



ORIGINAL RESEARCH ARTICLE

Vintage Port prediction and climate change scenarios

Helder Fraga^{1*}, Nathalie Guimarães¹ and João A. Santos¹

¹ Centre for the Research and Technology of Agro-Environmental and Biological Sciences (CITAB), Institute for Innovation, Capacity Building, and Sustainability of Agri-Food Production (Inov4Agro), University of Trás-os-Montes e Alto Douro (UTAD), P.O. Box 1013, 5000-801 Vila Real, Portugal



*correspondence:
hfraga@utad.pt

Associate editor:
Benjamin Bois



Received:
25 July 2023

Accepted:
29 November 2023

Published:
7 December 2023



This article is published under the **Creative Commons licence (CC BY 4.0)**.

Use of all or part of the content of this article must mention the authors, the year of publication, the title, the name of the journal, the volume, the pages and the DOI in compliance with the information given above.

ABSTRACT

The Douro region is renowned for its quality wines, particularly for the famous Port Wine. Vintage years, declared approximately 2–3 times per decade, signify exceptional quality linked to optimum climatic conditions driving grape quality attributes. Climate change poses challenges, as rising temperatures and extreme weather events impact viticulture. This study uses machine learning algorithms to assess the climatic influence on vintage years and climate change impacts for the next decades. Historical vintage data were collected from 1850 to 2014. Monthly climatic data for the same period were obtained, including temperature, precipitation, humidity, solar radiation, and wind components. Various machine-learning algorithms were selected for classification, and a statistical analysis helped identify relevant climate variables for differentiation. Cross-validation was used for model training and evaluation, with the hits and misses (confusion matrix) as the performance metric. The best-performing model underwent hyperparameter tuning. Subsequently, future climate projections were acquired for four regional climate models from 2030 until 2099 under different socio-economic scenarios (IPCC SSP2, SSP3, and SSP5). Quantile mapping bias adjustment was applied to correct future climate data and reduce model biases. Past data revealed that vintages occurred 23.6 % of the years, with an average of two vintage years per decade, with a slightly positive trend. Climate variables such as precipitation in March, air temperatures in April and May, humidity in March and April, solar radiation in March, and meridional wind in June were identified as important factors influencing vintage year occurrence. Machine-learning models were employed to predict vintage years based on the climate variables, with the XGBClassifier achieving the highest performance with 76 %/88 % hits for the vintage/non-vintage classes, respectively, and an ROC score of 0.86, demonstrating strong predictive capabilities. Future climate change scenarios under different socio-economic pathways were assessed, and the results indicated a decrease in the occurrence vintage years until 2099 (10.3 % for SSP2, 9.1 % for SSP3, and 5.8 % for SSP5). The study provides valuable insights into the relationship between climate variables and wine vintage years, enabling winemakers to make informed decisions about vineyard management and grape cultivation. The predictions suggest that climate change may challenge the wine industry, emphasising the need for adaptation strategies.

KEYWORDS: Machine learning, statistical models, wine quality, Douro, Portugal, future projections

INTRODUCTION

The Douro region of Portugal is one of the world's oldest and most renowned winemaking regions (Fraga *et al.*, 2017). Located in north-eastern Portugal (Figure 1), the region is known for its steep slope and terraced vineyards along the Douro River valley (Brochado *et al.*, 2021), yielding exceptional table wines that are gaining recognition worldwide (Rebelo *et al.*, 2015). The Douro is also the birthplace of the Port Wine, a fortified wine known for its high quality, which is one of Portugal's most famous products, accounting for approximately 50 % of the total wines exported (IVV, 2021). Similar to other wines and regions, Port Wine's quality attributes can vary from one year to the other, largely driven by the prevailing climatic conditions during the grapevine growing season. In the Douro, a year with top-quality Port Wine production is traditionally referred to as a "vintage year". Each company usually declares vintage years, depending on the quality of the grapes and, consequently, of the wines produced. If all (or nearly all) companies declare it, this consensus translates into a generalised Port vintage year. According to the Port and Douro Wines Institute (IVDP), generalised vintage years are typically declared around 2 or 3 times per decade (on average) from the existing records dating back to 1759. Vintage Port bottles available in the market make up approximately 3 % of the total Port, typically with high prices (Macedo *et al.*, 2021; Panzone and Simões, 2009). Some of the most famous vintage years include 1945, 1963, 1977, 1994, and 2011.

As mentioned above, the term vintage year directly relates to the growing season in which the grapes were harvested (IVV, 2021; Mayson and Duff, 2018). Therefore, the quality

of a wine is directly related to the weather patterns during the grapevine's annual cycle, particularly the temperature, amount of precipitation, and solar radiation, amongst other atmospheric variables (Robinson *et al.*, 2013; Smart *et al.*, 1991). It is traditional knowledge that vintage years are typically declared in years with warm and dry growing seasons, as these conditions allow the grapes to ripen fully and develop complex flavours resulting in high-quality grapes, whereas cooler or wetter growing seasons may produce grapes that are less ripe and less flavourful, resulting in lower quality wines (Magalhães, 2008; Mayson and Duff, 2018). The Douro region presents typical Mediterranean climatic characteristics, with cool, wet winters followed by warm-to-hot dry summers. These summers are ideal for ripening grapes, conditions that are typically preferred for producing a high-quality vintage. These climatic features allow the grapes to ripen fully and develop complex flavours and aromas. For example, the 2011 vintage in the Douro Valley was considered exceptional, attributed to the hot and dry growing season, conferring complex flavours to the wines. Similarly, the Bordeaux vintages are characterised by warm and dry growing seasons, resulting in high-quality wines (Robinson *et al.*, 2013).

Despite the general linkage between warm and dry conditions and high-quality wines, it is expected that this relationship is not linear. The changing climate conditions associated with anthropogenic forcing can have profound implications for the characteristics and quality of wines. Rising temperatures, changing precipitation patterns, and altered growing season lengths can impact grapevine development and grape berry sugar accumulation, acidity, phenolic composition, and flavour profiles (Van Leeuwen and Schultz, 2018).

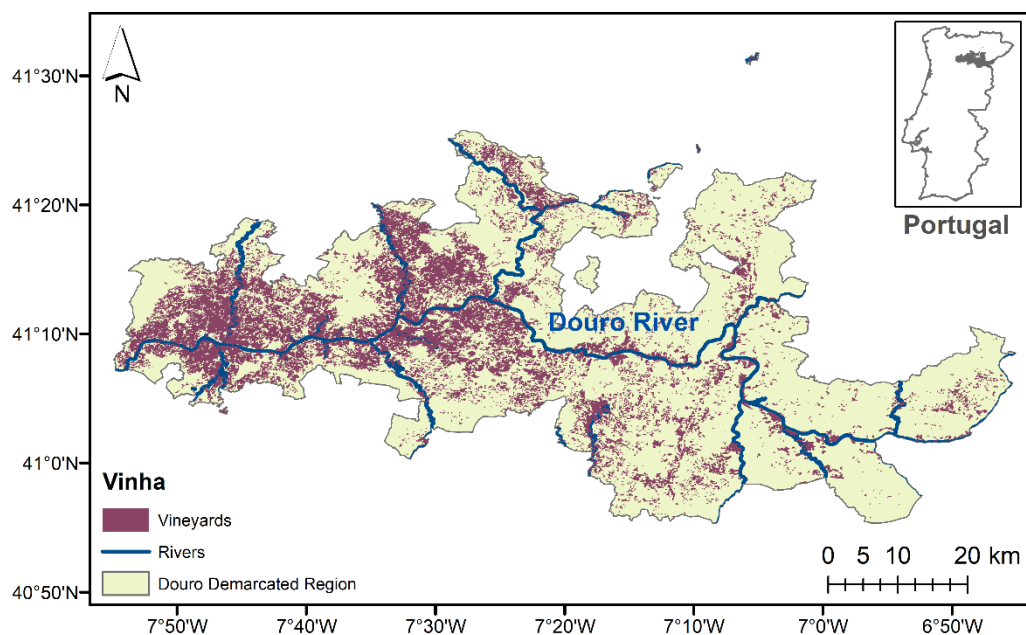


FIGURE 1. The geographical location of the Douro Demarcated Region in mainland Portugal, along with the vineyard land cover areas and main rivers.

The Douro winemaking region is located in a “climate change hotspot”, meaning that the impacts of climate change in this region may be particularly severe (Fraga *et al.*, 2020) in terms of viticultural productivity (Fraga *et al.*, 2022), particularly due to the increase in extreme weather events (Fonseca *et al.*, 2023). Currently, in the Douro, temperatures occasionally exceed 40 °C in the summer, with an intensification projected for the next decades, which may threaten quality wine production (Gambetta and Kurtural, 2021). As in many other regions worldwide, the region is also experiencing increasingly frequent extreme weather events, including heavy precipitation and hailstorms, late frosts, and heatwaves (C3S, 2023; Clemente *et al.*, 2022; Costa *et al.*, 2019; Fraga and Santos, 2017; Jones, 2012). These events have a vast destructive potential for viticulture, leading to losses in yield and quality attributes. As such, under climate change, “vintage years” may become infrequent and less predictable.

While significant research has been conducted on the impacts of climate change on wine production, several knowledge gaps still need to be addressed. While some studies attempt to assess the linkage between climatic conditions vintage/wine quality and (Biss and Ellis, 2021; Real *et al.*, 2017; Davis *et al.*, 2019; Salinger *et al.*, 2015), the knowledge of the occurrence of vintage years under climate change is still vastly unexplored. One of the main challenges is identifying the specific patterns that distinguish the binary vintage year time series (0—non-vintage, 1—vintage). Traditional statistical (linear) models have been widely used for time series prediction, but they often fail to capture the complex non-linear relationships in the data (Hastie *et al.*, 2013). In recent years, machine-learning models have emerged as a powerful alternative for time series prediction due to their ability to learn complex patterns (Hastie *et al.*, 2013). These models offer several advantages over traditional statistical models for predicting time series, such as high-dimensional data handling and automatic learning of hierarchical representations (Hastie *et al.*, 2013). Their ability to learn intricate patterns and relationships in data has led to their superior performance, making them particularly suitable to handle real-world problems, such as analysing vintage year drivers and classifying wine quality.

The current study aims to overcome this knowledge gap by modelling the occurrence of vintage years in the Douro wine region using machine learning and, subsequently, projecting the potential impacts of climate change based on the developed models. Therefore, the objectives of the present study are five-fold: 1) to examine a long time series of vintage years in the Douro Wine Region (1850–2014, 165 years); 2) to evaluate the existence of periodic behaviour in this series; 3) to train machine-learning models that can be used to explain the variability of the vintage time series; 4) to validate these models and analyse the uniqueness of a vintage year in climatic terms; and 5) to develop projections of vintage years taking into account several future scenarios and assess potential climate change impacts.

MATERIALS AND METHODS

1. Vintage data collection and pre-processing

In the present study, we collected vintage classification data from 1850 to 2014 (165 years) from the dataset available on the Portuguese IVDP (“Instituto dos Vinhos do Douro e Porto, I.P.”) at <https://www.ivdp.pt/pt/vinhos/vinhos-do-porto/vintages/> (while vintage data extends back to 1759, this period was selected for consistency, see section 2.5). The data contained information about the vintages of port wine, including the year of production, a short description of the meteorological conditions of each given year, some tasting notes, and the classification. In our approach, we opted to select only classical vintage years that were declared by the majority of producers for the subsequent analysis. We pre-processed the data by cleaning the text and converting it into structured binary data, where 1 corresponds to a vintage year and 0 corresponds to a non-vintage year. Discrete Fourier Transform analysis was then applied to this binary dataset to isolate possible periodicities in the data. An autocorrelation analysis was also performed to investigate the patterns and relationships within the vintage year time series.

2. Climatic data

As potential predictors of vintage wine quality, we used climatic data retrieved from the 20th Century Reanalysis dataset v3 (henceforth 20CR) from the National Oceanic and Atmospheric Administration (NOAA) over the period from 1850 to 2014 (while 20CR data began in 1836, this period was selected for consistency, see section 2.5). Reanalysis is a methodology that uses state-of-the-art earth models, archived atmospheric analysis, and updated data assimilation techniques to reconstruct past weather and climate conditions. The output of climate reanalysis is a dataset that provides information about atmospheric variables over a specified period, which can then be used to study past weather and climatic trends, variability, and extremes and calibrate and validate climate models for future projections. Climate reanalysis datasets are widely used in a variety of fields, including climate science and meteorology (Saha *et al.*, 2010), environmental science (Fuka *et al.*, 2014), and agriculture (Uniyal *et al.*, 2019). They are particularly useful for studying long-term climate variability, as they provide a consistent and comprehensive record of past climate conditions that can be used to identify trends and patterns over time (Bengtsson *et al.*, 2004). The 20CR gridded dataset contains data from the model NCEP GFS v14.0.1 (Compo *et al.*, 2011) at a resolution of ~75 km at the equator. Data over the Douro Region (40.75°–41.25° N; 6.0°–8.0° W; Figure 1) were selected and spatially averaged (corresponds to 2 grid boxes). The following 20CR variables were retrieved over the Douro region and used as features in the machine-learning models: monthly minimum, mean and maximum temperatures (TN, TM and TX, °C), precipitation (PR, mm), relative humidity (RH, %), solar radiation (RD, W/m²), zonal and meridional wind components (UW and VW, ms⁻¹), from January to September (months: 01, 02, ..., 09), during grapevine annual growing cycle.

The diurnal temperature range (DTR) was also computed for each month (TX minus TN).

3. Machine-learning models

We selected several classical machine-learning classification algorithms included in the scikit-learn v1.3.2 library (Fabian, 2011) in Python v3.11, comprising Logistic Regression (LogisticRegression; (Cox, 1958)), Decision Trees (DecisionTreeClassifier; (Fisher, 1936)), Random Forest (RandomForestClassifier; (Breiman, 2001)), Support Vector Machines (SVC; (Cortes and Vapnik, 1995)), K-Nearest Neighbor (KNeighborsClassifier; (Fix and Hodges, 1989)), Gaussian Naive Bayes (GaussianNB; (Bayes, 1958)), Multi-layer Perceptron (MLPClassifier; (Rumelhart *et al.*, 1987)), AdaBoost (AdaBoostClassifier; (Freund and Schapire, 1997)), Gradient Boosting (GradientBoostingClassifier; (Friedman, 2001)), Quadratic Discriminant Analysis (QuadraticDiscriminantAnalysis; (Fisher, 1936)), and XGBoost (XGBClassifier; (Chen *et al.*, 2016)). These algorithms were chosen based on previous research outcomes, documented in the specialised literature,

and their suitability for the nature of our data. A comparison between the strengths and weaknesses of each model can be found in Table 1.

4. Model Training

From the historical climate dataset, a set of 64 features are potential candidates for model selection (8 climatic variables \times 8 months). A statistical analysis was conducted to identify the most important climate variables that differentiate vintage years from non-vintage years. The Kolmogorov–Smirnov test was applied to compare the distribution of each climate variable between the two groups (Massey, 1951). Variables that showed statistical significance at a 5 % significance level ($p < 0.05$) were considered features with discrimination power. Subsequently, all models were run using only this sub-group of features. This preliminary approach is important, taking into account the computational resources needed for the modelling. To assess all possible combinations of the resulting sub-group of features, the bestFeatures v1.0 python package was used (Fraga, 2023). To evaluate the performance of each algorithm, this script uses a cross-validation scheme

TABLE 1. Each model/algorithm used in the present study along with the strengths and weaknesses.

Model	Nature	Strengths	Weaknesses
Logistic Regression	Linear	Simple and interpretable; Efficient for linear relationships	Limited to linear relationships; May not perform well with complex data patterns
Decision Trees	Non-Linear	Intuitive and easy to understand; Handles non-linearity well	Prone to overfitting; Sensitive to small variations in data
Random Forest	Non-Linear	Reduces overfitting through ensemble; Handles non-linearity well	Lack of interpretability; Computationally expensive for large datasets
Support Vector Machines	Non-Linear	Effective in high-dimensional spaces; Versatile kernel functions	Can be sensitive to choice of kernel parameters; Memory-intensive for large datasets
K-Nearest Neighbour	Non-Linear	Simple and easy to implement; Non-parametric and adaptable	Computationally expensive for large datasets; Sensitivity to irrelevant features
Gaussian Naive Bayes	Non-Linear	Simple and computationally efficient; Works well with high-dimensional data	Assumes independence between features; May not handle complex relationships well
Multi-layer Perceptron	Non-Linear	Suitable for complex relationships; Can learn hierarchical features	Prone to overfitting; Requires careful tuning of hyperparameters
AdaBoost	Non-Linear	Combines weak learners for improved accuracy; Robust to overfitting	Sensitive to noisy data; Can be computationally expensive
Gradient Boosting	Non-Linear	Sequential improvement over weak learners; Handles complex relationships	Prone to overfitting; Sensitive to hyperparameter tuning
Quadratic Discriminant Analysis	Non-Linear	Effective for non-linear relationships; Can handle multivariate normal distributions	Assumes normality and equal covariance for classes; May be sensitive to outliers
XGBoost	Non-Linear	Scalable and efficient implementation; Improved regularisation for better performance	Requires careful parameter tuning; Can be computationally expensive

when running each model for each combination of features. Specifically, a k -fold cross-validation technique with 5 folds means each algorithm is effectively evaluated multiple times with the split dataset. The k -fold method entails partitioning a given dataset into k distinct subsets, with one of these subsets designated as the validation set and the remaining $k-1$ subsets serving as the training data (Stone, 1974). This process is then repeated k times, with each subset being different. As such, there are always parts of the data that are unseen/withheld from the algorithm. The outcomes are then averaged or other prescribed methods to derive a singular estimate of the model's performance. As such, this methodology provides several benefits over other ways of statically train-test splitting the dataset. k -fold cross-validation plays a critical role in addressing the issue of overfitting, ensuring a robust evaluation of the model's capacity to generalise to previously unseen data.

During training, we used the default hyperparameters for each algorithm, which can be found in the scikit-learn library (Fabian, 2011). Cross-validation helped us to avoid model overfitting and obtain a more accurate estimate of the model performance. A confusion matrix (hits vs misses) was used to test the performance of each algorithm. This metric is commonly used for evaluating the performance of classification models. Furthermore, confusion matrixes depicting hits and misses for each model were also analysed. The best-performing model was selected based on the overall performance of the metrics. Hyperparameter Tuning was subsequently applied to the best-performing model, as determined by the evaluation metrics. Grid search techniques were employed to find the optimal combination of hyperparameters that maximised the model performance (Murphy, 2012). A comparison between the strengths and weaknesses of each model can be found in Table 1.

5. Future projections

The future climate data used in this research was obtained from the Copernicus Climate Change Service (C3S) platform (Service, 2023), which provides access to a comprehensive range of climate data and information. The climate data were sourced from four widely recognised climate models: CNRM-CM6-1, CNRM-CM6-1-HR, CNRM-ESM2-1, and IPSL-CM6A-LR (Boucher *et al.*, 2020; S  ferian *et al.*, 2019; Voldoire *et al.*, 2019). These models are part of the Earth System Model ensemble provided by the Copernicus Climate Data Store (CDS) (Service, 2023). The Copernicus platform is a collaborative initiative of the European Union and the European Space Agency, offering free and open access to climate data and information for research and decision-making purposes. All climatic variables were obtained from each model for the periods 1850–2014 and 2030–2099 (the former period was obtained for bias-correction reasons, which is explained below). Similarly to the historical climate data, the data from each model were selected and spatially averaged over the Douro Region (40.75°–41.25° N; 6.0°–8.0° W; Figure 1) (spatial resolution differs for each model, resulting in a different number of gridboxes extracted).

Furthermore, the future climate data were obtained under three different future scenarios. The Shared Socioeconomic Pathway (SSP) was developed by the Intergovernmental Panel on Climate Change (IPCC), namely SSP2, SSP3, and SSP5. These scenarios represent different plausible socio-economic futures and their associated greenhouse gas emissions trajectories (IPCC, 2021; Voldoire *et al.*, 2019). SSP2 is a middle-of-the-road scenario that assumes moderate population growth, intermediate levels of technological progress, and a balanced approach to economic and environmental goals (Riahi *et al.*, 2017). SSP3 portrays a world where regional competition takes precedence over global cooperation (Gidden *et al.*, 2019). It assumes high population growth, slow technological progress, and fragmented efforts to address environmental challenges. SSP5 is considered a severe scenario, which projects rapid economic growth, high population, and heavy reliance on fossil fuels (Kriegler *et al.*, 2017). Environmental concerns are typically disregarded in favour of economic development.

It is important to acknowledge that future climate projections are subject to uncertainties inherent to climate modelling, which typically result in bias from the observed climate (IPCC, 2021). As such, a quantile mapping bias adjustment method was applied to correct future climate data (Thrasher *et al.*, 2012). The quantile mapping bias correction is a widely used statistical method for adjusting biases in climate model projections, including future climate data. It aims to align the empirical distributions and corresponding statistical moments of the model-simulated data with observed data, thereby reducing systematic errors and improving the reliability of climate projections. This bias correction technique involves comparing the cumulative distribution functions of the model-simulated data with those of observed data, as this represents the probability distribution of a variable and provides information about its relative frequency of occurrence at different values. To apply quantile mapping, the cumulative distribution functions of the model-simulated data and observed data are first calculated for the same period (1850–2014). The bias correction is then performed by adjusting the model-simulated data to match the observed CDF and then applying the same principle to the modelled future data. This method is currently widespread in climate research studies (Lafon *et al.*, 2012; Martins *et al.*, 2021).

RESULTS

1. Historical Vintage Year Analysis

The analysis of the historical dataset revealed an occurrence of a vintage in approximately 23.6 % of the years from 1850 to 2014 (Figure 2). Analysing the occurrence of vintage years per decade (Figure 3), it is clear that there are, on average, two vintage years per decade. The 1920s and 2000s decades had the highest number of vintages (4), while the 1860s, 1880s, and 1950s had no vintage years. A slight positive trend is apparent (+0.01 vintages per decade), indicating a higher occurrence of vintages in the most recent decades, likely being a manifestation of the warming and drying trends in

Vintage years

1850	1851	1852	1853	1854	1855	1856	1857	1858	1859
1860	1861	1862	1863	1864	1865	1866	1867	1868	1869
1870	1871	1872	1873	1874	1875	1876	1877	1878	1879
1880	1881	1882	1883	1884	1885	1886	1887	1888	1889
1890	1891	1892	1893	1894	1895	1896	1897	1898	1899
1900	1901	1902	1903	1904	1905	1906	1907	1908	1909
1910	1911	1912	1913	1914	1915	1916	1917	1918	1919
1920	1921	1922	1923	1924	1925	1926	1927	1928	1929
1930	1931	1932	1933	1934	1935	1936	1937	1938	1939
1940	1941	1942	1943	1944	1945	1946	1947	1948	1949
1950	1951	1952	1953	1954	1955	1956	1957	1958	1959
1960	1961	1962	1963	1964	1965	1966	1967	1968	1969
1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
2010	2011	2012	2013	2014	2015	2016	2017	2018	2019

FIGURE 2. Occurrence of each vintage year between 1850 and 2019.

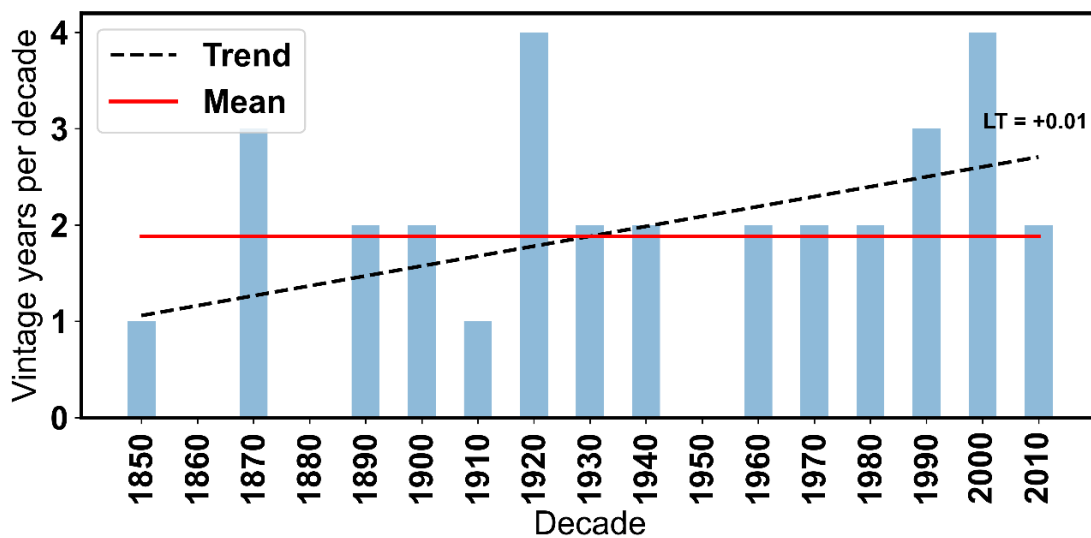


FIGURE 3. Sum of vintage years for each decade from 1850 and 2014. The mean number of vintages for the full period is also shown (red line), along with the linear regression trend line (LT, linear trend).

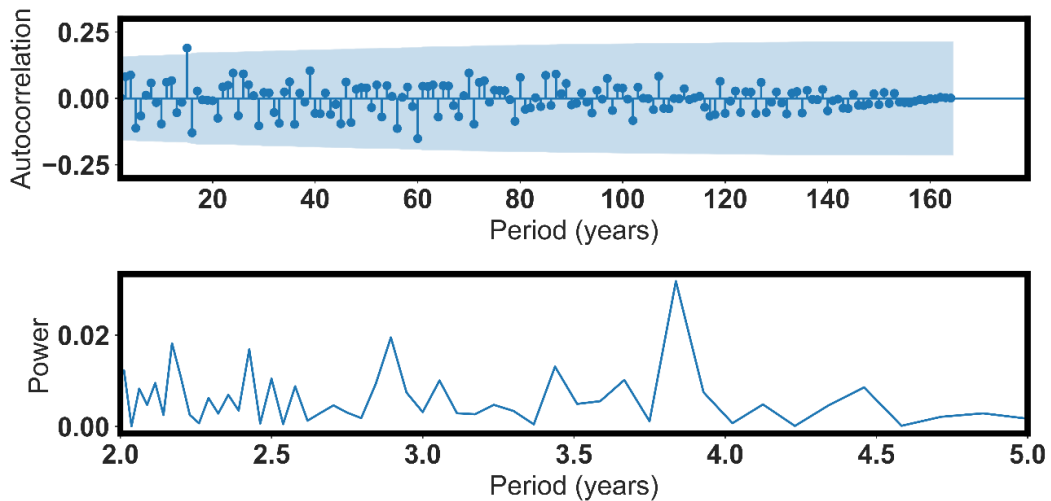


FIGURE 4. (Top panel) Autocorrelation between the lagged timeseries 1850–2014. (Bottom panel) Power spectral density of the occurrence of a vintage year.

the Douro Wine Region, but also may be due to viticultural and oenological advances since 1990. Regarding the spectral analysis (Figure 4 top), it reveals a leading four-year cycle, i.e., vintage years tend to occur every four years, which is in line with the two vintages per decade. However, the spectral power density is significantly spread over other frequencies, thus also highlighting that vintage years are not a cyclic occurrence. Regarding the autocorrelation (Figure 4 bottom), no clear periodic behaviour is found, which also stresses the irregularity in the occurrence of vintage years.

2. Climate influence

The Kolmogorov–Smirnov test was performed to identify the climate variables that significantly differentiate vintage years from non-vintage years. The results indicated that several variables showed statistical significance ($p < 0.05$). These important climate variables include March precipitation (PR03), May mean temperature (TM05), May minimum temperature (TN05), April maximum temperature (TX04), March and April relative humidity (RH03, RH04), March

solar radiation (RD03), and June meridional (north-south) wind component (VW06) (Table 2).

Figure 5 shows the distribution (left panels), as well as the annual time series (right panels) (1850–2014) of the vintage vs non-vintage years for the significantly different features abovementioned. From the empirical distributions associated with vintage/non-vintage years, it can be concluded that settled and relatively dry weather conditions in early spring tend to be favourable to the occurrence of vintage years. Anomalously low precipitation in March (PR03), anomalously low values of relative humidity in March and April (RH03, RH04), anomalously high values in solar radiation (RD03), and anomalously high values of April maximum temperature (TX04) are coherently associated with the occurrence of vintage years. Conversely, the prevalence of moist air masses driving cloudy, humid, and rainy conditions tends to be unfavourable to wine quality. In May, when grapevine flowering typically develops, mean (TM05) and minimum (TN05) temperatures above average are generally favourable. For June, the meridional (north-south) wind component (VW06) reveals that vintage

TABLE 2. Climatic variables with statistically significant empirical probability distributions between vintage and non-vintage years, using the Kolmogorov–Smirnov test (p -value < 0.05).

Climatic variable	Abbreviation	p -value
March precipitation	PR03	0.019
May mean temperature	TM05	0.004
May minimum temperature	TN05	0.017
April maximum temperature	TX04	0.011
March relative humidity	RH03	0.008
April relative humidity	RH04	0.003
March solar radiation	RD03	0.003
June meridional (north-south) wind component	VW06	0.006

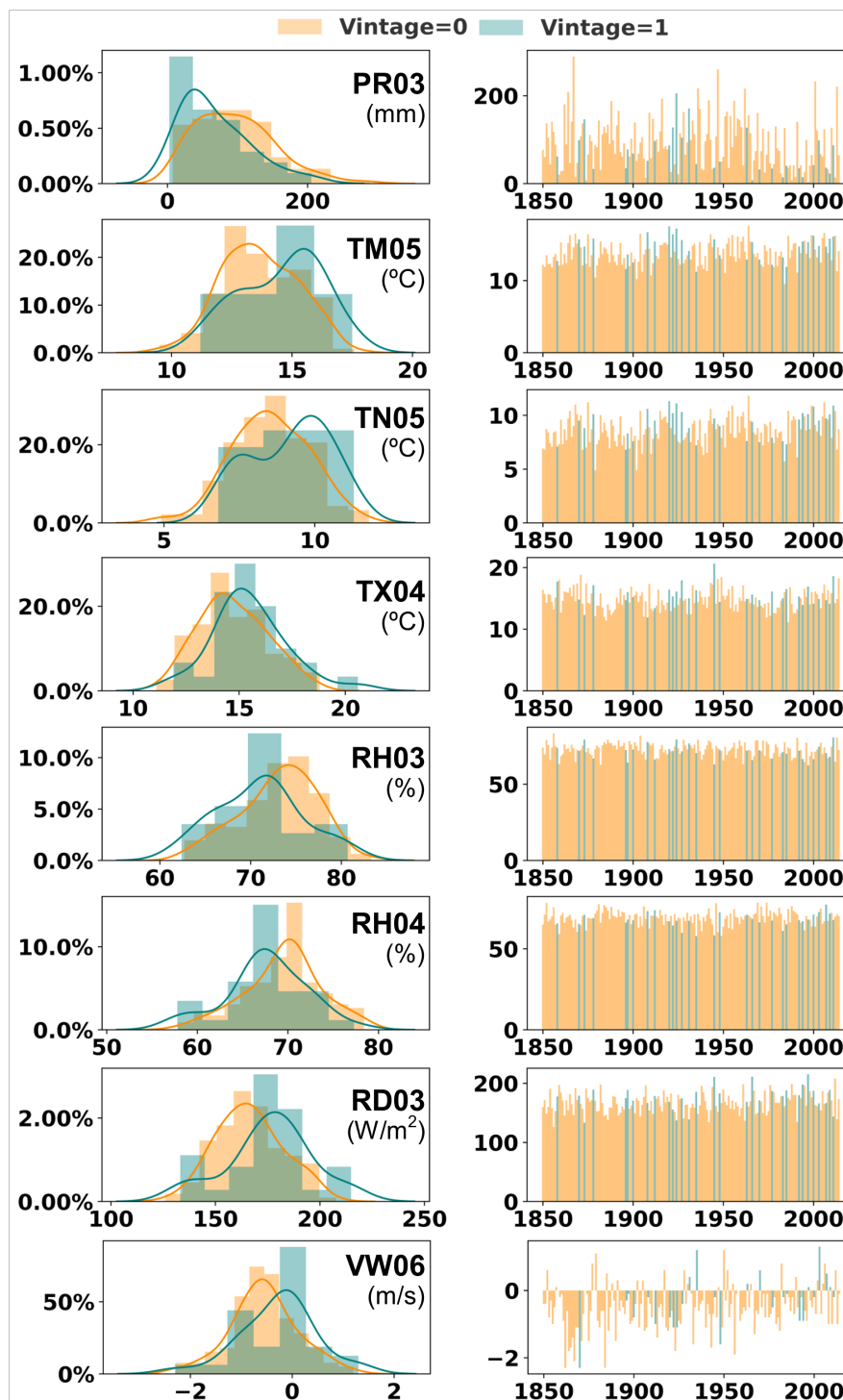


FIGURE 5. Distribution plots for the climatic variables with statistically significant differences (p -value < 0.05) in their empirical probability distributions between vintage and non-vintage years (Table 2). The plots include: March precipitation (PR03), May mean temperature (TM05), May minimum temperature (TN05), April maximum temperature (TX04), March and April relative humidity (RH03, RH04), March solar radiation (RD03), and June meridional (north–south) wind component (VW06).

years are commonly linked to weak winds, which also agrees with the occurrence of settled weather conditions. Northerly winds prevail in Portugal during summer, but strong northerly winds may bring cool and relatively moist air masses from the North Atlantic, thus not being the optimal conditions for grape berry ripening and synthesis of complex compounds.

3. Machine-Learning Model Performance Comparison

Various machine-learning classification models were trained and evaluated using the historical vintage year dataset to predict vintage years. As previously explained, model

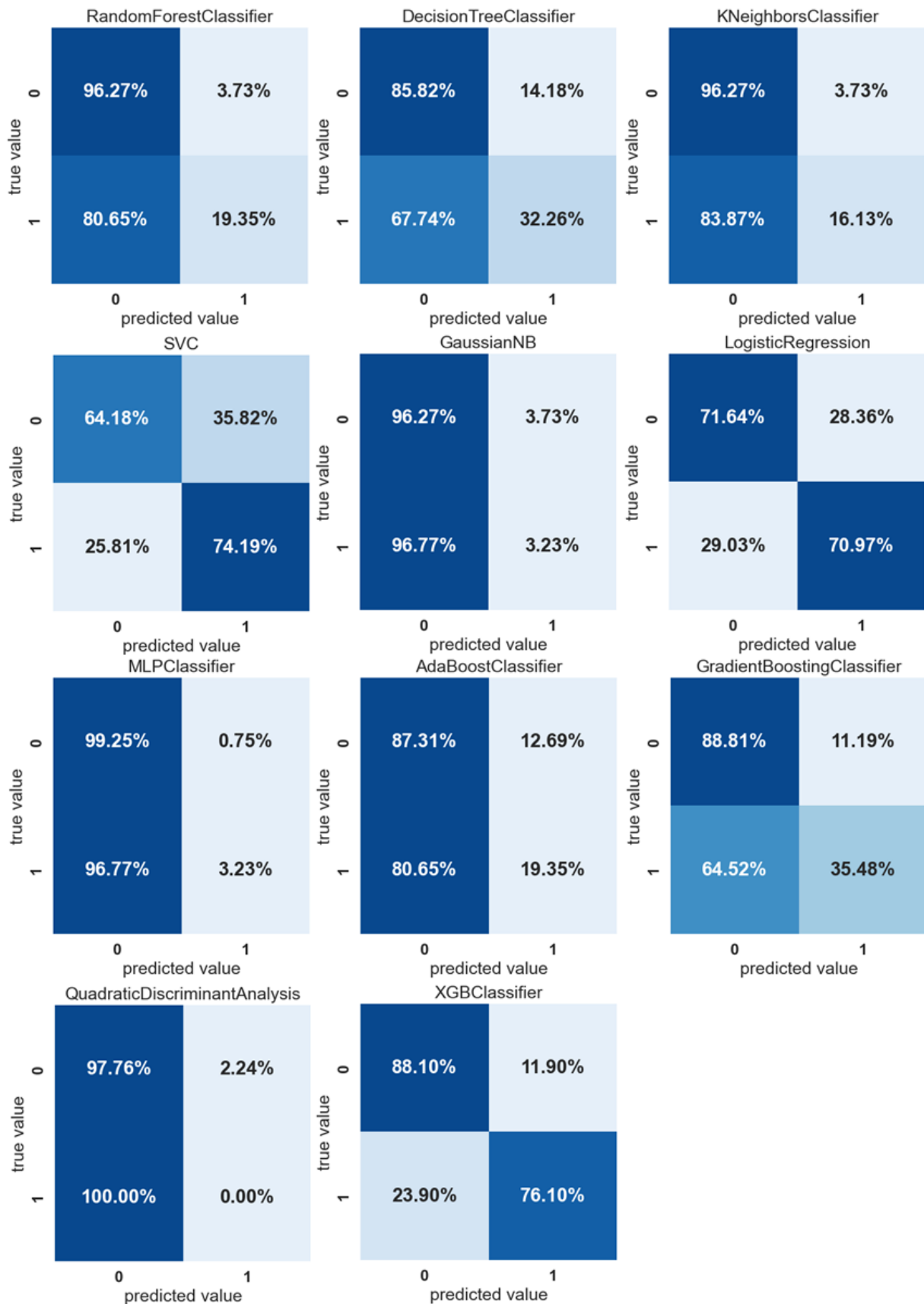


FIGURE 6. Confusion matrix for each model. The percentage of each plot considers all folds in the k-fold validation.

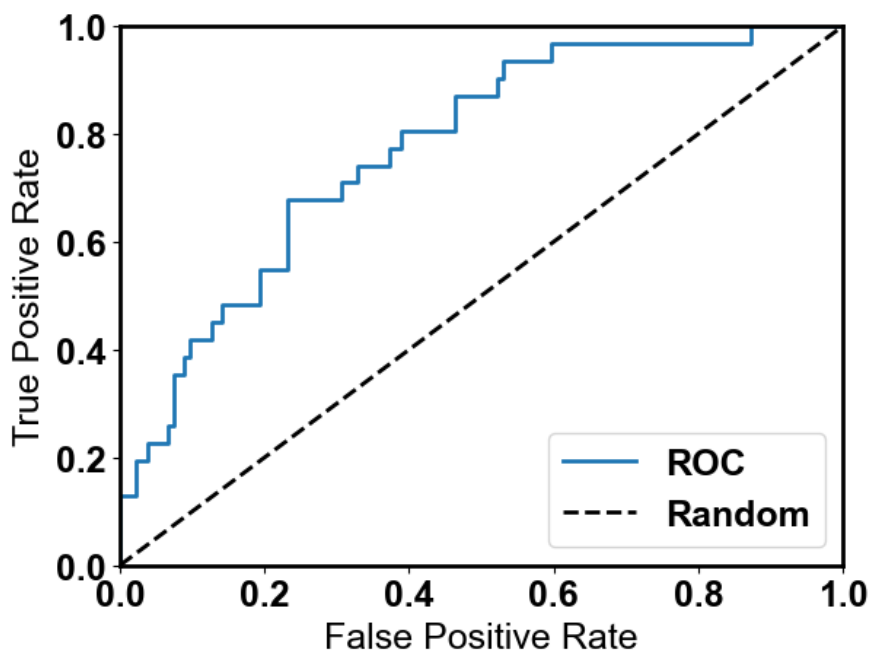


FIGURE 7. ROC AUC curve (true versus false positive rate) for the XGBClassifier model.

evaluation was performed using a 5-fold cross-validation, while the performance of each model was assessed by analysing the respective confusion matrices, representing the hits and misses. Each matrix details the hits and misses for each vintage class, taking into account the k -fold methodology (average percentage for the results of each fold). The results demonstrated that the machine-learning models achieved varying performance levels in predicting vintage years (Figure 6). The XGBClassifier exhibited the highest performance among the models, with 76 % hits for the vintage class and 88 % hits for the non-vintage class. SVC and LogisticRegression also performed adequately, with 74 %/64 % and 71 %/72 %, respectively. All other models performed poorly, not adequately simulating the vintage class. For example, QuadraticDiscriminantAnalysis placed all years as non-vintage, resulting in 100 %/0 % scores. Summarising, the XGBClassifier shows higher performance, predicting both class 0 and 1 (non-vintage and vintage, respectively, with LogisticRegression and SVC also showing relatively good performances. All other models show problems in assessing class 1, which indicates lower performances. Given these results, hyperparameter tuning was applied to the XGBClassifier to optimise its performance further. The hyperparameters were then tuned using the grid search technique (Murphy, 2012). The importance of the XGBClassifier feature indicated that the model selects TM05, RH04, and VW06 as the most important predictors (not shown).

A ROC curve was then created by varying the classification threshold (0 to 1) of the XGBClassifier model and calculating TPR (true-positive rate) and FPR (false-positive rate) at each threshold. This curve represents the ability to distinguish between two classes by displaying the trade-off between

sensitivity and specificity at various decision thresholds, with a better model producing a curve closer to the top-left corner and having a higher Area Under the Curve (AUC) value. The ideal classifier would have a TPR of 1 and an FPR of 0, while a random classifier would have a score of 0.5. XGBClassifier achieved an ROC score of 0.86. From Figure 7, the curve starts at the bottom left corner of the graph and moves steadily upwards and to the right, indicating that as the FPR increases, the TPR also increases. This is what we would expect to see for a classifier that is good at distinguishing between vintage and non-vintage years.

4. Predictions for Future Climate Change Scenarios

Considering the most important climatic features for the XGBClassifier algorithm (TM05, RH04, and VW06), these were analysed regarding their future anomalies (differences between future and present). Three IPCC Shared Socioeconomic Pathway (SSP) scenarios: SSP2, SSP3, and SSP5, were analysed, and changes were assessed for each scenario. Figure 8 shows that the climatic conditions will change, with stronger modifications in the most severe future scenarios (SSP5). TM05 points to a warming trend from 3.4 °C (SSP2) to 4.9 °C (SSP5). RH4 will decrease in all future scenarios, ranging from -7 to -8 %. VW06 suggests a strong northerly wind influence in the future (except in SSP2, where changes are small).

The tuned XGBClassifier model was then applied to future climate change projections under these scenarios. These projections aimed to assess the potential impact of climate change on vintage year occurrence from 2030 to 2099. The results indicated a decrease in the occurrence of vintage years under all three climate change scenarios

compared to the historical period (Figure 9). Considering the Ensemble mean of the 4 climate models, the vintage year occurrence rates were estimated to be 10.3 % for SSP2, 9.1 % for SSP3, and 5.8 % for SSP5. The CNRM-CM6-1-HR model generally presents the highest percentages

of vintage years per decade, while the IPSL-CM6A-LR shows the lowest. These outputs show the importance of using future climate data sources to consider the uncertainty tied to models. Nonetheless, all models point to a decrease in vintage occurrence in relation to the historical data.

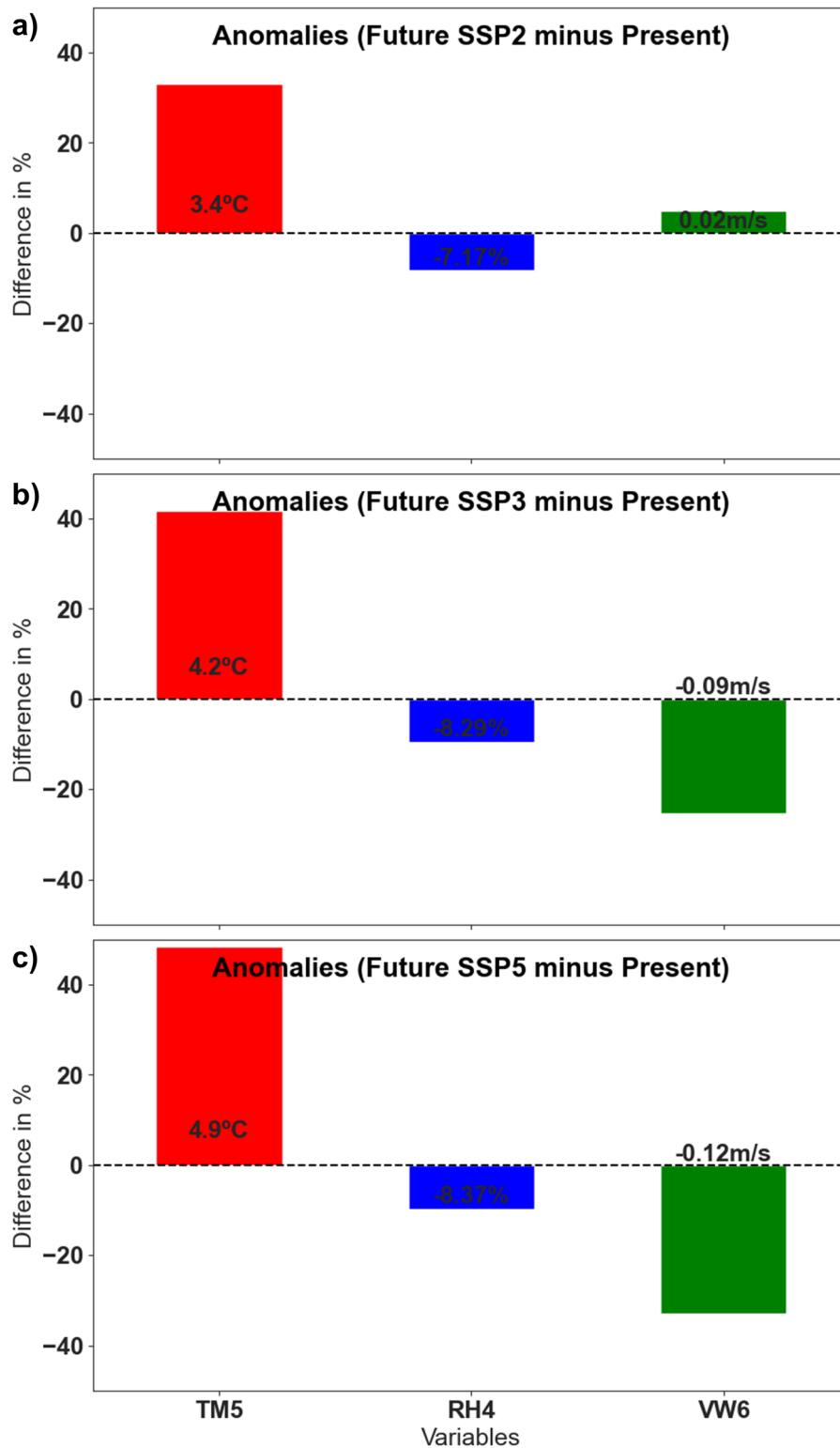


FIGURE 8. Differences between future (a) SSP2, (b) SSP3, (c) SSP5 and present (in % relative to the present) for each variable selected by XGBclassifier model. Differences in original units are also shown.

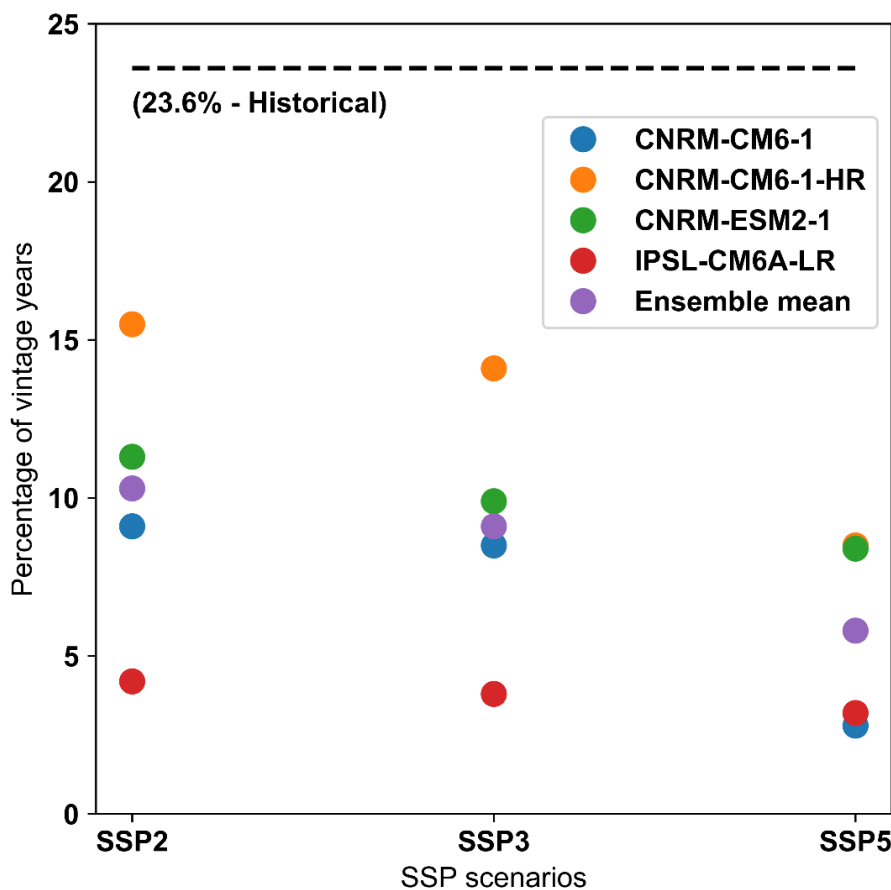


FIGURE 9. Projections for the percentage of vintage years in the future period (2030–2100), for the three selected SSP scenarios (SSP2, SSP3 and SSP5) and the four outlined climate models (see legend).

These findings suggest that climate change is expected to reduce the number of vintage years, posing challenges to the wine industry and emphasising the need for adaptation strategies. Overall, the results demonstrate the effectiveness of machine-learning models in predicting vintage years based on climate variables. The tuned XGBClassifier model exhibited the highest performance and captured the relationships between climate variables and vintage year occurrence. The predictions for future climate change scenarios highlight the potential impacts of climate change on wine vintage patterns.

DISCUSSION

The results of this study provide valuable insights into the relationship between climate variables and wine vintage years. The identification of important features (potential predictors), such as mean temperature in May (TM05), relative humidity in April (RH04), and meridional wind component in June (VW06), highlights the key climate factors that contribute to vintage year occurrence. These findings align with previous research that has emphasised the significance of these variables in determining wine quality and vintage suitability (Biss and Ellis, 2021;

Real *et al.*, 2017; Davis *et al.*, 2019). Jones (2005) conducted a study in the Douro Valley, Portugal, and identified mean temperature in May as a crucial factor influencing vintage quality. Similarly, Tonietto and Carbonneau (2004) emphasised the importance of temperature and precipitation during the growing season in determining wine quality and vintage characteristics. Real *et al.* (2017) identified that the growing season mean temperatures (April–September) above the region’s average, warm winters, cool July through veraison, and cool temperatures during ripening are important factors for vintage quality. Nonetheless, the previous study was based on a 30-year time window (1980–2009), and longer periods should be analysed to understand and confirm the relationships between vintage quality and climatic parameters. Interestingly, Davis *et al.* (2019) indicate that the most important climatic factor in distinguishing high-quality Burgundy vintages is the growing season temperature, especially the high diurnal temperature range (for red wines) and high average maximum temperatures (for white wines). Biss and Ellis (2021) modelled the Chablis vintage score using the growing season mean temperature, minimum temperature, and rainfall during the ripening period. Our results enhance the understanding of the relationship and connection between vintage quality and climate factors, using a much longer time series than the mentioned studies.

Three machine-learning models employed in this study, SVC, LogisticRegression, and particularly the XGBClassifier, demonstrated strong predictive capabilities, detecting both classes (vintage and non-vintage). This high performance indicates the effectiveness of the selected climate variables in distinguishing between vintage and non-vintage years. Similar studies using modelling techniques have reported promising results in predicting wine quality (Biss and Ellis, 2021; Cortez *et al.*, 2009). Cortez *et al.* (2009) employed machine-learning algorithms to predict wine quality parameters. Their study highlighted the effectiveness of machine-learning models in assessing wine quality.

The findings of this research may have significant implications for the wine industry and vineyard management. Understanding the influence of climate variables on vintage years enables winemakers to make informed decisions regarding grape cultivation, harvesting, and winery management. The predicted decrease in vintage year occurrence under future climate change scenarios raises concerns about the potential challenges wine producers face. The prospect of higher temperatures in May, lower humidity in April, and stronger northerly winds in June present a complex set of challenges for viticulture in the future. Traditionally, grapevines have thrived within specific atmospheric conditions, particularly optimum temperatures, ensuring ideal ripening. However, as temperatures increase, the pace of ripening accelerates, potentially pushing grapes ahead of their optimum conditions for flavour and balance (White *et al.*, 2006). Elevated May temperatures may also pose risks of sunburn and heat stress for the vines and grape berries if not carefully managed. Lower humidity in April may increase water stress, affecting vine health and grape development. Furthermore, the combination of lower humidity and stronger northerly winds in June may contribute to increased evaporation rates and water loss from the soil, exacerbating the risk of drought conditions.

These shifts may also lead to imbalances in sugar content, acidity, and the development of flavour compounds, which are the hallmarks of high-quality wines (White *et al.*, 2006). Our study indicates that a tipping point may be close or already passed, manifesting in terms of fewer vintage years in the future. These results align with Jones *et al.* (2005), which projected that in regions currently producing high-quality wines, such as the Douro, climate change impacts may point toward unbalanced ripening of grapes, resulting in an overall quality loss and difficulty maintaining wine styles. Hence, by considering the identified important climatic factors, winemakers can adapt their practices to optimise wine quality in specific vintages. Adaptation strategies, such as implementing new viticultural techniques, adjusting grape varieties, and exploring alternative wine regions, may be necessary to maintain wine quality and sustainability under changing climatic conditions (Hannah *et al.*, 2013; van Leeuwen and Darriet, 2016; van Leeuwen *et al.*, 2019). For example, Bramley (2005) suggested that grape quality sensing technology should be invested in. These findings support the growing consensus among researchers and

industry professionals that climate change poses significant risks to wine production (Fraga *et al.*, 2016; Jones and Alves, 2012).

The present study builds upon these earlier works by modelling wine vintage years over a broader temporal scale and incorporating future climate change scenarios. By using a wider range of machine-learning models and assessing the impact of different climate scenarios, this research provides a comprehensive understanding of vintage year prediction and its implications for wine production under changing climatic conditions. Despite the valuable insights gained from this study, some limitations should be acknowledged. Firstly, it is important to recognise the potential limitations of using reanalysis data for climate analysis, as discussed by Thorne and Vose (2010). While reanalysis data are not the ideal choice for trend analysis, our research primarily focuses on characterising the climate disparities between vintage and non-vintage years, minimising the impact of this issue on our research. Additionally, the analysis focused primarily on monthly climate variables, overlooking other possible important factors, such as soil characteristics, viticultural practices, and vineyard management techniques. Future research may consider incorporating these additional variables to enhance the accuracy and robustness of the predictive models. Moreover, this study specifically focuses on the Douro wine region and its historical and future climate data. The extrapolation of the findings to other wine-growing regions may be hampered by variations in climate, grape varieties, and viticultural practices. Future research may explore the applicability of the developed models and their performance across different wine regions on a global scale. Additionally, the analysis considered a limited number of climate models for future climate projections. Expanding the scope to include a broader range of climate models and scenarios would provide a more comprehensive assessment of the potential impacts of climate change on wine vintage occurrence, thus incorporating a wider spectrum of uncertainties.

Despite the above-stated limitations, this study contributes to the existing body of knowledge by advancing our understanding of the relationship between climate variables and wine vintage years. The findings also underscore the importance of proactive measures to adapt to changing climatic conditions, as climatic variables play a key role in wine quality. These measures should envision the maintenance of the long-term socioeconomic sustainability of the wine sector, promoting the production of high-quality wines, but always allied with environmental protection and sustainable use of natural resources, such as soils, water, and biodiversity.

CONCLUSIONS

This research successfully modelled wine vintage years using monthly climate variables as predictors. The machine-learning models, particularly the XGBClassifier, demonstrated strong predictive performance, enabling a

clear differentiation between vintage and non-vintage years. The identified driving climate variables, namely mean temperature in May (TM05), relative humidity in April (RH04), and northerly wind in June (VW06), significantly contribute to the vintage year occurrence. The findings have practical implications for winemakers and vineyard managers, aiding in decision-making related to grape cultivation, harvesting, and wine production. Furthermore, the predictions for future climate change scenarios suggest a decrease in vintage year occurrence, highlighting the need for adaptation strategies to mitigate the potential impacts of climate change on wine production. This research expands upon previous studies by incorporating machine-learning techniques, assessing a wider range of climate models and scenarios, and modelling vintage years over a longer historical period. Despite its limitations, the study improves our understanding of wine vintage prediction potential and its implications for the wine industry under changing climatic conditions.

ACKNOWLEDGEMENTS

This work was financed by the CoaClimateRisk “O impacto das alterações climáticas e medidas de adaptação para as principais culturas agrícolas na região do Vale do Côa” project (COA/CAC/0030/2019) financed by National Funds by the Portuguese Foundation for Science and Technology (FCT). We thank the project UIDB/04033/2020, LA/P/0126/2020 and 2022.04553.PTDC. NG thanks the financial support provided by national funds through FCT – Portuguese Foundation for Science and Technology (UI/BD/150727/2020), under the Doctoral Programme “Agricultural Production Chains – from fork to farm” (PD/00122/2012) and from the European Social Funds and the Regional Operational Programme Norte 2020. HF thanks the FCT for 2022.02317.CEECIND.

REFERENCES

- Bayes, T. (1958). An Essay Towards Solving A Problem In The Doctrine Of Chances. *Biometrika*, 45(3-4): 296-315. <https://doi.org/10.1093/biomet/45.3-4.296>
- Bengtsson, L., Hagemann, S., & Hodges, K.I. (2004). Can climate trends be calculated from reanalysis data? *Journal of Geophysical Research-Atmospheres*, 109(D11). <https://doi.org/10.1029/2004JD004536>
- Biss, A., & Ellis, R. (2021). Modelling Chablis vintage quality in response to inter-annual variation in weather. *OENO One*, 55(3): 209-228. <https://doi.org/10.20870/oeno-one.2021.55.3.4709>
- Boucher, O., Servonnat, J., Albright, A. L., Aumont, O. [...] (2020). Presentation and Evaluation of the IPSL-CM6A-LR Climate Model. *Journal of Advances in Modeling Earth Systems*, 12(7): e2019MS002010.
- Bramley, R.G.V. (2005). Understanding variability in winegrape production systems - 2. Within vineyard variation in quality over several vintages. *Australian Journal of Grape and Wine Research*, 11(1): 33-42. <https://doi.org/10.1111/j.1755-0238.2005.tb00277.x>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1): 5-32. <https://doi.org/10.1023/A:1010933404324>
- Brochado, A., Stoleriu, O., & Lupu, C. (2021). Wine tourism: a multisensory experience. *Current Issues in Tourism*, 24(5): 597-615. <https://doi.org/10.1080/13683500.2019.1649373>
- C3S, 2023. Copernicus Climate Change Service (C3S) (2023): European State of the Climate 2022, Summary.
- Real, A. C., Borges, J., Cabral, J.S., & Jones, G.V. (2017). A climatology of Vintage Port quality. *International Journal of Climatology*, 37(10): 3798-3809. <https://doi.org/10.1002/joc.4953>
- Chen, T.Q., Guestrin, C., & Assoc Comp, M. (2016). XGBoost: A Scalable Tree Boosting System, 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), San Francisco, CA, pp. 785-794. <https://doi.org/10.1145/2939672.2939785>
- 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. *Agronomy*, 12(6). <https://doi.org/10.3390/agronomy12061404>
- Compo, G.P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., [...] (2011). The Twentieth Century Reanalysis Project. *Quarterly Journal of the Royal Meteorological Society*, 137(654): 1-28. <https://doi.org/10.1002/qj.776>
- Cortes, C., & Vapnik, V. (1995). SUPPORT-VECTOR NETWORKS. *Machine Learning*, 20(3): 273-297. <https://doi.org/10.1007/BF00994018>
- Cortez, P., Cerdeira, A., Almeida, F., Matos, T., & Reis, J. (2009). Modeling wine preferences by data mining from physicochemical properties. *Decision Support Systems*, 47(4): 547-553. <https://doi.org/10.1016/j.dss.2009.05.016>
- Costa, R., Fraga, H., Fonseca, A., García de Cortázar-Atauri, I., Val, M. C. [...] (2019). Grapevine Phenology of cv. Touriga Franca and Touriga Nacional in the Douro Wine Region: Modelling and Climate Change Projections. *Agronomy-Basel*, 9(4). <https://doi.org/10.3390/agronomy9040210>
- Cox, D.R. (1958). THE REGRESSION-ANALYSIS OF BINARY SEQUENCES. *Journal of the Royal Statistical Society, Series B*, 20(2): 215-242. <https://doi.org/10.1111/j.2517-6161.1958.tb00292.x>
- Davis, R.E., Dimon, R.A., Jones, G.V., & Bois, B. (2019). The effect of climate on Burgundy vintage quality rankings. *OENO One*, 53(1): 60-74. <https://doi.org/10.20870/oeno-one.2019.53.1.2359>
- Fabian, P. (2011). Scikit-learn: Machine learning in Python. *Journal of machine learning research*, 12: 2825.
- Fisher, R.A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7: 179-188. <https://doi.org/10.1111/j.1469-1809.1936.tb02137.x>
- Fix, E., & Hodges, J.L. (1989). DISCRIMINATORY ANALYSIS - NONPARAMETRIC DISCRIMINATION - CONSISTENCY PROPERTIES. *International Statistical Review*, 57(3): 238-247. <https://doi.org/10.2307/1403797>
- Fonseca, A., Fraga, H., & Santos, J.A. (2023). Exposure of Portuguese viticulture to weather extremes under climate change. *Climate Services*, 30. <https://doi.org/10.1016/j.cliser.2023.100357>
- Fraga, H. (2023). bestFeatures - A python script that exhaustively searches through all possible combinations of features for assessing model performance.
- Fraga, H., Costa, R., & Santos, J.A. (2017). Multivariate Clustering of Viticultural Terroirs in the Douro Winemaking Region. *Ciência e Técnica Vitivinícola*, 32(2): 142-153. <https://doi.org/10.1051/ctv/20173202142>

- 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. *Global Change Biology*, 22(11): 3774-3788. <https://doi.org/10.1111/gcb.13382>
- Fraga, H., Guimarães, N., Freitas, T.R., Malheiro, A.C., & Santos, J.A. (2022). Future Scenarios for Olive Tree and Grapevine Potential Yields in the World Heritage Côa Region, Portugal. *Agronomy*, 12(2): 350. <https://doi.org/10.3390/agronomy12020350>
- Fraga, H., Molitor, D., Leolini, L., & Santos, J.A. (2020). What Is the Impact of Heatwaves on European Viticulture? A Modelling Assessment. *Applied Sciences*, 10(9): 3030. <https://doi.org/10.3390/app10093030>
- Fraga, H., & Santos, J.A. (2017). Daily prediction of seasonal grapevine production in the Douro wine region based on favourable meteorological conditions. *Australian Journal of Grape and Wine Research*, 23(2): 296 - 304. <https://doi.org/10.1111/ajgw.12278>
- Freund, Y., & Schapire, R.E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1): 119-139. <https://doi.org/10.1006/jcss.1997.1504>
- Friedman, J.H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5): 1189-1232. <https://doi.org/10.1214/aos/1013203451>
- Fuka, D.R., Todd Walter, M., MacAlister, C., Degaetano, A. T. [...] (2014). Using the Climate Forecast System Reanalysis as weather input data for watershed models. *Hydrological Processes*, 28(22): 5613-5623. <https://doi.org/10.1002/hyp.10073>
- Gambetta, G.A., & Kurtural, S.K. (2021). Global warming and wine quality: are we close to the tipping point? *OENE One*, 55(3): 353-361. <https://doi.org/10.20870/oeno-one.2021.55.3.4774>
- Gidden, M. J., Riahi, K., Smith, S. J., Fujimori, S. [...] (2019). Global emissions pathways under different socioeconomic scenarios for use in CMIP6: a dataset of harmonized emissions trajectories through the end of the century. *Geosci. Model Dev.*, 12(4): 1443-1475. <https://doi.org/10.5194/gmd-12-1443-2019>
- Hannah, L., Roehrdanz, P. R., Ikegami, M., & Hijmans, R. J. (2013). Climate change, wine, and conservation. *Proceedings of the National Academy of Sciences of the United States of America*, 110(17): 6907-12. <https://doi.org/10.1073/pnas.1210127110>
- Hastie, T., Tibshirani, R., & Friedman, J. (2013). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. *Springer New York*.
- IPCC (2021). Climate Change 2021: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Barros, V.R., C.B. Field, D.J. Dokken, M.D. Mastrandrea, K.J. Mach, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- IVV (2021). Vinhos e Aguardentes de Portugal, Instituto da Vinha e do Vinho, I.P., Lisboa, Portugal.
- Jones, G.V. (2005). Climate change in the western united states grape growing regions. Proceedings of the Seventh International Symposium on Grapevine *Physiology and Biotechnology*, 689: 41-59. <https://doi.org/10.17660/ActaHortic.2005.689.2>
- Jones, G.V. (2012). A Climate Assessment for the Douro Wine Region: An Examination for the Past, Present and Future Conditions for Wine Production. ADVID, Peso da Régua, Portugal.
- Jones, G.V., & Alves, F. (2012). Impact of climate change on wine production: a global overview and regional assessment in the Douro Valley of Portugal. *International Journal of Global Warming*, 4(3/4): 383-406. <https://doi.org/10.1504/IJGW.2012.049448>
- Jones, G.V., White, M.A., Cooper, O.R., & Storchmann, K. (2005). Climate Change and Global Wine Quality. *Climatic Change*, 73(3): 319-343. <https://doi.org/10.1007/s10584-005-4704-2>
- Kriegler, E., Bauer, N., Popp, A., Humpenöder, F. [...] (2017). Fossil-fueled development (SSP5): An energy and resource intensive scenario for the 21st century. *Global Environmental Change*, 42: 297-315. <https://doi.org/10.1016/j.gloenvcha.2016.05.015>
- Lafon, T., Dadson, S., Buys, G., & Prudhomme, C. (2012). Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods. *International Journal of Climatology*: n/a-n/a. <https://doi.org/10.1002/joc.3518>
- Macedo, A., Gouveia, S., Rebelo, J., Santos, J., & Fraga, H. (2021). International trade, non-tariff measures and climate change: insights from Port wine exports. *Journal of Economic Studies*, 48(6): 1228-1243. <https://doi.org/10.1108/JES-04-2020-0161>
- Magalhães, N. (2008). Tratado de viticultura: a videira, a vinha e o terroir. Chaves Ferreira, Lisboa, Portugal, 605 pp.
- Martins, J., Fraga, H., Fonseca, A., & Santos, J.A. (2021). Climate Projections for Precipitation and Temperature Indicators in the Douro Wine Region: The Importance of Bias Correction. *Agronomy*, 11(5). <https://doi.org/10.3390/agronomy11050990>
- Massey, F.J. (1951). THE KOLMOGOROV-SMIRNOV TEST FOR GOODNESS OF FIT. *Journal of the American Statistical Association*, 46(253): 68-78. <https://doi.org/10.1080/01621459.1951.10500769>
- Mayson, R., & Duff, L. (2018). Port and the Douro. Infinite Ideas Limited.
- Murphy, K.P. (2012). Machine Learning: A Probabilistic Perspective. MIT Press.
- Panzone, L.A., & Simões, O.M. (2009). The Importance of Regional and Local Origin in the Choice of Wine: Hedonic Models of Portuguese Wines in Portugal. *Journal of Wine Research*, 20(1): 27-44. <https://doi.org/10.1080/09571260902978527>
- Rebelo, J., Caldas, J., & Guedes, A. (2015). The Douro Region: Wine and Tourism. *Almatourism-Journal of Tourism Culture and Territorial Development*, 6(11): 75-90.
- Riahi, K., van Vuuren, D. P., Kriegler, E. [...] (2017). The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, 42: 153-168. <https://doi.org/10.1016/j.gloenvcha.2016.05.009>
- Robinson, J., Harding, J., & Vouillamoz, J. (2013). Wine Grapes: A complete guide to 1,368 vine varieties, including their origins and flavours. Penguin Books Limited.
- Rumelhart, D.E., McClelland, J.L., & Group, P.D.P.R. (1987). Parallel Distributed Processing, Volume 1: Explorations in the Microstructure of Cognition: Foundations. MIT Press. <https://doi.org/10.7551/mitpress/5237.001.0001>
- Saha, S., Moorthi, S., Pan, H.-L. [...] (2010). THE NCEP CLIMATE FORECAST SYSTEM REANALYSIS. *Bulletin of the American Meteorological Society*, 91(8): 1015-1057.
- Salinger, M.J., Baldi, M., Grifoni, D., Jones, G. [...] (2015). Seasonal differences in climate in the Chianti region of Tuscany and the relationship to vintage wine quality. *International Journal of Biometeorology*, 59(12): 1799-1811. <https://doi.org/10.1007/s00484-015-0988-8>

- Séférian, R., Nabat, P., Michou, M., Saint-Martin, D. [...] (2019). Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate. *Journal of Advances in Modeling Earth Systems*, 11(12): 4182-4227. <https://doi.org/10.1029/2019MS001791>
- Service, C.C.C. (2323). Copernicus Climate Data Store (CDS). Retrieved from <https://cds.climate.copernicus.eu/>.
- Smart, R.E., Robinson, M.D., & Robinson, M. (1991). Sunlight Into Wine: A Handbook for Winegrape Canopy Management. Winetitles Pty Limited, 88 pp.
- Stone, M. (1974). Cross-Validatory Choice and Assessment of Statistical Predictions. *Journal of the Royal Statistical Society, Series B*, 36(2): 111-147. <https://doi.org/10.1111/j.2517-6161.1974.tb00994.x>
- Thorne, P.W., & Vose, R.S. (2010). Reanalyses Suitable for Characterizing Long-Term Trends. *Bulletin of the American Meteorological Society*, 91(3): 353-362. <https://doi.org/10.1175/2009BAMS2858.1>
- Thrasher, B., Maurer, E.P., McKellar, C., & Duffy, P.B. (2012). Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrology and Earth System Sciences*, 16(9): 3309-3314. <https://doi.org/10.5194/hess-16-3309-2012>
- Tonietto, J., & Carbonneau, A. (2004). A multicriteria climatic classification system for grape-growing regions worldwide. *Agricultural and Forest Meteorology*, 124(1-2): 81-97. <https://doi.org/10.1016/j.agrformet.2003.06.001>
- Uniyal, B., Dietrich, J., Vu, N.Q., Jha, M.K., & Arumi, J.L. (2019). Simulation of regional irrigation requirement with SWAT in different agro-climatic zones driven by observed climate and two reanalysis datasets. *Science of the Total Environment*, 649: 846-865. <https://doi.org/10.1016/j.scitotenv.2018.08.248>
- van Leeuwen, C., & Darriet, P. (2016). The Impact of Climate Change on Viticulture and Wine Quality. *Journal of Wine Economics*, 11(1): 150-167. <https://doi.org/10.1017/jwe.2015.21>
- van Leeuwen, C., Destrac-Irvine, A., Dubernet, M., Duchêne, E. [...] (2019). An Update on the Impact of Climate Change in Viticulture and Potential Adaptations. *Agronomy*, 9(9). <https://doi.org/10.3390/agronomy9090514>
- Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B. [...] (2019). Evaluation of CMIP6 DECK Experiments With CNRM-CM6-1. *Journal of Advances in Modeling Earth Systems*, 11(7): 2177-2213. <https://doi.org/10.1029/2019MS001683>
- White, M.A., Diffenbaugh, N.S., Jones, G.V., Pal, J.S., & Giorgi, F. (2006). Extreme heat reduces and shifts United States premium wine production in the 21st century. *Proceedings of the National Academy of Sciences of the United States*, 103(30): 11217-22. <https://doi.org/10.1073/pnas.0603230103>