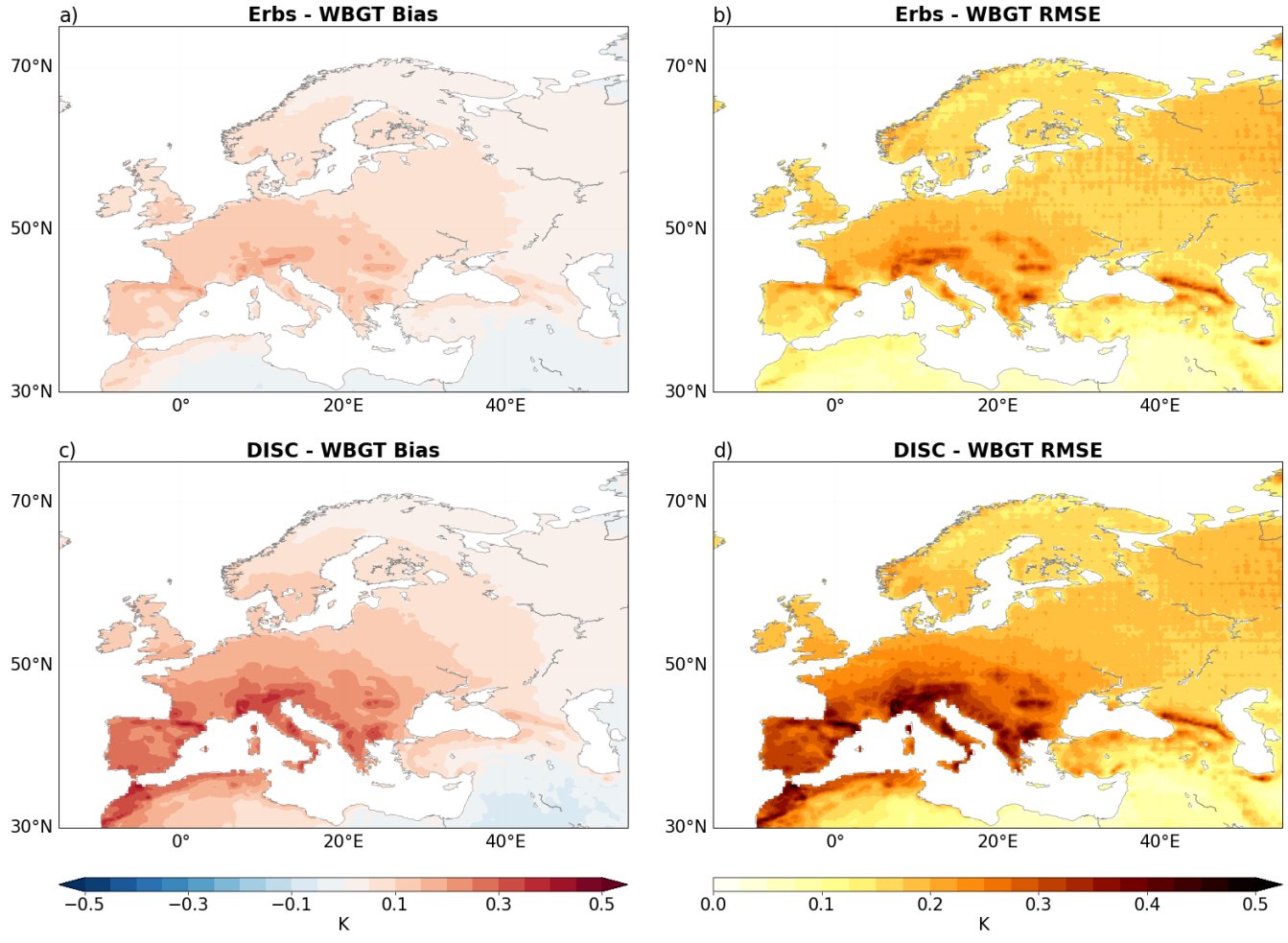

Figure 2. WBGT bias from using approximated solar and thermal radiation components. Same as Figure 1 but for WBGT.

about 0.3–0.4 K. For the DISC model, errors are larger, with bias and RMSE reaching about 0.4 and 0.5 K, respectively. The lower errors found for WBGT are consistent with the weaker influence of radiation inputs on WBGT.

Overall, these findings suggest that the proposed methodology provides reasonably accurate estimates of thermal stress indices from the available SPS outputs, with the Erbs model performing better than the DISC model. This is consistent with results by Weihs et al. (2012), which showed that uncertainties in radiation estimates only moderately impact final UTCI values. Given that the thermal stress categories span 6–8 K for the UTCI and 2–3 K for WBGT, the errors introduced by the proposed methodology are smaller than the width of these categories, particularly when using the Erbs model. Therefore, these errors are expected to have a limited impact on the classification of days exceeding moderate heat stress thresholds (HSD_{UTCI} and HSD_{WBGT}). Nevertheless, misclassifications may occur for days with index values close to the categories boundaries. According to results shown in Figures 1 and 2, we adopt the Erbs model as the reference empirical decomposition model

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

throughout the study. In addition, we apply the same methodology to compute the UTCI and WBGT to both the SEAS5 ensemble and ERA5, ensuring consistency between the two datasets. This strategy enables, for the first time, the derivation of thermal stress indices from SPS outputs, representing a novelty in literature. However, alternative approaches to address current limitations in deriving thermal stress indices from SPS outputs may exist, and are briefly discussed in section 4.

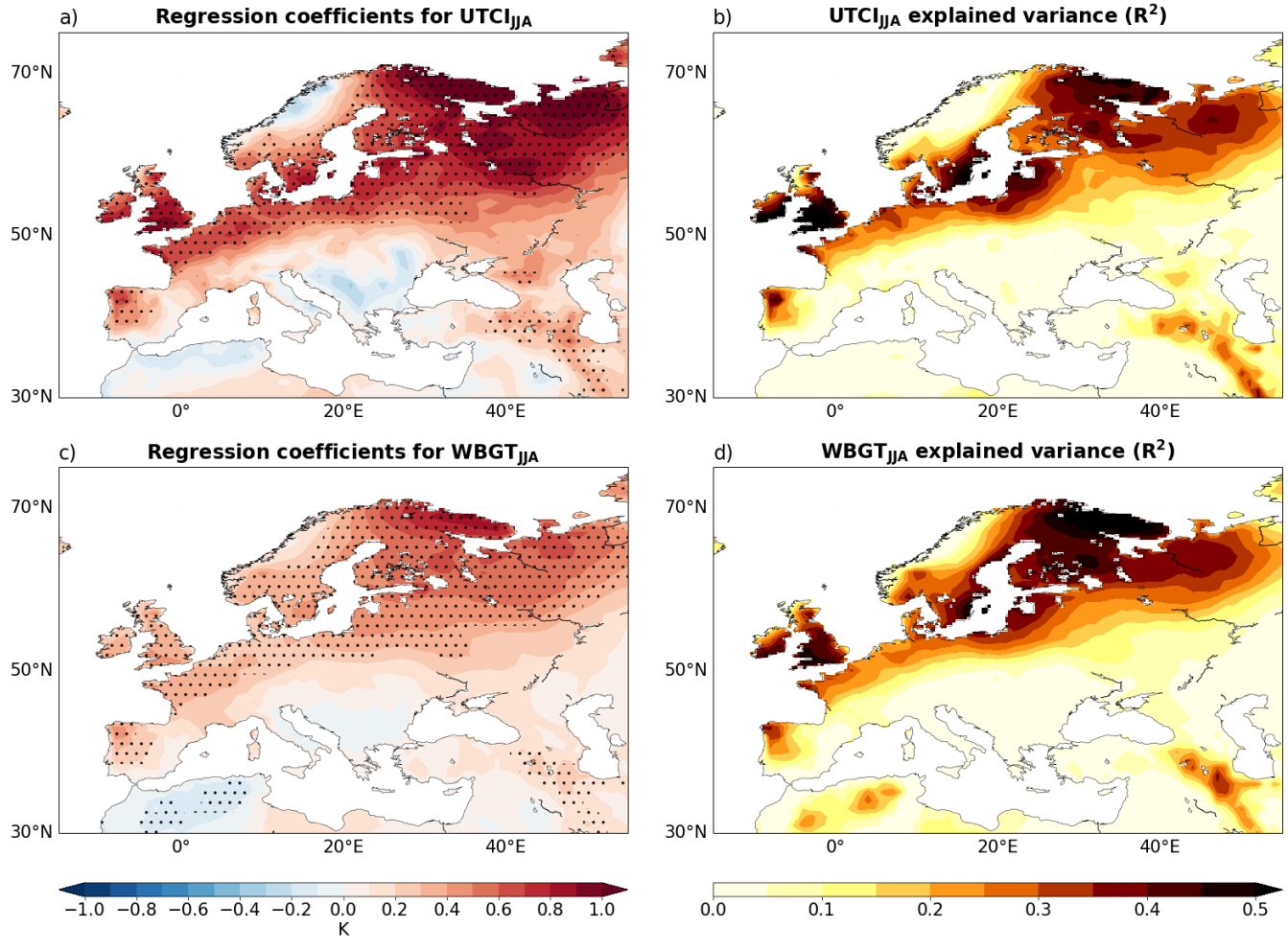

Figure 3. Influence of the summer NAO on summer UTCI and WBGT. a) Linear regression for summer UTCI onto the summer NAO. ERA5 data during 2003–2024 are used. Black dots denote linear regression coefficients that are statistically significant against the null hypothesis of coefficients equal to zero at the 90% confidence level. b) R^2 coefficients for summer UTCI onto the summer NAO. The R^2 coefficient here is adopted as a measure of the fraction of the summer UTCI variance explained by the summer NAO (DelSole and Tippett, 2022). c) and d) Same as a) and b) but for summer WBGT.

Figure 3a indicates that a positive (negative) summer NAO is associated with higher (lower) summer UTCI across much of Europe between 45°N and 70°N, with maxima extending from France to the Baltic region and western Russia. In this

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

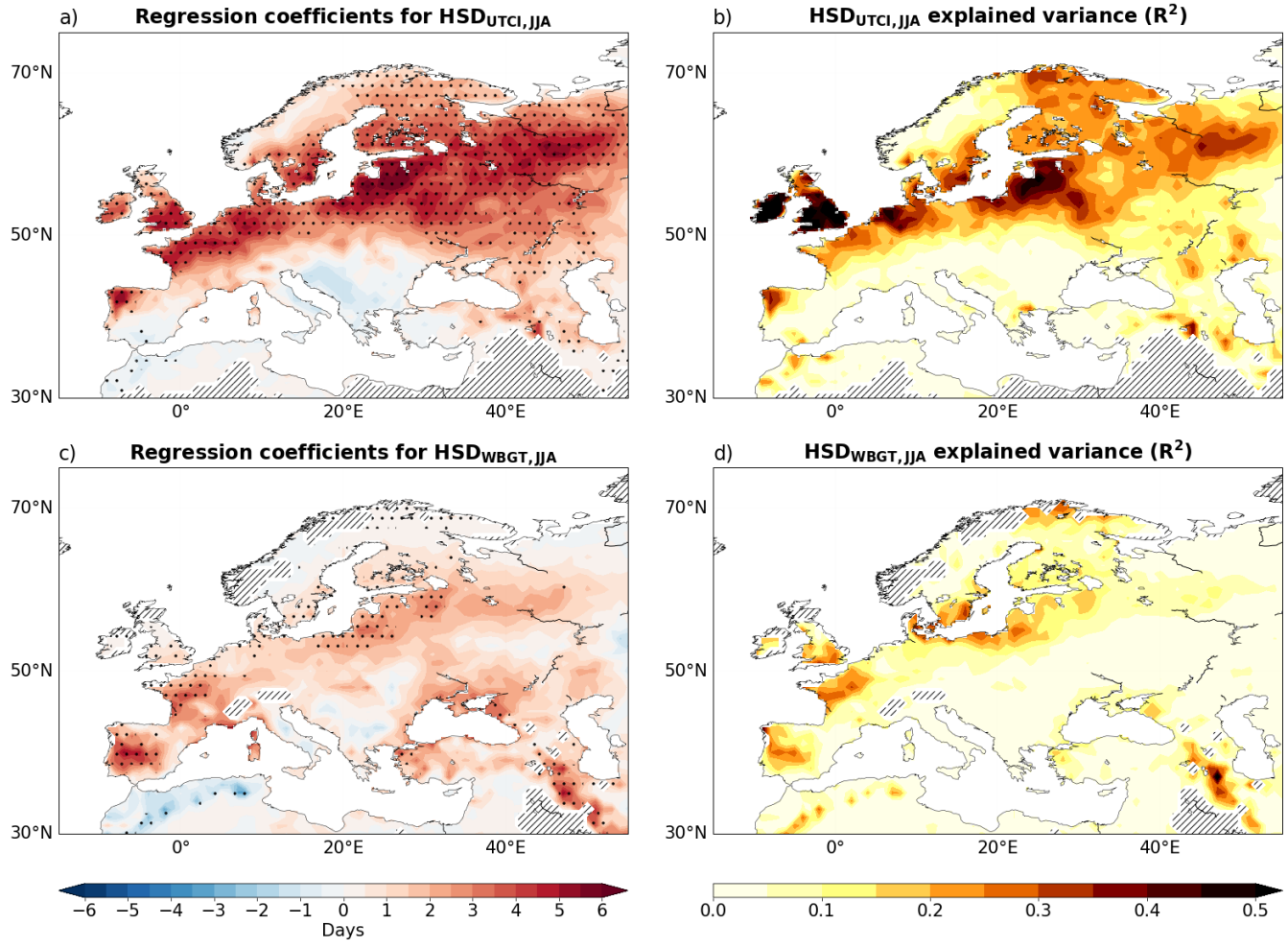

Figure 4. Influence of the summer NAO on summer HSD_{UTCI} and HSD_{WBGT}. Same as Figure 3 but for summer HSD_{UTCI} (a and b) and HSD_{WBGT} (c and d). Hatched areas indicate grid points where HSD_{UTCI} and HSD_{WBGT} show no interannual variability (i.e., the number of summer days exceeding a certain threshold is constant), meaning that the linear regression and R^2 coefficients can not be quantified.

context, the summer NAO accounts for a large portion of summer UTCI variance, with R^2 values exceeding 0.4–0.5 in several regions (Figure 3b). Consistent with this signal, Figure 4a shows that a positive (negative) summer NAO is also associated with more (fewer) HSD_{UTCI} across the same areas, explaining a large portion of its variability (Figure 4b). A similar relationship between the summer NAO and thermal stress conditions is found for WBGT, although with some notable differences. Indeed,
305 Figure 3 (c,d) shows that a positive (negative) summer NAO is associated with increased (reduced) summer WBGT from western to eastern Europe, with explained variance comparable to that of the UTCI. However, the response in HSD_{WBGT} is weaker and more spatially confined to western and central Europe, despite maintaining the same sign (Figure 4c,d). This suggests that, while large-scale circulation associated with the summer NAO effectively modulates mean WBGT, it is not

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

always sufficient to push conditions beyond heat stress thresholds, which may depend more strongly on local or sub-seasonal processes. Nevertheless, the results in Figures 3 and 4 highlight the important influence of the summer NAO on both the intensity and frequency of heat stress conditions across large parts of Europe. This NAO imprint on heat stress conditions is partly in line with previous evidence at the local scale (Głogowski et al., 2020), and supports the NAO-based subsampling strategy to enhance seasonal predictions of thermal conditions that can be hazardous to human health.

Having established the imprint of the summer NAO on European heat stress conditions, we next evaluate how well the SEAS5 represents this mode of variability, focusing on its spatial structure, the fraction of summer low-frequency atmospheric variability it explains, and the interannual variability of its corresponding climate index. Figures 5, S2 and S3 show that SEAS5 reproduces the broad spatial structure of the summer NAO pattern seen in ERA5, although the Z500 anomalies are weaker than in ERA5. SEAS5 also underestimates the fraction of summer low-frequency atmospheric variability explained by the NAO, accounting for 20.2–22.2% over 2003–2024, compared with 27.5–45.2% in ERA5. A further difference concerns the temporal evolution of the summer NAO. ERA5 exhibits substantial temporal variability in both the NAO pattern and its explained variance, consistent with previous studies highlighting the non-stationarity in NAO variability (Famooss Paolini et al., 2024; Outten and Davy, 2024). In particular, the positive NAO centre of action alternates between being mainly confined over Scandinavia and extending toward the central North Atlantic, while the NAO explained variance exhibits a substantial interannual variability (with the largest values in recent years). By contrast, SEAS5 exhibits a more stationary representation of the summer NAO, with comparatively little temporal evolution in either the spatial pattern or the explained variance. Finally, Figure 6a shows that SEAS5 captures part of the interannual variability of the observed summer NAO index, with a correlation of 0.40 between the full ensemble mean and ERA5 indices. However, SEAS5 still underestimates the amplitude of observed NAO index, resulting in a low standard deviation (0.28) and a relatively large RMSE (1.10) for the full ensemble mean index.

Consistent with previous studies (Dunstone et al., 2023; Sun et al., 2024), these results indicate that SEAS5 has limited ability to represent summer low-frequency atmospheric variability over the North Atlantic sector. In particular, it underestimates NAO-related Z500 anomalies and the associated explained variance, fails to reproduce their observed temporal evolution, and only partially reproduces the interannual variability of the summer NAO index. These limitations should be taken into account when interpreting the subsampling results presented next. While subsampling may improve the prediction of the summer NAO phase, residual errors in its spatial structure and explained variance may still affect the dynamical link between the summer NAO and European heat stress conditions, potentially limiting the added value of the post-processing approach adopted here.

Figure 6a shows that the SEAS5 prediction skill improves by applying the teleconnection-based subsampling technique adapted for a real-time forecast environment. Specifically, the correlation between the ERA5 summer NAO and the corresponding SEAS5 index increases from 0.40 (93% confidence level) in the full ensemble to 0.63 (99% confidence level) in the subsampled ensemble, while the RMSE decreases from 1.10 to 0.91. The interannual variability of the summer NAO index is also better captured, rising from 0.28 to 0.67. However, the improvement is smaller than in our previous study (Famooss Paolini et al., 2025), where the subsampling methodology was applied over the 1993–2016 hindcast period. This is consistent with the fact that hindcast-based evaluations may overestimate skill by exploiting information not available in real-time forecasting

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

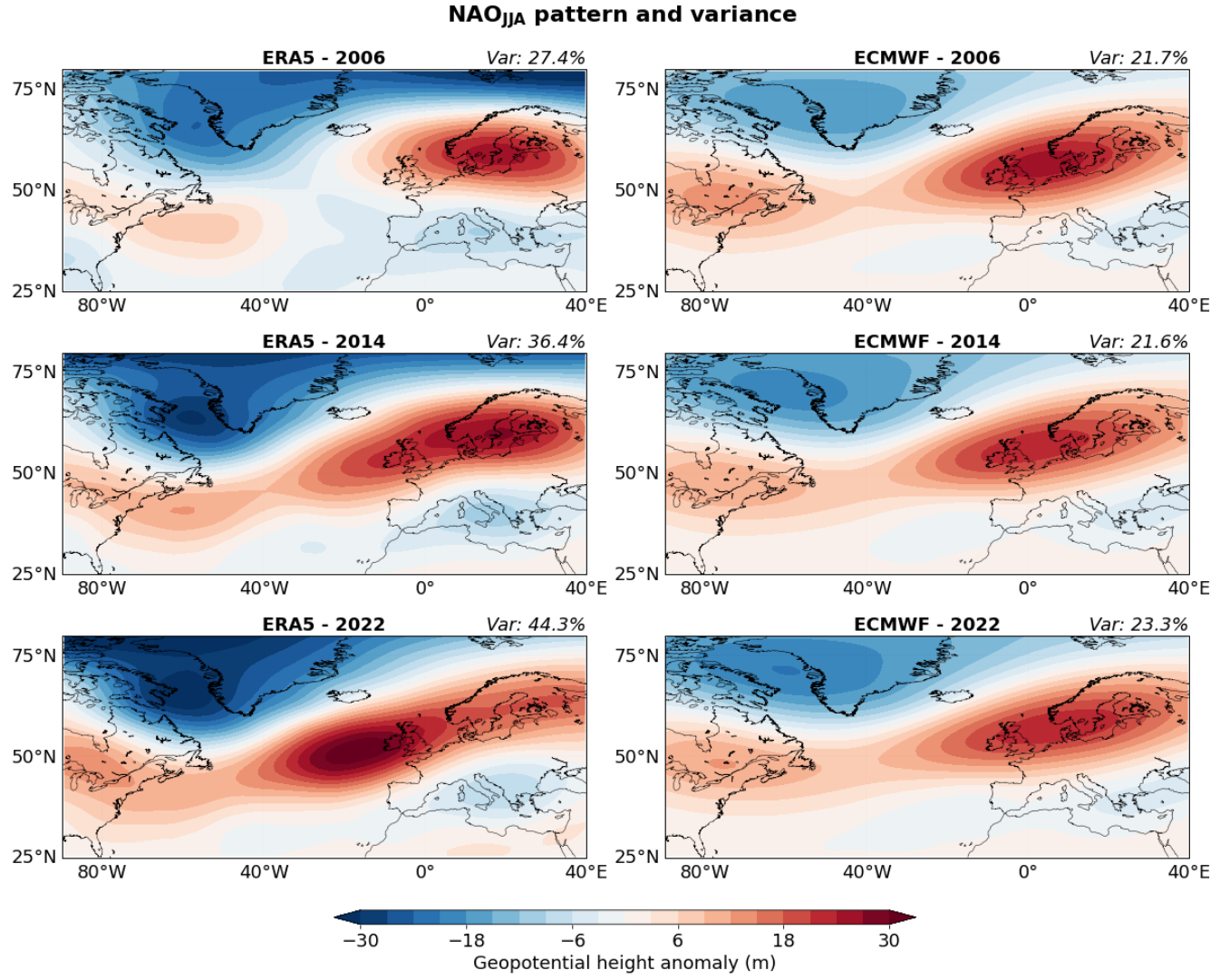

Figure 5. Summer NAO pattern and variance in ERA5 and ECMWF SPS SEAS5 for selected forecast years. The variance of the summer low-frequency variability explained by the NAO is shown at the top right of each panel. The summer NAO is defined as the leading EOF of summer Z500 anomalies over the North Atlantic sector (25°–80°N, 90°W–40°E). For each forecast year, the pattern is estimated using only the preceding 22 years, and then used to project that year's Z500 anomalies to compute the NAO index. Only a subset of 2003–2024 forecast years are shown here. Refer to Figures S2 and S3 for all forecast years.

(Risbey et al., 2021). Differences may also arise from the analysed period and from the use of a single-model ensemble rather than a multi-model ensemble (Hagedorn et al., 2005; Weisheimer et al., 2019, 2020).

345 Results in Figure 6 also show that the performance of the subsampling depends on the number of ensemble members selected for each predictor. When all predictors are used, the correlation between the ERA5 summer NAO and the corresponding

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

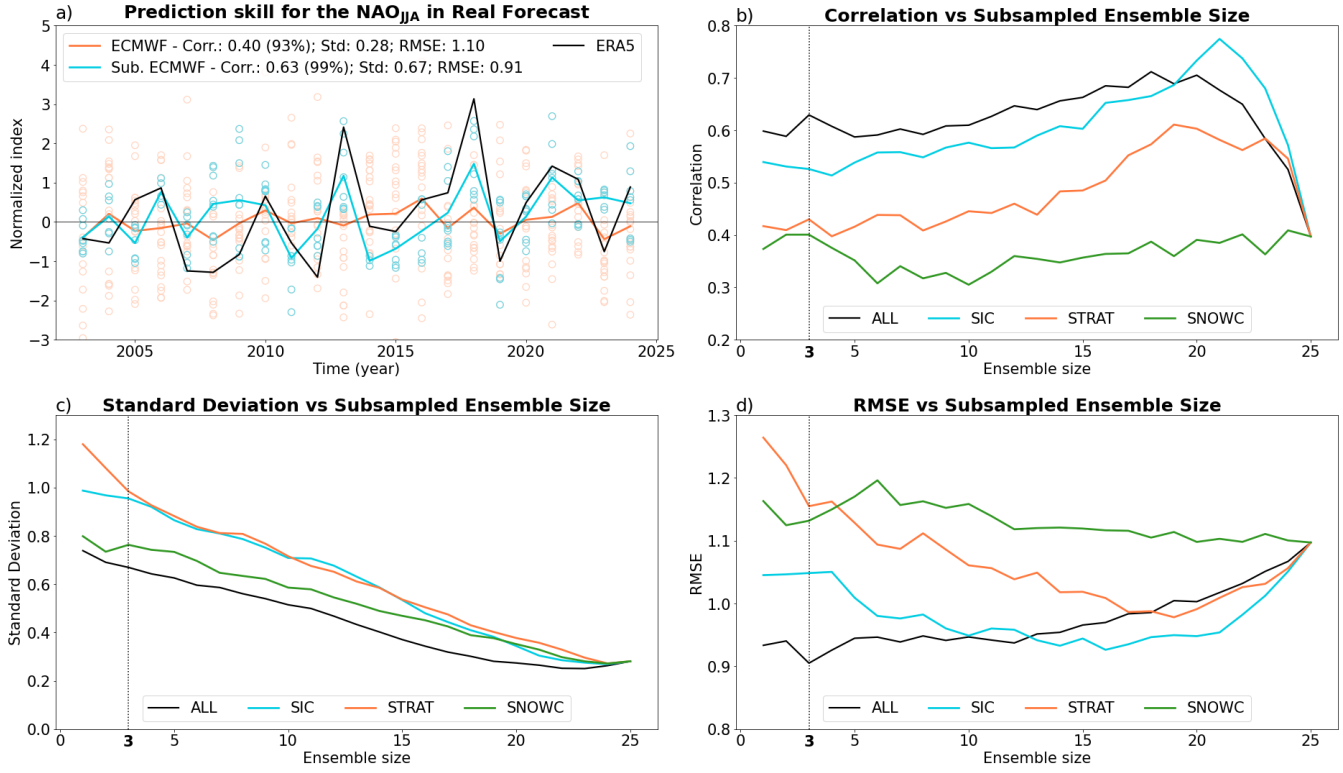

Figure 6. ECMWF SEAS5 prediction skill for summer NAO through teleconnection-based subsampling. a) The summer NAO index in ERA5 (black), full ECMWF ensemble (orange), and subsampled ECMWF ensemble (cyan). The orange (cyan) circles represent each member of the full (subsampled) ensemble. The orange (cyan) solid line represents the ensemble mean of the full (subsampled) ensemble. The subsampling is performed using all NAO predictors and selecting 3 members for every predictor. The correlation between the summer NAO index in ERA5 and the corresponding index as ensemble mean of the full and subsampled ECMWF ensemble is presented in the legend, along with its confidence level in parentheses. The variability of the summer NAO index as ensemble mean of the full and subsampled ensemble is also shown in the legend as standard deviation (Std). Finally, the RMSE for the full and subsampled ensemble is provided. b) Correlation between the summer NAO index in ERA5 and the corresponding index as ensemble mean of the subsampled ECMWF ensemble as a function of members selected per predictor. The subsampling of the ECMWF ensemble is performed using all the NAO predictors (black) and using the Arctic SIC (cyan), the zonally averaged zonal wind in the lower stratosphere (STRAT; orange), and the Northern Hemisphere SNOWC (green), individually. c) Same as b) but for standard deviation. d) Same as b) but for RMSE.

index in the subsampled ECMWF ensemble gradually increases with ensemble size, before dropping rapidly as the subsample approaches the full ensemble (Figure 6b). At the same time, the interannual variability of the summer NAO index decreases (Figure 6c), while the RMSE increases (Figure 6d). This highlights the importance of carefully assessing the number of selected members in order to balance predictive skill and the amplitude of the predicted signal.

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

Furthermore, Figure 6 shows that using all predictors generally yields better performance than single predictors. This is not the case for the summer NAO interannual variability, which remains larger when single predictors are used. This is consistent with the fact that combining all predictors typically results in a larger number of ensemble members, which smooths the signal and reduces the variability of the ensemble mean index. Among individual predictors, Arctic SIC emerges as the most effective, outperforming the configuration with all predictors for relatively large subsamples, particularly in terms of correlation and RMSE (Figure 6 and Figure S4). This may suggest that using SIC alone with a large number of members could further improve SEAS5 prediction skill. However, such a choice would also increase the smoothing effect of ensemble, thereby damping the amplitude of the summer NAO signal (Figure 6c). This would act against the need to increase the predictable component of summer low-frequency atmospheric variability over the North Atlantic sector in systems affected by signal-to-noise problems (Scaife and Smith, 2018; Dunstone et al., 2023), and would likely reduce the subsampling benefits for predictions of thermal stress indices and heat stress conditions.

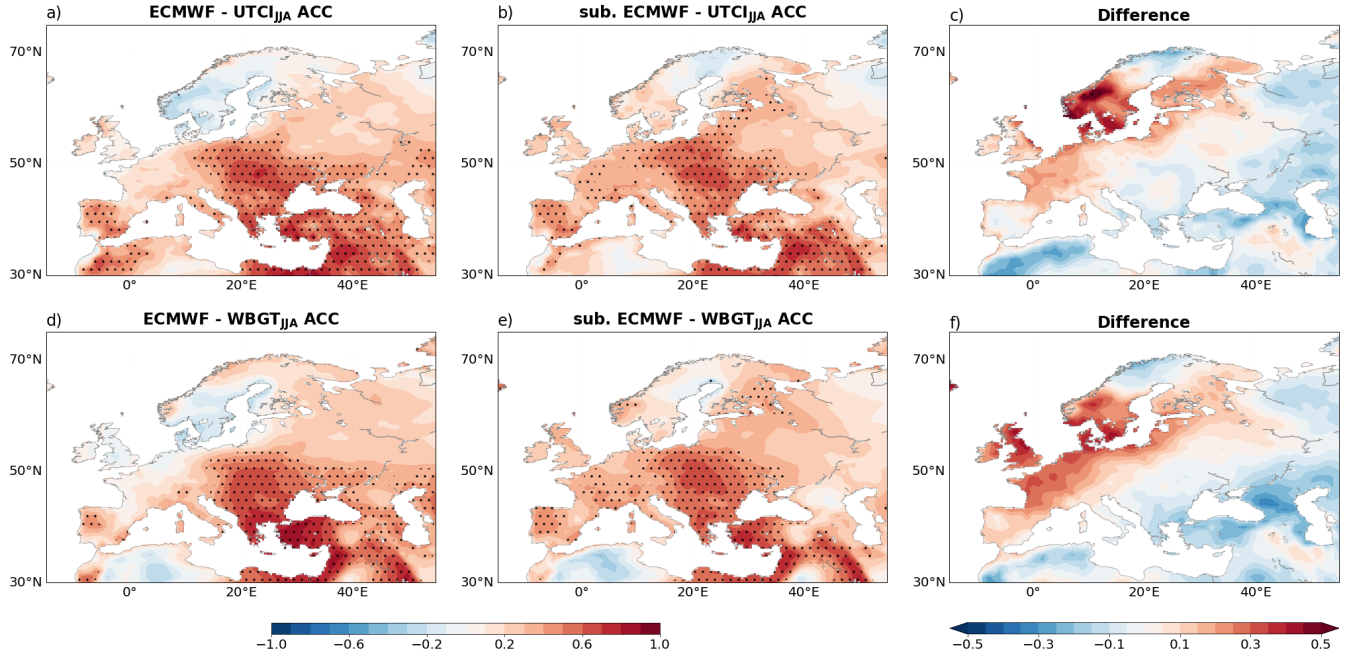

Figure 7. Impact of the summer NAO-based subsampling on the ECMWF SPS SEAS5 prediction skill for summer UTCI and WBGT. a) ACCs computed for the full ECMWF ensemble mean with respect to ERA5 for summer UTCI. b) Same as a) but for the subsampled ECMWF mean. The subsampling is performed using all the NAO predictors and selecting 3 members for every predictor. c) Difference between the ACCs computed for the subsampled ECMWF mean and the ACCs computed for the full ECMWF mean for summer UTCI. Black dots denote ACCs and differences that are statistically significant at the 90% confidence level. d), e) and f) Same as a), b), and c) but for summer WBGT.

Based on the results above, we next assess the seasonal prediction skill for both the summer thermal stress indices and heat stress conditions hazardous to human health, using the full and subsampled ECMWF ensemble. The full ECMWF ensemble exhibits moderate skill in predicting summer UTCI and WBGT over southern and eastern Europe, with maxima over eastern

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

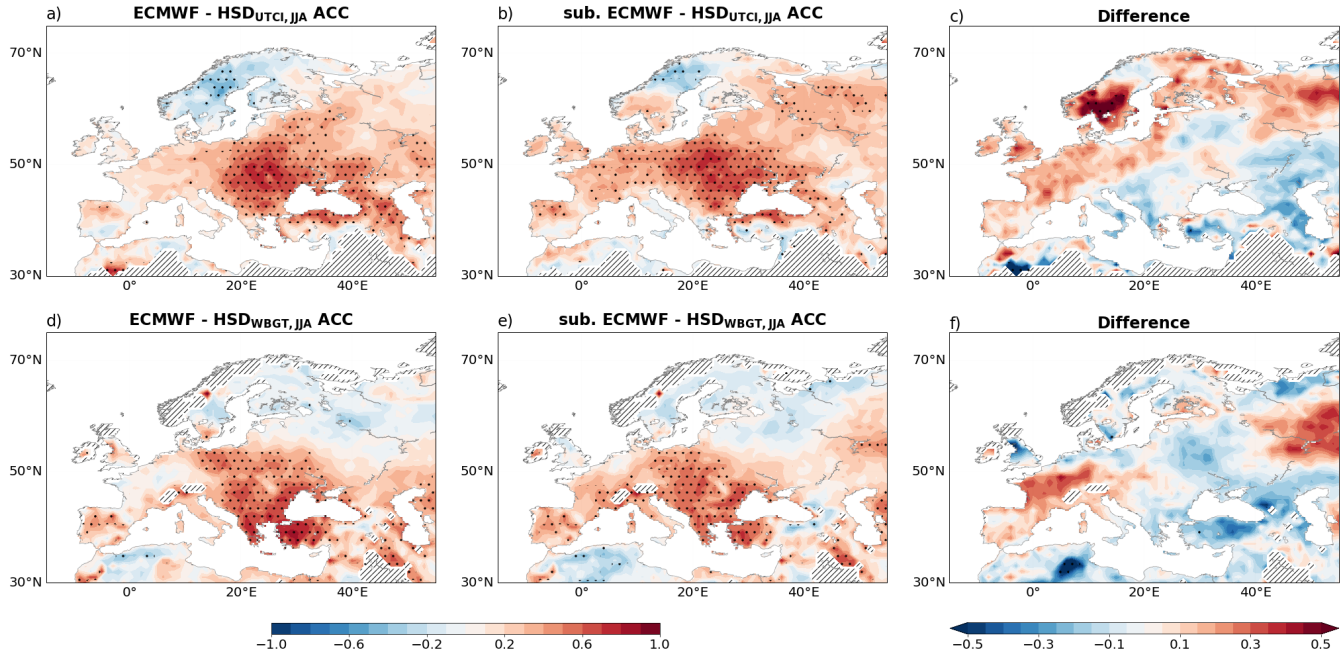

Figure 8. Impact of the summer NAO-based subsampling on the ECMWF SPS SEAS5 prediction skill for summer HSD_{UTCI} and HSD_{WBGT}. Same as Figure 7 but for summer HSD_{UTCI} (a, b and c) and HSD_{WBGT} (d, e, and f). Hatched areas indicate grid points where HSD_{UTCI} and HSD_{WBGT} show no interannual variability (i.e., the number of summer days exceeding a certain threshold is constant), similarly to what shown in Figure 4.

Mediterranean and Balkans regions (Figure 7; left column). However, limited or no statistically significant prediction skill is shown over large portions of western and northern Europe. This lack of skill is partly consistent with limited ability of SEAS5 to represent the summer NAO, in terms of its interannual variability, spatial patterns, and the fraction of summer low-frequency atmospheric variability it explains (Figures 5–6 and S2–S3). The summer NAO is the leading mode of the summer low-frequency atmospheric variability over the North Atlantic sector, influencing climate processes in several European regions. Thus, its suboptimal representation is expected to lead to a reduced predictive performance, especially over those regions such as western and northern Europe strongly influenced by this pattern. Moreover, it should be taken into account that the SEAS5 shows similar issues also in capturing other modes of summer low-frequency atmospheric variability over the North Atlantic sector, such as the Eastern Atlantic pattern (not shown). This aspect further deteriorate the SEAS5 predictive performance for summer season over Europe.

Consistently with our hypothesis, Figure 7 (middle and right columns) shows that retaining only a subset of ensemble members that better represent the summer NAO teleconnections improves the SEAS5 prediction skill of summer UTCI and WBGT across several European regions. This is especially true in western, central and northern Europe, where the summer NAO plays an important role in influencing thermal stress indices and their variance (Figure 3). Although the differences

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

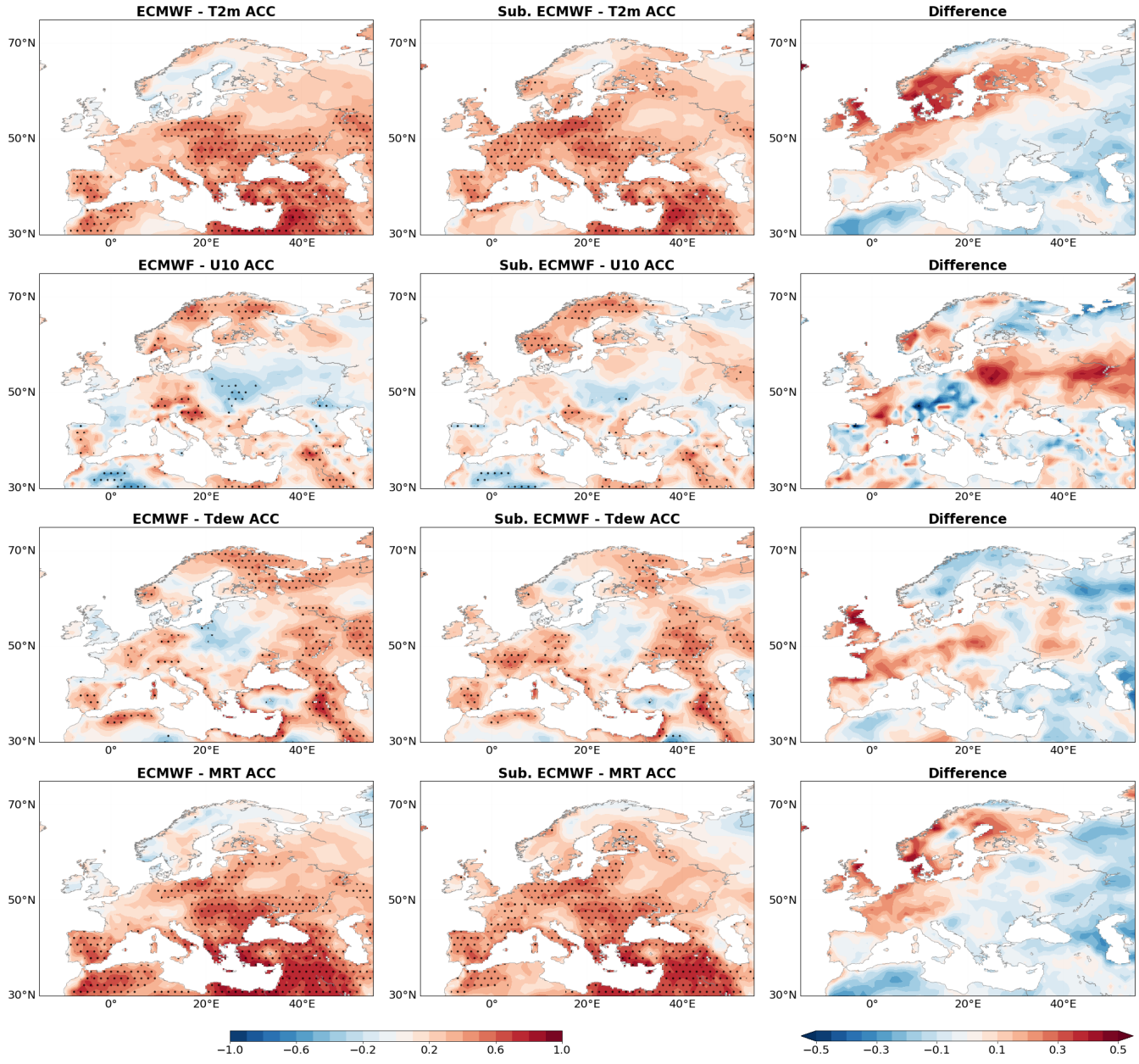

Figure 9. Impact of the summer NAO-based subsampling on the ECMWF SPS SEAS5 prediction skill for climatic variables used to compute summer UTCI and WBGT. Same as Figure 7 but for the climatic variables required to compute the UTCI and WBGT, that is T2m (first row), U10 (second row), Tdew (third row), and MRT (fourth row).

between the subsampled and full ECMWF ensemble are not always statistically significant, the subsampled ensemble displays ACCs which are positive and often statistically significant over a broader area, providing robustness to the improvements

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

achieved through real-forecast subsampling. Differently, the skill decreases over the Balkans and eastern Europe, where the summer NAO exerts a weaker influence on summer UTCI and WBGT. This skill reduction may reflect biases in the spatial pattern of the summer NAO within the ECMWF ensemble. As mentioned above, although the model captures the overall structure of this weather regime, the amplitude of Z500 anomalies is weaker, and the centers of action are slightly displaced compared to ERA5. Moreover, SEAS5 underestimates the fraction of summer low-frequency variability explained by the summer NAO, which likely contributes further to the reduced performance in some portions of Europe. Finally, as mentioned above, it should be noted that summer low-frequency atmospheric variability over Europe is not governed by the summer NAO alone. Other circulation modes, not included in the present subsampling strategy, may exert a stronger influence on thermal stress variability in specific regions and could therefore contribute to the regional skill degradation observed after NAO-based subsampling (Cassou et al., 2005; Bednorz et al., 2019; Sun et al., 2024).

A similar picture emerges for the predictions of heat stress conditions hazardous for human health (Figure 8). The full ECMWF ensemble exhibits moderate skill in predicting both summer $\mathrm{HSD}_{\mathrm{UTCI}}$ and $\mathrm{HSD}_{\mathrm{WBGT}}$ over eastern Europe and the Balkans. By contrast, skill remains weak or not statistically significant over most of western and northern Europe, with areas of low skill that are broader and more spatially fragmented than those found for summer UTCI and WBGT. Consistent with Figure 7, the prediction skill of summer $\mathrm{HSD}_{\mathrm{UTCI}}$ increases after subsampling over large portions of western and northern Europe. A similar behaviour is found for the summer $\mathrm{HSD}_{\mathrm{WBGT}}$, although in this case the improvements are more localized and mainly confined over western Europe. At the same time, subsampling reduces skill over parts of eastern Europe and the Mediterranean, further indicating that the benefits of NAO-based subsampling are not spatially uniform. In this context, it is worth noting that threshold-based metrics are generally harder to predict than seasonal-mean indices, as they are more sensitive to short-term variability and local processes. Moreover, their prediction can be affected by systematic model biases (Crespi et al., 2021), as these biases may alter the number of days exceeding fixed temperature thresholds commonly used to define hazardous heat stress conditions (Zare et al., 2018; Brimicombe et al., 2023). This issue is less relevant for indices based on percentile thresholds (Sulikowska and Wypych, 2020; Famooss Paolini et al., 2026), which depend on the distribution of values within each dataset rather than their absolute intensity.

In this context, Figure 9 shows that the improvements in the heat stress skill are primarily driven by better predictions of the thermal variables used as inputs to the UTCI and WBGT indices. SEAS5 already exhibits relatively high prediction skill for T2m and MRT over large portions of southern, eastern, and central Europe, and this skill is further enhanced after subsampling, especially over western and northern Europe. This spatial pattern is consistent with the improvements found for summer UTCI and WBGT, as well as for their corresponding HSD indices (Figures 7 and 8). By contrast, summer Tdew only partly benefits from subsampling, with skill improvements mainly confined to western Europe. Similarly, the prediction skill of U10 is spatially heterogeneous, showing some improvement over western and northern Europe, but a reduction over central Europe after subsampling. Overall, these results suggest that the enhanced prediction skill of summer UTCI, WBGT, and corresponding HSD indices over western, central and northern Europe arise primarily from better prediction of their thermal components, while the contribution from humidity and wind remains more modest and/or regionally confined.

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

4 Discussion

The results presented in section 3 show that thermal stress indices such as the UTCI and WBGT can be reliably derived from SPS outputs with the methodological framework proposed here, and that the seasonal prediction skill of thermal stress conditions can be improved through summer NAO-based subsampling over several European regions. Despite this promising evidence, several aspects of the present study warrant further discussion.

First, the proposed framework to calculate thermal stress indices from raw SPS outputs relies on simplifying assumptions that may not be always valid across regions and meteorological conditions. While the most robust solution would be the routine provision of UTCI and WBGT within SPS archives, or at least the routine provision of all necessary variables for their direct computation, future research may evaluate to improve the framework presented here. For example, recent studies have shown that the direct component of downwards solar radiation (one of the variables missing from SPS outputs to compute the UTCI and WBGT) can be estimated through explainable machine learning models that outperform widely used empirical decomposition models (Rajagukguk and Lee, 2025). Adapting such approaches to seasonal forecast datasets represents a promising avenue to reduce uncertainties in derived UTCI and WBGT and, ultimately, to strengthen their predictions in SPSs. Second, the improvements in prediction skill for UTCI, WBGT, and HSDs achieved through the operational subsampling are not spatially uniform. In particular, skill decreases over eastern Europe and Balkan region (Figures 7 and 8), consistent with the weaker influence of the summer NAO on thermal stress conditions in these regions (Figures 3 and 4). This suggests that other circulation patterns, such as the Eastern Atlantic pattern, may play a more important role in shaping summer climate variability in regions less strongly influenced by the NAO (Cassou et al., 2005; Bednorz et al., 2019; Sun et al., 2024). Furthermore, these regional differences may partly reflect biases in the simulated summer NAO pattern and in the fraction of atmospheric variability it explains (Figures 5 and S2–S3). Although subsampling improves the prediction of the summer NAO phase, such errors may still affect the dynamical link between the large-scale circulation and surface climate (Głogowski et al., 2020), thereby limiting the benefits of subsampling. These findings motivate further research aimed at better understanding and reducing model biases in the representation of North Atlantic circulation patterns, while extending teleconnection-based subsampling to additional modes of summer atmospheric variability.

Third, the limited improvements found in some European regions may partly reflect the single-model framework adopted here, which makes the results more sensitive to the potential deficiencies of ECMWF SPS SEAS5. We focus on SEAS5 ensemble because it currently provides the longest and most consistent SPS available online, with outputs produced with the same model version and available from 1981 onward. Extending the approach to multi-model SPS ensembles would be the natural next step as soon as sufficiently long and configuration-consistent datasets become available over a comparable period.

Finally, it is noted that the operational value of subsampling depends both on an adequate representation of forecast uncertainty and on the accuracy of the statistical predictions used to select ensemble members. Regarding the former, retaining only a limited number of members around the statistically predicted NAO state may reduce the ensemble spread and lead to overconfident probabilities if the raw subsampled distribution is used directly in downstream applications (e.g., cost–loss decision frameworks; Weisheimer and Palmer, 2014). This issue, which is not assessed here and does not affect the ensemble-mean

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

skill improvements shown in this study, could be addressed through a posteriori recalibration based on historical spread–error relationships, or by transferring uncertainty information from the full ensemble to the subsampled forecast (Gneiting et al., 2005). Regarding the latter, improving the statistical prediction of the summer NAO is key to fully exploiting the potential of the subsampling technique. This highlights the need to consolidate the predictors of summer North Atlantic variability, which remain far less explored than those of its winter counterpart. Furthermore, it motivates the use of more advanced statistical approaches capable of capturing potentially non-stationary, non-linear, and multivariate relationships with these predictors. In this context, recent progress in machine learning offers a promising avenue for improving the prediction of large-scale climate variability modes on seasonal timescales (Sun et al., 2024; Bachèlery et al., 2025). Integrating such approaches to better inform teleconnection-based subsampling could further improve the representation of summer low-frequency variability within the subsampled model ensemble and, in turn, enhance seasonal predictions of European summer climate, including thermal stress conditions.

5 Conclusions

A key innovation of this study is the development of a practical methodological framework to estimate thermal stress indices such as the UTCI and WBGT from SPS outputs. Indeed, current SPSs do not provide these indices directly and their computation is hindered by limitations in both the availability and temporal resolution of radiation data (Table 1). To overcome these limitations, we estimated the missing radiation component using empirical decomposition models and reconstructed instantaneous radiation fluxes from their daily aggregates using climatological relationships derived from ERA5. Validation against ERA5 indicates that the proposed methodology produces spatially consistent UTCI and WBGT fields across Europe, with errors generally below 2 K (Figures 1 and 2). Because the UTCI and WBGT thermal stress categories are defined over intervals larger than these errors, the latter are expected to only marginally affect the numbers of summer days exceeding fixed heat stress thresholds (i.e., $\mathrm{HSD}_{\mathrm{UTCI}}$ and $\mathrm{HSD}_{\mathrm{WBGT}}$; Zare et al., 2018; Brimicombe et al., 2023). These findings demonstrate the feasibility of deriving thermal stress indices from existing SPS outputs, extending their potential applications to seasonal forecasts of human heat stress.

Furthermore, this study show that the seasonal prediction of European summer thermal stress indices and heat stress conditions can be improved through a NAO-based subsampling approach designed to emulate a real-time, operational forecast setting. In particular, the post-processing technique increases the summer NAO prediction skill of the ECMWF SPS SEAS5 from about 0.40 for the full ensemble to about 0.63 for subsampled ensemble during 2003–2024 (Figure 6). In this context, the analysis indicates that combining all predictors generally outperforms single-predictor configurations for NAO-based subsampling, providing a better compromise between improving summer NAO prediction skill and preserving the amplitude of its interannual variability. The improvement in NAO prediction skill translates into enhanced prediction skill for the summer-mean UTCI and WBGT (Figure 7) as well as for the number of summer days exceeding UTCI and WBGT thresholds associated with at least moderate heat stress conditions ($\mathrm{HSD}_{\mathrm{UTCI}}$ and $\mathrm{HSD}_{\mathrm{WBGT}}$; Figure 8). The largest benefits emerge over western, central and northern Europe, where the full SEAS5 ensemble exhibits the lowest prediction skill, and consistent with the dynamical

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

role of the summer NAO in shaping European summer climate and thermal stress conditions (Figures 3 and 4). Results suggest that these skill improvements are primarily driven by a better representation of the UTCI and WBGT thermal components over these regions, while improvements in humidity and wind conditions are weaker and/or more regionally confined (Figure 9).

To conclude, this study represents a first demonstration on how SPSs can be used to deliver forecasts of thermal stress indices and heat stress conditions across Europe, representing an important novelty for health-related climate services. In addition, the study shows that these predictions can be improved through an operational NAO-based subsampling strategy, which enhances skill over several European regions compared with the full dynamical ensemble. The proposed framework could therefore provide the basis for operational post-processing chains, in which raw SPS outputs are transformed into seasonal outlooks of thermal stress indices and heat stress conditions, supporting early preparedness for heat-related health risks. Building on these promising results, future work should aim to further improve the derivation of thermal stress indices from raw SPS outputs, and to increase the gains achievable with subsampling-based hybrid approaches, for example by (i) extending the approach presented here to multi-model SPS ensembles, (ii) broadening the member selection to additional modes of summer low-frequency variability (alongside efforts to reduce biases in their simulated patterns and explained variance), and (iii) leveraging machine learning approaches to improve statistical predictions of the climate indices used to guide ensemble member selection.

Code availability. The underlying code for this study is not publicly available but may be made available to qualified researchers on reasonable request from the corresponding author.

Data availability. The ERA5 dataset is available from Copernicus Climate Data Store (CDS) at the link https://cds.climate.copernicus.eu/datasets?q=era5, while the ECMWF SPS SEAS5 dataset is available from CDS at the link https://cds.climate.copernicus.eu/datasets?q=Seasonal+prediction+seasonal.

Author contributions. LFP conceived the study, designed the methodology, performed the analyses, and wrote the manuscript. PR, CDN, and FW contributed to conceptualization and methodology and reviewed and edited the manuscript. SP, EB, and SDS reviewed and edited the manuscript. CDN and FW provided resources, including hosting LFP at ECMWF (Reading) during part of the work. CDN, FW, PR, SP, EB, and SDS contributed to funding acquisition that supported this study.

Competing interests. All authors declare no financial or non-financial competing interests.

Acknowledgements. All authors were supported by Horizon Europe Project “SoluTion foR mItiGatinG climate-induced hEalth thReats" - TRIGGER (CUP J33C22002420005).

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

References

Achatz, U., Alexander, M. J., Becker, E., Chun, H.-Y., Dörnbrack, A., Holt, L., Plougonven, R., Polichtchouk, I., Sato, K., Sheshadri, A., et al.: Atmospheric gravity waves: Processes and parameterization, Journal of the Atmospheric Sciences, 81, 237–262, https://doi.org/10.1175/JAS-D-23-0210.1, 2024.

Bachèlery, M.-L., Brajard, J., Patacchiola, M., Illig, S., and Keenlyside, N.: Predicting Atlantic and Benguela Niño events with deep learning, Science Advances, 11, eads5185, https://doi.org/10.1126/sciadv.ads5185, 2025.

Bednorz, E., Czernecki, B., Tomczyk, A. M., and Półrolniczak, M.: If not NAO then what?—regional circulation patterns governing summer air temperatures in Poland, Theoretical and Applied Climatology, 136, 1325–1337, https://doi.org/10.1007/s00704-018-2562-x, 2019.

Beverley, J. D., Woolnough, S. J., Baker, L. H., Johnson, S. J., and Weisheimer, A.: The northern hemisphere circumglobal teleconnection in a seasonal forecast model and its relationship to European summer forecast skill, Climate Dynamics, 52, 3759–3771, https://doi.org/10.1007/s00382-018-4371-4, 2019.

Błażejczyk, K., Twardosz, R., Wałach, P., Czarnecka, K., and Błażejczyk, A.: Heat strain and mortality effects of prolonged central European heat wave—an example of June 2019 in Poland, International Journal of Biometeorology, 66, 149–161, https://doi.org/10.1007/s00484-021-02202-0, 2022.

Brimicombe, C., Di Napoli, C., Quintino, T., Pappenberger, F., Cornforth, R., and Cloke, H. L.: Thermofeel: A python thermal comfort indices library, SoftwareX, 18, 101 005, https://doi.org/10.1016/j.softx.2022.101005, 2022.

Brimicombe, C., Lo, C. H. B., Pappenberger, F., Di Napoli, C., Maciel, P., Quintino, T., Cornforth, R., and Cloke, H. L.: Wet Bulb Globe Temperature: Indicating Extreme Heat Risk on a Global Grid, GeoHealth, 7, e2022GH000 701, https://doi.org/10.1029/2022GH000701, 2023.

Bröde, P., Fiala, D., Błażejczyk, K., Holmér, I., Jendritzky, G., Kampmann, B., Tinz, B., and Havenith, G.: Deriving the operational procedure for the Universal Thermal Climate Index (UTCI), International journal of biometeorology, 56, 481–494, https://doi.org/10.1007/s00484-011-0454-1, 2012.

Bruno Soares, M., Daly, M., and Dessai, S.: Assessing the value of seasonal climate forecasts for decision-making, Wiley Interdisciplinary Reviews: Climate Change, 9, e523, https://doi.org/10.1002/wcc.523, 2018.

Budd, G. M.: Wet-bulb globe temperature (WBGT)—its history and its limitations, Journal of science and medicine in sport, 11, 20–32, https://doi.org/10.1016/j.jsams.2007.07.003, 2008.

Cassou, C., Terray, L., and Phillips, A. S.: Tropical Atlantic Influence on European Heat Waves, Journal of Climate, 18, 2805–2811, https://doi.org/10.1175/JCLI3506.1, 2005.

Colón-González, F. J., Soares Bastos, L., Hofmann, B., Hopkin, A., Harpham, Q., Crocker, T., Amato, R., Ferrario, I., Moschini, F., James, S., et al.: Probabilistic seasonal dengue forecasting in Vietnam: A modelling study using superensembles, PLoS medicine, 18, e1003 542, https://doi.org/10.1371/journal.pmed.1003542, 2021.

Crespi, A., Petitta, M., Marson, P., Viel, C., and Grigis, L.: Verification and bias adjustment of ECMWF SEAS5 seasonal forecasts over Europe for climate service applications, Climate, 9, 181, https://doi.org/10.3390/cli9120181, 2021.

Dalelane, C., Dobrynin, M., and Fröhlich, K.: Seasonal forecasts of winter temperature improved by higher-order modes of mean sea level pressure variability in the north Atlantic sector, Geophysical Research Letters, 47, e2020GL088 717, https://doi.org/10.1029/2020GL088717, 2020.

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

De Felice, M., Alessandri, A., and Catalano, F.: Seasonal climate forecasts for medium-term electricity demand forecasting, Applied Energy, 137, 435–444, https://doi.org/10.1016/j.apenergy.2014.10.030, 2015.

DelSole, T. and Tippett, M.: Statistical methods for climate scientists, Cambridge University Press, 2022.

Di Napoli, C., Pappenberger, F., and Cloke, H. L.: Verification of Heat Stress Thresholds for a Health-Based Heat-Wave Definition, Journal of Applied Meteorology and Climatology, 58, 1177–1194, https://doi.org/10.1175/JAMC-D-18-0246.1, 2019.

Dobrynin, M., Domeisen, D. I. V., Müller, W. A., Bell, L., Brune, S., Bunzel, F., Düsterhus, A., Fröhlich, K., Pohlmann, H., and Baehr, J.: Improved Teleconnection-Based Dynamical Seasonal Predictions of Boreal Winter, Geophysical Research Letters, 45, 3605–3614, https://doi.org/10.1002/2018GL077209, 2018.

Dobrynin, M., Düsterhus, A., Fröhlich, K., Athanasiadis, P., Ruggieri, P., Müller, W. A., and Baehr, J.: Hidden Potential in Predicting Wintertime Temperature Anomalies in the Northern Hemisphere, Geophysical Research Letters, 49, https://doi.org/10.1029/2021GL095063, 2022.

Dunstone, N., Smith, D. M., Hardiman, S. C., Hermanson, L., Ineson, S., Kay, G., Li, C., Lockwood, J. F., Scaife, A. A., Thornton, H., et al.: Skilful predictions of the Summer North Atlantic Oscillation, Communications Earth & Environment, 4, 409, https://doi.org/10.1038/s43247-023-01063-2, 2023.

Düsterhus, A.: Seasonal statistical–dynamical prediction of the North Atlantic Oscillation by probabilistic post-processing and its evaluation, Nonlinear Processes in Geophysics, 27, 121–131, https://doi.org/10.5194/npg-27-121-2020, 2020.

Eade, R., Smith, D., Scaife, A., Wallace, E., Dunstone, N., Hermanson, L., and Robinson, N.: Do seasonal-to-decadal climate predictions underestimate the predictability of the real world?, Geophysical Research Letters, 41, 5620–5628, https://doi.org/10.1002/2014GL061146, 2014.

Erbs, D., Klein, S., and Duffie, J.: Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation, Solar energy, 28, 293–302, https://doi.org/10.1016/0038-092X(82)90302-4, 1982.

Famooss Paolini, L., Athanasiadis, P. J., Ruggieri, P., and Bellucci, A.: The atmospheric response to meridional shifts of the Gulf Stream SST front and its dependence on model resolution, Journal of Climate, 35, 6007–6030, https://doi.org/10.1175/JCLI-D-21-0530.1, 2022.

Famooss Paolini, L., Omrani, N.-E., Bellucci, A., Athanasiadis, P. J., Ruggieri, P., Patrizio, C. R., and Keenlyside, N.: Nonstationarity in the NAO–Gulf Stream SST front interaction, Journal of Climate, 37, 1629–1650, https://doi.org/10.1175/JCLI-D-23-0476.1, 2024.

Famooss Paolini, L., Ruggieri, P., Pascale, S., Brattich, E., and Di Sabatino, S.: Hybrid statistical–dynamical seasonal prediction of summer extreme temperatures in Europe, Quarterly Journal of the Royal Meteorological Society, 151, e4900, https://doi.org/10.1002/qj.4900, 2025.

Famooss Paolini, L., Pascale, S., Ruggieri, P., Brattich, E., and Di Sabatino, S.: Drivers of summer extreme temperature trends in Europe: insight from three major heat waves using flow analogues, Climate Dynamics, 64, 235, https://doi.org/10.1007/s00382-026-08171-7, 2026.

Fisher, R.: On the “probable error" of a coefficient of correlation deduced from a small sample, Metron, 1, 3–32, 1920.

Gargiulo, F., Ruggieri, P., Famooss Paolini, L., and Di Sabatino, S.: A Hybrid Statistical–Dynamical Approach for Seasonal Prediction of the Boreal Winter Stratosphere, Accepted in Climate Dynamics, 2026.

Głogowski, A., Bryś, K., and Bryś, T.: Influence of NAO on forming the UTCI index in Kłodzko, Theoretical and Applied Climatology, 142, 1555–1567, https://doi.org/10.1007/s00704-020-03340-y, 2020.

Gneiting, T., Raftery, A. E., Westveld III, A. H., and Goldman, T.: Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation, Monthly weather review, 133, 1098–1118, https://doi.org/10.1175/MWR2904.1, 2005.

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

Gualdi, S., Camino, E. R., Guisado, E. R., Toreti, A., Pasqui, M., Ardilouze, C., Badi, W., Batté, L., Driouech, F., Fröhlich, K., García-Serrano, J., Terzago, S., and Torralba, V.: Seasonal Predictions and Their Applications in the Mediterranean Region: Part II—Prediction-Based Services, International Journal of Climatology, 46, e70 300, https://doi.org/10.1002/joc.70300, 2026.

Hagedorn, R., Doblas-Reyes, F. J., and Palmer, T. N.: The rationale behind the success of multi-model ensembles in seasonal forecasting—I. Basic concept, Tellus A: Dynamic Meteorology and Oceanography, 57, 219–233, https://doi.org/10.1111/j.1600-0870.2005.00103.x, 2005.

Hardiman, S. C., Dunstone, N. J., Scaife, A. A., Smith, D. M., Comer, R., Nie, Y., and Ren, H.-L.: Missing eddy feedback may explain weak signal-to-noise ratios in climate predictions, Npj Climate and Atmospheric Science, 5, 57, https://doi.org/10.1038/s41612-022-00280-4, 2022.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G. D., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P. d., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hewitt, C., Mason, S., and Walland, D.: The global framework for climate services, Nature Climate Change, 2, 831–832, https://doi.org/10.1038/nclimate1745, 2012.

Hinkle, D. E., Wiersma, W., and Jurs, S. G.: Applied Statistics for the Behavioral Sciences, Houghton Mifflin, 2003.

Jendritzky, G., De Dear, R., and Havenith, G.: UTCI—why another thermal index?, International journal of biometeorology, 56, 421–428, https://doi.org/10.1007/s00484-011-0513-7, 2012.

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G., Keeley, S. P. E., Mogensen, K., Zuo, H., and Monge-Sanz, B. M.: SEAS5: the new ECMWF seasonal forecast system, Geoscientific Model Development, 12, 1087–1117, https://doi.org/10.5194/gmd-12-1087-2019, 2019.

Krüger, E. L.: Literature Review on UTCI Applications, Applications of the Universal Thermal Climate Index UTCI in Biometeorology, pp. 23–65, https://doi.org/10.1007/978-3-030-76716-7_3, 2021.

Mărmureanu, L., Antonescu, B., Ene, D., Andrei, S., and Turcu, R.: Changes in the thermal stress across Europe between 1940–2023, International Journal of Biometeorology, 69, 3569–3586, https://doi.org/10.1007/s00484-025-03043-x, 2025.

Matsumura, S., Zhang, X., and Yamazaki, K.: Summer Arctic atmospheric circulation response to spring Eurasian snow cover and its possible linkage to accelerated sea ice decrease, Journal of Climate, 27, 6551–6558, https://doi.org/10.1175/JCLI-D-13-00549.1, 2014.

Maxwell, E. L.: A quasi-physical model for converting hourly global horizontal to direct normal insolation, Tech. rep., Solar Energy Research Inst., Golden, CO (USA), 1987.

Mishra, N., Prodhomme, C., and Guemas, V.: Multi-model skill assessment of seasonal temperature and precipitation forecasts over Europe, Climate Dynamics, 52, 4207–4225, https://doi.org/10.1007/s00382-018-4404-z, 2019.

Neddermann, N.-C., Müller, W. A., Dobrynin, M., Düsterhus, A., and Baehr, J.: Seasonal predictability of European summer climate re-assessed, Climate Dynamics, 53, 3039–3056, https://doi.org/10.1007/s00382-019-04678-4, 2019.

O'Reilly, C. H.: Signal-to-noise errors in early winter Euro-Atlantic predictions linked to weak ENSO teleconnections and pervasive jet biases, Quarterly Journal of the Royal Meteorological Society, 151, e4952, https://doi.org/10.1002/qj.4952, 2025.

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

Outten, S. and Davy, R.: Changes in the North Atlantic Oscillation over the 20th century, Weather and Climate Dynamics, 5, 753–762, https://doi.org/10.5194/wcd-5-753-2024, 2024.

Parajuli, S. P., Biggs, T., Galvez, N. L., Quintana, P., Ciborowski, H., De Sales, F., Zavala Perez, M. A., Jones, C., Thompson, C., Hurtado-Gonzalez, S., et al.: Heat-related rest-break recommendations for farmworkers in California based on wet-bulb globe temperature, Communications Earth & Environment, 6, 359, https://doi.org/10.1038/s43247-025-02327-9, 2025.

Patrizio, C. R., Athanasiadis, P. J., Frankignoul, C., Iovino, D., Masina, S., Famooss Paolini, L., and Gualdi, S.: Improved extratropical North Atlantic atmosphere–ocean variability with increasing ocean model resolution, Journal of Climate, 36, 8403–8424, https://doi.org/10.1175/JCLI-D-23-0230.1, 2023.

Pepler, A. S., Díaz, L. B., Prodhomme, C., Doblas-Reyes, F. J., and Kumar, A.: The ability of a multi-model seasonal forecasting ensemble to forecast the frequency of warm, cold and wet extremes, Weather and Climate Extremes, 9, 68–77, https://doi.org/10.1016/j.wace.2015.06.005, 2015.

Perez, R., Seals, R., Zelenka, A., and Ineichen, P.: Climatic evaluation of models that predict hourly direct irradiance from hourly global irradiance: prospects for performance improvements, Solar Energy, 44, 99–108, https://doi.org/10.1016/0038-092X(90)90071-J, 1990.

Prodhomme, C., Materia, S., Ardilouze, C., White, R. H., Batté, L., Guemas, V., Fragkoulidis, G., and García-Serrano, J.: Seasonal prediction of European summer heatwaves, Climate Dynamics, 58, 2149–2166, https://doi.org/10.1007/s00382-021-05828-3, 2022.

Rajagukguk, R. A. and Lee, H.: Application of explainable machine learning for estimating direct and diffuse components of solar irradiance, Scientific Reports, 15, 7402, https://doi.org/10.1038/s41598-025-91158-x, 2025.

Risbey, J. S., Squire, D. T., Black, A. S., DelSole, T., Lepore, C., Matear, R. J., Monselesan, D. P., Moore, T. S., Richardson, D., Schepen, A., et al.: Standard assessments of climate forecast skill can be misleading, Nature Communications, 12, 4346, https://doi.org/10.1038/s41467-021-23771-z, 2021.

Roberts, C., Vitart, F., and Balmaseda, M.: Hemispheric impact of North Atlantic SSTs in subseasonal forecasts, Geophysical Research Letters, 48, e2020GL0911 446, https://doi.org/10.1029/2020GL091446, 2021.

Scaife, A. A. and Smith, D.: A signal-to-noise paradox in climate science, npj Climate and Atmospheric Science, 1, 28, https://doi.org/10.1038/s41612-018-0038-4, 2018.

Screen, J. A.: Influence of Arctic sea ice on European summer precipitation, Environmental Research Letters, 8, 044 015, https://doi.org/10.1088/1748-9326/8/4/044015, 2013.

Slater, L. J., Arnal, L., Boucher, M.-A., Chang, A. Y.-Y., Moulds, S., Murphy, C., Nearing, G., Shalev, G., Shen, C., Speight, L., et al.: Hybrid forecasting: blending climate predictions with AI models, Hydrology and earth system sciences, 27, 1865–1889, https://doi.org/10.5194/hess-27-1865-2023, 2023.

Smith, D. M., Eade, R., Andrews, M., Ayres, H., Clark, A., Chripko, S., Deser, C., Dunstone, N., García-Serrano, J., Gastineau, G., et al.: Robust but weak winter atmospheric circulation response to future Arctic sea ice loss, Nature communications, 13, 727, https://doi.org/10.1038/s41467-022-28283-y, 2022.

Sulikowska, A. and Wypych, A.: Summer temperature extremes in Europe: how does the definition affect the results?, Theoretical and Applied Climatology, 141, 19–30, https://doi.org/10.1007/s00704-020-03166-8, 2020.

Sun, Y., Simpson, I., Wei, H.-L., and Hanna, E.: Probabilistic seasonal forecasts of North Atlantic atmospheric circulation using complex systems modelling and comparison with dynamical models, Meteorological Applications, 31, e2178, https://doi.org/10.1002/met.2178, 2024.

https://doi.org/10.5194/egusphere-2026-6640
Preprint. Discussion started: 30 October 2026
© Author(s) 2026. CC BY 4.0 License.

Wang, L. and Ting, M.: Stratosphere-Troposphere Coupling Leading to Extended Seasonal Predictability of Summer North Atlantic Oscillation and Boreal Climate, Geophysical Research Letters, 49, e2021GL096 362, https://doi.org/10.1029/2021GL096362, 2022.

Weihs, P., Staiger, H., Tinz, B., Batchvarova, E., Rieder, H., Vuilleumier, L., Maturilli, M., and Jendritzky, G.: The uncertainty of UTCI due to uncertainties in the determination of radiation fluxes derived from measured and observed meteorological data, International journal of biometeorology, 56, 537–555, https://doi.org/10.1007/s00484-011-0416-7, 2012.

Weisheimer, A. and Palmer, T.: On the reliability of seasonal climate forecasts, Journal of the Royal Society Interface, 11, 20131 162, https://doi.org/10.1098/rsif.2013.1162, 2014.

Weisheimer, A., Decremer, D., MacLeod, D., O'Reilly, C., Stockdale, T. N., Johnson, S., and Palmer, T. N.: How confident are predictability estimates of the winter North Atlantic Oscillation?, Quarterly Journal of the Royal Meteorological Society, 145, 140–159, https://doi.org/10.1002/qj.3446, 2019.

Weisheimer, A., Befort, D. J., MacLeod, D., Palmer, T., O'Reilly, C., and Strømmen, K.: Seasonal forecasts of the twentieth century, Bulletin of the American Meteorological Society, 101, E1413–E1426, https://doi.org/10.1175/BAMS-D-19-0019.1, 2020.

Weisheimer, A., Baker, L. H., Bröcker, J., Garfinkel, C. I., Hardiman, S. C., Hodson, D. L., Palmer, T. N., Robson, J. I., Scaife, A. A., Screen, J. A., et al.: The signal-to-noise paradox in climate forecasts: revisiting our understanding and identifying future priorities, Bulletin of the American Meteorological Society, 105, E651–E659, https://doi.org/10.1175/BAMS-D-24-0019.1, 2024.

Williams, N. C., Scaife, A. A., and Screen, J. A.: Underpredicted ENSO teleconnections in seasonal forecasts, Geophysical Research Letters, 50, e2022GL101 689, https://doi.org/10.1029/2022GL101689, 2023.

Wu, R. W.-Y., Büeler, D., Pyrina, M., Day, J., Imamovic, A., Humphrey, V., and Domeisen, D. I. V.: Subseasonal prediction of compound heat and drought events, EGUsphere, 2026, 1–26, https://doi.org/10.5194/egusphere-2026-2465, 2026.

Zare, S., Hasheminejad, N., Shirvan, H. E., Hemmatjo, R., Sarebanzadeh, K., and Ahmadi, S.: Comparing Universal Thermal Climate Index (UTCI) with selected thermal indices/environmental parameters during 12 months of the year, Weather and climate extremes, 19, 49–57, https://doi.org/10.1016/j.wace.2018.01.004, 2018.