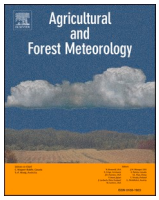
Agricultural and Forest Meteorology 331 (2023) 109342

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet



Performance of seasonal forecasts for the flowering and veraison of two major Portuguese grapevine varieties

Chenyao Yang [a,*], Andrej Ceglar [b], Christoph Menz [c], Joana Martins [d], Helder Fraga [a], João A. Santos [a]

[a] Centre for the Research and Technology of Agro-Environmental and Biological Sciences (CITAB)/Inov4Agro (Institute for Innovation, Capacity Building and Sustainability of Agri-Food Production), University of Trás-os-Montes and Alto Douro (UTAD), 5000-801, Vila Real, Portugal
[b] Climate Change Centre of the European Central Bank, Sonnemannstrasse 20, 60314 Frankfurt am Main, Germany
[c] Potsdam Institute for Climate Impact Research e. V. (PIK), Telegrafenberg A 31, 14473 Potsdam, Germany
[d] Centre for Environmental and Marine Studies (CESAM), University of Aveiro (UA), 3810-193 Aveiro, Portugal

ARTICLE INFO

Keywords:
Mediterranean climate
Seasonal climate forecast
Phenology forecast
Phenology modelling
Forecast performance
Climate risk

ABSTRACT

Seasonal phenology forecasts are becoming increasingly demanded by winegrowers and viticulturists. Forecast performance needs to be investigated over space and time before practical applications. We assess seasonal forecast performance (skill, probability and accuracy) in predicting flowering and veraison stages of two representative varieties in Portugal over 1993–2017. The state-of-the-art forecast system ECMWF-SEAS5 provides 7-month seasonal forecasts and is coupled with a locally adapted phenology model. Overall, findings illustrate the dependence of forecast performance on initialization timings, regions and predicting subjects (stages and varieties). Forecast performance improves by delaying the initialization timing and only forecasts initialized on April 1st show better skills than climatology on predicting phenology terciles (early/normal/late). The considerable bias of daily values of seasonal climate predictions can represent the main barrier to accurate forecasts. Better prediction performance is consistently found in Central-Southern regions compared to Northern regions, attributing to an earlier phenology occurrence with a shorter forecast length. Comparable predictive skills between flowering and veraison for both varieties imply better predictability in summer. Consequently, promising seasonal phenology predictions are foreseen in Central-Southern wine regions using forecasts initialized on April 1st with approximately 1–2/3–4 months lead time for flowering/veraison: potential prediction errors are ~2 weeks, along with an overall moderate forecast skill on categorical events. However, considerable inter-annual variability of forecast performance over the same classified phenology years reflects the substantial influence of climate variability. This may represent the main challenge for reliable forecasts in Mediterranean regions. Recommendations are suggested for methodological innovations and practical applications towards reliable regional phenology forecasts.

1. Introduction

Climate variability and extreme weather events have profound impacts on agricultural systems (Ceglar et al., 2016; Challinor et al., 2014; Lobell and Field, 2007). In recent decades, the agricultural sector has become increasingly exposed to unfavourable climate conditions and weather extremes (Zhu and Troy, 2018). For instance, climate-related crop failures are widely reported across Europe in 2003, 2010 and 2018 (Bento et al., 2022). These climate-related risks can be exacerbated under global warming with the increased frequency and magnitude of extreme weather events such as heatwaves (Fraga et al., 2020). To enhance the resilience of the agricultural sector, the development of a reliable seasonal forecast product that can accurately capture the signal of seasonal climate anomalies, may aid in effectively and dynamically adapting to climate variability and climate change (Bento et al., 2022; Ceglar et al., 2017). Early season forecasts have been shown to provide useful information on anticipated climate events of interest to farmers (Ceglar and Toreti, 2021; Chen and Tao, 2022; Jha et al., 2019). They can proactively adjust their agronomic management in advance to minimize potentially negative impacts of adverse weather events while

* Corresponding author.
E-mail address: cyang@utad.pt (C. Yang).

https://doi.org/10.1016/j.agrformet.2023.109342
Received 10 August 2022; Received in revised form 16 January 2023; Accepted 22 January 2023
Available online 28 January 2023

maximising the benefits when favourable conditions are anticipated (Bento et al., 2022; Brown et al., 2018; Tao et al., 2018; van der Velde and Nisini, 2019).

Grapevine (*Vitis vinifera* L.) is a fruit crop of economic importance worldwide. The viticulture and winemaking sectors are greatly influenced by local climatic conditions and annual weather patterns (Santos et al., 2020b). Providing forecasts of phenology development during a given growing season is a promising avenue for enhancing the climate resilience of the sector (Taylor and White, 2020). Plant phenology, a periodically recurring pattern of development stages during the growing season (Piao et al., 2019), is generally of central importance for other physiological processes. For instance, in process-based crop modelling, the phenology can affect vegetation dynamics, soil water balance and plant water use, as well as biomass and yield formation (Brisson et al., 2003; Holzworth et al., 2014). Phenology development is mostly temperature-driven. A generally advanced phenology stage is projected in response to rising temperatures for different varieties and regions, which in turn may affect berry composition, wine quality and typicity (Clemente et al., 2022; Fraga et al., 2016; Ramos et al., 2021; Yang et al., 2022a).

The flowering/bloom timing (BBCH65) and veraison/ripening onset (BBCH81) are pivotal grapevine stages (Lorenz et al., 1995). Available seasonal forecasts at these stages can help monitor vine growth and better inform decision-making for winegrowers and viticulturists in vineyard management. This approach promotes the reduction of risks resulting from climate change and variability. Accurate predictions of the flowering stage can warrant timely implementation of field interventions, such as supplemental irrigation in dry conditions to avoid inflorescence necrosis and ensure adequate fruit set. For the veraison stage, early-season predictions can help anticipate harvest (plan material and labour costs) and inform measures to avoid undesirable impacts on wine quality. For instance, climate warming can expose local varieties to excessively high temperatures during grape ripening, resulting in high sugar content and low acidity, anthocyanin and aromas (Mosedale et al., 2016; van Leeuwen et al., 2019). Timely implementation of measures, e.g. canopy pruning to reduce the leaf area to fruit weight ratio to delay the veraison stage, can avoid high ripening temperatures, thus being beneficial for a balanced grape composition (van Leeuwen et al., 2019; van Leeuwen and Destrac-Irvine, 2017). Additionally, seasonal forecasts can become a viable tool for financial entities, such as banks, asset management and insurance companies, to hedge potential economic losses and secure farmers' income.

One of the key components of performing seasonal forecasts on phenology is to deploy a mechanistic phenology model that reliably simulates the observed phenology of a specific variety. Phenology modelling often relies on the concept of "degree-days", assuming the phenology occurrence when a variety-dependant accumulation of heat units is achieved (Costa et al., 2019; Parker et al., 2013; Reis et al., 2020; Rodrigues et al., 2022; Santos et al., 2019). These models generally share the characteristics of being simple in structure and providing good efficiency and accuracy for phenology predictions. For instance, the Grapevine Flowering Veraison model, following the same simple structure as the Growing Degree-Days (GDD) model, can well (e.g. RMSE$<$7 days) simulate flowering and veraison stages of ~80 varieties at 123 different sites spanning over 48 years (Parker et al., 2013).

Seasonal forecasts are created by numerical climate models using a complex set of hydrodynamic equations to describe the evolution of the atmosphere, ocean and land surface, for a few weeks or months from a given initial state of the Earth system (Johnson et al., 2019). Due to uncertainties in the initial state, ensemble forecasts are performed, with individual ensemble members being run with slightly different perturbations of the initial conditions (Du et al., 2012; Johnson et al., 2019; Stockdale, 2021; Wang et al., 2019). These ensemble members reflect a distribution of forecast outcomes, corresponding to nonlinear atmospheric evolution to small perturbations in initial conditions. The forecast system has limited ability to accurately predict daily weather variations at a local scale beyond approximately two weeks (Stockdale, 2021). Since the phenology modelling depends on the accurate input of daily weather, accurate predictions of phenology stages driven by daily output of seasonal climate predictions are therefore constrained. Nonetheless, there are several predictable components (e.g. sea surface temperature) that do not constrain the atmosphere enough to allow reasonable predictions of seasonal climate (Stockdale, 2021). Modifications of "usual" atmospheric conditions by these components are the essence of seasonal forecasts. Note that the seasonal forecast is different from the weather forecast, where the latter has a particular emphasis on the accuracy of daily weather conditions.

The performance of seasonal forecasts is regionally dependant. Higher forecast skills are observed in the tropical pacific due to the relatively high predictability of ENSO (Johnson et al., 2019). In contrast, limited predictabilities are widely reported in Southern Europe, such as Portugal (Ceglar et al., 2018; Ceglar and Toreti, 2021; Crespi et al., 2021; van der Velde and Nisini, 2019). For instance, van der Velde and Nisini (2019) assess the performance of the MARS Crop Yield Forecasting System (MCYFS) in predicting yields of many important crops in Europe, highlighting a relatively low prediction accuracy in Portugal. Therefore, more research efforts shall be devoted to regions where the current forecast performance is limited.

In Portugal, viticulture is a key social-economic sector, with a vineyard area of roughly 194,000 ha and annual wine production of around 6.4 million hectolitres in 2020 (OIV, 2021). Globally, Portugal is currently the 11th wine producer and the 10th wine exporter (OIV, 2021). Implementation of reliable seasonal forecasts is expected to bring considerable benefits to the sector facing uncertainties from increasing climate variability. Yet, few studies have attempted to assess the performance of seasonal forecasts in providing early-season predictions on grapevine phenology across Portuguese wine regions. Therefore, we have carried out a study to evaluate the seasonal forecast performance for predicting flowering and veraison stages of two representative varieties of these wine regions, namely cv. *Touriga Franca* (TF) and cv. *Touriga Nacional* (TN) (Fraga et al., 2017; Reis et al., 2020; Yang et al., 2021). Specifically, we aim to (i) evaluate the seasonal forecast skill and probability in predicting the tercile-based phenology category: early/normal/late event; (ii) and quantify the potential prediction bias/error in phenology timings (DOY). The assessment is carried out for the re-forecast period (1993–2017), which can be compared to historical records that can provide a measure of performance expected in real-time forecasts. Particular emphasis is given to the spatial-temporal characteristics of the forecast performance.

2. Data and methods

2.1. Studied wine regions

The present study covers 47 main wine regions in Portugal (Table S1), designated as the Protected Designation of Origin (PDO) which represents the highest classification level of viticulture and winemaking regions (Candiago et al., 2022). Climate conditions in these PDOs are characterized by Mediterranean-type climates, with warm dry summers and mild wet autumns and winters (Fraga et al., 2017; Yang et al., 2022b). TF and TN are widely cultivated throughout these PDOs. However, there are also a large number of autochthonous and international varieties grown in Portuguese vineyards, resulting in either mono-varietal or blended wines, though the latter can be more common (Fraga et al., 2017). The multi-varietal phenology stages can span a long period over the season, which was hard to be captured by the phenology modelling technique. Therefore, vineyards with blended varieties are currently not considered.

2.2. Meteorological datasets

2.2.1. Observational datasets

The meteorological variables considered were daily surface (2-m) minimum (T_{min}, °C), mean (T_{mean}, °C) and maximum temperatures (T_{max}, °C). These variables were retrieved from E-OBS (v24.0e), a European-wide high-resolution (0.1°) daily gridded weather dataset, which has been used over the last decade in various studies on climate monitoring and climate model evaluations (Cornes et al., 2018; Haylock et al., 2008; Jacob et al., 2014). In Portugal, the dataset was previously shown to be useful in climate change impact assessments for several important cropping systems (Yang et al., 2020, 2019, 2018, 2017). Observed temperature data over 1993–2017 were extracted for study regions, and mainly used to correct raw predictions of seasonal climate obtained from ECMWF-SEAS5 (2.2.3), and implement grid-based phenology simulations to provide spatial reference data (2.3.3). E-OBS played an important role in the real-time forecast as it could provide a monthly update, which can be synchronized with the monthly initialized forecast dataset.

2.2.2. Seasonal forecast datasets

The seasonal forecast data on T_{min} and T_{max} were collected from the fifth generation of the seasonal forecast system (SEAS5) of the European Centre for Medium-Range Weather Forecast (ECMWF) (Johnson et al., 2019), available via the Copernicus Climate Change Service (C3S, https://cds.climate.copernicus.eu/). The SEAS5 became operational (real-time forecasts) in November 2017, representing a state-of-the-art forecast system with improved atmosphere and ocean model configurations and horizontal spatial resolutions, i.e. ~36 km in atmospheric components (Johnson et al., 2019), but being up-scaled to 1° in C3S (Ceglar and Toreti, 2021). A set of re-forecasts over 1993–2017 (hindcasts) were chosen, consisting of 25 ensemble members that were initialized on the 1st day of each month and integrated for 7 months every year. Within this study, three forecast initialization timings were chosen: February 1st, March 1st and April 1st, considering the typical flowering and veraison periods of May–June and July–August respectively in Portugal (Fraga et al., 2017, 2016). These initialization dates were well ahead of the earliest possible dates of locally observed flowering and veraison stages (with forecast values), while at the same time guaranteeing that the latest possible dates of both stages were still within the lead time of the earliest issued forecast (February 1st). To simulate the phenological cycle, the observational (E-OBS) and forecast data were combined: before the forecast initialization date of each year, the observational gridded data was utilized and using seasonal forecasts from the initialization timing onwards, following the approach described by Ceglar and Toreti (2021). This aimed to emulate the real predictions, assuming observed data availability before each initialization. A similar procedure has also been reported in other studies (Ceglar and Toreti, 2021; Chen and Tao, 2022; Jha et al., 2019; Semenov, 2007).

2.2.3. Evaluations of bias-adjusted forecast datasets

The quantile mapping method (Gudmundsson et al., 2012; Panofsky et al., 1968; Piani et al., 2010), was applied to reduce bias in the forecast temperatures using E-OBS data as the reference dataset. It is a reliable and robust method to minimize the bias in the outputs of different forecast models (Ceglar et al., 2018; Ceglar and Toreti, 2021). As observed (0.1°) and forecast (1°) datasets were at different resolutions, a flow chart was presented to illustrate how observed datasets are upscaled to the forecast mesh for bias correction (Fig. S1). The bias-adjusted forecast dataset was then supplied to the phenology model to perform the phenology forecast. To evaluate how well it can reproduce observed seasonal climate conditions, GDD with a base temperature of 0°C (Parker et al., 2013) was computed between referenced and forecasted series (ensemble median) at each initialization timing over 1993–2017. Their correspondence was evaluated using the coefficient of determination (R^2), while the assessment of whether there was a significant difference ($p<=0.05$) in their distributions was performed using the two-sample Kolmogorov-Smirnov test (Hodges, 1958). Mean Bias Error (MBE) and Root Mean Squared Error (RMSE) were calculated to assess the difference in seasonal total GDD between the two series for 1993–2017. In particular, MBE can give indications of whether the forecast over- (positive) or under-estimate (negative) the observed GDD.

2.3. Phenology simulations

2.3.1. Choice of phenology model for local conditions

Numerous models exist for grapevine phenology modelling, within this study the sigmoid model ($f_{sigmoid}$) was chosen. Previous analysis indicated $f_{sigmoid}$ showed a better trade-off between model complexity and prediction accuracy in simulating the flowering and veraison stages of TF and TN, as compared to other commonly used phenology models (GDD, Richardson and Wang) (Costa et al., 2019). Besides, $f_{sigmoid}$ had been successfully evaluated for simulating the flowering and veraison stages of another 34 grapevine varieties at 19 vineyard plots over four major winemaking regions in Portugal (Reis et al., 2020). The $f_{sigmoid}$ shown in Eq. (1) only contained two parameters as proposed by Hänninen (1990):

$$f_{sigmoid} = 1 \,/\, \left(1 + e^{(d*(T_{mean}-b))}\right) \quad (1)$$

where T_{mean} represented the daily mean temperature, d denoted the sharpness of the curve; b was the mid-response temperature. For forecast runs, T_{mean} was calculated as the arithmetic average of bias-adjusted T_{min} and T_{max}, while observed T_{mean} was directly used as inputs for reference phenology simulations; $f_{sigmoid}$ simulated the occurrence of a target phenology stage at the date when the state of forcing temperatures reached a threshold value of F^* (variety-dependant thermal forcing) from a user-specified starting date (t_0). To model the flowering stage, t_0 was set to be on January 1st of each year, being consistent with previous studies (Costa et al., 2019; Reis et al., 2020). To model the veraison stage, its simulation directly started from the simulated flowering stage to preserve the continuity of phenology development. Such a sequential scheme highlighted that a good veraison simulation was also dependant on the accurate simulation of the flowering stage (Yang et al., 2022a).

2.3.2. Calibration of the phenology model using gridded weather data

The sigmoid model, despite being extensively calibrated and evaluated before (Costa et al., 2019; Reis et al., 2020), was re-calibrated using the gridded E-OBS climate dataset. This mainly aimed to adapt the phenology model simulations to the spatial scale of the forecast runs. The observed phenology (flowering and veraison) data was obtained for TF in 1995–2014, and TN in 1990–2000 and 2006–2014 from a vineyard located within the Lisboa wine region at 39.041°N, 9.181°W (Reis et al., 2020). The parameter calibration used the simulated annealing optimization algorithm of Metropolis (Metropolis et al., 1953), which was implemented in the Phenology Modelling Platform (v5.5) (Chuine et al., 1998) under the leave-one-out cross-validation scheme. Final calibrated parameters (Fig. S2) are obtained based on the parameter vector with minimum RMSE during cross-validation. Overall satisfactory performance was obtained, with predicted stages within a 5-day difference (MAE and RMSE) and negligible mean bias with respect to observed stages for both varieties (Fig. 1). The inter-annual variability of observed flowering and veraison stages were well reproduced (Fig. S3), with predictions explaining 73–75% and 63–79% of observed variability for TF and TN, respectively (Fig. 1).

2.3.3. Reference phenology simulations

To evaluate the forecast performance, long-term high-quality observed phenology data is required throughout each grid point of these wine regions (or at least for each region). However, this is currently not possible due to limited data availability. To mitigate the gap, reference

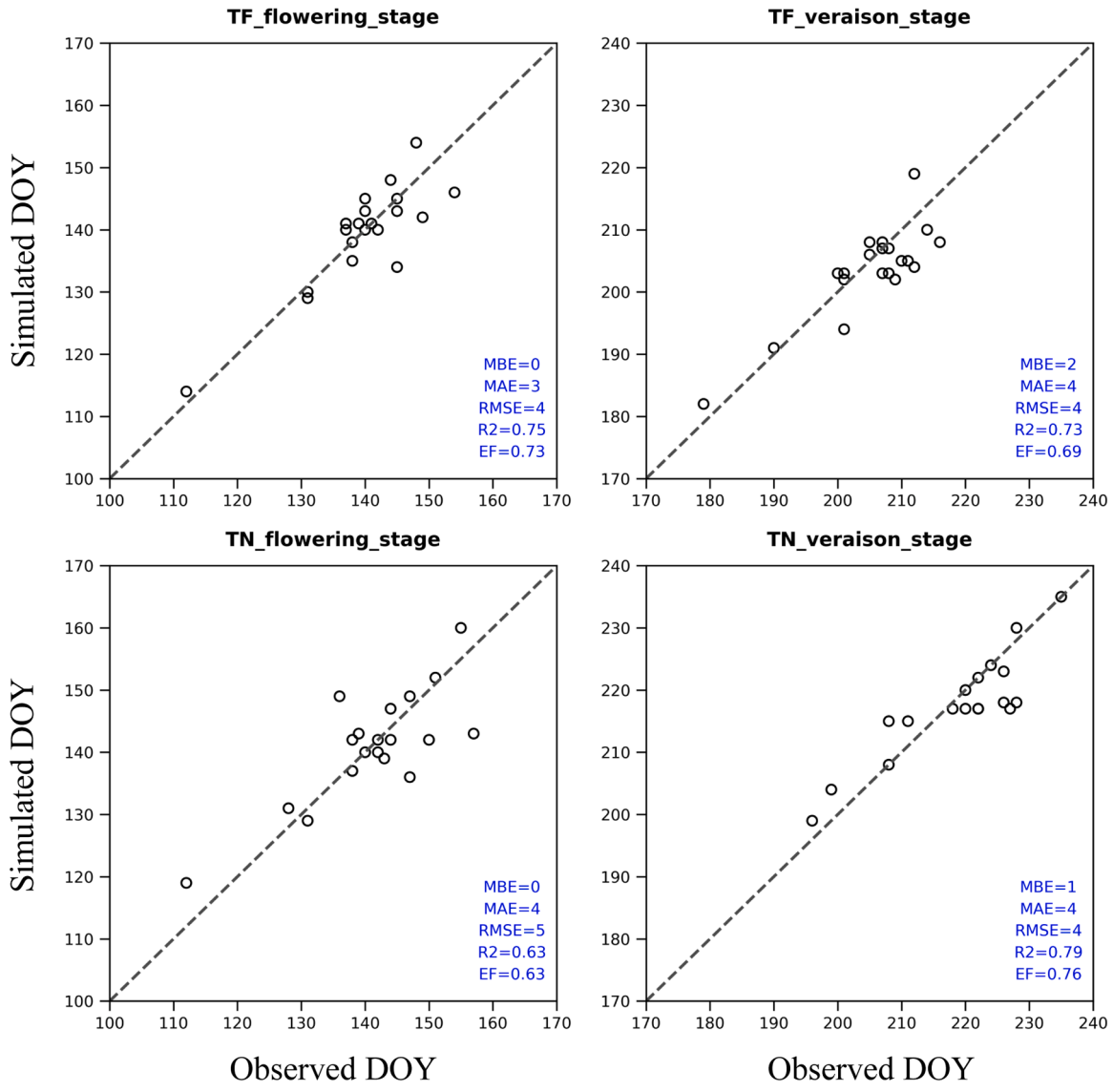


Fig. 1. Comparison of the observed and simulated Day Of Year (DOY) between the flowering and veraison stages of Touriga Franca (TF) and Touriga Nacional (TN). The evaluation metrics are Mean Bias Error (MBE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Coefficient of Determination (R^2) and Efficiency of simulations (EF).

phenology has been simulated with the gridded climate dataset (E-OBS) at its original resolution (0.1°) using the re-calibrated sigmoid model. The reliability and applicability of the E-OBS dataset to reproduce locally observed vine phenology has been successfully demonstrated (with an acceptable level of prediction errors) in different Portuguese wine regions (Rodrigues et al., 2022; Yang et al., 2022b, 2021). Using simulations with observed weather as reference data is a common practice for seasonal forecast studies (Ceglar et al., 2018; Ceglar and Toreti, 2021; Ferrise et al., 2015; Jha et al., 2019; Semenov, 2007). The resulting reference phenology simulations indicated a slightly earlier flowering occurrence for TN than TF, but a markedly late veraison/ripening onset for TN than TF (Fig. S4–S5). These are highly consistent with the actual phenology characteristics of the two varieties, as reported by Costa et al. (2019) and Reis et al. (2020).

2.4. Evaluation for seasonal forecast on categorical phenology events

The simulated phenology stages (DOY) were classified into three categories (tercile classes) relevant for agricultural decision-making: early (<33 percentile), normal (between 33 and 66 percentiles) and late (>66 percentile) occurrence (Ceglar and Toreti, 2021). The classification was performed based on the reference phenology DOY series at each grid point in 1993–2017. The resulting reference categorization thresholds (DOY) were also applied to categorize the forecast DOY series, which was undertaken for each forecast ensemble member at each initialization date.

2.4.1. Fair ranked probability skill score (FRPSS)

To evaluate the overall forecast skill/predictability on the three categorical events (early/normal/late), FRPSS was computed. FRPSS measures how well the probability forecast predicts the observed category (Ferro, 2014). FRPSS is derived from the Ranked Probability Score (RPS). RPS computed the mean squared distance between the cumulative probabilities of n forecast-reference pairs along the defined time series for k equally probable forecast categories while correcting for ensemble size (Ferro, 2014) (Eq. (2)):

$$\mathrm{RPS} = \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{K}\left[\left(\left(\frac{fc_{i,k}}{m} - ob_{i,k}\right)^2 - \frac{fc_{i,k}\left(m - fc_{i,k}\right)}{m^2(m-1)}\right)\right] \quad (2)$$

where $fc_{i,k}$ denoted the number of ensemble members that forecasted the k^{th} category at i^{th} forecast; $ob_{i,k}$ denoted the observation in k^{th} category (1 for occurrence, 0 for no occurrence) at i^{th} forecast; m denoted the total ensemble members; $\frac{fc_{i,k}}{m}$ thus indicating the ensemble forecast probability of occurrence of a given category; and $\frac{fc_{i,k}\ (m-fc_{i,k})}{m^2(m-1)}$ served to correct the ensemble size. RPS was evaluated between climatology (RPS_{clim}) and forecast (RPS_{fc}), leading to FRPSS $= 1 - \frac{RPS_{fc}}{RPS_{clim}}$. FRPSS ranged from $-\infty$ to 1: 1 for a perfect forecast, 0 for forecasts that do not perform any better than the reference forecast, and positive (negative) when the forecast is better (worse) than this reference (Ceglar and Toreti, 2021). The reference forecast reflected information only on the long-term mean agro-climate conditions.

2.4.2. Gilbert skill score (GSS)

To further assess the forecast skills on each tercile class separately, the GSS was computed involving several terms: hit (event forecast to occur and does occur), miss (event forecast not to occur but does occur), false alarm (event forecast to occur but does not occur) and correct negative (event forecast not to occur and does not occur) (Daniel S. Wilks, 2020; Schaefer, 1990). It quantifies the forecast quality by measuring how well the predicted events correspond to the observations, while adjusting for correct forecasts due to chance (Daniel S. Wilks, 2020; Schaefer, 1990). GSS ranges from −1/3 (random forecast with no skills) to 1 (perfect forecast). The score estimates the ability of a deterministic prediction to correctly forecast a dichotomous event (Ceglar et al., 2018). The probability threshold was defined at 60% for a given forecast event to occur, i.e. more than 60% of ensemble members predicted the category.

2.4.3. Ensemble forecast probability of the observed events (ens_prob)

Both skill score and forecast probability were important for evaluating the performance of the ensemble forecast system. The forecast probability of occurrence of the reference category was estimated based on the fraction of ensemble members predicting the event, denoted as *ens_prob*. It represented a simplified form to provide an indication of how certain forecasts agreed with the observed category based on the spread of the ensemble member predictions (the higher the probability, the more certainty the predicted outcome) (Ferro, 2014). It was similar to $\frac{fc_{i,k}}{m}$ in FRPSS, but computed exclusively for the reference category in each year at a given grid point. Since the reference category varied spatially and temporally, we further divided the period of 1993–2017 into early/normal/late phenology years when a reference tercile class appeared to be a spatial majority in a particular year. *ens_prob* was computed for the classified early/normal/late years accordingly.

2.5. Evaluation of the potential for seasonal forecast on phenology stages (DOY)

The magnitude of prediction errors between the reference and forecast DOY series of each ensemble member was quantified using MAE and RMSE at every grid point. The ensemble mean of MAE and RMSE was respectively adopted, which were then aggregated and mapped for each wine region. Compared to MAE, RMSE put additional weight on the large difference (Taylor, 2001). These were two commonly employed metrics that could give straightforward estimations of the magnitude of forecast errors, allowing direct comparisons to those of other phenology modelling studies (Costa et al., 2019; Parker et al., 2013; Reis et al., 2020; Rodrigues et al., 2022).

3. Results

3.1. Evaluation of bias-adjusted seasonal forecast performance

The forecast series can very well reproduce (R^2>0.95) the reference series over 1993–2017 for seasonal daily GDD from any forecast initialization dates until the empirically estimated latest possible flowering (July 31st) and veraison stage (September 30th), irrespective of cultivars (Fig. S6–S8). This is also consistent for individual years as little variability is found from year to year (Fig. S6–S8). Besides, a statistical difference (p<0.05) in the distribution between the full reference and forecast series is found in 16/47 and 22/47 regions for the empirical flowering and veraison stages respectively for February 1st forecasts (Fig. S6), which slightly decreases to 14/47 and 19/47 for April 1st forecasts (Fig. S8). Nonetheless, there is a considerable daily bias (day-to-day difference) between the reference and forecast series (the biggest difference is found in two northern regions 28 and 29), as evidenced by the considerable difference in their 90% variability range over 1993–2017 (a larger magnitude of variability is generally found for the reference series) (Fig. S6–S8). This consequently leads to their difference in the seasonal total GDD attained (Fig. S9), hence affecting the phenology forecasts. For forecasts issued on February 1st, the mean observed seasonal total GDD to reach the empirical flowering (July 31st) and veraison stages (September 30th) is respectively 2500 to 3300 and 3700 to >=4500°C days^{-1} for studied wine regions (Fig. S9). For both stages, overestimations of seasonal total GDD (indicated by MBE) tend to primarily occur in the central and southern (C-S) regions (with lower latitude than region 11) whereas most northern (N) regions (with equal or higher latitude than the region 11) show underestimation (Fig. 2a). Relatively higher forecast errors (RMSE>200°C days^{-1}, up to 600°C days^{-1}) are found in the N regions and lower errors (RMSE<=200°C days^{-1}) in the C-S regions (Fig. 2b). Since more observations are integrated in phenology estimations when the target stage is approaching, March 1st and April 1st forecasts show decreased RMSE than February 1st forecasts for most of the regions (Fig. 2b). However, the identified spatial pattern of underestimations/overestimations remains largely unchanged from February 1st to April 1st forecasts (Fig. 2a).

3.2. Performance of the seasonal forecast on categorical phenology events

3.2.1. Overall forecast performance

As shown in Fig. 3, the overall forecast skill (indicated by FRPSS) in predicting categorical flowering and veraison events of two varieties improves when using seasonal forecasts initialized later in the season from February 1st to April 1st. Regardless of stages and varieties, February 1st forecasts are not more skilful than climatology (FRPSS<0) and March 1st forecasts show low skills in most regions with FRPSS<0.2 (Fig. 3). Moderately skilful predictions are detected using April 1st forecasts in C-S regions (0.2<FRPSS<0.6), while N regions generally show marginally useful forecasts (0<FRPSS<0.4) (Fig. 3). In other words, April 1st forecasts in C-S regions tend to have a higher probability in correctly predicting early/normal/late category than those of N regions. Furthermore, comparable forecast skills are found between the flowering and veraison stages of both varieties (Fig. 3). This is more valuable for the latter given that the veraison stage is usually between 2 and 3 months later than the flowering stage. Moreover, variety-dependant forecast performance is found, especially noticeable in April 1st forecasts, where predictive skills of both stages are slightly higher for TF (Fig. 3a) than for TN (Fig. 3b). However, the differences in predictive skills are much smaller between varieties than between different initialization timings (Fig. 3).

3.2.2. Forecast performance for each tercile

GSS is computed to provide a measure of how the forecasts on each tercile class correspond to the observed tercile. Similar to FRPSS, the predictability of each phenology tercile indicated by GSS, also improves with initialization timings: barely useful for February 1st forecasts (<0.2, i.e. less than one fifth of cases could correctly predict normal events); marginally useful for March 1st forecasts (<0.4); moderately useful for April 1st forecasts (0–0.6, i.e. correct tercile predictions up to about two thirds of the cases) (Fig. 4). No discernible difference is detected

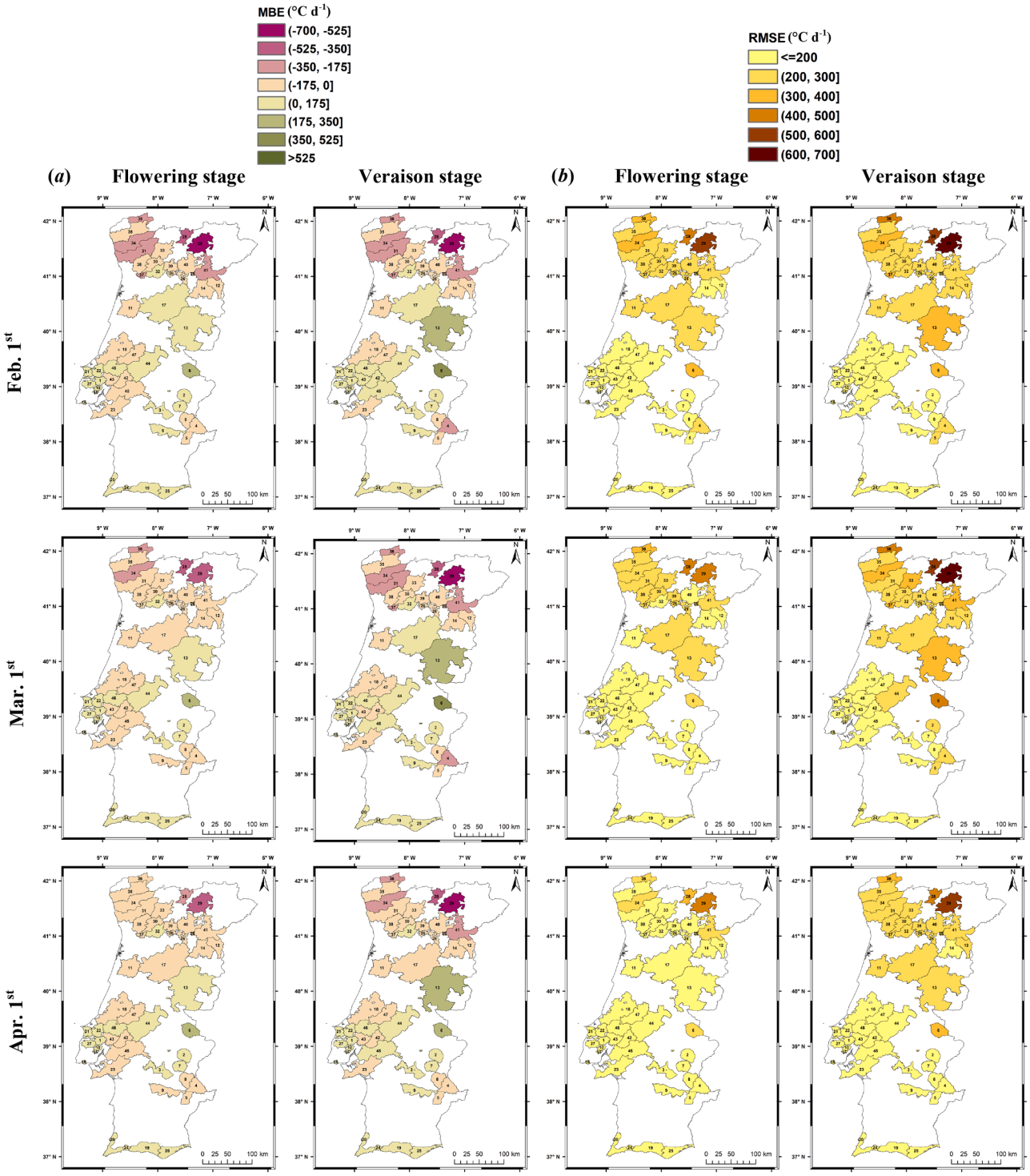


Fig. 2. The difference of seasonal total Growing Degree Days (GDD, °C d^{-1}) attained from a forecast initialization date (February 1st, March 1st and April 1st) to the date representing the late bounding value of the flowering (July 31st) and veraison stage (September 30th) irrespective of cultivars over 1993—2017 between reference and (ensemble median) forecast data. The difference is characterized as the (***a***) Mean Bias Error (MBE) and (***b***) Root Mean Squared Error (RMSE). For MBE, positive/negative values indicate overestimations/underestimations of total attained GDD. GDD is computed using daily mean temperature with a base temperature of 0°C.

amongst early/normal/late tercile using February 1st forecasts, while both March 1st and April 1st forecasts indicate a better performance for early and late tercile than for the normal one (constantly <0.2) irrespective of stages and varieties (exceptions only occur in very few regions) (Fig. 4). Spatially, better forecast skills in C-S than in N regions are mainly discovered for the early/late category using March 1st and April 1st forecasts, whereas this is absent for the normal category (Fig. 4). Both March 1st and April 1st forecasts show marginally useful forecasts (predominantly GSS<0.4) in the N regions, implying these are having limited prospects for potential applications (Fig. 4). Besides, it is difficult to distinguish the performance between the early and late terciles, as one outperforms the other in some regions and the reverse

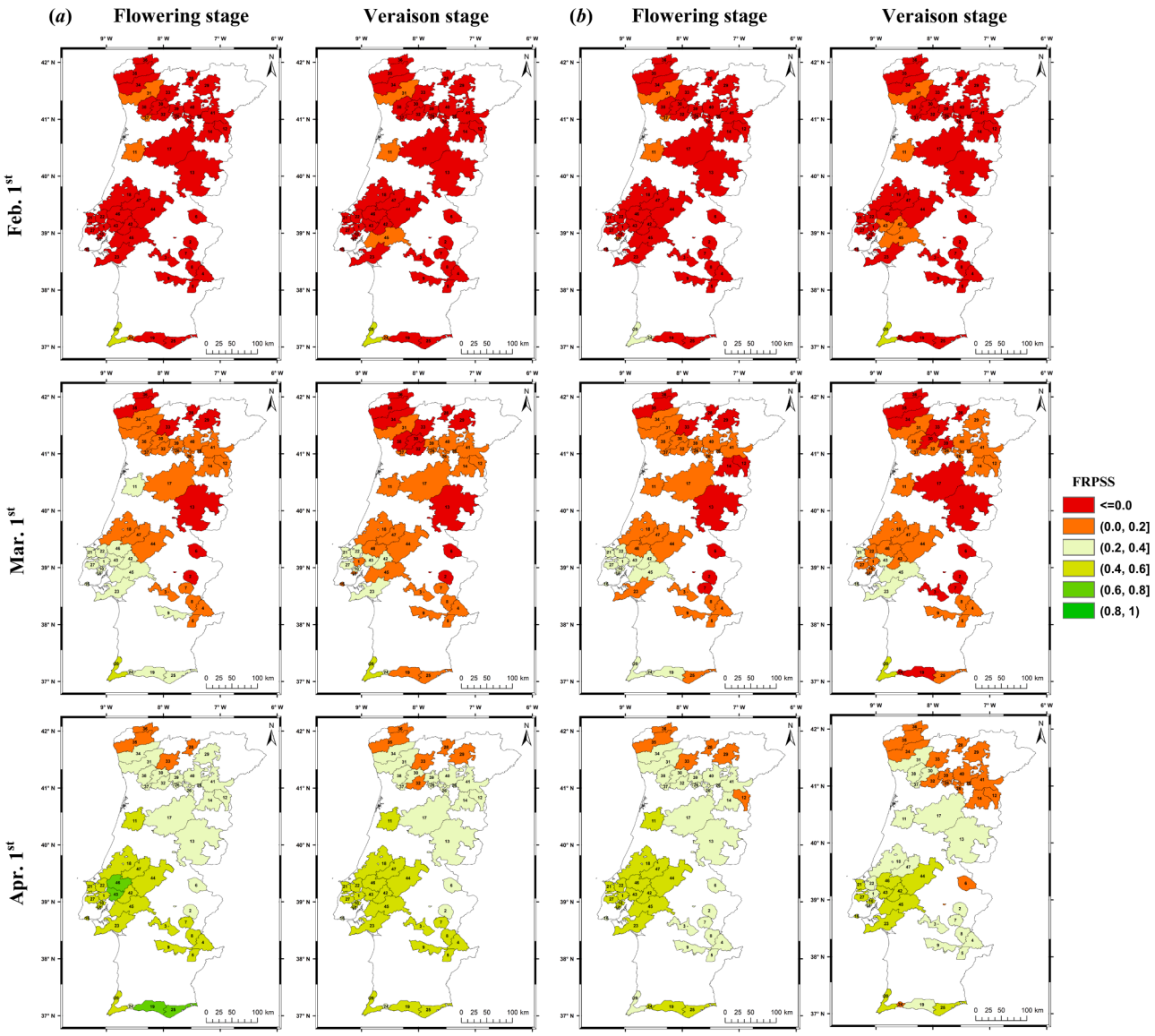


Fig. 3. The Fair Ranked Probability Skill Score (FRPSS) evaluated for the seasonal forecast on the flowering and veraison stage categorized in terciles (early/normal/late) over 1993—2017 for (***a***) Touriga Franca (TF) and (***b***) Touriga Nacional (TN) in studied wine regions in Portugal. It is evaluated for seasonal forecast initialized on February 1st, March 1st and April 1st with a 7-month lead time each year over 1993—2017. Higher/lower FRPSS values denote overall better/poorer predictive skills for the terciles.

occurs in other regions (Fig. 4). A slightly better performance in TF (Fig. 4a–b) than in TN (Fig. 4c–d) is also found, being consistent with results from FRPSS.

3.2.3. *Forecast probability of observed categories over 1993–2017*

The 1993–2017 period is classified into early/normal/late phenology years based on reference phenology simulations (Section 2.4.3). The forecast probabilities of the occurrence of observed categorical events in these classified years have been computed for studied phenology stages and varieties (Fig. S10–S13). As expected, *ens_prob* also increases from February 1st to April 1st forecasts under classified early/normal/late years (Fig. S10–S13). Normal flowering and veraison years account for most of the years during 1993–2017 for both varieties (Fig. S10–S13). However, the median *ens_prob* of early/normal/late years, indicates a consistently high forecast probability to correctly predict the occurrence of observed categories in early and/or late years than in normal years throughout studied wine regions, regardless of initialization timings, phenology stages and varieties (Fig. 5). This achieves similar results as those from GSS (Fig. 4). The median *ens_prob* of normal years is essentially well below 0.5 for all stages and varieties (<50% of correct prediction chances for the observed category in normal phenology years), whereas it is common to have >70% probability of correct predictions in early and/or late years (Fig. 5). In contrast, the extremely low (5th percentile) and high (95th percentile) *ens_prob* tends (except some of the April 1st forecasts) to follow a similar spatial pattern amongst early/normal/late years (Fig. 5). This can be associated with a considerable inter-annual variability of *ens_prob* within early/normal/late years (Fig. S14). Such a year-to-year variability of forecast performance tends to have a similarly large magnitude at different initialization timing (Fig. S14). For instance, despite the overall poor skills of February 1st forecasts (Fig. 3–4), prevailing high prediction probability (*ens_prob*>0.8) is still noticeable in some of the early and/or late years, e.g. 1993, 1994 and 1996 (Fig. S10–S13). For April 1st forecasts with promising application potentials, in a cluster of wine regions (4,5,8,23,42,45), the predicted probabilities of occurrence of reference flowering terciles for TN are quite well (>0.8) in 1997 but poor (<0.2) in 2014, where both years are classified as early flowering years (Fig. S12).

Fig. 4. The Gilbert Skill Score (GSS) evaluated for the seasonal forecast on the (***a***) flowering and (***b***) veraison stage of Touriga Franca (TF), and (***c***) flowering and (***d***) veraison stage of Touriga Nacional (TN) categorized into terciles (early/normal/late tercile) over 1993—2017. It is evaluated for the seasonal forecast initialized on February 1st, March 1st and April 1st with a 7-month lead time each year over 1993—2017. Higher/lower GSS values denote better/poorer predictive skills for a given tercile.

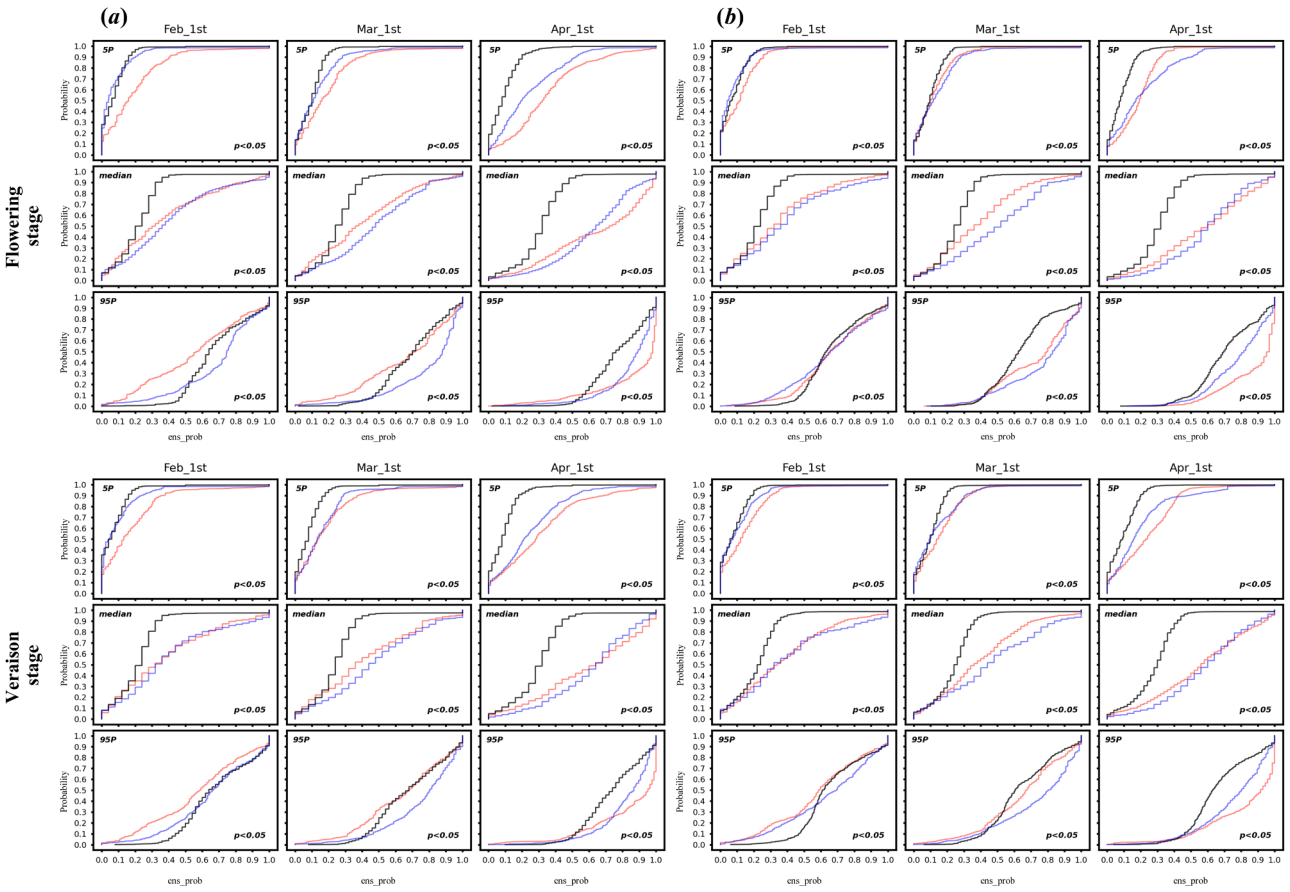


Fig. 5. The Empirical Cumulative Distribution Function (ECDF) of ensemble forecast probability (*ens_prob*) for the estimated occurrence of categorized reference flowering (upper panels) and veraison events (lower panels) of (***a***) Touriga Franca (TF) and (***b***) Touriga Nacional (TN) over classified early/normal/late years. Early years: pink line; Normal years: grey line; Late years: blue line. ECDF is used to characterize the spatial variations of *ens_prob* for the 5th percentile (5P), median and 95th percentile (95P) of classified years (see Fig. S10—Fig. S13). At any given *ens_prob* for different lines, a higher/lower ECDF value denotes a higher/lower likelihood of having more wine areas with *ens_prob* lower than the given value. A significant difference ($p<0.05$) for each paired distribution of *ens_prob* amongst Early/Normal/Late years is labelled using the Kolmogorov-Smirnov test. It is evaluated for seasonal forecast initialized on February 1st, March 1st and April 1st with a 7-month lead time each year over 1993—2017 in studied wine regions of Portugal.

3.3. Assessing the potential of seasonal forecast on phenology stages (DOY)

The quantified magnitude of prediction errors of the seasonal forecast of phenology stages also gradually decreases from February 1st to April 1st forecasts for studied stages and varieties (Fig. 6 and S15). Overall, better prediction performance is also found in C-S than in N wine regions, but the prediction accuracy is relatively higher in TN than in TF for both stages (in contrast to those in categorical prediction performance) (Fig. 6 and S15). For detailed information, a range of prediction errors expressed as RMSE (MAE) between C-S and N regions for both stages of the two varieties are summarized in Table S2.

4. Discussion and conclusions

4.1. Difference between bias-adjusted seasonal forecast and observed climate

The performance of seasonal phenology forecast, using a well-adapted phenology model with a seasonal forecast system (ECMWF-SEAS5), is evaluated for two important phenology stages of two representative varieties throughout wine regions in Portugal in 1993–2017. The reliability of phenology forecasts would depend on how the raw data of seasonal forecasts is being processed. Various bias-adjustment methods are commonly applied (Ceglar and Toreti, 2021; Chen and Tao, 2022; Wang et al., 2019). We have chosen the quantile mapping method because it has previously demonstrated its suitability in predicting wheat flowering timings across Europe (Ceglar and Toreti, 2021).

To evaluate the resulting bias-adjusted seasonal climate against observations, a pack of diagnostic metrics are applied (R^2, MBE, RMSE and Kolmogorov-Smirnov test). The findings reveal that the within-season climate variability (seasonal evolution pattern) is well reproduced ($R^2>0.95$) by forecast series at different initialization timings and for each year over 1993–2017 (Fig. S6–S8). Conversely, a considerable daily bias of predicted seasonal climate is found (Fig. S6–S8), which results in the difference in the seasonal total GDD between reference and forecast series (Fig. 2 and Fig. S9). This can primarily explain the lack of skill (Fig. 3) and substantial errors (Fig. 6) for February 1st and March 1st forecasts. For April 1st forecasts, such a daily bias (Fig. S6) and resulting prediction errors in the seasonal total GDD (Fig. 2b) are still considerable. The errors in the daily output of seasonal forecasts can lead to large prediction errors for variables of interest (Chen and Tao, 2022; Jha et al., 2019). Chen and Tao (2022) show that seasonal climate predictions perform well in reproducing reference climate anomalies, but are poor to simulate daily weather conditions, leading to less skilful crop yield forecasts than those of climatology. To better translate the untapped seasonal climate predictability into information relevant to agricultural decision-making (e.g. phenology), we point out the need to improve the representation of intra-seasonal variability of climate variables, with a

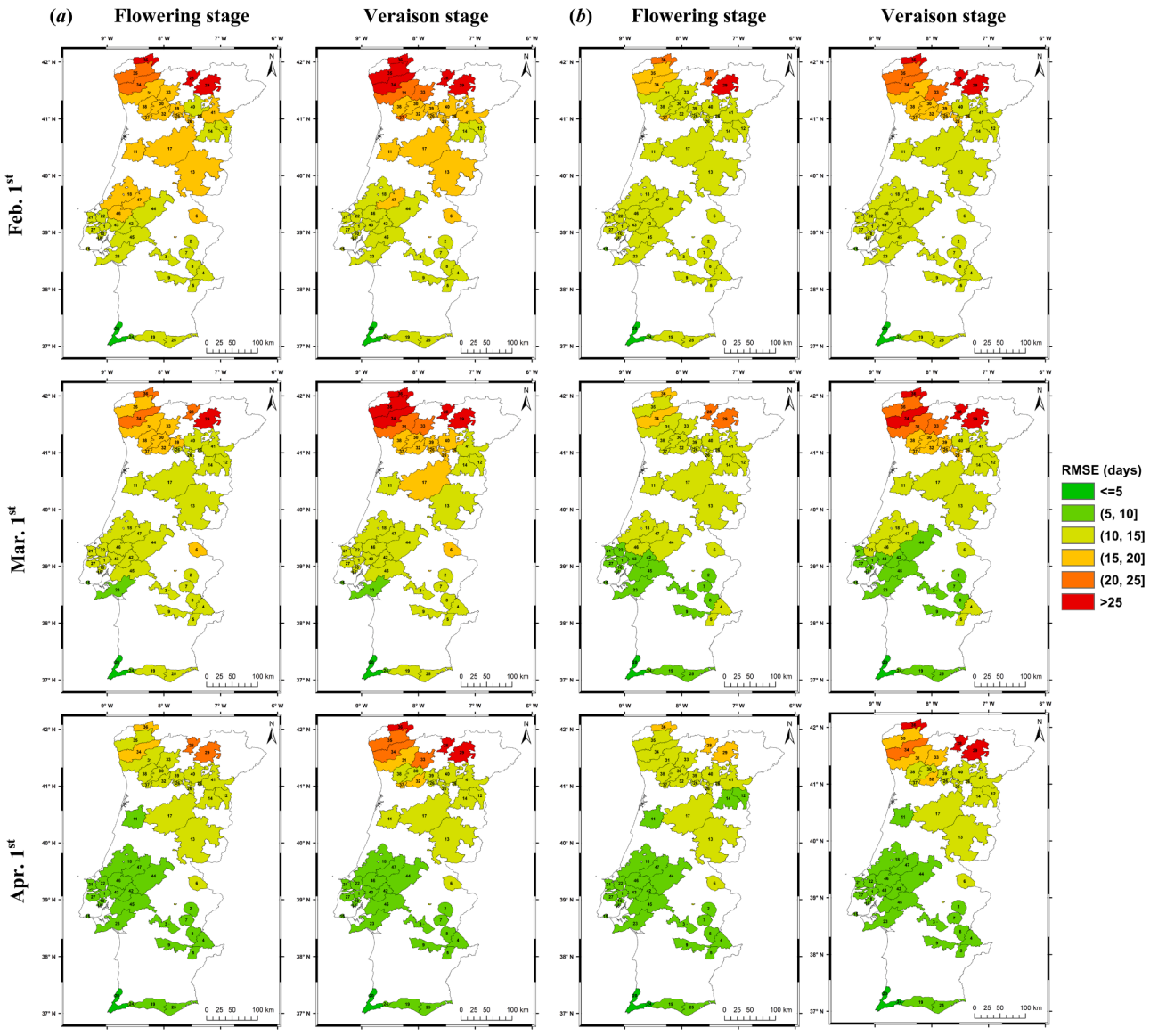


Fig. 6. The Root Mean Squared Error (RMSE) between the reference and seasonal forecast on the flowering and veraison stage (DOY) over 1993—2017 for (***a***) Touriga Franca (TF) and (***b***) Touriga Nacional (TN). It is evaluated for seasonal forecast initialized on February 1st, March 1st and April 1st with a 7-month lead time each year over 1993—2017.

particular emphasis to explore novel methods to reduce the day-to-day difference between predicted and observed seasonal climate. More comprehensive and standardized metrics (beyond metrics applied here) should be established to evaluate to what extent the difference between forecast and observed series could lead to generally poor forecast performance.

4.2. Predictability of categorical phenology events

4.2.1. Overall performance

FRPSS is a comprehensive measure to evaluate the predictive skills of categorical events from probabilistic forecasts (Ferro, 2014). Our findings demonstrate that the predictability of the phenology tercile classes mainly depends on the forecast initialization timings, regions and the nature of predicting subjects (i.e. phenology stages and varieties). Following other studies (Bento et al., 2022; Ceglar and Toreti, 2021; Jha et al., 2019), the forecast performance improves with the initialization timing along the growing season, as more observed data has been integrated. Notably, this pattern is consistent throughout this study. For FRPSS, February 1st and March 1st forecasts show no better performance than climatology for most of the wine regions, whereas moderately skilful predictions are obtained using April 1st forecasts in C-S wine regions (Fig. 3). This may reflect the bottleneck issue in seasonal forecasts, i.e. a tendency to better inform farmer's decision-making with a shorter lead-time (Nyamekye et al., 2021), but with fewer added-values (in terms of early information availability) relative to those with a longer lead time. A better prediction performance in C-S regions than in N regions might be essentially due to a shorter forecast length in C-S regions as a result of higher temperatures, with typically earlier phenology occurrence (about 20–40 days earlier as shown in Fig. S4–S5). This spatial pattern can be further enhanced by the tendency of underestimations/overestimations for seasonal total GDD in N/C-S regions (Fig. 2a), corresponding to later/earlier phenology predictions. Therefore, substantially less amount of forecast information could be utilized in C-S than in N regions for the seasonal phenology forecasts of both varieties. The spatial pattern of better forecast performance in C-S than in N regions remains unaltered for either predicting categorical events or actual phenology stages (Fig. 3–4 and Fig. 6). Slightly better predictive skills for TF than for TN can be partially attributed to the earlier phenology occurrence of TF, particularly for the veraison stage

(Fig. S4–S5) (Costa et al., 2019; Reis et al., 2020). Nevertheless, since observed phenology data is not available throughout all wine regions, uncertainties exist concerning how the model adapts to the two varieties using locally available observations. This emphasizes the need to build a network of phenology observations to facilitate extensive model calibrations and evaluations.

4.2.2. Forecast skills between early/late and normal phenology tercile

To gain insights into the predictive skills of each phenology tercile, GSS is computed to measure how well the forecasts of anomalies (early/late) and normal events correspond to observations. GSS complements FRPSS, as the former additionally reveals that early and late flowering and veraison events are more predictable than normal ones, especially when using April 1st forecasts (Fig. 4). Therefore, the identified overall moderate forecast skills (0.2<FRPSS<0.6) (Fig. 3) result from interactions between comparably higher predictability in early/late categories and lower predictability in the normal one. This agrees with Ceglar and Toreti (2021) showing more reliable predictions for early/late than normal wheat flowering terticles in Europe. Anomalously early/late phenology occurrence can be reasonably predicted (0.2<GSS<0.6) only using April 1st forecasts in C-S regions (Fig. 4), which again suggests the role of daily bias of seasonal climate on limiting the predictive skills of anomalous events. However, this also depends on how the phenology model responds to the bias. As previously discussed, a tendency for early phenology occurrence is associated with a shorter forecast length, thus less error-prone. Indeed, for both the early and late categories, higher predictability is always obtained in C-S than in N regions (Fig. 4). Nonetheless, it remains difficult to accurately pinpoint the reasons underlying the constantly poor forecast skills (GSS<0.2) of the normal category (Fig. 4).

Comparable predictive skills between flowering and veraison for both varieties (Fig. 3–4) can be essentially attributed to the known strength of SEAS5 to capture anomalous signals in summer months, where it tends to better capture both the warm (third tercile of T_{mean}) and cold anomalies (first tercile of T_{mean}) during June–August than during March-May in Iberian Peninsula (Crespi et al., 2021). Similarly, the ECMWF forecast system coupled with a regression model also shows that the predictability of different wine production categories is higher in the summer months (May-August) than in the spring months (February–April) (Santos et al., 2020a). Such a performance is particularly critical for predicting the late veraison event, which can occur until the end of August in C-S regions (Fig. S4–S5).

4.2.3. Forecast probability of observations in early/normal/late years

The inter-annual variability of forecast performance is an important feature of the Mediterranean climate. In an attempt to minimize such influence, the study period is classified into early/normal/late years. The classified years vary between phenology stages and varieties (Fig. S10–S13), which again highlights the dependence of forecast performance on the nature of predicting subjects (Ceglar and Toreti, 2021). The median confidence (median *ens_prob*) to correctly predict the reference category is generally higher in classified early/late phenology years than that in normal phenology years for most of the wine regions (Fig. 5). For a study on early-season predictions of durum wheat yields across the Mediterranean Basin, better estimations of observations are also obtained in years with either highest or lowest reference yields (Ferrise et al., 2015). Since the classified early/late phenology years imply the spatial majority of early/late reference categories in those years, the forecast probability results reinforce the skill assessment findings (Fig. 4), i.e. better predictability for early/late than for normal category. In contrast to the median values, a remarkable temporal variability of *ens_prob* over years belonging to the same classified phenology years, and for all three forecast timings (Fig. 5 and Fig. S14) confirm the substantial impacts of irregular Mediterranean climate conditions on seasonal phenology forecast. This can be the main challenging aspect to develop reliable seasonal forecasts in the Iberian Peninsula (Crespi et al., 2021). To address this issue, one possible approach is to use a better initialization scheme, meaning to assimilate more observations and interpolate them to best align with the seasonal forecast model, to minimize artefacts at initialization time. Ceglar et al. (2018) revealed realistic land-surface initializations in May and June can improve seasonal predictions of warm extremes and heat waves, leading to better capturing of the inter-annual variability of observed maize yield anomalies. A more explicit classification of anomalous events, e.g. by adding the extreme categories (<5th and >95th quantile), can be useful to better describe the anomalies arising from the year-to-year variability. However, this may lead to poorer forecast performance for each category, due to the trade-off between forecast skills and the specificity of the event to predict.

4.3. Potential of seasonal forecast on the phenology stages (DOY)

The categorical predictions only serve to alert winegrowers to the probability of having an anomalous event. This may have more practical applications at the regional level, as a warning can be issued for a certain wine region. However, for those with a clear need for accurate phenology predictions, it is necessary to deploy a more precise forecast system. For this purpose, our results demonstrate a promising operational forecast for some Portuguese wine regions. The quantified potential forecast errors (MAE and RMSE) (Fig. 6 and Fig. S15), show a generally similar pattern as those from categorical predictions (Fig. 3–4). However, better predictions are achieved in TN than in TF, which is contrary to those of categorical predictions. For April 1st forecasts, with identified moderate skills on categorical events, the prediction errors are generally <=10 days in C-S regions and 10–25 days in N regions, for both stages of the two varieties (Fig. 6 and Fig. S15). However, these errors need to be considered together with simulation errors of observed phenology using gridded climate inputs (up to 5 days of RMSE) (Fig. 1), leading to estimated cumulative prediction errors of about 2 weeks in C-S regions and 2–4 weeks (forecasts are up to 28 days apart from actual observations) in N regions using April 1st forecasts. To accurately monitor the phenology development and better inform the winegrowers, RMSE normally needs to be shorter than a week (<7 days) (Parker et al., 2013; Reis et al., 2020). Hence, the accuracy of the current seasonal forecast is insufficient to meet such a criterion, but it depends on whether the users are risk-adverse or risk-tolerant.

To further reduce prediction errors, the current framework could be directly combined with remote-sensed observations, as it can benefit from continuously evolving information on crop growth conditions (e.g. LAI and EVI) relating to phenology development (Chen and Tao, 2022; Peng et al., 2018). Incorporating remote-sensed observations might also provide one pathway to respond to the challenge of obtaining accurate data about the canopy temperature over large areas, which is the actual driver of vine phenology. Vineyards are associated with topographical niches in different regions, and fine-scale microclimate variations are greatly influenced by the topography and terrain (Mosedale et al., 2016). This is a challenging issue to capture by either the forecast (e.g. SEAS5 does not allow distinguish micro-climates) or the downscaling technique (each one can have its shortcomings). Thus, the operational forecast performance can substantially decrease from vineyards located within a flat terrain (e.g. Alentejo) to those of mountainous areas (e.g. Douro).

4.4. Recommendations for methodological improvements and practical applications

Our study provides the basis for practical applications, but the method itself can be improved. Prospective methodological innovations can result from improved forecast systems with higher spatial-temporal resolutions, in combination with enhanced observational systems that can better characterize the initial conditions. The size of the seasonal forecast ensemble and the number of phenology models can be

increased. Although the sigmoid mode is well adapted to local conditions, the use of a multi-phenology-model framework can be more robust to cope with the potential variability of phenology simulations associated with different model structures (Seidel et al., 2018; Tao et al., 2018; Wallach et al., 2021). The multi-model ensemble mean/median proves to better represent observations than any individual model (Martre et al., 2015). Comparably, different forecasting systems may also lead to different seasonal climate predictions and forecast performance of target variables (Chen and Tao, 2022). The use of various forecast systems with large ensembles is also recommended. Nonetheless, the choice of SEAS5 for the European continent is generally considered appropriate (Crespi et al., 2021; Johnson et al., 2019). Different calibration/bias-adjustment methods can result in different phenology/seasonal climate simulations, which emphasize the need to investigate their impacts. All these aspects attempt to address how to make the best use of available resources/observations to operationalize the seasonal forecast.

For practical applications of seasonal phenology forecast, it is recommended to always evaluate the relevant lead time and initialization for specific regions and events. Our results show promising operational forecasts in C-S wine regions (PDO in Alentejo, Lisboa, and Ribatejo) initialized on April 1st, i.e. approximately 1–2 and 3–4 months lead time for the flowering and veraison stages, respectively. In particular, it is early enough for the veraison stage, thus with higher added value. Winegrowers can have up to 4 months in advance to develop plans to reduce the risk of having low yield/quality. The pronounced ability to predict the early phenology tercile can be of foremost relevance to making informed decisions for winegrowers in these regions under climate change. As increased predictability can be expected, resulting from earlier phenology occurrences (Costa et al., 2019), in response to the projected warming, with more frequent extreme events in future climates for these vineyards (Santos et al., 2020). Moreover, the initiatives to build climate service tools in a co-design approach with end-users, including winegrowers, viticulturists, and service extension consultants, can significantly facilitate the implementation of a reliable forecast system (Ceglar and Toreti, 2021). The current approach can be extrapolated to other wine regions worldwide, mostly to those with similar climates.

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.

Data availability

Data will be made available on request.

Acknowledgements

This study is funded by the Clim4Vitis project—"Climate change impact mitigation for European viticulture: knowledge transfer for an integrated approach", funded by the European Union's Horizon 2020 Research and Innovation Programme, under grant agreement no. 810176. This work is also supported by National Funds by FCT – Portuguese Foundation for Science and Technology under the project UIDB/04033/2020 and funded by the CoaClimateRisk project (COA/CAC/0030/2019) of the FCT. H.F. thanks the FCT for contract CEECIND/00447/2017 and 2022.02317.CEECIND.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.agrformet.2023.109342.

References

Bento, V.A., Russo, A., Dutra, E., Ribeiro, A.F.S., Gouveia, C.M., Trigo, R.M., 2022. Persistence versus dynamical seasonal forecasts of cereal crop yields. Sci. Rep. 12, 7422. https://doi.org/10.1038/s41598-022-11228-2.

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., Bussière, F., Cabidoche, Y.M., Cellier, P., Debaeke, P., Gaudillère, J.P., Hénault, C., Maraux, F., Seguin, B., Sinoquet, H., 2003. An overview of the crop model stics. Eur. J. Agron. 18, 309–332. https://doi.org/10.1016/S1161-0301(02)00110-7.

Brown, J.N., Hochman, Z., Holzworth, D., Horan, H., 2018. Seasonal climate forecasts provide more definitive and accurate crop yield predictions. Agric. For. Meteorol. 260–261, 247–254. https://doi.org/10.1016/j.agrformet.2018.06.001.

Candiago, S., Tscholl, S., Bassani, L., Fraga, H., Egarter Vigl, L., 2022. A geospatial inventory of regulatory information for wine protected designations of origin in Europe. Sci. Data 9, 394. https://doi.org/10.1038/s41597-022-01513-0.

Ceglar, A., Toreti, A., 2021. Seasonal climate forecast can inform the European agricultural sector well in advance of harvesting. npj Clim. Atmos. Sci. 4, 42. https://doi.org/10.1038/s41612-021-00198-3.

Ceglar, A., Toreti, A., Lecerf, R., Van der Velde, M., Dentener, F., 2016. Impact of meteorological drivers on regional inter-annual crop yield variability in France. Agric. For. Meteorol. 216, 58–67. https://doi.org/10.1016/j.agrformet.2015.10.004.

Ceglar, A., Toreti, A., Prodhomme, C., Zampieri, M., Turco, M., Doblas-Reyes, F.J., 2018. Land-surface initialisation improves seasonal climate prediction skill for maize yield forecast. Sci. Rep. 8, 1322. https://doi.org/10.1038/s41598-018-19586-6.

Ceglar, A., Turco, M., Toreti, A., Doblas-Reyes, F.J., 2017. Linking crop yield anomalies to large-scale atmospheric circulation in Europe. Agric. For. Meteorol. 240–241, 35–45. https://doi.org/10.1016/j.agrformet.2017.03.019.

Challinor, A.J., Watson, J., Lobell, D.B., Howden, S.M., Smith, D.R., Chhetri, N., 2014. A meta-analysis of crop yield under climate change and adaptation. Nat. Clim. Chang. 4, 287–291. https://doi.org/10.1038/nclimate2153.

Chen, Y., Tao, F., 2022. Potential of remote sensing data-crop model assimilation and seasonal weather forecasts for early-season crop yield forecasting over a large area. F. Crop. Res. 276, 108398 https://doi.org/10.1016/j.fcr.2021.108398.

Chuine, I., Cour, P., Rousseau, D.D., 1998. Fitting models predicting dates of flowering of temperate-zone trees using simulated annealing. Plant. Cell Environ. 21, 455–466. https://doi.org/10.1046/j.1365-3040.1998.00299.x.

Clemente, N., Santos, J.A., Fontes, N., Graça, A., Gonçalves, I., Fraga, H., 2022. Grapevine sugar concentration model (GSCM): a decision support tool for the douro superior winemaking region. Agron. https://doi.org/10.3390/agronomy12061404.

Cornes, R.C., van der Schrier, G., van den Besselaar, E.J.M., Jones, P.D., 2018. An ensemble version of the E-OBS temperature and precipitation data sets. J. Geophys. Res. Atmos. 123, 9391–9409. https://doi.org/10.1029/2017JD028200.

Costa, R., Fraga, H., Fonseca, A., De Cortázar-Atauri, I.G., Val, M.C., Carlos, C., Reis, S., Santos, J.A., 2019. Grapevine phenology of cv. touriga franca and touriga nacional in the douro wine region: modelling and climate change projections. Agronomy. https://doi.org/10.3390/agronomy9040210.

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

Wilks, Daniel S., 2020. Statistical Methods in the Atmospheric Sciences, 4th ed. Elsevier, NY. https://doi.org/10.1016/C2017-0-03921-6.

Du, H., Doblas-Reyes, F.J., García-Serrano, J., Guemas, V., Soufflet, Y., Wouters, B., 2012. Sensitivity of decadal predictions to the initial atmospheric and oceanic perturbations. Clim. Dyn. 39, 2013–2023. https://doi.org/10.1007/s00382-011-1285-9.

Ferrise, R., Toscano, P., Pasqui, M., Moriondo, M., Primicerio, J., MA, S., Bindi, M., 2015. Monthly-to-seasonal predictions of durum wheat yield over the Mediterranean Basin. Clim. Res. 65, 7–21.

Ferro, C.A.T., 2014. Fair scores for ensemble forecasts. Q. J. R. Meteorol. Soc. 140, 1917–1923. https://doi.org/10.1002/qj.2270.

Fraga, H., García de Cortázar Atauri, I., Malheiro, A.C., Moutinho-Pereira, J., Santos, J. A., 2017. Viticulture in Portugal: a review of recent trends and climate change projections. OENO One 51, 61–69. https://doi.org/10.20870/oeno-one.2017.51.2.1621.

Fraga, H., García de Cortázar Atauri, I., Malheiro, A.C., Santos, J.A., 2016. Modelling climate change impacts on viticultural yield, phenology and stress conditions in Europe. Glob. Chang. Biol. 22, 3774–3788. https://doi.org/10.1111/gcb.13382.

Fraga, H., Molitor, D., Leolini, L., Santos, J.A., 2020. What is the impact of heatwaves on European viticulture? A modelling assessment. Appl. Sci. https://doi.org/10.3390/app10093030.

Gudmundsson, L., Bremnes, J.B., Haugen, J.E., Engen-Skaugen, T., 2012. Technical Note: downscaling RCM precipitation to the station scale using statistical transformations - a comparison of methods. Hydrol. Earth Syst. Sci. 16, 3383–3390. https://doi.org/10.5194/hess-16-3383-2012.

Hänninen, H., 1990. Modelling bud dormancy release in trees from cool and temperate regions. Acta For. Fenn. 213, 1–47. https://doi.org/10.14214/aff.7660.

Haylock, M.R., Hofstra, N., Klein Tank, A.M.G., Klok, E.J., Jones, P.D., New, M., 2008. A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. J. Geophys. Res. Atmos. 113 https://doi.org/10.1029/2008JD010201.

Hodges, J.L., 1958. The significance probability of the Smirnov two-sample test. Ark. för Mat. 3, 469–486.

Holzworth, D.P., Huth, N.I., deVoil, P.G., 2014. APSIM – Evolution towards a new generation of agricultural systems simulation. Environ. Model. Softw. 62, 327–350. https://doi.org/10.1016/j.envsoft.2014.07.009.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O.B., Bouwer, L.M., Braun, A., Colette, A., Déqué, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., Yiou, P., 2014. EURO-CORDEX: new high-resolution climate change projections for European impact research. Reg. Environ. Chang. 14, 563–578. https://doi.org/10.1007/s10113-013-0499-2.

Jha, P.K., Athanasiadis, P., Gualdi, S., Trabucco, A., Mereu, V., Shelia, V., Hoogenboom, G., 2019. Using daily data from seasonal forecasts in dynamic crop models for yield prediction: a case study for rice in Nepal's. Terai. Agric. For. Meteorol. 265, 349–358. https://doi.org/10.1016/j.agrformet.2018.11.029.

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., Monge-Sanz, B.M., 2019. SEAS5: the new ECMWF seasonal forecast system. Geosci. Model Dev. 12, 1087–1117. https://doi.org/10.5194/gmd-12-1087-2019.

Lobell, D.B., Field, C.B., 2007. Global scale climate–crop yield relationships and the impacts of recent warming. Environ. Res. Lett. 2, 14002. https://doi.org/10.1088/1748-9326/2/1/014002.

Lorenz, D.H., Eichhorn, K.W., Bleiholder, H., Klose, R., Meier, U., Weber, E., 1995. Growth stages of the grapevine: phenological growth stages of the grapevine (Vitis vinifera L. ssp. vinifera)—codes and descriptions according to the extended BBCH scale†. Aust. J. Grape Wine Res. 1, 100–103. https://doi.org/10.1111/j.1755-0238.1995.tb00085.x.

Martre, P., Wallach, D., Asseng, S., Ewert, F., Jones, J.W., Rötter, R.P., Boote, K.J., Ruane, A.C., Thorburn, P.J., Cammarano, D., Hatfield, J.L., Rosenzweig, C., Aggarwal, P.K., Angulo, C., Basso, B., Bertuzzi, P., Biernath, C., Brisson, N., Challinor, A.J., Doltra, J., Gayler, S., Goldberg, R., Grant, R.F., Heng, L., Hooker, J., Hunt, L.A., Ingwersen, J., Izaurralde, R.C., Kersebaum, K.C., Müller, C., Kumar, S.N., Nendel, C., O'leary, G., Olesen, J.E., Osborne, T.M., Palosuo, T., Priesack, E., Ripoche, D., Semenov, M.A., Shcherbak, I., Steduto, P., Stöckle, C.O., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Travasso, M., Waha, K., White, J.W., Wolf, J., 2015. Multimodel ensembles of wheat growth: many models are better than one. Glob. Chang. Biol. 21, 911–925. https://doi.org/10.1111/gcb.12768.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. J. Chem. Phys. 21, 1087–1092. https://doi.org/10.1063/1.1699114.

Mosedale, J.R., Abernethy, K.E., Smart, R.E., Wilson, R.J., Maclean, I.M.D., 2016. Climate change impacts and adaptive strategies: lessons from the grapevine. Glob. Chang. Biol. 22, 3814–3828. https://doi.org/10.1111/gcb.13406.

Nyamekye, A.B., Nyadzi, E., Dewulf, A., Werners, S., Van Slobbe, E., Biesbroek, R.G., Termeer, C.J.A.M., Ludwig, F., 2021. Forecast probability, lead time and farmer decision-making in rice farming systems in Northern Ghana. Clim. Risk Manag. 31, 100258 https://doi.org/10.1016/j.crm.2020.100258.

OIV, 2021. State of the world vitivinicultural sector in 2020. Paris, France.

Panofsky, H.A., Brier, G.W., Best, W.H., 1968. Some Application of Statistics to meteorology. Mineral Industries Extension Services, School of Mineral Industries. Pennsylvania State College.

Parker, A., de Cortázar-Atauri, I.G., Chuine, I., Barbeau, G., Bois, B., Boursiquot, J.-M., Cahurel, J.-Y., Claverie, M., Dufourcq, T., Gény, L., Guimberteau, G., Hofmann, R. W., Jacquet, O., Lacombe, T., Monamy, C., Ojeda, H., Panigai, L., Payan, J.-C., Lovelle, B.R., Rouchaud, E., Schneider, C., Spring, J.-L., Storchi, P., Tomasi, D., Trambouze, W., Trought, M., van Leeuwen, C., 2013. Classification of varieties for their timing of flowering and veraison using a modelling approach: a case study for the grapevine species Vitis vinifera L. Agric. For. Meteorol. 180, 249–264. https://doi.org/10.1016/j.agrformet.2013.06.005.

Peng, B., Guan, K., Pan, M., Li, Y., 2018. Benefits of seasonal climate prediction and satellite data for forecasting U.S. Maize yield. Geophys. Res. Lett. 45, 9662–9671. https://doi.org/10.1029/2018GL079291.

Piani, C., Haerter, J.O., Coppola, E., 2010. Statistical bias correction for daily precipitation in regional climate models over. Europe. Theor. Appl. Climatol. 99, 187–192. https://doi.org/10.1007/s00704-009-0134-9.

Piao, S., Liu, Q., Chen, A., Janssens, I.A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., Zhu, X., 2019. Plant phenology and global climate change: current progresses and challenges. Glob. Chang. Biol. 25, 1922–1940. https://doi.org/10.1111/gcb.14619.

Ramos, M.C., Go, D.T.H.C., Castro, S., 2021. Spatial and temporal variability of cv. Tempranillo response within the Toro DO (Spain) and projected changes under climate change. OENO One 55, 349–366. https://doi.org/10.20870/oeno-one.2021.55.1.4493.

Reis, S., Fraga, H., Carlos, C., Silvestre, J., Eiras-Dias, J., Rodrigues, P., Santos, J.A., 2020. Grapevine phenology in four portuguese wine regions: modeling and predictions. Appl. Sci. https://doi.org/10.3390/app10113708.

Rodrigues, P., Pedroso, V., Reis, S., Yang, C., Santos, J.A., 2022. Climate change impacts on phenology and ripening of cv. Touriga Nacional in the Dão wine region. Portugal. Int. J. Climatol. https://doi.org/10.1002/joc.7633.

Santos, João A, Ceglar, A., Toreti, A., Prodhomme, C., 2020a. Performance of seasonal forecasts of Douro and Port wine production. Agric. For. Meteorol. 291, 108095 https://doi.org/10.1016/j.agrformet.2020.108095.

Santos, J.A., Costa, R., Fraga, H., 2019. New insights into thermal growing conditions of Portuguese grapevine varieties under changing climates. Theor. Appl. Climatol. 135, 1215–1226. https://doi.org/10.1007/s00704-018-2443-3.

Santos, João A., Fraga, H., Malheiro, A.C., Moutinho-Pereira, J., Dinis, L.T., Correia, C., Moriondo, M., Leolini, L., Dibari, C., Costafreda-Aumedes, S., Kartschall, T., Menz, C., Molitor, D., Junk, J., Beyer, M., Schultz, H.R., 2020b. A review of the potential climate change impacts and adaptation options for European viticulture. Appl. Sci. 10 https://doi.org/10.3390/app10093092.

Santos, M., Fonseca, A., Fraga, H., Jones, G.V., Santos, J.A., 2020. Bioclimatic conditions of the Portuguese wine denominations of origin under changing climates. Int. J. Climatol. 40, 927–941. https://doi.org/10.1002/joc.6248.

Schaefer, J.T., 1990. The critical success index as an indicator of warning skill. Weather Forecast 5, 570–575. https://doi.org/10.1175/1520-0434(1990)005<0570:TCSIAA>2.0.CO;2.

Seidel, S.J., Palosuo, T., Thorburn, P., Wallach, D., 2018. Towards improved calibration of crop models – where are we now and where should we go? Eur. J. Agron. https://doi.org/10.1016/j.eja.2018.01.006.

Semenov, M.A., 2007. Utility of dynamical seasonal forecasts in predicting crop yield. Clim. Res. 34, 71–81.

Stockdale, T., 2021. SEAS5 user guide. https://doi.org/10.21957/2y67999y.

Tao, F., Rötter, R.P., Palosuo, T., Gregorio Hernández Díaz-Ambrona, C., Mínguez, M.I., Semenov, M.A., Kersebaum, K.C., Nendel, C., Specka, X., Hoffmann, H., Ewert, F., Dambreville, A., Martre, P., Rodríguez, L., Ruiz-Ramos, M., Gaiser, T., Höhn, J.G., Salo, T., Ferrise, R., Bindi, M., Cammarano, D., Schulman, A.H., 2018. Contribution of crop model structure, parameters and climate projections to uncertainty in climate change impact assessments. Glob. Chang. Biol. 24, 1291–1307. https://doi.org/10.1111/gcb.14019.

Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. Atmos. 106, 7183–7192. https://doi.org/10.1029/2000JD900719.

Taylor, S.D., White, E.P., 2020. Automated data-intensive forecasting of plant phenology throughout the United States. Ecol. Appl. 30, e02025. https://doi.org/10.1002/eap.2025.

van der Velde, M., Nisini, L., 2019. Performance of the MARS-crop yield forecasting system for the European Union: assessing accuracy, in-season, and year-to-year improvements from 1993 to 2015. Agric. Syst. 168, 203–212. https://doi.org/10.1016/j.agsy.2018.06.009.

van Leeuwen, C., Destrac-Irvine, A., 2017. Modified grape composition under climate change conditions requires adaptations in the vineyard. Oeno One. https://doi.org/10.20870/oeno-one.2016.0.0.1647.

van Leeuwen, C., Destrac-Irvine, A., Dubernet, M., Duchêne, E., Gowdy, M., Marguerit, E., Pieri, P., Parker, A., de Rességuier, L., Ollat, N., 2019. An update on the impact of climate change in viticulture and potential adaptations. Agron. https://doi.org/10.3390/agronomy9090514.

Wallach, D., Palosuo, T., Thorburn, P., Gourdain, E., Asseng, S., Basso, B., Buis, S., Crout, N., Dibari, C., Dumont, B., Ferrise, R., Gaiser, T., Garcia, C., Gayler, S., Ghahramani, A., Hochman, Z., Hoek, S., Hoogenboom, G., Horan, H., Huang, M., Jabloun, M., Jing, Q., Justes, E., Kersebaum, K.C., Klosterhalfen, A., Launay, M., Luo, Q., Maestrini, B., Mielenz, H., Moriondo, M., Nariman Zadeh, H., Olesen, J.E., Poyda, A., Priesack, E., Pullens, J.W.M., Qian, B., Schütze, N., Shelia, V., Souissi, A., Specka, X., Srivastava, A.K., Stella, T., Streck, T., Trombi, G., Wallor, E., Wang, J., Weber, T.K.D., Weihermüller, L., de Wit, A., Wöhling, T., Xiao, L., Zhao, C., Zhu, Y., Seidel, S.J., 2021. How well do crop modeling groups predict wheat phenology, given calibration data from the target population? Eur. J. Agron. 124, 126195 https://doi.org/10.1016/j.eja.2020.126195.

Wang, Q.J., Shao, Y., Song, Y., Schepen, A., Robertson, D.E., Ryu, D., Pappenberger, F., 2019. An evaluation of ECMWF SEAS5 seasonal climate forecasts for Australia using a new forecast calibration algorithm. Environ. Model. Softw. 122, 104550 https://doi.org/10.1016/j.envsoft.2019.104550.

Yang, C., Fraga, H., Ieperen, W.Van, Santos, J.A., 2017. Assessment of irrigated maize yield response to climate change scenarios in Portugal. Agric. Water Manag. 184, 178–190. https://doi.org/10.1016/j.agwat.2017.02.004.

Yang, C., Fraga, H., van Ieperen, W., Santos, J.A., 2020. Assessing the impacts of recent-past climatic constraints on potential wheat yield and adaptation options under Mediterranean climate in southern Portugal. Agric. Syst. https://doi.org/10.1016/j.agsy.2020.102844.

Yang, C., Fraga, H., Van Ieperen, W., Santos, J.A., 2018. Modelling climate change impacts on early and late harvest grassland systems in Portugal. Crop. Pasture Sci. 69, 821–836. https://doi.org/10.1071/CP17428.

Yang, C., Fraga, H., van Ieperen, W., Trindade, H., Santos, J.A., 2019. Effects of climate change and adaptation options on winter wheat yield under rainfed Mediterranean

conditions in southern Portugal. Clim. Change 154, 159–178. https://doi.org/10.1007/s10584-019-02419-4.

Yang, C., Menz, C., De Abreu Jaffe, M.S., Costafreda-Aumedes, S., Moriondo, M., Leolini, L., Torres-Matallana, A., Molitor, D., Junk, J., Fraga, H., van Leeuwen, C., Santos, J.A., 2022a. Projections of climate change impacts on flowering-veraison water deficits for riesling and Müller-thurgau in Germany. Remote Sens. https://doi.org/10.3390/rs14061519.

Yang, C., Menz, C., Fraga, H., Costafreda-Aumedes, S., Leolini, L., Ramos, M.C., Molitor, D., van Leeuwen, C., Santos, J.A., 2022b. Assessing the grapevine crop water stress indicator over the flowering-veraison phase and the potential yield lose rate in important European wine regions. Agric. Water Manag. 261, 107349 https://doi.org/10.1016/j.agwat.2021.107349.

Yang, C., Menz, C., Fraga, H., Reis, S., Machado, N., Malheiro, A.C., Santos, J.A., 2021. Simultaneous calibration of grapevine phenology and yield with a soil–plant–atmosphere system model using the frequentist method. Agron. https://doi.org/10.3390/agronomy11081659.

Zhu, X., Troy, T.J., 2018. Agriculturally relevant climate extremes and their trends in the world's major growing regions. Earth's Futur 6, 656–672. https://doi.org/10.1002/2017EF000687.