Statistical downscaling of seasonal spring-frost predictions

Keywords: Blocking events; Statistical downscaling; Perfect-prognosis; Extreme events; Climate services

Sebastiano Roncoroni[a, *] and Panos J. Athanasiadis[a]

[a]CMCC Foundation - Euro-Mediterranean Center on Climate Change, Italy
[*]Corresponding author: sebastiano.roncoroni@cmcc.it

March 2025

Highlights

- Spring-frost events in Catalonia are statistically associated with large-scale patterns of anomalous blocking frequency.
- Seasonal predictions skilfully reproduce variability of atmospheric blocking events in March at one month lead time.
- A statistical downscaling method based on large-scale patterns of blocking frequency is trained to predict frost risk at high spatial resolution.
- Skill of statistically downscaled seasonal predictions of spring-frost occurrence can improve over existing prediction benchmarks.

This preprint research paper has not been peer reviewed.

Abstract

Spring frost occurring after bud break of grapevines can hinder the plants' growth, and cause large economic losses to the winemaking sector. On-field protection strategies exist at the seasonal horizon, but their cost-effectiveness depends on the availability of accurate risk predictions at an appropriate spatial resolution. In this work, we investigate the predictability of spring-frost events in the important winemaking region of Catalonia, at a high spatial resolution of 10 km, and at the seasonal horizon. This problem setup caters for the specific requirements of one of the case studies considered within the European project ASPECT (Adaptation-oriented Seamless Predictions of European ClimaTe), where researchers and end-users co-develop strategies of climate information production and delivery. To this end, we present a statistical downscaling method which leverages the predictable components of the large-scale atmospheric variability, and builds upon a robust statistical link identified in observations between spring-frost occurrence and atmospheric blocking events. The method is trained with reanalysis and observational data, and subsequently tested using 1-month lead-time retrospective predictions of blocking frequency from several operational seasonal prediction systems. The statistically-downscaled predictions are evaluated probabilistically against standard verification benchmarks, such as the climatological and dynamical predictions. It is found that blocking frequency variability is sufficiently well reproduced by the seasonal prediction systems over the areas important to Catalan frost, and that the skill of the downscaled predictions can improve over the prediction references, although the enhancement is not always statistically significant. We conclude by discussing limitations and perspectives for this approach.

1 Introduction

Spring frost occurs when near-surface air temperature falls below zero degrees during the spring months. If spring frost takes place after bud burst of woody plants —namely, after plants have started producing new leaves— the formation of ice crystals can damage exposed plant organs (Snyder and de Melo-Abreu, 2005), with catastrophic effects on agriculture and the environment. For example, it was estimated that spring frost caused more economic losses than any other climate-related hazard in Greece over the last few decades (Papagiannaki et al., 2014).

The winemaking sector is particularly vulnerable to frost risk. The latest figures indicate that a number of countries worldwide suffered from a decrease

This preprint research paper has not been peer reviewed.

in wine production in 2024 due, at least partly, to frost damage (International Organization of Vine and Wine, 2024): most notably, New Zealand's production dropped by 23% compared to 2023 after frosts in the Marlborough region lead to reduced harvest. Within Europe, Germany, Czechia, and Romania anticipate severe fall-offs, too. In the context of grapevine cultivation, on-field frost protection measures exist on a range of time horizons (Lamichhane, 2021), from seasonal (i.e., one to a few months in advance) to decadal and beyond: their cost-effectiveness, however, depends critically on the availability of accurate, high-resolution frost risk predictions at the relevant time scale.

The goal of this work is to assess the predictability of spring-frost events in Catalonia, at a spatial resolution adequate for on-field management (approximately 10 km) and at the seasonal horizon. This choice tailors the specific needs emerged from a process of interaction and co-production of climate information between scientists involved in the ASPECT (Adaptation-oriented Seamless Predictions of European ClimaTe) project and representatives from a Catalan winemaking company (to provide an account of the interaction process is beyond the scope of this manuscript). Catalonia is a traditional winemaking region in Spain whose winemaking sector was worth approximately 326 millions of euros in 2023 according to recent estimates (Institut Català de la Vinya i el Vi, 2023).

From a scientific perspective, the problem involves the transfer of climate information across spatial scales to predict high-impact events. Global seasonal predictions of the atmosphere and the oceans are produced operationally by national and international centres, thus offering large-ensemble, multi-system predictions. Due to computational costs, however, state-of-the-art seasonal prediction systems run at a horizontal resolution of approximately 100 km, often too coarse to represent surface climate at a spatial scale meaningful to decision makers. One way to circumvent the problem is to use a combination of model and observational data to establish a statistical relationship between coarse-grid model output and local climate: this process of statistical learning is at the very core of statistical downscaling (Maraun et al., 2010). However, it in turn prompts the question of which physical variables should the existence of such a relationship be sought for, and whether the former are predictable at the forecast horizon of interest.

European surface climate is strongly influenced by low-frequency patterns of atmospheric variability (Hurrell et al., 2003), of which leading-order modes such as the North Atlantic Oscillation (NAO) on one side, and atmospheric blocking on the other side, are different but related incarnations (Woollings et al., 2008). Atmospheric blocking events, in particular, are quasi-stationary circulation anomalies associated with high-pressure systems: they occur in

This preprint research paper has not been peer reviewed.

the mid-high latitudes, can last more than a week, and can be thought of as atmospheric "traffic jams" obstructing the westerly flow and the progression of synoptic weather disturbances (Nakamura and Huang, 2018). It has been suggested (Bieli et al., 2015, Brunner et al., 2017, Kautz et al., 2022, Sousa et al., 2018, Trigo et al., 2004) that blocking over the Euro-Atlantic domain has the potential to drive extreme cold conditions in Europe during winter and spring, prominently through anomalous southward advection of cold air. Furthermore, significant progress has recently been reported in the ability of the seasonal prediction systems to capture atmospheric blocking variability over the northern hemisphere (Athanasiadis et al., 2014). Thus, evidence that blocking events may help predict European cold extremes at the seasonal time horizon accrues.

In this work, we seek to leverage the seasonal predictability of atmospheric blocking events over the Euro-Atlantic domain to produce accurate, high-resolution predictions of spring-frost occurrence in Catalonia. To this end, we develop a statistical downscaling method which builds upon a statistical relationship between blocking events and frost we identify using observational and reanalysis data. This approach, known as perfect-prognosis, has considerable provenance in the literature when applied to the NAO and other high-order modes of atmospheric variability to predict surface mean or extreme values (Cionni et al., 2022, Ramon et al., 2021) but, to the best of our knowledge, it has not been tested on blocking variability so far. As a second element of novelty, we adopt a blocking detection algorithm (Sousa et al., 2021) more recent than that used by Athanasiadis et al. (2014), which opens the possibility of considering physically distinct blocking types individually. Alongside the statistical downscaling method, we analyse raw dynamical predictions of frost occurrence based on the predicted values of near-surface daily minimum temperature. The dynamical predictions represent the skill of the seasonal prediction systems taken at face value, and offer a term of comparison to measure any prediction improvements made with the statistical downscaling method.

The structure of the paper is as follows: the data employed in the study and information to retrieve them from online archives are presented in sections 2.1–2.3. Sections 2.4–2.8 contain details on the methods used to build the statistical downscaling method and to evaluate the goodness of its predictions. Our results and the related discussion are offered in Section 3. We conclude in Section 4.

This preprint research paper has not been peer reviewed.

2 Materials and methods

2.1 Observational data

We use near-surface daily minimum air temperature data (T_n) to study spring-frost events in Catalonia. The data were retrieved from the E-OBS dataset version 28.0 (Copernicus Climate Change Service, Climate Data Store, 2020, Cornes et al., 2018) through the Copernicus Climate Change Service (C3S) Climate Data Store portal. E-OBS is a high-resolution observational gridded product based on in-situ measurements over Europe from a distributed network of stations. We downloaded daily minimum temperature fields on a regular grid at 0.1° horizontal resolution over the 1950–2021 period. The horizontal resolution corresponds to a horizontal spacing of approximately 10 km, adequate for on-field management.

Daily frost occurrence is defined as $T_n < 0°$ C (Liu et al., 2018, Meehl et al., 2004). Adoption of a fixed threshold to define extreme cold events is appropriate as plant physiology is directly impacted by temperatures falling below a critical value, which depends on the species and the phenological phase (Augspurger, 2009). Surface wind conditions and humidity modulate frost damage on vegetation (Snyder and de Melo-Abreu, 2005), but in this work we neglect multivariate facets of the problem and concentrate on temperature effects only. Monthly frost occurrence (MFO) is defined based on daily frost events: for a given month, MFO = 1 if frost occurs within that month, and MFO = 0 otherwise. We use MFO to characterise frost hazard at the monthly time scale: this choice is meant to provide a risk outlook which, at the cost of forgoing information on the duration and intensity of events, is accessible to non-specialised users and may feed straightforwardly into simple cost/loss ratio models of economic value (Richardson, 2003).

Frost events are not equally probable in Catalonia over the three spring months (Figure S1 in the Supplementary Material): in some of the southern coastal regions, frost occurs almost exclusively during March —the coldest spring month— and never in May. Elsewhere, the discrepancy between the probability of frost occurrence in March and April–May is not so stark, but a clear difference remains. Frosts taking place late into the growing season pose a greater threat of vegetation damage (Snyder and de Melo-Abreu, 2005), and are thus particularly important to end-users. On the other hand, the climatological start of the growing season is drifting earlier in the year due to global warming (Liu et al., 2018), potentially increasing the impact of early frost events. Considering also that the bulk of frost events within our sample takes place during March, and that robust statistics are key to statistical downscaling, in the remainder of this work we focus on the task

This preprint research paper has not been peer reviewed.

of predicting March MFO.

2.2 Reanalysis data

The large-scale circulation patterns associated with frost events are characterised through analysis of 500 hPa daily geopotential height data (z500) over the 1950–2021 period for February and March. The geopotential data were retrieved from the ERA5 dataset (Copernicus Climate Change Service, Climate Data Store, 2023, Hersbach et al., 2023) through the C3S Climate Data Store portal, where they are archived at 1° spatial resolution and hourly frequency. ERA5 is an atmospheric reanalysis of the Earth's climate produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) employing the Integrated Forecast System combined with various observational datasets to reproduce past atmospheric states of the last decades. Prior to downloading the data, we (i) sub-selected them to the northern hemisphere, (ii) resampled them to a horizontal spatial resolution of 2.5° by bilinear interpolation, and (iii) resampled them to daily temporal resolution by taking daily means. The spatial resolution was chosen to match that used for the blocking detection algorithm of Sousa et al. (2021), which we adopt in this study (Section 2.4).

2.3 Seasonal prediction data

Seasonal predictions are numerical predictions of the Earth's climate system which extend up to a few months ahead in time. They are produced operationally using a variety of prediction systems, each relying on different choices of numerical model, discrete grid, physical parametrisations, and initialisation techniques. Combining the information from multiple prediction systems is essential to sample the uncertainty related to errors in the systems themselves. In addition, ensembles of predictions are created to enhance the signal-to-noise ratio (Siegert et al., 2016), where the noise is due to the chaotic nature of the Earth's climate, while the predictable component of the variability (signal) arises from its slowly-evolving components, including the oceans (Scaife and Smith, 2018). Retrospective predictions, named hindcasts, are provided by each centre for the 1993–2016 period to allow evaluation against observational data.

In this work, we retrieved daily minimum near-surface temperature and 500 hPa geopotential height data from the seasonal forecast portal of the C3S Climate Data Store (Copernicus Climate Change Service, Climate Data Store, 2018a,b). The predictions used here were initialised each year in the 1993–2016 hindcast period on February 1st (burst mode systems) or around

This preprint research paper has not been peer reviewed.

Centre	System	C3S ID	Ensemble size	Initialisation	Documentation
CMCC	SPS3.5	3.5	40	Burst	Gualdi et al. (2020)
DWD	GCFS2.1	2.1	30	Burst	DWD (2024)
ECMWF	SEAS5.1	5.1	25	Burst	Stockdale (2021)
MF	System8	8	25	Lagged	MF (2024)
UKMO	GloSea6-GC3.2	602	21	Lagged	UKMO (2024)

Table 1: Overview of the seasonal prediction systems considered in this study.

the beginning of February (lagged mode systems), and extend until the end of March. A verification period longer than 1993–2016 could be obtained by considering forecasts produced from 2016 onwards alongside the retrospective predictions, but we restricted our analysis to the 1993–2016 hindcast period to adhere to the research standards (Lledó et al., 2020, Ramon et al., 2021, Thornton et al., 2023, Torralba et al., 2024) and allow inter-study comparison. We retrieved data from 5 prediction systems (see Table 1 for a summary): the Euro-Mediterranean Centre on Climate Change SPS3.5 (CMCC), the German Climate Forecast System 2.1 (DWD) the European Centre for Medium-Range Weather Forecasts SEAS5.1 (ECMWF), the Météo-France System8 (MF), and the UK Met Office GloSea6-GC3.2 version 602 (UKMO). For the UKMO, in particular, we considered predictions initialised on February 1st and 9th and on January 25th, for a total of 21 members. Additionally, we collated a multi-model ensemble of predictions (MM) by combining all ensemble members from the 5 prediction systems, for a total of 141 members.

The temperature data were retrieved at daily temporal resolution on a $1° \times 1°$ grid, over the [10°W–30°E, 34°–55°N] European domain. z500 data were retrieved at daily temporal resolution, interpolated to 2.5° horizontal resolution on a regular lat/lon grid, and sub-selected to the northern hemisphere.

Note that, throughout the manuscript, we refer to "seasonal predictions" with a slightly different meaning than in usual applications: here, we use February-initialised forecasts to predict frost occurrence in March (Section 2.1), whereas one would normally target March-April-May (MAM) aggregated variables instead. On the one hand, the shorter forecast horizon brings our problem closer to that of sub-seasonal forecasts (weeks ahead), with expected skill benefits. On the other hand, temporal aggregation becomes less effective (one month available to compute averages instead of three), with anticipated negative impacts. The two main differences with the traditional setup, therefore, wield competing effects on prediction skill.

This preprint research paper has not been peer reviewed.

2.4 Blocking detection

We adopt the blocking detection algorithm of Sousa et al. (2021), to which we refer for the complete discussion. The algorithm performs two tasks: it detects instantaneous blocking structures from daily 500 hPa geopotential height fields, and it implements a tracking procedure to monitor how events evolve in time.

One of the key features of the detection part is that daily structures are classified by type, each type corresponding to a separate stage of a conceptual blocking lifecycle [see Figure 1 in Sousa et al. (2021) for a diagram]. The four types are subtropical ridges (anomalous extensions of the sub-tropical belt, hereafter referred to as *ridge*), omega blocking (characterised by distinctive Ω-shaped geopotential height contours, hereafter referred to as *omega*), Rex blocking (cut-off lows, *Rex*) and hybrid (an intermediate type between *omega* and *Rex*). The tracking part of the algorithm enforces a persistence requirement, and re-classifies each event according to which blocking type occurred most frequently during the blocking lifecycle.

Here, we follow the procedure of Sousa et al. (2021) exactly for the detection algorithm, but opt for a simplified version of the tracking procedure which only implements a lean persistence requirement, similar to that of Athanasiadis et al. (2020). After identifying which daily structures belong to the same event by means of a labelling algorithm from the dask-image Python library (dask-image, 2024), we retain those blocking events such that at least 10 grid points within each instantaneous structure are "blocked" (irrespective of the blocking type), for a minimum of 5 consecutive days. We do not require that exactly the same grid points be blocked throughout the 5-day periods. We tried making reasonable changes to the minimum number of grid points and consecutive days and found that, albeit quantitative details may vary, the main thrust of our results does not change appreciably.

In order to isolate the relationship between each blocking type and frost occurrence, we do not re-classify events according to which blocking type occurs most frequently during their life-cycle. However, we recombine *omega*- and hybrid-type structures into a single blocking type (which we refer to simply as *omega*), as inspection of a representative number of cases revealed a tendency to fluctuate between the two types. In the case of the C3S seasonal predictions, we experimented with applying lead-time dependent bias adjustment to the 500 hPa geopotential height fields prior to blocking detection [see discussion in Athanasiadis et al. (2014) and references therein], and found that including that step does not significantly improve over the results presented here (not shown).

In summary, the outcome of our blocking detection procedure is a daily

This preprint research paper has not been peer reviewed.

catalogue of persistent blocking events in the northern hemisphere, classified on a daily basis into three categories: subtropical *ridges*, *omega*, and *Rex*. For each grid point and blocking type, we compute monthly blocking frequencies as the percentage of blocked days in the month.

2.5 Composite and predictive power analysis

Composite analysis is a technique used to assess whether the statistical properties of one or more physical variables change appreciably when conditioned on the occurrence of an event.

Here, the variables considered are the anomalous monthly frequencies of the three blocking types, where anomalies are computed from ERA5 data relative to the respective climatologies over the 1950–2021 period. The event we condition on is the occurrence of March frost in Catalonia (MFO = 1), as determined by E-OBS data. Specifically, in order to capture local differences in the relationship between frost and blocking events, the composite analysis is performed separately for each target grid point in the domain: the composites themselves are defined as the average of the anomalous blocking frequencies over those years when frost occurs in March within 1950–2021.

To discern whether the composites are significantly different from the climatology we follow the test of Terray et al. (2003): at each grid point, the test evaluates a statistic u and compares its value with a normal distribution. The composites are deemed statistically significant at the chosen significance level when u is outside the acceptance interval of the test's null hypothesis, as computed based on a standard 2-tailed test.

We complement the composite analysis by gauging the predictive power the composites are endowed with. This step is informative because, as noted by Boschat et al. (2016) whose methodology we follow, it may well be that the composite patterns are not strongly associated with the event of interest even if they themselves are statistically significant. First, we compute the pattern correlation between the monthly anomalous blocking frequency fields and the blocking composites over the Euro-Atlantic domain [80°E–40°W, 30°N–90°N]. The pattern correlation is a measure of similarity between the spatial anomalies of two fields: it is defined as their weighted and normalised spatial covariance (Madonna et al., 2017), where the weights account for the poleward convergence of meridians. Pattern correlation is continuously valued within -1 and 1.

After computing the pattern correlations we bin the $[-1, 1]$ interval, and we count the total number of years (samples) and the number of years with frost (events) within each bin. The ratio L = events / samples is such that $0 \leq L \leq 1$, and can be interpreted as the likelihood of frost for a given

This preprint research paper has not been peer reviewed.

correlation bin. The composite patterns wield predictive power to the extent that the associated pattern correlations resolve low and high values of the likelihood: the ideal scenario is such that $L \simeq 0$ when the pattern correlation is negative or zero (i.e., there is little chance of frost in years when the anomalous blocking frequencies are dissimilar to the composite patterns), and $L \simeq 1$ when the pattern correlation is close to 1 (high likelihood of frost when the anomalies resemble the composites).

2.6 Statistical downscaling with perfect-prognosis methods

Statistical downscaling is a mathematical technique which aims at transferring climate information from coarse to finer spatial scales. Statistical downscaling methods divide into different classes, but all share some common ingredients (Maraun et al., 2010). There is a target, fine-grid variable y named predictand, and one or more coarse-grid variables $X = (X_1, X_2, ...)$ used to compute it (the predictors). The statistical downscaling problem consists in learning an hypothesised relationship $y = f(X)$ between predictors and predictand.

The method developed in this manuscript belongs to the perfect-prognosis class. In this case, the relationship $y = f(X)$ is learnt from observational or reanalysis data: typically, X is physically distinct from y and describes the large-scale circulation. If the relationship learnt from observations is later applied to climate model data, the approach is sometimes termed hybrid downscaling (Ramon et al., 2021). Key requirements on the predictors are that X be informative, as there must be some kind of association between predictors and predictand, and reasonably well reproduced by the climate models (Maraun et al., 2010).

Figure 1 offers a schematic of the statistical downscaling method we present in this work. The predictand y is March MFO, defined on a $0.1° \times 0.1°$ grid (approximately 10 km horizontal resolution). $X = (X_1, X_2, X_3)$, on the other hand, is a multivariate large-scale variable on a $2.5° \times 2.5°$ grid (approximately 250 km horizontal resolution). X is defined as the pattern correlation between monthly anomalous blocking frequencies and composite patterns over the Euro-Atlantic domain for the three blocking types. The relationship between y and X is learnt from a set of predictand-predictors tuples (y_t, X_t), where t runs over the 1950–2021 period, y is obtained from E-OBS data, and X from ERA5 data. Subsequently, we apply it to retrospectively predict y using either ERA5 data (perfect-prognosis predictions) or ensemble-mean, 1-month lead-time predictions of X obtained from the

This preprint research paper has not been peer reviewed.

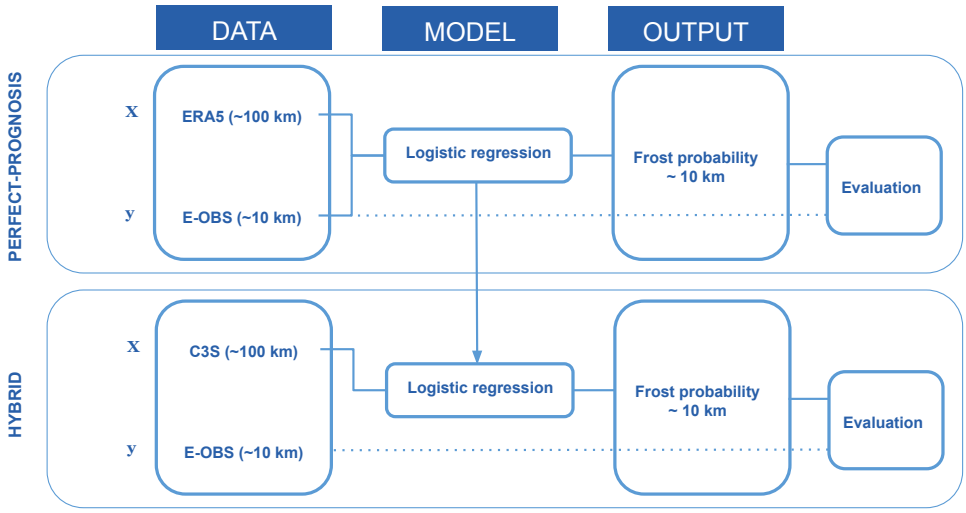


Figure 1: Schematic of the statistical downscaling method.

C3S multi-model seasonal prediction ensemble (hybrid predictions).

Mathematically, the relationship $y = f(X_1, X_2, X_3)$ is modelled with a logistic regression. The logistic regression predicts the probability of $y = 1$ according to the logistic function (Abu-Mostafa et al., 2012):

$$f(X) = \frac{1}{1 + \exp\left(w_0 - wX\right)}\,, \tag{1}$$

where the weights w and w_0 are chosen so as to minimise an L2-regularised cost function. Note that the model issues probabilistic forecasts, i.e. predictions $\hat{y}$ are continuously valued between 0 and 1. For a perfect model, $\hat{y} = 1$ whenever frost occurs in reality and $\hat{y} = 0$ otherwise. We adopt the Leave-One-Out (LOO) cross-validation protocol in order to avoid overfitting of the training data (details in Section S1 in the Supplementary Material).

2.7 Statistical downscaling with raw dynamical predictions

Alongside predictions obtained with the statistical downscaling method, we analyse raw dynamical predictions of frost occurrence. In this case, the coarse-grid variable X used to predict March MFO is MFO itself, and the downscaling step reduces to a simple interpolation to the fine-scale target grid, as detailed below.

This preprint research paper has not been peer reviewed.

For each prediction system (Table 1), we interpolate February-initialised daily minimum temperature fields from the $1° \times 1°$ C3S grid to the $0.1° \times 0.1°$ EOBS grid by bilinear interpolation. Next, we compute MFO for each member separately, and predict frost probability as the ensemble average of MFO. Because MFO is a dichotomous variable, this corresponds to the fraction of ensemble members with MFO = 1. We have experimented with applying post-processing steps further to the bilinear interpolation, such as simple or lead-dependent bias adjustment of daily minimum temperature fields, and found that the predictions do not improve significantly (Figures S2 and S3 in the Supplementary Material). It is beyond the scope of this work to evaluate more complex post-processing methods.

2.8 Forecast verification

We evaluate the goodness of MFO predictions based on two scores for probabilistic forecasts of binary events: the Brier score and the ROC score.

The Brier score is the mean square error of the forecasts, and can be decomposed into a reliability and a resolution term (Wilks, 2006). The reliability term measures the calibration of the forecasts, and puts a penalty on the score when the forecasts are biased compared to the observations. The resolution term measures the discrimination ability of the forecasts, and puts a premium on the score when they can sort observations into subsamples whose relative frequencies of occurrence depart from the climatology.

The ROC score is the area under the Receiving Operating Characteristic curve: it measures the forecasts' discrimination, and is independent of calibration: thus, biased forecasts can attain high ROC scores. Both the Brier and the ROC scores can be related to the economic value of the forecasts through simple cost/loss ratio models (Richardson, 2003).

To place our results into context, we contrast MFO predictions against the climatology $\hat{y}_c$, which is the relevant reference prediction for our problem setting. The climatology invariably predicts the average probability of event occurrence: for example, if March frost occurs once every three years on average, then $\hat{y}_c = 0.33$ identically. To facilitate the comparison between predictions and benchmark, the respective scores can be combined to form a skill score. Given the score s of a prediction system, a benchmark score s_b, and the perfect score s_p for the metric considered, the percent skill score σ is defined as:

$$\sigma = \frac{s - s_b}{s_p - s_b} \cdot 100\,. \tag{2}$$

Skill scores are continuously valued in $(-\infty, 100]$. Perfect predictions have skill 100, while predictions with the same score as the benchmark register

This preprint research paper has not been peer reviewed.

$\sigma = 0$. Negative skill implies that the candidate predictions achieve a worse score than the reference one.

We use a bootstrap method (Cionni et al., 2022) to assess the statistical significance of skill scores: we generate $N = 500$ random realisations of plausible prediction time series by extracting $\hat{y}$ from the actual predictions with resampling. For each realisation, we compute the associated skill score. The original skill score is significant at the α-th significance level if its value is greater than the α-th percentile of the bootstrap skill score distribution. This test evaluates how likely it would be to obtain the actual prediction score with random predictions, regardless of whether there is skill improvement over the benchmark: therefore, it is possible that predictions with negative skill attain statistical significance.

3 Results and discussion

3.1 Climatology of spring frost in Catalonia

Figure 2 (left) shows the climatological frequency of spring-frost occurrences in Catalonia over the reference period 1950–2021 from the E-OBS data. The climatology ranges from less than one event per season in the warm southern coastal areas to several tens of events per season in the north, along the main crest of the Pyrenees mountain range (surface elevation is represented by background colour shading). The highest values, however, are attained at very high surface elevations, i.e. on terrain not suitable for grapevine cultivation. The frost statistics of the corresponding grid points are therefore not relevant to the case study considered in this manuscript, and it is thus desirable to exclude them from the analysis.

To this aim, we consider the climatology of March MFO in Catalonia over the period 1950–2021, shown in Figure 2 (right). To guarantee a minimal balance between the "event" and "non-event" categories, grid points where the 1993–2016 climatology (corresponding to the verification period) is outside the 0.15–0.85 interval are filtered out in grey and will be neglected in the following. Note that the 0.85 probability boundary roughly follows the 700 m elevation contour, highlighting that the chosen criterion effectively selects low- and mid- elevation grid points typically suitable for grape cultivation. Consistently with Figure 2 (left), we observe that MFO is influenced by factors such as terrain and proximity to the sea: the highest climatological probabilities are found at high elevations away from the coast, while the lowest values cluster around low-elevation coastal areas (mild climate). Capturing these local variations is one of the challenges inherent to

This preprint research paper has not been peer reviewed.

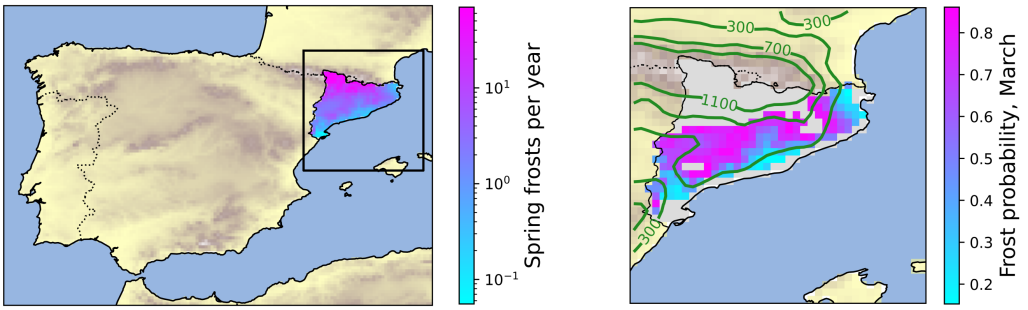


Figure 2: Left: climatological frequency of MAM frost occurrence in Catalonia. Right: climatological probability of frost in March. Surface elevation is shown in the background in shades of brown. In the right-hand-side panel, the 300 m, 700 m, and 1100 m contours are marked by thick green lines (contours have been smoothed with a Gaussian filter for ease of visualisation), and grid points with frost probability outside the 0.15–0.85 range over the 1993–2016 verification period are masked in gray. Data from E-OBS.

the downscaling problem.

Forecasting skill can be artificially inflated if the target variable is non-stationary due to climate change (Wulff et al., 2022). Here we checked for the existence of linear trends in March minimum temperature, of which MFO is a distilled descriptor (see Section S2 in the Supplementary Material for details). We found scarce evidence of robust linear trends in the Catalan domain, which does not warrant that temperature time series be detrended prior to further analysis.

3.2 Skill of the raw dynamical predictions

Figures 3 and 4 show the skill of the raw dynamical predictions against the climatological reference, as measured by the ROC and Brier skill scores, respectively. The seasonal prediction systems tend to perform poorly compared to the climatological benchmark: for DWD, ECMWF, MF and MM, both skill scores are negative over most of the domain. The situation is less clear for the CMCC predictions, which exhibit hints of positive skill scores (particularly ROCSS) in the southern coastal sector, in some cases with statistical significance (dotted grid points). However, even in this case both the ROCSS and the BSS are negative over most of the region, indicating that there is no appreciable improvement over the climatology.

The UKMO system stands out as it consistently performs better than the

This preprint research paper has not been peer reviewed.

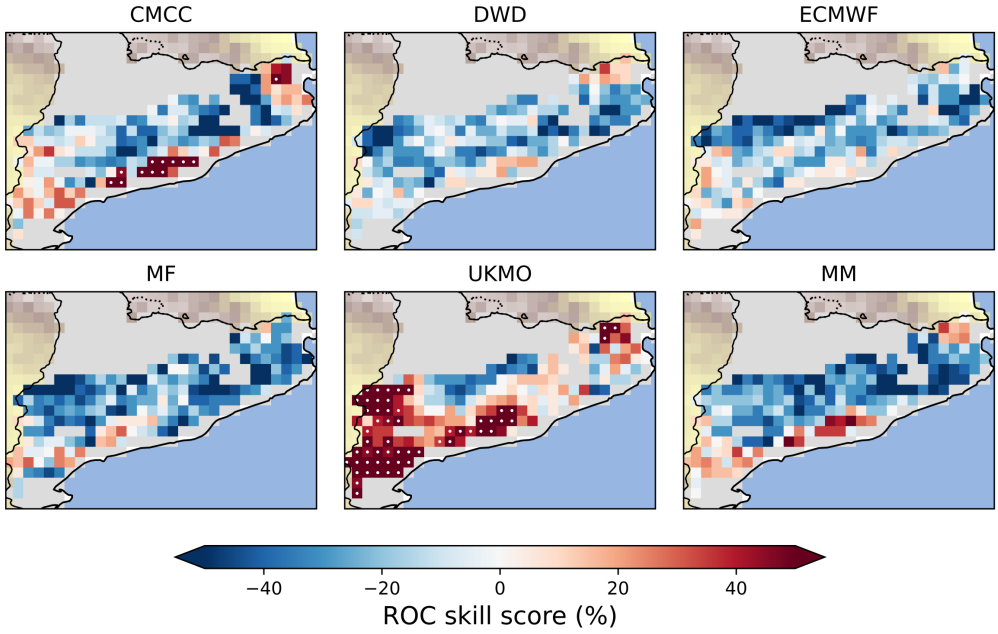


Figure 3: ROC skill score of the raw dynamical predictions against the climatological reference for the individual seasonal prediction models (CMCC, DWD, ECMWF, MF, UKMO) and for the multi-model ensemble (MM). Dots mark grid points where the skill score is statistically significant at the 95% significance level.

climatology over the south-western area, where the majority of grid points attain statistically significant values. It is noteworthy that an individual predictive system can outshine the multi-model ensemble average, as one would normally expect to obtain better predictions from the largest possible ensemble (Scaife and Smith, 2018). This non-trivial result suggests that model errors in predicting MFO are so large for some of the remaining systems that the advantage of picking the best prediction system (here UKMO) outweighs the benefits of maximising the ensemble size. Note also that, overall, the ROCSS takes higher positive values than the BSS. Considering that the ROC score is insensitive to probability biases in the predictions, while the Brier score is affected by calibration errors (Section 2.8), we conclude that the raw dynamical predictions are to a varying extent affected by probability biases.

In the following sections, we will compare the MM and UKMO raw dynamical predictions of March MFO against the predictions obtained with the statistical downscaling model. The MM ensemble was chosen to represent the

This preprint research paper has not been peer reviewed.

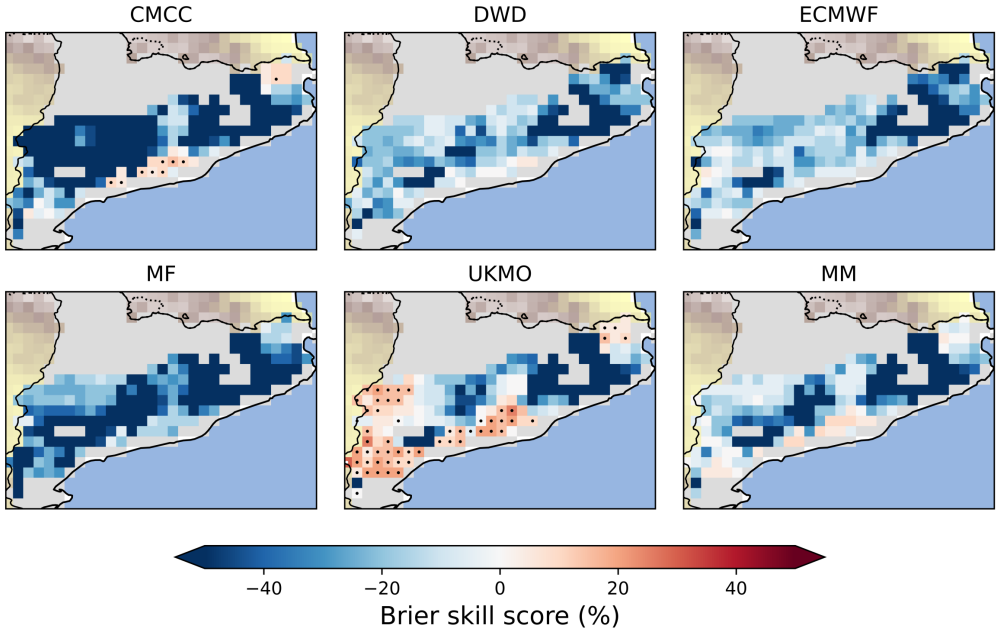


Figure 4: As in Figure 3, but for the Brier skill score.

overall skill of the prediction systems, while the UKMO system was selected by virtue of its good performance.

3.3 Relationship between blocking and March frost

We investigate whether a statistical link exists between blocking occurrence over the Euro-Atlantic domain and March frost in Catalonia. After inspecting the synoptics associated with a few individual frost events (Section S3 in the Supplementary Materials for an example), we turned to composite and predictive power analysis to carry out the task.

3.3.1 Composite analysis

Climate composites of March monthly blocking frequency anomalies against March frost occurrence (i.e., months with MFO = 1) are computed for each of the three blocking types according to the method described in Section 2.5. Figure 5 (left) shows the composites indexed by frost occurrence at the E-OBS grid point nearest to the ASPECT end-user's vineyards (marked by a pink star in the plots). The blocking types shown are *ridge* (top), *omega* (middle), and *Rex* (bottom), and blocking frequency is expressed as the percentage of blocked days per month. Grid points statistically significant

This preprint research paper has not been peer reviewed.

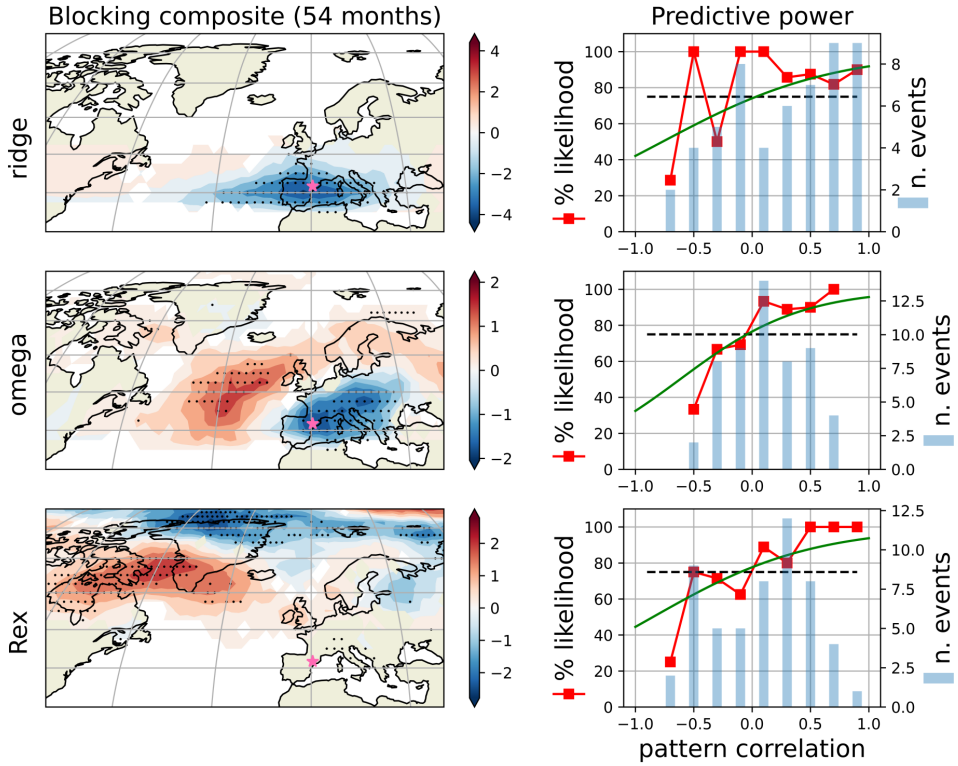


Figure 5: Left: composites of anomalous monthly blocking frequency (colour shading; note the different scale of the colorbars) against March frost occurrence at the E-OBS grid point nearest to the end-user's vineyards (pink star). Dots mark grid points where the blocking composites are statistically significant at the 90% significance level. Right: predictive power associated with the blocking composites: number of frost events (cyan bars) and frost likelihood (magenta lines) per correlation bin of pattern correlation. The dashed black line represents the climatological probability of March frost occurrence, the continuous green line shows the probability of frost occurrence predicted by the logistic regression (Eqn. 1). Blocking events are differentiated by type: *ridge* (top), *omega* (middle), and *Rex* (bottom).

at the 90% significance level according to the test of Terray et al. (2003) are marked by dots.

The composites reveal clear, coherent, and statistically significant patterns of blocking frequency anomalies. Specifically, frost occurrence in Catalonia is associated with: (i) reduced frequency of *ridge*-type blocking (top) over the western Mediterranean and the Iberian peninsula; (ii) increased frequency of *omega*-type blocking (middle) in the central Atlantic combined

This preprint research paper has not been peer reviewed.

with negative anomalies over the western Mediterranean; and (iii) increased frequency of *Rex*-type blocking (bottom) in a zonally-oriented band extending from South Greenland to the Labrador Sea, combined with negative anomalies at higher latitudes over Greenland.

Being constructed from monthly-averaged fields, such patterns of blocking anomalies underscore a *statistical*, rather than *physical*, association between large-scale blocking events and frost occurrence in the target domain. Investigation of the physical mechanisms actually producing the surface temperature extremes warrants use of daily or sub-daily data (Bieli et al., 2015, Sousa et al., 2018) which, for anything beyond the simple example shown in Section S3 in the Supplementary Materials, is outside the scope of this manuscript. Nevertheless, we highlight that the composites of Figure 5 are consistent with well-studied mechanisms of large-scale forcing: *omega*-type blocking events, for example, are associated with cold surface conditions at the southeastern flank of the block due to anomalous southward advection along tilted geopotential contours (Bieli et al., 2015, Brunner et al., 2017), which accords with the positive *omega* anomaly over the central Atlantic region shown in Figure 5 (middle). Conversely, wintertime blocks and subtropical ridges were found to be associated with warm surface anomalies near the centre of the block and at its northern flank due to a combination of anomalous advection of warm Atlantic air, adiabatic heating (Sousa et al., 2018), and —marginally— diabatic processes (Brunner et al., 2017, Sousa et al., 2018). This is well reflected by the negative *ridge* and *omega* composite anomalies over the western Mediterranean. It is more difficult to interpret the *Rex*-type anomalies directly in terms of advective or thermal processes: however, Greenland blocking projects onto the negative phase of the NAO (Woollings et al., 2008), which tends to drive cold conditions over Europe [albeit only weakly so during Spring (Cionni et al., 2022)]. Another possibility is that, by virtue of its large-scale nature, Greenland blocking exerts an indirect influence on cold extremes in Catalonia by statistically reducing the chance of *ridge* occurrence over Iberia which, as shown above, tends to anti-correlate with cold temperature anomalies.

We checked that the results above apply beyond the selected grid point by inspecting individual composites for a number of grid points located in different parts of the region. We found that the quantitative details may change (as desirable if one is to capture local differences in climate), but the general thrust of our arguments stands regardless of the exact point considered: the composites tend to be noisier and less significant, though, at grid points with low probability of March frost (near the 0.15 threshold) due to the smaller number of events they are based upon.

This preprint research paper has not been peer reviewed.

3.3.2 Predictive power analysis

To complement the investigation of the relationship between March frost in Catalonia and blocking events, we present the results of the predictive power analysis alongside the composites. Figure 5 (right) shows the number of events (cyan bars) and percent frost likelihood (magenta line) for each correlation bin of pattern correlation of blocking anomalies. The predictive power plots are shown separately for the three blocking types: *ridge* (top), *omega* (middle), and *Rex* (bottom).

The number of event distributions (cyan bars) indicate that there is a modest preference for frost occurrence at positive correlations, more markedly so for *ridges* (top). A non-negligible fraction of the total events, however, occurs at negative correlations, implying that frost can take place even when the monthly anomalies are not similar to the composite patterns. The *likelihood* of occurrence (L), however, is expressed not by the number of frost events alone, but by its ratio with the total number of samples per bin (magenta lines). If the anomalous patterns of Figure 5 (left) bore no relevance to the occurrence of frost, the likelihood would be the same for each correlation bin, and equal to the climatological probability of frost (dashed black line). Instead, we observe that L drops well below the climatology at negative correlations (around 30% for *omega* and *Rex*), and rises to 100% at high positive correlations. Thus, the predictive power analysis reveals that the composite patterns are endowed with valuable statistical information, which can be leveraged to make frost predictions.

For illustration purposes, we have fitted the logistic relation (Eqn. 1) used by the statistical downscaling method individually to each blocking type, and superimposed the results to the predictive power plots. The logistic regressions (green lines) capture the sampled likelihood L (magenta lines) accurately. This underscores that the probabilities predicted by the logistic regression model can be interpreted simply as a multivaried and regularised version of the likelihood curves, which adds to the explainability of our downscaling method. Similar to the composite analysis, these results generalise well to other points in the domain except at very low climatological probabilities, where noise begins to hinder the interpretation of the predictive power plots.

3.4 Perfect-prognosis predictions

Perfect-prognosis predictions are obtained by applying Eqn. 1 to predictors X computed from ERA5 reanalysis data. The training period is 1950–2021 and, to remain consistent with the dynamical and hybrid predictions, the

This preprint research paper has not been peer reviewed.

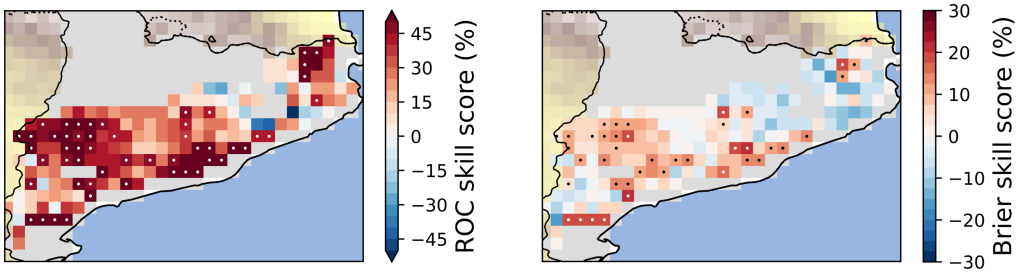


Figure 6: ROC (left) and Brier (right) skill scores for the perfect-prognosis predictions obtained with the statistical downscaling method (note the different scale of the colorbars). The training period is 1950–2021, and the verification period is 1993–2016. Points where the skill score is statistically significant at the 95% significance level are marked by dots. Skill scores are computed against the climatological reference.

verification period is 1993–2016.

The skill associated with perfect-prognosis predictions can be interpreted as the statistical model's potential skill, namely, the skill which would be obtained under perfect knowledge of the predictors. Such a perfect knowledge cannot be achieved in reality due to (i) errors in the seasonal prediction systems, and (ii) the fact that only a (typically small) fraction of the observed signal is predictable: nevertheless, perfect-prognosis predictions are an important complement to hybrid predictions (discussed in Section 3.6 below), as they allow to estimate the impact that imperfect knowledge of X bears on model performance.

Figure 6 shows the ROC (left) and Brier (right) skill scores for the perfect-prognosis predictions (note the different scale of the colorbars), where the skill scores are computed against the climatological reference. Remarkably, the ROC skill score is positive over the majority of the domain, indicating that the statistical downscaling model performs overall better than the climatology for this metric. The highest scores (with peaks of above 75% improvement over the climatology) are attained in the west, in the central coastal areas, and in the north-east. Furthermore, the ROC skill score is significant for about 29% of the grid points at the 95% significance level.

The Brier skill score displays a spatial pattern similar to the ROCSS, but the values attained are generally lower (22% at best), positive over a smaller fraction of the domain, and less significant (22% of the grid points). The Brier score contains a penalty term associated with the forecasts' reliability (Section 2.8): therefore, similar to the raw dynamical predictions,

This preprint research paper has not been peer reviewed.

perfect-prognosis predictions are likely affected by probability biases. Nevertheless, the consistently positive values registered for the ROCSS show that the perfect-prognosis predictions retain the ability to resolve frost occurrence: to the best of our knowledge, this is the first time that the possibility of producing high-resolution, skilful probabilistic predictions of spring-frost events based on atmospheric blocking events is explicitly demonstrated.

Perfect-prognosis skill is highly dependent on whether or not blocking anomalies and composites are computed in Leave-One-Out mode (a test year characterised by strongly anomalous blocking frequency fields can impact the climatology appreciably). If LOO cross-validation does not include anomaly computations (Figure S4 in the Supplementary Material), the maximum ROCSS rockets to 90%, with 75% of the target grid points being statistically significant (compared to a maximum skill score of 75% and 29% of significant grid points with LOO). The BSS is similarly greatly enhanced. This skill, however, does not represent the potential skill which could be achieved in a real operational setting fairly. In view of this, we caveat against adoption of the statistical downscaling method presented here without full LOO coverage.

Finally, sensitivity tests we performed indicate that *omega* blocking is the strongest predictor for March MFO in Catalonia (not shown). Due to the paucity of training data, however, it is not possible to robustly determine whether the finding reflects a real signal or is due to statistical fluctuations within our training sample. Furthermore, the seasonal prediction systems tend to have higher skill for *ridge*- and *Rex*- blocking events than for *omega*-type ones (see Section 3.5 below). For these reasons, we do not attempt to optimise the model by sub-selecting its predictors, and retain information from all three blocking types in the following.

3.5 Skill of blocking frequency predictions

Having established that perfect knowledge of the predictors would translate into skill improvements over the climatological benchmark, the ability of the statistical downscaling method to issue actual frost predictions ultimately rests on how well the seasonal prediction systems capture large-scale blocking frequency anomalies.

Figure 7 shows the Pearson correlation coefficient between the observed (ERA5 data) and predicted (ensemble-mean C3S multi-model) blocking frequency anomalies for each blocking type over the Euro-Atlantic domain (colour shading), computed with LOO cross-validation. As before, the verification period is 1993–2016, and grid points where the correlation is statistically significant at the 95% significance level are marked by dots. The

This preprint research paper has not been peer reviewed.

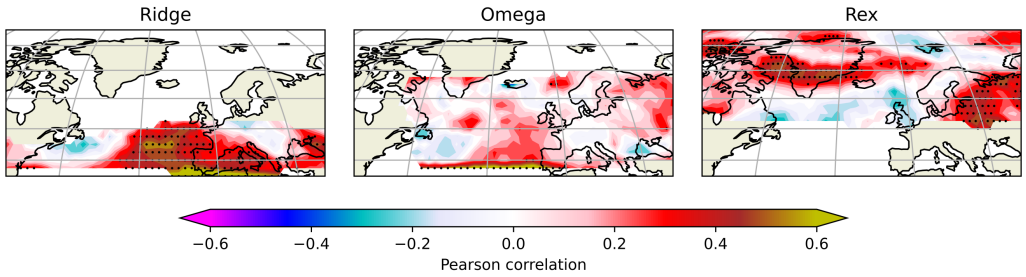


Figure 7: Pearson correlation coefficient between predicted (ensemble-mean C3S multi-model) and observed (ERA5) monthly blocking frequency anomalies for *ridge* (left), *omega* (middle), and *Rex* (right) blocking over the Euro-Atlantic domain (colour shading). Grid points where the correlation is statistically significant at the 95% significance level are marked by dots.

statistical significance is based on a one-tailed test with correction for temporal autocorrelation (Bretherton et al., 1999). Correlations are shown only where the climatological frequency of blocking occurrence is greater than 1%.

In order to maximise the number of available ensemble members and thereby increase the signal to noise ratio (Siegert et al., 2016), the correlations shown in Figure 7 are based on multi-model, ensemble-mean predictions of monthly blocking frequency anomalies (note, though, that the blocking detection algorithm is applied to each member individually). We highlight that, contrary to the MFO predictions shown in Section 3.5, no individual prediction system consistently achieves higher skill for blocking frequency than the multi-model across the domain considered (figures S7–S11 in the Supplementary Material). By the same token, we do not analyse predictions based on individual ensemble members: this foregoes information on the ensemble spread, but we rely on the evaluation metrics presented in Section 2.8 and on comprehensive comparison with relevant benchmarks to provide a measure of prediction uncertainty (see Section 3.6 below).

The Pearson correlation coefficient between predicted and observed *ridge*-type blocking events is shown in 7 (left). Frost occurrence in Catalonia is associated with negative *ridge* anomalies over the south-eastern Atlantic, the Iberian peninsula, and the western Mediterranean (Figure 5, top). Figure 7 (left) demonstrates that the C3S multi-model predictions consistently attain positive correlation coefficients over that area, with a peak value of 0.6. The predictions achieve statistical significance at the 95% significance level over the eastern part of the sector considered. The assessment is less clear for *omega*-type blocking, shown in Figure 7 (middle). Here, the regions

This preprint research paper has not been peer reviewed.

where blocking variability is important to frost occurrence in Catalonia are the central Atlantic (where *omega* blocking favours cold conditions in Catalonia) and the Iberian area (where, similarly to *ridges*, *omega* occurrence is conducive to warm conditions in the target domain), as shown in Figure 5 (middle). In the central Atlantic sector and on the western flank of the Iberian sector correlation is positive, although the strength of the association is weaker than for *ridges* (with coefficients of 0.3 at most) and not statistically significant. Hints of positive and significant skill can be found within a narrow latitudinal band south of 37.5° N, but these are probably associated with fluctuations of the average subtropical belt position (which is used by the blocking detection algorithm to identify daily structures), and only marginally relevant to predictions. There are areas of positive *omega* skill over and west of Greenland too, possibly associated with *omega* blocking maturing into *Rex*-type events. Correlation, however, degrades starkly over central Europe and the Italian peninsula, where the Pearson coefficient is null or weakly negative. Finally, *Rex* variability is well represented over northern and southern Greenland, with peak coefficients of 0.5 and statistical significance in the latter case. *Rex* correlation is also statistically significant over the Baltic region and north-western Russia, although that is less relevant to the case study considered here.

Our analysis underscores that, with the notable exception of *omega* blocking over central and southern Europe, monthly anomalies of blocking occurrence frequency are skilfully reproduced by the C3S multi-model ensemble over the areas important to Catalan frost.

3.6 Hybrid predictions

To what extent does the potential skill associated with the statistical downscaling model translate into predictive skill? We address this question by analysing the predictions obtained by feeding the multi-model, 1-month lead-time predictions of X into Eqn. 1. Figure 8 shows the hybrid predictions' ROC (left) and Brier (right) skill scores (note the different scale of the colorbars) computed against the climatology (top), and the MM (middle) and UKMO (bottom) raw dynamical predictions.

The hybrid predictions perform overall better than the climatology for the ROCSS, the skill score being positive over a large fraction of the domain (top left). The BSS, on the other hand, attains positive values in parts of the central and western sectors only (top right), again pointing at the existence of probability biases in the predictions. For both metrics, though, skill is on average lower and less significant than in the perfect-prognosis case, probably due the comparatively weak correlation coefficients found between predicted

This preprint research paper has not been peer reviewed.

and observed *omega*-type anomalies over the Euro-Atlantic area (Figure 7).

Careful comparison of the top panels of figures 8 and 6, however, reveals that the hybrid skill can locally surpass that of the perfect-prognosis predictions. This counterintuitive effect is likely the result of errors in the predicted blocking anomalies combined with the small size of the 1993–2016 hindcast period used for verification. Such fluctuations are expected to be negligible on average when the verification period is sufficiently long, but remain appreciable here. This highlights that perfect-prognosis skill can be strictly interpreted as the potential skill for the statistical predictions in the limit of large verification datasets only.

Hybrid predictions compare favourably against the multi-model raw dynamical predictions for both metrics (middle panels): the ROCSS and the BSS are consistently positive over most of the domain but, in spite of the large skill scores achieved, only few grid points attain statistical significance. Remember that the skill score is a relative measure of goodness which pits a predictive system against a benchmark: when the benchmark is not performant (as it is not here, see figures 3 and 4), large skill scores do not guarantee statistical significance.

The bottom panels of figure 8 show that hybrid predictions improve over the UKMO raw dynamical predictions only locally. The regions of improvement are the central and eastern parts of Catalonia, where UKMO predictions are inferior to the climatology (figures 3 and 4). In this case too, though, statistical significance is achieved only sparsely. In the south-western part of the region, instead, it is the UKMO which asserts itself as the best-performing system, reflecting the high levels of skill found in that sector in Section 3.2.

Our results underscore the complexity of the prediction landscape for March frost risk in Catalonia. No individual prediction system delivers optimal forecasting performances across the geographical domain and the evaluation metrics considered. The climatology is a robust (particularly in the absence of strong trends, Section S2 in the Supplementary Material) but suboptimal choice, although it should still be preferred in selected locations if prediction success is measured by BSS. Dynamical predictions made with the UKMO system are clear winners in the south-eastern areas, but their skill plummets elsewhere. Hybrid predictions made with the statistical downscaling method lack strong statistical significance, which hinders a conclusive assessment of their robustness in a real operational setting. On the other hand, the associated skill scores are more spatially coherent and, especially for ROCSS, consistently improve over the climatology.

Figure 8, overall, demonstrates that there is great promise in leveraging the predictability of atmospheric blocking events to forecast surface climate hazards at the seasonal horizon. Further work is needed, however, to con-

This preprint research paper has not been peer reviewed.

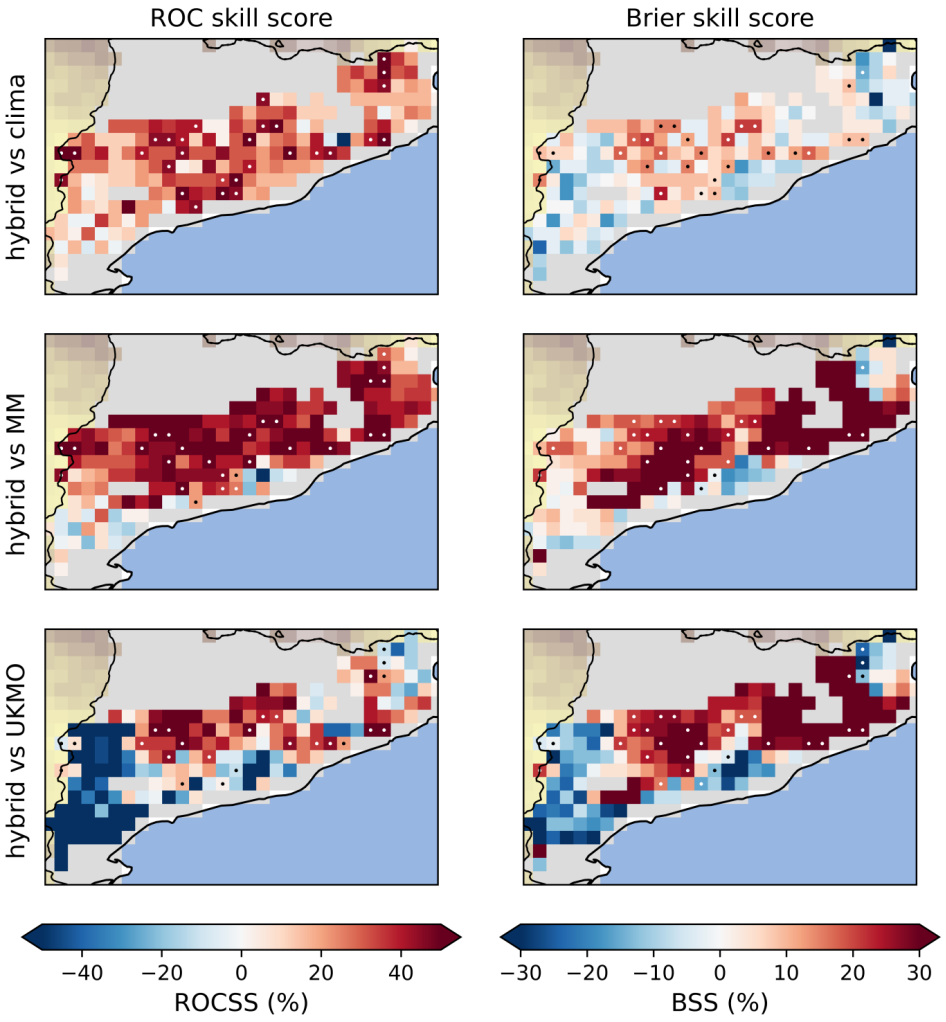


Figure 8: As in Figure 6, but for the hybrid predictions. The skill score is computed against the climatology (top), and the MM (middle) and UKMO (bottom) raw dynamical predictions.

solidate and extend the results presented here. Key tasks looking ahead are: (i) to improve the representation of atmospheric blocking anomalies in the seasonal prediction systems (particularly for *omega*-type blocking), when ensemble sub-sampling techniques [e.g., Dobrynin et al. (2018)] are an attractive way forward. (ii) To fully explore the method's hyper-parameter space, for example by considering larger spatial domains than the Euro-Atlantic domain considered here to compute pattern correlations. (iii) To evaluate the statistical downscaling method for different datasets, target periods (e.g.,

This preprint research paper has not been peer reviewed.

April), lead times, and —potentially— climate hazards (e.g. early warming events in late winter, summer heat waves). Robustness under varying scenarios might serve as a partial surrogate for statistical significance (which may or may not be within reach given current constraints on blocking predictability and hindcast length), and help make a stronger case for the operational applicability of the method.

4 Conclusions

By establishing a firm statistical link between atmospheric blocking frequency anomalies and frost events, we developed a statistical downscaling method which predicts the probability of March frost occurrence in the winemaking region of Catalonia one month in advance. The method leverages the predictability of large-scale blocking variability at the seasonal time horizon, and outputs frost risk information at a spatial resolution relevant to on-field management. Predictions made with perfect knowledge of the large-scale drivers (perfect-prognosis) outperform the climatological reference, and part of this skill translates to predictions made using actual forecasted values of the drivers (hybrid predictions). In addition, we have analysed raw dynamical predictions of frost occurrence, which represent the skill associated with unprocessed seasonal predictions. Our analysis highlighted that no individual prediction strategy is optimal across the domain and metrics considered, and that a portfolio of predictions is therefore needed to best meet the demands of the frost protection problem.

Our findings corroborate and extend earlier results: on the one hand, we have found that some aspects of seasonal blocking variability are predictable with significant skill, as reported by Athanasiadis et al. (2014) for seasonal predictions in wintertime. By adopting the recent blocking detection algorithm of Sousa et al. (2021), though, we were able for the first time to prove skill for different types of blocking events on an individual basis: compared to traditional methods [e.g. Scherrer et al. (2006)], this allows ourselves to stratify the circulation anomalies associated with frost events according to the blocking type, and —as a consequence— to make accurate use of blocking predictability. On the other hand, while previous studies tended to rely on teleconnection indices such as the NAO (Cionni et al., 2022, Ramon et al., 2021) to predict surface climate with perfect-prognosis methods, here we explicitly showed that atmospheric blocking similarly wields considerable predictive power.

To demonstrate scalability of the prototype model proposed here is a priority going forward, while further work is needed to extend the range of

This preprint research paper has not been peer reviewed.

applicability to different lead times, regional domains, and surface extremes. Bolstering the statistical robustness of the results, in addition, would increase the reliability of the method and pave the way to quicker application in operational scenarios, and is another key element for future research.

Spring frost is a natural hazard which every year causes great damage to the environment and the economy (Papagiannaki et al., 2014). This work was catered towards a specific case study considered within the ASPECT project but, we believe, our results have implications which potentially extend beyond that remit.

CRediT authorship contribution statement

Sebastiano Roncoroni: Conceptualization, Software, Formal analysis, Writing – Original Draft, Visualization. **Panos J. Athanasiadis:** Conceptualization, Writing – Review & Editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors are thankful to Verónica Torralba, Carlos Delgado-Torres, Marta Terrado, and all the ASPECT partners involved in the case study considered in this manuscript for their work. This research was supported by the ASPECT project, which has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081460. Views and opinions expressed are, however, those of the author(s) 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.

References

Abu-Mostafa, Y. S., Magdon-Ismail, M., and Lin, H.-T. (2012). *Learning From Data*. AMLbook.

This preprint research paper has not been peer reviewed.

Athanasiadis, P. J., Bellucci, A., Hermanson, L., Scaife, A. A., MacLachlan, C., Arribas, A., Materia, S., Borrelli, A., and Gualdi, S. (2014). The Representation of Atmospheric Blocking and the Associated Low-Frequency Variability in Two Seasonal Prediction Systems. *J. Clim.*, 27(24):9082 – 9100.

Athanasiadis, P. J., Yeager, S., Kwon, Y.-O., Bellucci, A., Smith, D. W., and Tibaldi, S. (2020). Decadal predictability of North Atlantic blocking and the NAO. *npj Clim. Atmos. Sci.*, 3(1):2397–3722.

Augspurger, C. K. (2009). Spring 2007 warmth and frost: phenology, damage and refoliation in a temperate deciduous forest. *Funct. Ecol.*, 23:1031–1039.

Bieli, M., Pfahl, S., and Wernli, H. (2015). A Lagrangian investigation of hot and cold temperature extremes in Europe. *Q.J.R. Meteorol. Soc.*, 141(686):98–108.

Boschat, G., Simmons, I., Purich, A., Cowan, T., and Bernardes Pezza, A. (2016). On the use of composite analyses to form physical hypotheses: An example from heat wave-SST associations. *Sci. Rep.*, 6:29599.

Bretherton, C. S., Widmann, M., Dymnikov, V. P., Wallace, J. M., and Bladé, I. (1999). The effective number of spatial degrees of freedom of a time-varying field. *J. Clim.*, 12(7):1990 – 2009.

Brunner, L., Hegerl, G. C., and Steiner, A. K. (2017). Connecting atmospheric blocking to european temperature extremes in spring. *J. Clim.*, 30(2):585–594.

Cionni, I., Lledó, L., Torralba, V., and Dell'Aquila, A. (2022). Seasonal predictions of energy-relevant climate variables through Euro-Atlantic Teleconnections. *Clim. Serv.*, 26:100294.

Copernicus Climate Change Service, Climate Data Store (2018a). Seasonal forecast daily and subdaily data on single levels. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Last accessed on 05-08-2024).

Copernicus Climate Change Service, Climate Data Store (2018b). Seasonal forecast subdaily data on pressure levels. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Last accessed on 30-08-2024).

Copernicus Climate Change Service, Climate Data Store (2020). E-OBS daily gridded meteorological data for Europe from 1950 to present derived from

This preprint research paper has not been peer reviewed.

in-situ observations. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Last accessed on 05-08-2024).

Copernicus Climate Change Service, Climate Data Store (2023). ERA5 hourly data on pressure levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Last accessed on 05-08-2024).

Cornes, R., van der Schrier, G., van den Besselaar, E., and Jones, P. (2018). An Ensemble Version of the E-OBS Temperature and Precipitation Datasets. *J. Geophys. Res. Atmos.*, 123.

dask-image (2024). Image processing with Dask Arrays. https://image.dask.org/en/latest/.

Dobrynin, M., Domeisen, D. I. V., Müller, W. A., Bell, L., Brune, S., Bunzel, F., Düsterhus, A., Fröhlich, K., Pohlmann, H., and Baehr, J. (2018). Improved Teleconnection-Based Dynamical Seasonal Predictions of Boreal Winter. *Geophys. Res. Lett.*, 45(8):3605–3614.

DWD (2024). Description of GCFS2.1-v20200320 C3S contribution. https://confluence.ecmwf.int/display/CKB/Description+of+GCFS2.1-v20200320+C3S+contribution.

Gualdi, S., Borrelli, A., Cantelli, A., Davoli, G., Chaves Montero, M. d. M., Masina, S., Navarra, A., Sanna, A., and Tibaldi, S. (2020). The new CMCC Operational Seasonal Prediction System. Technical report, CMCC.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N. (2023). ERA5 hourly data on pressure levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Last accessed on 05-08-2024).

Hurrell, J. W., Kushnir, Y., Ottersen, G., and Visbeck, M. (2003). An Overview of the North Atlantic Oscillation. In Hurrell, J. W. and Kushnir, Y. and Ottersen, G. and Visbeck, M., editor, *The North Atlantic Oscillation: Climatic Significance and Environmental Impact*, pages 1–35. American Geophysical Union.

Institut Català de la Vinya i el Vi (2023). Dades de comercialització del vi. https://incavi.gencat.cat/ca/coneix-vi-catala/vi-catala-xifres/dades-comercialitzacio/.

This preprint research paper has not been peer reviewed.

International Organization of Vine and Wine (2024). World Wine Production Outlook. OIV First Estimates.

Kautz, L.-A., Martius, O., Pfahl, S., Pinto, J. G., Ramos, A. M., Sousa, P. M., and Woollings, T. (2022). Atmospheric blocking and weather extremes over the Euro-Atlantic sector – a review. *Weather and Clim. Dyn.*, 3(1):305–336.

Lamichhane, J. (2021). Rising risks of late-spring frosts in a changing climate. *Nat. Clim. Chang.*, 11:554–555.

Liu, Q., Piao, S., Janssens, I., Fu, Y., Peng, S., Lian, X., Ciais, P., Myneni, R., J., P., and Wang, T. (2018). Extension of the growing season increases vegetation exposure to frost. *Nat. Commun.*, 9(426).

Lledó, L., Cionni, I., Torralba, V., Bretonnière, P.-A., and Samsó, M. (2020). Seasonal prediction of Euro-Atlantic teleconnections from multiple systems. *Environ. Res. Lett.*, 15(7):074009.

Madonna, E., Li, C., Grams, C., and Woollings, T. (2017). The link between eddy-driven jet variability and weather regimes in the North Atlantic-European sector. *Q.J.R. Meteorol. Soc.*, 143:2960–2972.

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeá, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I. (2010). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Rev. Geophys.*, 48(3).

Meehl, G., Tebaldi, C., and Nychka, D. (2004). Changes in frost days in simulations of twentyfirst century climate. *Clim. Dyn.*, 23:495–511.

MF (2024). Description of System8-v20210101 C3S contribution. https://confluence.ecmwf.int/display/CKB/Description+of+System8-v20210101+C3S+contribution.

Nakamura, N. and Huang, C. S. Y. (2018). Atmospheric blocking as a traffic jam in the jet stream. *Sci.*, 361:42–47.

Papagiannaki, K., Lagouvardos, K., Kotroni, V., and Papagiannakis, G. (2014). Agricultural losses related to frost events: use of the 850 hpa level temperature as an explanatory variable of the damage cost. *Nat. Hazards and Earth Syst. Sci.*, 14(9):2375–2386.

This preprint research paper has not been peer reviewed.

Ramon, J., Lledó, L., Bretonnière, P.-A., Samsó, M., and Doblas-Reyes, F. J. (2021). A perfect prognosis downscaling methodology for seasonal prediction of local-scale wind speeds. *Environ. Res. Lett.*, 16:054010.

Richardson, D. S. (2003). Economic value and skill. In Jolliffe, I. T. and Stephenson, D. B., editor, *Forecast Verification*, pages 165–187. Wiley.

Scaife, A. and Smith, D. A. (2018). A signal-to-noise paradox in climate science. *npj Clim. Atmos. Sci.*, 1(28).

Scherrer, S. C., Croci-Maspoli, M., Schwierz, C., and Appenzeller, C. (2006). Two-dimensional indices of atmospheric blocking and their statistical relationship with winter climate patterns in the Euro-Atlantic region. *Int. J. Climatol.*, 26(2):233–249.

Siegert, S., Stephenson, D. B., Sansom, P. G., Scaife, A. A., Eade, R., and Arribas, A. (2016). A Bayesian Framework for Verification and Recalibration of Ensemble Forecasts: How Uncertain is NAO Predictability? *J. Clim.*, 29(3):995–1012.

Snyder, R. L. and de Melo-Abreu, J. P. (2005). *Frost Protection: fundamentals, practice and economics*, volume 1. FAO.

Sousa, P. M., Barriopedro, D., García-Herrera, R., Woollings, T., and Trigo, R. M. (2021). A new combined detection algorithm for blocking and subtropical ridges. *J. Clim.*, 34(18):7735–7758.

Sousa, P. M., Trigo, R. M., Barriopedro, D., Soares, P. M. M., and Santos, J. A. (2018). European temperature responses to blocking and ridge regional patterns. *Clim. Dyn.*, 50:457–477.

Stockdale, T. (2021). SEAS5 user guide. Technical report, ECMWF.

Terray, P., Delécluse, P., Labattu, S., and Terray, L. (2003). Sea surface temperature associations with the late Indian summer monsoon. *Clim. Dyn.*, 21(7):593–618.

Thornton, H. E., Smith, D. M., Scaife, A. A., and Dunstone, N. J. (2023). Seasonal Predictability of the East Atlantic Pattern in Late Autumn and Early Winter. *Geophys. Res. Lett.*, 50(1):e2022GL100712.

Torralba, V., Materia, S., Cavicchia, L., Álvarez Castro, M. C., Prodhomme, C., McAdam, R., Scoccimarro, E., and Gualdi, S. (2024). Nighttime heat waves in the Euro-Mediterranean region: definition, characterisation, and seasonal prediction. *Environ. Res. Lett.*, 19(3):034001.

This preprint research paper has not been peer reviewed.

Trigo, R., Trigo, I., DaCamara, C., and Osborn, T. J. (2004). Climate impact of the European winter blocking episodes from the NCEP/NCAR Reanalyses. *Clim. Dyn.*, 23:17–28.

UKMO (2024). Description of System8-v20210101 C3S contribution. https://confluence.ecmwf.int/display/CKB/Description+of+System8-v20210101+C3S+contribution.

Wilks, D. (2006). *Statistical methods in the atmospheric sciences*. Elsevier, 2nd edition.

Woollings, T., Hoskins, B., Blackburn, M., and Berrisford, P. (2008). A New Rossby Wave–Breaking Interpretation of the North Atlantic Oscillation. *J. Atmos. Sci.*, 65(2):609 – 626.

Wulff, C. O., Vitart, F., and Domeisen, D. I. V. (2022). Influence of trends on subseasonal temperature prediction skill. *Q.J.R. Meteorol. Soc.*, 148(744):1280–1299.

This preprint research paper has not been peer reviewed.