Ref. Ares(2020)4007827 - 29/07/2020

H2020-SC5-01-2017

MED-GOLD

Turning climate-related information into added value for traditional **MED**iterranean **G**rape, **OL**ive and **D**urum wheat food systems

Deliverable 3.2

Report on the methodology followed to implement the wine pilot services

This project has received funding from the European Union's Horizon 2020 research and innovation programme under gran agreement No 776467.

Document Status Sheet

Deliverable Title	***Report on the methodology followed to implement the wine pilot services.***	
Brief Description	This document reports the methodologies involved in the development of the wine pilot service tool. It contains the description of a range of techniques, strategic approaches and datasets that have been analyzed and implemented by the scientific partners (seasonal forecasts and climate projections), the technical partners (dashboard implementation) and the scientific partners in collaboration with the wine champion (case study characterization, narratives and economic impact assessment). The results obtained along with the feedback gathered ([D3.6] and [D3.7]]) has driven the co-development phase up to the current version of the wine pilot service tool.	
WP number		WP 3
Lead Beneficiary	Barcelona Supercomputing Center	
Contributors	*Raül Marcos-Matamoros, Nube González-Reviriego, Antonio Graça, Alessandro Del Aquilla, Ilaria Vigo, Sara Silva, Konstantinos V. Varotsos, Michael Sanderson*	
Creation Date	17/01/2018	
Version Number	1.1	
Version Date	17/06/2020	
Deliverable Due Date	31/07/2020	
Actual Delivery Date	30/07/2020	
Nature of the Deliverable	R	*R - Report* *P - Prototype* *D - Demonstrator* *O - Other*
Dissemination Level/ Audience	PU	*PU - Public* *PP - Restricted to other programme participants, including the Commission services* *RE - Restricted to a group specified by the consortium, including the Commission services* *CO - Confidential, only for members of the consortium, including the Commission services*

Revision History Log

Version	Date	Created / Modified by	Pages	Comments
1.0	17-01-2018	GMV	14	Initial Draft
1.1	15-07-2020	BSC	78	Final version

All partners involved in the production/implementation of the deliverable should comment and report (if needed) in the above table. The above table should support the decisions made for the specific deliverable in order to include the agreement of all involved parties for the final version of the document.
Finally, after the peer review process, the deliverable should be modified accordingly to the comments and the reflections to the comments should be reported in the above table.

Table of Contents

Executive Summary **7**

1. Introduction **9**
1.1 Purpose 9
1.2 Scope 9
1.3 Definitions and Acronyms 10
1.3.1 Definitions 10
1.3.2 Acronym 11

2. References **14**
2.1 Reference Documents 14

3. Data **17**
3.1 Gridded Observational Datasets 17
3.1.1 E-OBS 17
3.1.2 PTHRES 17
3.2 Reanalysis Datasets 18
3.2.1 ERA5 18
3.2.2 ERA5 Land 18
3.3 Satellite Datasets 18
3.3.1 SENTINEL2 18
3.3.2 MODIS 19
3.3.3 LANDSAT8 19
3.4 Seasonal Forecasts 19
3.4.1 ECMWF SEAS5 19
3.5 Climate Projection Datasets 20

4. Methodology **21**
4.1 Identification of Bioclimatic Indicators Relevant for the Wine Sector 22
4.1.1 Spring Total Precipitation (SprR) 22
4.1.2 Harvest Total Precipitation (HarvestR) 23
4.1.3 Growing Season Temperature (GST) 23
4.1.4 Growing Degree Days (GDD) 24
4.1.5 Number of Heat Stress Days (SU35) 24
4.1.6 Warm Spell Duration Index (WSDI) 25
4.2 Identification of 'Good' and 'Bad' Years for the Wine Sector 26
4.2.1 'Bad' Year 1988 26
4.2.2 'Bad' Year 1993 26
4.2.3 'Bad' Year 2002 27

4.2.4 'Good' Year 2007 27
4.2.5 'Good' Year 2011 27
4.3 'Good' and 'Bad' Years in Terms of Bioclimatic Indicators 29
4.2 Development of the MED-GOLD Compound Wine Risk Indices 31
4.2.1 MED-GOLD compound Sanitary Risk Index 31
4.2.2 MED-GOLD Compound Heat Risk Index 33
4.4 Methods for Development of the Climate Service based on Seasonal Predictions 34
4.3.1 Simple Bias Correction (SBC) 38
4.3.2 Data Interpolation 39
4.3.3 Combination of Bias Correction and Interpolation 39
4.3.4 Quality assessment of seasonal predictions 42
Pearson's correlation 42
Fair Ranked Probability Skill Score (FRPSS) 42
4.3.5 Real-Time Seasonal Predictions for the Year 2020 43
4.3.6 Methodology for Economic Assessment of Using Retrospective Seasonal Predictions in a Decision-Making Context 44
4.5 Methods for Development of the Climate Service Based on Climate Change Projections 48
4.5.1 Regional Climate Models Evaluation 48
4.5.2 Testing Bias Adjustment Methods 48
4.5.3 PTHRES Integration 49
4.6 Development of MED-GOLD Dashboard Interactive Tool 50
4.7 Development of the Climate Service Based on Satellite Data 54

5. Conclusions 60

Annex A. Dashboard Rest API Reference 62

LIST OF TABLES AND FIGURES

Table 1-1 Definitions 10
Table 1-2 Acronyms 11
Table 2-1 Reference Documents 14
Table 3-1 List of Regional Climate models used 20
Table 4-1 Possible combinations of tercile forecasts and observations 47
Table 4-2 Data availability on the first release of the MED-GOLD Dashboard (May 2020) 51
Table 4-3 indices calculation 55

Figure 4-1 Spring sanitary pressure indicators for the Douro wine region in chosen years (JulDay Last = last julian calendar day of the spraying period). 28
Figure 4-2 Total Douro declared wine production in chosen years. 28
Figure 4-3 Grape quality parameters for a Douro wine region representative grape variety in chosen years as recorded in SOGRAPE vineyards (dashed lines indicate target values). 29
Figure 4-4 Ranking of bad and good years in terms of percentiles of the selected bioclimatic indicators for the representative IPMA weather station of Santa Barbara. 30
Figure 4-5 Values of Sanitary Risk Index for the reference period 1979-2018 obtained from the representative IPMA Santa Barbara weather station 33
Figure 4-6 Values of Heat Risk Index for the reference period 1979-2018 obtained from the representative IPMA Santa Barbara weather station data. 34
Figure 4-7 Most likely tercile category map for Sanitary Risk Index foreseen for year 2002. The dataset used to compute this map corresponds with the seasonal predictions issued in February coming from CDS-C3S ECMWF SEAS5. 35
Figure 4-8 Temporal evolution of the most likely tercile category of the Sanitary Risk Index obtained from the seasonal predictions issued in February from the CDS-C3S ECMWF SEAS5 (coloured bars) to the closest grid point of the weather station of Santa Barbara (Douro Valley, Portugal). The stars show the observed value (in percentile) of the Sanitary Risk Index obtained with the ERA5 reanalysis dataset. 36
Figure 4-9 Probabilistic distribution of Sanitary Risk Index for the year 2002 obtained from the seasonal predictions issued in February from the CDS-C3S ECMWF SEAS5 to the closest grid point of the weather station Santa Barbara (Douro Valley, Portugal) in comparison with observed value in the ERA5 reanalysis dataset. 37
Figure 4-10 Seasonal predictions of the most likely tercile category of SprR (first row, panels a, b and c) and map of FRPSS values (second row, panels d, e and f) over the Douro Valley. 41
Figure 4-11 Step-by-step parameters selection control for Long Term Projections. 50
Figure 4-12 MED-GOLD Dashboard's data and image export options. 52
Figure 4-13 As shown in the diagram below, MED-GOLD Dashboard (dark red path) utilizes all sub-modules of the ICT Platform's data pipeline (in orange) in order to provide value to users. 52
Figure 4-14 Sanitary Risk Index from high-resolution observational dataset PTHRES for 2015 as reported in the MED-GOLD Dashboard front-end with the time series for a specific grid-point. 53
Figure 4-15 Seasonal forecast of Sanitary Risk Index for 2020: Map of the most likely tercile as reported in the MED-GOLD Dashboard front-end with the time series for a specific grid-point. 53
Figure 4-16 SU35 over the Douro Valley: Ensemble mean change in the number of days with TX > 35ºC during AMJJASO between 2031-2060 and 1971-2000 under RCP4.5. 54
Figure 4-17 Example of the NDWI index variations (calculated from MODIS) variations during 2018 over the Douro Valley. 56
Figure 4-18 EO imagery covered area. Green area corresponds to Sentinel 2. 57
Figure 4-19 Methodology proposal for MED-GOLD validation using EO dataset. 58
Figure 4-20 NDVI average (blue line) and temperature average (orange line) comparison. 59

EXECUTIVE SUMMARY

This report accounts for deliverable 3.2, "Report on the methodology followed to implement the wine pilot services''. It contains the work performed in this regard during the period comprising M4 and M30. It is mainly focused on the methods implemented within the MED-GOLD WP3 for the co-development of the wine climate service, whereas the results derived from them will be included in the forthcoming D3.3.

The first part of the document provides an introduction together with a summary of the datasets that have been used in the methodologies applied during the co-development of the pilot service. Some of these datasets were already presented in D1.3 [RD.1] and D1.4 [RD.2] while others have been identified after the submission of these deliverables. The second part of the document is focused on the methodologies implemented during the different stages of the development of the wine pilot service. It begins by describing the more relevant bioclimatic indicators (and associated essential climate variables) selected by the MED-GOLD Champion SOGRAPE. Afterwards, the compound risk indices, the Sanitary and Heat Risk indices, are introduced as a combination of some of the aforementioned bioclimatic indicators. This methodological work is framed within the climate characterization of several 'bad' and 'good' years concerning wine production outcomes: phenology, sanitary pressure, maturation, and harvest. In this work, the strategy for selecting these years is presented. Then, the seasonal component of the pilot service, which is based on seasonal predictions of essential climate variables, bioclimatic indicators and risk indices, is described, including the methods applied for bias adjusting the seasonal predictions, downscaling them for having the spatial resolution required by the end-user and assessing the quality of the predictions. To do so, different bias correction and zero-order downscaling combinations have been tested using reanalysis data (ERA5) and the PTHRES dataset. Finally, a methodology to assess the economic impact of the use of retrospective seasonal predictions in a decision-making context is also described. Turning to the long-term future climate, the essential climate variables and the bioclimatic indicators for three EURO-CORDEX simulation periods are obtained: (i) the 1971-2000 used as a reference period (ii) and the 2031-2060 and 2071-2100, under RCP4.5 and RCP8.5, for future scenarios. The native horizontal resolution of the climate projections is 0.11º (~12km) and has been downscaled to 1km horizontal resolution with the use of the PTHRES dataset.

The next step is the development of the MED-GOLD Dashboard interactive tool. This interactive tool consists of an ICT platform with a Dashboard front end (so-called MED-GOLD Dashboard) that displays historical climate data, seasonal predictions and climate change projections according to user needs collected along WP3, in particular, task 3.1 and the feedback received along task 3.3. Last point within the methods covers the aspects relative to satellite data. It involves the downloading and postprocessing of 923 satellite scenes (at different spectral and spatial resolution from MODIS, LANDSAT8 and SENTINEL2) covering the area of interest and providing 9 derived products (e.g. vegetation indices). Finally, a summary of the document together with the main conclusions is presented.

With this deliverable, the project has contributed to the achievement of the following objectives (DOA, PartB Table1.1):

No.	Objective	Yes
1	To co-design, co-develop, test, and assess the added value of proof-of-concept climate services for olive, grape, and durum wheat	X
2	To refine, validate, and upscale the three pilot services with the wider European and global user communities for olive, grape, and durum wheat	
3	To ensure replicability of MED-GOLD climate services in other crops/climates (e.g., coffee) and to establish links to policy making globally	X
4	To implement a comprehensive communication and commercialization plan for MED-GOLD climate services to enhance market uptake	
5	To build better informed and connected end-user communities for the global olive oil, wine, and pasta food systems and related policy making	X

1. Introduction

MED-GOLD seeks to create innovative climate services to assist the adaptation of agricultural management for supporting climate information coming from seasonal predictions and long-term climate change projections. The methods applied for tailoring, postprocessing and adapting the information provided to the end-users aim to ease several aspects of decision-making in the wine sector (such as pest management and vineyard plantation planning, among others) that are critical for the wine industry's resilience to climate change.

1.1 Purpose

The purpose of this deliverable is to report the different methodologies and strategies used for developing the wine climate service. It includes the description of the datasets selected and assesses the methods applied for the computation of the user-selection of bioclimatic indicators and risk indices at different temporal scales (the ones corresponding to historical climate, seasonal predictions and climate change projections). It also contains the technical details of the bias-adjustment techniques used to reduce the systematic errors of the climate models, the downscaling approaches followed to increase the spatial resolution of the climate information and the verification metrics implemented to provide estimates of the seasonal climate uncertainty and performance. This deliverable also aims to provide a description of the methods used for the development of the MED-GOLD Dashboard interactive tool, for the economic assessment of using retrospective seasonal forecasts in the context of the wine sector and for the use of satellite data in the context of wine climate service.

1.2 Scope

MED-GOLD seeks to fill the gap of climate driven information between its 'interpreters' and the potential users in the agricultural sector by the co-development of the climate services tools involving both scientific partners and decision-makers in the Mediterranean region. Within this framework, MED-GOLD WP3, and in particular the work done in task 3.2 starts from the results of the critical needs obtained in task 3.1 for developing a complex methodology, which integrate climate statistical methods, economic assessment and climate data visualisation to shape a climate service that meets the needs of the wine industry in terms of adaptation and mitigation to climate change.

1.3 Definitions and Acronyms

1.3.1 Definitions

Concepts and terms used in this document and needing a definition are included in the following table (some of these definitions along with others relevant for the project can be also found in the MED-GOLD glossary):

Table 1-1 Definitions. Click to see more definitions.

Concept / Term	Definition
Bias	The model error(s) in simulated climate, when compared with observations.
Bias-adjustment/correction	A collective term for statistical methods applied to climate model data to reduce the biases.
Climate change projections	Climate scenarios for the future, commonly, for the next century
(Agro-) Climate indices	Parameters describing a specific characteristic of climate (e.g. intensity, frequency or distribution) for the agricultural sector
Climate memory	Personal memory from climate conditions in recent years
Climate models	Climate models are mathematical algorithms that simulate the interactions of the important drivers of climate (e.g. atmosphere, oceans, land surface, ice).
Dashboard	Integrated visualization display
Downscaling	Methods for creating climate information at higher spatial and/or temporal scales from coarser resolution simulation.
Essential Climate Variables	A group of variables that critically contributes to the characterization of Earth's climate. For the surface, the variables include air temperature, precipitation, humidity, wind speed and direction, etc.
Indicator	A parameter describing a reality, i.e., synthesizing the effects of future climate change with relevance to a specific sector and business
Percentile	Division of the population distribution in 100 categories
Seasonal forecasts	Predictions of the climatic conditions for the coming months
(Predictive) Skill	A statistical measure of the accuracy of seasonal forecasts
Tercile	Division of the population distribution into three categories

1.3.2 Acronym

Acronyms used in this document and needing a definition are included in the following table:

Table 1-2 Acronyms and abbreviations.

Acronym	Definition
AoI	Area of interest
ASCAT	Advanced SCATterometer
BaU	Buisiness-as-Usual
CDF	Cumulative Distribution Function
CDS-C3S	Climate Data Store - Copernicus Climate Change Service
CMIP5	Coupled Model Intercomparison Project Phase 5
CORDEX	Coordinated Regional Climate Downscaling Experiment
DV	Douro Valley
ECA&D	European Climate Assessment & Dataset
ECMWF	European Centre for Medium-Range Weather Forecast
ECMWF SEAS5	ECMWF Seasonal Forecasting System 5
ECV(s)	Essential Climate Variable(s)
ERA5	ECMWF Atmospheric ReAnalysis 5
EQM	Empirical Quantile Mapping
E-OBS	ENSEMBLES daily gridded observational dataset
ETCCDI	Expert Team on Climate Change Detection and Indices
(F)RPSS	(Fair) Rank Probability Skill Score
GDD	Growing Degree Days

GST	Growing Season average Temperature
HarvestR	Harvest total precipitation
ICT	Information and Communication Technology
IDW	Inverse Distance Weighting
IFS	Integrated Forecast System
IP	Iberian Peninsula(r)
IPMA	Instituto Português do Mar e da Atmosfera
LST	Land Surface Temperature
LTDN	Local solar Time on Descending Node
MSI	Multi Spectral Instrument
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NEMO	Nucleus for European Modelling of the Ocean
NMDI	Normalized Multiband Drought Index
ORAS5	Ocean ReAnalysis (in ECMWF SEAS5)
PTHRES	PorTugal High RESolution data set
RCM	Regional Climate Model
RCP	Representative Concentration Pathways
SBC (SBCMW)	Simple Bias Correction (with mMoving window)
SprR	Spring total precipitation
SST	Sea Surface Temperature

SSO	Sun-synchronous orbit
SU35	Number of heat stress days
TVX	Temperature Vegetation Index
TXx	Interannual maximum temperature
WSDI	Warm Spell Duration Index

2. References

2.1 Reference Documents

The following documents, although not part of this document, amplify or clarify its contents. Reference documents are those not applicable and referenced within this document. They are referenced in this document in the form [RD.x]:

Table 2-1 Reference Documents.

Ref.	Title	Code	Version	Date
[RD.1]	MED-GOLD Deliverable 1.3: Assessment of quality of European climate observations and their appropriateness for use in climate services for each sector			2019
[RD.2]	MED-GOLD Deliverable 1.4: Mapping uncertainty and skill in seasonal forecasts and climate data			2019
[RD.3]	MED-GOLD Deliverable 3.1: Report on the two case studies at seasonal- and long-term timescales for the wine sector			2019
[RD.4]	MED-GOLD Deliverable 3.6: First feedback report from users on wine pilot service development			2019
[RD.5]	MED-GOLD Deliverable 3.7: Second feedback report from users on wine pilot service development			2020
[RD.6]	Payload-Mass Trends for Earth-Observation and Space-Exploration Satellites. Rast et al. (1999) http://www.esa.int/esapub/bulletin/bullet97/rast.pdf			1999
[RD.7]	Monthly precipitation in mm at 1 km resolution based on SM2RAIN-ASCAT 2007-2018, IMERGE, CHELSA Climate and WorldClim. Brocca et al. (2019) https://zenodo.org/record/3256275#.XuC15UUzZPY			2018
[RD.8]	Research paper: High-Resolution Temperature Datasets in Portugal from a Geostatistical Approach: Variability and Extremes. Fonseca, A.R., and Santos, J.A. (2017). J. Appl. Meteor. Climatol. 57, 627–644. https://doi.org/10.1175/JAMC-D-17-0215.1			2017
[RD.9]	Research paper: A statistical analysis of the relationship between climatic factors and the Normalized Difference Vegetation Index in China. Song, Yi & Mingguo, Ma. (2011). International Journal of Remote Sensing. https://www.researchgate.net/publication/232062766_A_statistical_analysis_of_the_relationship_between_climatic_factors_and_the_Normalized_Difference_Vegetation_Index_in_China			2011
[RD.10]	Research paper: The ERA5 global reanalysis. Hersbach et al. (2020). Quarterly Journal of the			2020

	Royal Meteorological Society. https://doi.org/10.1002/qj.3803			
[RD.11]	Research paper: Simulations of the ENSO Hydroclimate Signals in the Pacific Northwest Columbia River Basin. Leung et al. (1999) Bull. Amer. Meteor. Soc. 80 (11): 2313–2330. https://doi.org/10.1175/1520-0477(1999)080<2313:SOTEHS>2.0.CO;2			1999
[RD.12]	Book: Statistical methods in the atmospheric sciences. Wilks, Daniel S. (2011) Academic press Vol. 100. shorturl.at/crvy6			2011
[RD.13]	Book: Forecast verification: a practitioner's guide in atmospheric science. Jolliffe, Ian T., and David B. Stephenson, eds. (2012) John Wiley & Sons. https://doi.org/10.1002/9781119960003			2012
[RD.14]	Research paper: Three recommendations for evaluating climate predictions.Fricker et al. (2013) Meteorological Applications 20.2: 246-255. https://doi.org/10.1002/met.1409			2013
[RD.15]	Research paper: Fair scores for ensemble forecasts. Ferro, C. A. T. (2014) Quarterly Journal of the Royal Meteorological Society 140.683: 1917-1923. https://doi.org/10.1002/qj.2270			2014
[RD.16]	Research paper: A climate projection dataset tailored for the European energy sector. Bartók, Blanka, et al. (2019) Climate services 16: 100138. https://doi.org/10.1016/j.cliser.2019.100138			2019
[RD.17]	Research paper: The next generation of scenarios for climate change research and assessment. Moss RH et al. (2010) Nature 463:747-756. https://doi.org/10.1038/nature08823			2010
[RD.18]	Research paper: The representative concentration pathways: an overview. van Vuuren DP et al. (2011) Climatic Change 109:5. https://doi.org/10.1007/s10584-011-0148-z			2011
[RD.19]	Research paper: The R-based climate4R open framework for reproducible climate data access and post-processing. Iturbide M. et al. (2019) Environmental Modelling & Software 111:42-54. https://doi.org/10.1016/j.envsoft.2018.09.009			2019
[RD.20]	Research paper: Testing bias adjustment methods for regional climate change applications under observational uncertainty and resolution mismatch. Casanueva A. et al. (2020) Atmospheric Science Letters e978. https://doi.org/10.1002/asl.978			2020
[RD.21]	Research paper: An integrated assessment of climate change impacts for Greece in the near future. Giannakopoulos et al. (2011) Regional Environmental Change 11:829-843. https://doi.org/10.1007/s10113-011-0219-8			2011
[RD.22]	Research paper: Ozone-temperature relationship during the 2003 and 2014 heatwaves in Europe. Varotsos et al. (2019) Regional Environmental Change 19:1653-1665. https://doi.org/10.1007/s10113-019-01498-4			2019
[RD.23]	Research paper: Mapping model agreement on future climate projections. Tebaldi et al. (2011)			2011

	Geophysical Research Letters 38. https://doi.org/10.1029/2011gl049863			
[RD.24]	Research paper: Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. Kotlarski, S. et al. (2014) Geosci. Model Dev., 7, 1297–1333, https://doi.org/10.5194/gmd-7-1297-2014			2014
[RD.25]	Research paper: Statistical bias correction for daily precipitation in regional climate models over Europe. C. Piani et al. (2009) Theoretical and Applied Climatology, 99, 187-192. https://doi.org/10.1007/s00704-009-0134-9			2009
[RD.26]	Research paper: Statistical downscaling with the downscaleR package (v3.1.0): contribution to the VALUE intercomparison experiment. Bedia J et al. (2020) Geoscientific Model Development. https://doi.org/10.5194/gmd-13-1711-2020			2020
[RD.27]	Research paper: Probabilistic downscaling approaches: Application to wind cumulative distribution functions. A. Michelangeli, M. Vrac, H. Loukos. (2009) Geophys. Res. Lett., https://doi.org/10.1029/2009GL038401			2009
[RD.28]	Research paper: An intercomparison of a large ensemble of statistical downscaling methods over Europe: Results from the VALUE perfect predictor cross-validation experiment. Gutiérrez, JM et al. (2019) Int. J. Climatol. 2019; 39: 3750– 3785. https://doi.org/10.1002/joc.5462			2019
[RD.29]	Research paper: The interplay between atmospheric conditions and grape berry quality parameters in Portugal. C. Costa, A. Graça, N. Fontes, M. Teixeira, H. Gerós, JA Santos. (2020) *Appl. Sci.* **2020**, *10*, 4943.			2020

3. Data

Deliverables D1.3 [RD.1] and D1.4 [RD.2] made a first assessment of, respectively, the potential of different observational reference and climate datasets (seasonal forecasts and climate projections) that could be useful for the project. In this section, a summary is given of the actual datasets used in the methodologies applied so far for the development of the wine pilot service tool. This includes some useful datasets that were identified after the submission of D1.3 [RD.1]. These datasets have been useful in assessing the quality of the climate service (either directly, e.g. PTHRES and ERA5, or indirectly, e.g. satellite data) as well as an extra information source for the end-users. Some of these datasets have a global coverage (e.g. ERA5 or ERA5Land) whereas others are high resolution regional datasets (PTHRES) or local registers (non-climatic datasets).

3.1 Gridded Observational Datasets

3.1.1 E-OBS

E-OBS is a dataset that provides the Essential Climate Variables of temperature and rainfall for Europe and is created by gridding, i.e., interpolation of irregularly-spaced surface weather stations from the ECA&D database onto a regular grid. E-OBS is widely used for monitoring extremes across Europe and for validating climate models with the daily data covering the period 1950-2019. In the early stages of the project, E-OBSv17 (horizontal resolution ~25km) was evaluated against data from nine independent weather stations operated by Sogrape Vinhos SA, SOGRAPE, over Douro Valley for the years 2011-2017 (the nine weather stations are part of the network used to create the E-OBS dataset). The analysis revealed that E-OBSv17 could be used for evaluating and bias adjusting the Regional Climate Model (RCM) simulations [RD.1].

3.1.2 PTHRES

Portugal high-resolution observational temperature and precipitation dataset (denoted as PTHRES hereafter, [RD.8]) provides daily temperature (maximum, mean and minimum) and daily precipitation for Portugal from 1951 to 2015 with approximately 1 km horizontal resolution. The data has been developed based on the 25-km resolution E-OBS dataset, and has been later validated with observational daily temperatures from 23 Portuguese stations not included in E-OBS. The validation process involved the mean estimate, R^2 (together with the corresponding p values), bias and RMSE of extreme indices (e.g., maximum temperature, inter-annual maximum, TXx and summer days, SU). The dataset was further validated by SOGRAPE against data from their own weather station network showing high determination coefficients for temperature (> 90%) and less for precipitation (51-92%) [RD.29] This dataset was provided to MED-GOLD partners from University of Trás-os-Montes and Alto Douro (UTAD). The dataset can be openly used in public visualization platforms, but the access to the raw data should be kept restricted to the MED-GOLD partners.

3.2 Reanalysis Datasets

3.2.1 ERA5

ERA5 is the latest climate reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF), providing hourly data on many atmospheric, land-surface and oceanic climate variables for the period from 1979 to the present. This reanalysis replaces the ERA-Interim reanalysis which was started in 2006. ERA5 is based on the Integrated Forecasting System (IFS) Cy41r2 (operational in 2016). ERA5 benefits from a decade of developments in model physics, core dynamics and data assimilation. In addition to a significantly enhanced horizontal resolution of 31 km (compared to 80 km for ERA-Interim), ERA5 has hourly output, and an uncertainty estimate from an ensemble (3-hourly at half the horizontal resolution). The reanalysis includes 137 levels in the atmosphere (up to 0.01hPa) on a regular latitude-longitude grid. For more information on this reanalysis see [RD.10].

3.2.2 ERA5 Land

ERA5-Land is a reanalysis dataset that provides a consistent view of the evolution of land variables over several decades at an enhanced resolution compared to ERA5 (9km). ERA5-Land has been produced by replaying the land component of the ECMWF ERA5 climate reanalysis. This reanalysis combines model data with observations from across the world into a globally complete and consistent dataset using the laws of physics.

3.3 Satellite Datasets

3.3.1 SENTINEL2

Copernicus is the European Union's Earth Observation programme aimed at providing remote sensing data services and Copernicus services information for the benefit of European citizens though products and applications, and to support policy and decision-making for social, economic and environmental benefits.

Sentinel-2 is a multispectral operational imaging mission formed by two satellites. This mission compliments the SPOT and Landsat missions. The satellites are in the same orbital plane phased 180 degrees. The orbit is an SSO with a 10-day repeat cycle. The mean altitude is 786 km and the LTDN is 10:30. It carries a MSI, Multi Spectral Instrument. The acquisition strategy consists of systematic push-broom acquisitions, plus lateral mode capability for monitoring emergency events. It samples 13 spectral bands: four bands at 10 m, six bands at 20 m and three bands at 60 m spatial resolution. The orbital swath width is 290 km.

3.3.2 MODIS

MODIS (or Moderate Resolution Imaging Spectroradiometer) is a key instrument aboard the Terra (originally known as EOS AM-1) and Aqua (originally known as EOS PM-1) satellites. Terra's orbit around the Earth is timed to pass from north to south across the equator in the morning, while Aqua passes south to north over the equator in the afternoon. Terra MODIS and Aqua MODIS are viewing the entire Earth's surface every 1 to 2 days, acquiring data in 36 spectral bands, or groups of wavelengths. Spatial Resolution: 250m (bands 1-2), 500m (bands 3-7), 1000m (bands 8-36).

3.3.3 LANDSAT8

The Landsat 8 satellite orbits the Earth in a sun-synchronous, near-polar orbit, at an altitude of 705 km (438 mi), inclined at 98.2 degrees, and circles the Earth every 99 minutes. The satellite has a 16-day repeat cycle with an equatorial crossing time: 10:00 a.m. +/- 15 minutes. Landsat 8 acquires about 740 scenes a day on the Worldwide Reference System-2 (WRS-2) path/row system, with a swath overlap (or sidelap) varying from 7 percent at the Equator to a maximum of approximately 85 percent at extreme latitudes. The scene size is 185 km x 180 km (114 mi x 112 mi).

3.4 Seasonal Forecasts

3.4.1 ECMWF SEAS5

The ECMWF SEAS5 seasonal predictions dataset obtained from the Climate Data Store of the Copernicus Climate Change Service (CDS-C3S) is the fifth generation of seasonal prediction which replaces the former System 4 and uses the Integrated Forecast System (IFS) Cycle 43r1. There are 25 ensemble members for the hindcast period from 1993 to 2016.

SEAS5 includes enhancements in the land-surface initialization, atmospheric resolution and the ocean model when compared to S4. For instance, regarding the land-surface initialization, the former includes a new offline recalculation at the native atmospheric resolution with an improved precipitation forcing. The performance tests show a good degree of consistency between the initialization of SEAS5 re-forecast and the real-time predictions while the initialization is not perfect (i.e., the real-time assimilation is not identical as reanalysis). In addition, the SEAS5 uses the new version of ocean model NEMO (Nucleus for European Modelling of the Ocean) with upgraded ocean physics and resolution. Finally, the ocean and sea-ice initial conditions are provided by the new ocean analysis and reanalysis ensemble (ORAS5).

3.5 Climate Projection Datasets

The climate projection work employs daily maximum, daily minimum and daily mean temperatures, as well as daily precipitation for the Douro valley from a five member GCM-RCM sub-ensemble

simulations (Table 3-1) from the EURO-CORDEX modelling experiment (http://www.euro-cordex.net) with the model sub-ensemble used in this study being similar to the five member sub-ensemble identified and proposed in [RD.16]. The horizontal resolution of the models is 0.11° (~12km) with the simulated data in this project covering three periods: the 1971-2000 which is used as the reference period and the periods 2031-2060 and 2071-2100 under two Representative Concentration Pathways (RCP) scenarios, the RCP4.5 and RCP8.5 [RD.17-18].

Table 3-1 List of Regional Climate models used.

Institute	RCM	GCM
SMHI	RCA4	HadGEM2-ES
SMHI	RCA4	CNRM-CERFACS-CNRM-CM5
IPSL-INERIS	WRF331F	IPSL-IPSL-CM5A-MR
KNMI	RACMO22E	ICHEC-EC-EARTH
MPI-CSC	REMO2009	MPI-M-MPI-ESM-LR

4. METHODOLOGY

The co-development of the MED-GOLD wine pilot service is a multidisciplinary effort encompassing decision makers, scientific and technical partners. Consequently, the methodology employed is varied and tailored to each feature and time-scale. In all the cases the steps taken have followed either a co-design approach or a fluid communication scheme to involve the wine champion SOGRAPE at all the stages of the tool's development.

The methodologies that have been developed in WP3 are focused on the Douro Valley, in Portugal, covering an approximate extension of 2430km^2 framed in the following geographic coordinates, N=41.55, S=40.92, W=-7.92, E=-6.75. The relevance of this strategic region is due to its microclimate and soil which allows the production of high quality grapes for the wine industry (e.g. for the production of the famous Port wine). However, it is worth noting that although the methods used are framed within the Douro region, the workflows developed make use of global coverage datasets tailored to a specific area through regional and local data. This enables the transferability of the work performed to other areas and sectors in a seamless and straightforward way.

This section begins by introducing the most relevant bioclimatic indicators for the wine industry. The next step is to characterize the 'good' and 'bad' years selected by the wine champion and that have led to the construction of two wine risk indices. The seasonal predictions methods employed by the BSC are presented afterwards. They include the bias correction selected as well as the analysis to discern the effect of the order of application of the downscaling, the computation of the bioclimatic indicator and the bias correction method. This section also contains the metrics used in the verification process to check the performance of the predictions. Additionally, the interaction with the MED-GOLD Champion SOGRAPE presented the difficulties found by end-users in using retrospective forecasts to estimate the economic impact of using seasonal forecasts compared to their customary methods. The methodology developed by the BSC and SOGRAPE on this regard is also detailed in this section. In the last section devoted to seasonal forecasting, the workflow for the computation of the 2020 seasonal forecasts of SprR, HarvestR and GST has been included.

The next methods described in this section are linked to the climate change time scale and they have been developed by NOA. It involves the evaluation of the "raw" RCM simulations against E-OBSv17 using a selected number among the 27 ETCCDI 27 (Expert Team on Climate Change Detection and Indices) indices for the period 1971-2000 as well as the bias adjustment of the raw daily data using the empirical quantile mapping, EQM, and E-OBS as reference climate data. After that, the projections for 2031-2060 and 2071-2100 essential climate variables (hereafter, ECVs) and bioclimatic indicators have been computed from the daily bias adjusted simulations under two Representative Concentration Pathways (RCP) scenarios from EURO-CORDEX, the RCP4.5 and RCP8.5. It is mentioned that the results have an horizontal resolution of 1x1km since PTHRES was eventually used to bias adjust the climate model simulations.
In parallel with these developments, BeetoBit, ENEA and UTH were working on the consolidation of the MED-GOLD front-end and backend of the wine pilot service interactive tool. Finally, the last

section reviews the methodological approaches relative to satellite data. It involves the postprocessing of 923 satellite scenes from different satellite products (e.g. MODIS, LANDSAT8 or SENTINEL2) covering the area of interest and providing 9 derived products.

4.1 Identification of Bioclimatic Indicators Relevant for the Wine Sector

Six bioclimatic indicators related to vine growth , grape yields and harvest dates have been identified by the champion SOGRAPE since the early stages of the project (see [RD.2] and references therein). Their definitions and mathematical formula together with their viticultural meanings are summarised in the following lines. These indicators are: spring total precipitation (SprR), harvest total precipitation (HarvestR), growing season average temperature (GST), growing degree days (GDD), warm spell duration index (WSDI) and number of heat stress days (SU35). However, as the GDD is highly correlated with GST, it is not included in further analysis regarding Risk Indices.

4.1.1 Spring Total Precipitation (SprR)

SprR is the total precipitation from 21st April to 21st June (for the Northern Hemisphere). The wetness of spring represented by this index will affect the level of vigour which is associated with the fungal disease and, in turn, affects the amount of the cost due to the changes of protective treatments or operations.

$$Total\ precipitation = \sum_{start\ date}^{end\ date} prlr$$

where *prlr* is the daily average precipitation in mm. The start (end) date is the first (last) day of the period considered for the specific index. For instance, SprR takes the period from 21st April to 21st June,

$$sprR = \sum_{21st\ April}^{21st\ June} prlr$$

Dry springs will delay vegetative growth and reduce vigour and leaf area total surface. Fungal disease pressure will be lower and therefore there will be less need for protective and/or curative treatments, which translates into lower costs. Wet springs will promote greater vigour, increase the risk of fungal disease and disrupt vineyard operations, as they may prevent machinery from entering in the vineyard due to mud. They are usually associated with higher costs.

4.1.2 Harvest Total Precipitation (HarvestR)

HarvestR is the total precipitation from 21st August to 21st October (the usual harvest period in the Northern Hemisphere). Winegrowers and winemakers experience the major risks during the harvest season because the precipitation received during this period not only influences the quality and quantity of berries through the physiological processes, but also likely impedes the movement of machines and people. A (fully) dry condition is preferred over this period.

$$HarvestR = \sum_{21st\ August}^{21st\ October} prlr$$

Wet harvests are one of the major risks for both winegrowers and winemakers. Scarce and moderate rainfall values during the summer can be positive, especially in dry and semi-arid to arid areas, as they provide the necessary water and humidity for the physiological processes of the grapevine to occur, thus avoiding the need for irrigation. However, heavy rain downpours are detrimental to quality as berries will absorb water and dilute quality compounds such as sugars, acids, polyphenols, color and aroma precursors. They may also reduce quantity if hail events occur. Continuous rainfall during the harvesting period sets the conditions for widespread fungal infections (*Botrytis* being the most prevalent, but also *Armilaria*, *Pennicillium*, etc.) destroying berries, causing grapes to develop acetic acid bacteria and increasing levels of acetic, gluconic acids, ethyl acetate and other compounds very detrimental for wine quality. Additionally, in areas with deep soils, rain during harvest hampers mobility, hindering the movement of machines and people and accelerating the harvest before the quality is totally lost. Therefore, ideally, a harvest period should be wholly dry.

4.1.3 Growing Season Temperature (GST)

GST is the average of daily mean temperatures between 1st April and 31st October (for the Northern Hemisphere). The optimum range of GST varies with the grapevine variety, thus there are two ways of using GST (i) to select the varieties for a given location or (ii) to determine the location for a given variety. More importantly, GST gained popularity within the wine sector as more unprecedented values started to be observed in many regions, disrupting local varieties' phenology. The equation is shown as below.

$$GST = \frac{1}{n} \times \sum_{1st\ April}^{31st\ October} tas$$

where *n* is the number of days and *tas* is daily average temperature (°C). GST provides information on which varieties are best suited for a given site or inversely, which are the best places to grow a

specific variety. In a climate change scenario, it becomes an important index to use when making decisions about planting or replanting a vineyard. For existing vineyards, GST also provides information on the suitability of its varieties for the climate of specific years, explaining quality and production variation. This index became popular when climate change started becoming an issue, as a clear and intuitive way to have a general idea of which areas would gain or lose suitability to produce quality wines. Many grapevine varieties across the world have been characterized by their optimum GST. This indicator can also be used to characterize the climate drive for grape and wine quality across different years as it impacts the growth cycle and phenology.

4.1.4 Growing Degree Days (GDD)

GDD is the sum of differences between daily mean temperatures and 10°C (vegetative growth minimum temperature) between 1st April and 31st October (for the Northern Hemisphere). It is an index of characterization of the wine-growing areas towards their suitability to specific types of wines/grapes, and is highly associated with GST.

$$GDD = \sum_{1st\ April}^{31st\ October} maximum\,[(tas - 10), 0°C]$$

Since the negative values are discarded, the *GDD* is calculated by summing the positive differences between the temperature difference and 10°C over the defined period. However, GDD has been shown to be highly correlated with GST, and therefore seems to be redundant for risk indices already using GST.

4.1.5 Number of Heat Stress Days (SU35)

In MED-GOLD SU35 is defined as the 7-month count of days on which the daily temperature maximum exceeds 35°C between 1st April and 31st October (for the Northern Hemisphere). For grapevine, 35°C is an average threshold for photosynthesis to occur. In a practical situation, a higher index probably leads to a lower berry quality (more organic acids consumed) and, in turn, translates into a quality loss that may be offset with higher costs from acidity correction and/or demand for extra water (due to the needs of plant cooling).

$$SU35_{i,j} = \sum_{1st\ April}^{31st\ October} \left[tasmax^{mod,i,j} > 35°C\right]$$

where $tasmax^{mod,i,j}$ is the daily temperature maximum prediction at the ith day of the jth year taking into account all the ensemble members but excluding the *jth* year.

There is a modification applied to the definition of SU35 which is later used in the context of seasonal forecasts as follows. Instead of using the fixed 35°C, the observed percentile corresponding to 35°C is used. This modification has the bias correction implicitly included in the computation.

$$SU35'_{i,j} = \sum_{1st\ April}^{31st\ October} \left[percentile^{mod,i,j} > percentile_{35°C}^{obs,i,j}\right]$$

where $percentile^{mod,i,j}$ is the percentile corresponding to the predicted temperature at the ith day of the jth year taking into account all the ensemble members but excluding the *jth* year. The $percentile_{35°C}^{obs,i,j}$ is the observed percentile corresponding to 35°C at the ith day of the jth year (without the *jth* year considered). Above this threshold, the plant closes its stomata. If this situation occurs after veraison, maturation will be arrested for as long as the situation holds, decreasing sugar, polyphenol and aroma precursor levels, all essential for grape and wine quality. Deprived from its normal energy source, the plant will turn to use organic acids that will decrease the acidity levels of the berry, decreasing its quality. The plant will also use more water to cool its tissues, mainly after temperatures decrease in the evening. The higher the index, the lower the quality of the berry and its aptitude to produce quality grapes. The loss of acidity will mean, even for lower index levels, higher costs of correcting acidity (to balance wine taste and keep microbiological safety in the musts) and water needs (to support grapevine's integrity in the face of high temperatures). There are inter- and intra-varietal variations in grapevines towards this threshold.

4.1.6 Warm Spell Duration Index (WSDI)

In MED-GOLD WSDI is defined as the 7-month count of days on which the daily temperature maximum exceeds the temperature corresponding to its 90th percentile for at least 6 consecutive days between 1st April and 31st October (for the Northern Hemisphere). It signals the extreme heatwave in a warm region and in turn increases additional losses due to flowering disruption and/or the dehydration of berry and leaf, etc.

$$WSDI_{i,j} = \sum_{1st\ April}^{31st\ October} \left[tasmax^{mod,i,j} > tasmax_{90th\ percentile}^{mod,i,j}\right]_{6\ consecutive\ days\ or\ more}$$

where $tasmax^{mod,i,j}$ is prediction of the daily temperature maximum at the ith day of the jth year and $tasmax^{mod,i,j}_{90thpercentile}$ is the temperature corresponding to its 90th percentile at the ith day taking into account all years (except for the *jth* year) and all ensemble members.

Considered an index for heatwaves, the same considerations as SU35 apply here. This index, however, signals when warm regions start to become too extreme and cause additional losses because of flowering disruption (when too early in the season) or extreme berry and leaf dehydration and scalding (berry skin sunburn, leaf and shoot desiccation), in addition to excessive water depletion.

4.2 Identification of 'Good' and 'Bad' Years for the Wine Sector

In order to obtain an indication on how the selected bioclimatic indicators can provide a reliable overview of the behaviour of the productive season and the major features of interest of the grape-wine agrofood system in the Douro region, SOGRAPE has identified a number of 'good' and 'bad' years in recent decades, based on expert judgement and climate (and non-climatic) data already available [RD.1]. In general, the bad years were years when production was low, high sanitary pressure caused either crop or quality loss and/or incurred high protection costs. The 'bad' years identified by SOGRAPE were 1988,1993 and 2002. By opposition, 'good' years were years of fair production, high quality and moderate to low protection costs. The 'good' years identified by SOGRAPE were 2007 and 2011. The narrative descriptions of both 'good' and 'bad' years are detailed below and the characterisation of these years in terms of sanitary risk, wine production and grape quality parameters is shown, respectively, in figures 4-1, 4-2 and 4-3.

4.2.1 'Bad' Year 1988

Warmer than normal temperatures and higher than normal spring rainfall led to increased disease pressure (downy mildew - Plasmopara viticola), continuing well into the early summer (disease risk period from 19-Apr to 18-Jul, 90 days) which, through an uncontrollable outbreak caused major yield losses, making this one of the lowest yielding years in decades (76 ML against an average of 118 ML in the previous 10 years and 162 ML in the previous year). Favourable conditions during late summer and harvest period were not enough to improve quality, which was also greatly affected by powdery mildew in many areas.

4.2.2 'Bad' Year 1993

Dry winter conditions made for late bud-break. Temperature and precipitation levels within normal values, made for an average spring disease pressure, not too severe (disease risk period from 3-May to 28-Jun, 56 days). The warm summer temperatures anticipated a high-quality harvest until, in September and October, heavy and prolonged rains caused an outbreak of massive and

widespread grey rot (*Botrytis, Aspergillus, Penicillium*, etc.), completely destroying the harvest quality and considerably reducing yields and value.

4.2.3 'Bad' Year 2002

Late bud-break, due to some extreme episodes of low temperatures in winter and, especially to low winter and early spring precipitation, causing episodes of moderate to severe drought early in the season. Disease pressure was almost absent (disease risk period from 14-May to 27-May, 13 days). However, vegetative growth was thus stunted, which made for harvests to start in mid-September and last well until October, when the equinoctial rains severely damaged grape quality, diluting their contents and causing local outbreaks of rot (*Botrytis*) in some areas.

4.2.4 'Good' Year 2007

Winter with normal precipitation made for good water reserves that lasted well until summer due to milder temperatures, not far above normal values. Disease pressure was moderate to high, but still within control, as temperatures did not promote outbreaks (disease risk period from 30-Apr to 16-Jul, 77 days). Regular and moderate episodic rainfall events during late summer were not enough to cause disease outbreaks and helped to maintain a good hydric status in the grapevines, leading to extremely high quality grapes, especially in terms of aromatics and color intensity. Harvest started in mid-September without major rain events, sustaining the high quality of the grapes.

4.2.5 'Good' Year 2011

Mild temperatures throughout the growth cycle caused by a relatively late bud-break and a smooth, regular vegetative growth with almost no disease pressure (disease risk period from 20-Apr to 30-May, 40 days). The absence of the usual high temperature extremes during the maturation phase, together with small amounts of precipitation, not enough to cause disease pressure but sufficient to support the grapevine hydric status, led to the development of grapes very balanced and even with more concentration than in 2007, which were harvested untouched and in pristine conditions, due to the general dry harvest conditions. Considered as a perfect year as one might get.

Figure 4-1 Spring sanitary pressure indicators for the Douro wine region in chosen years (JulDay Last = last julian calendar day of the spraying period).

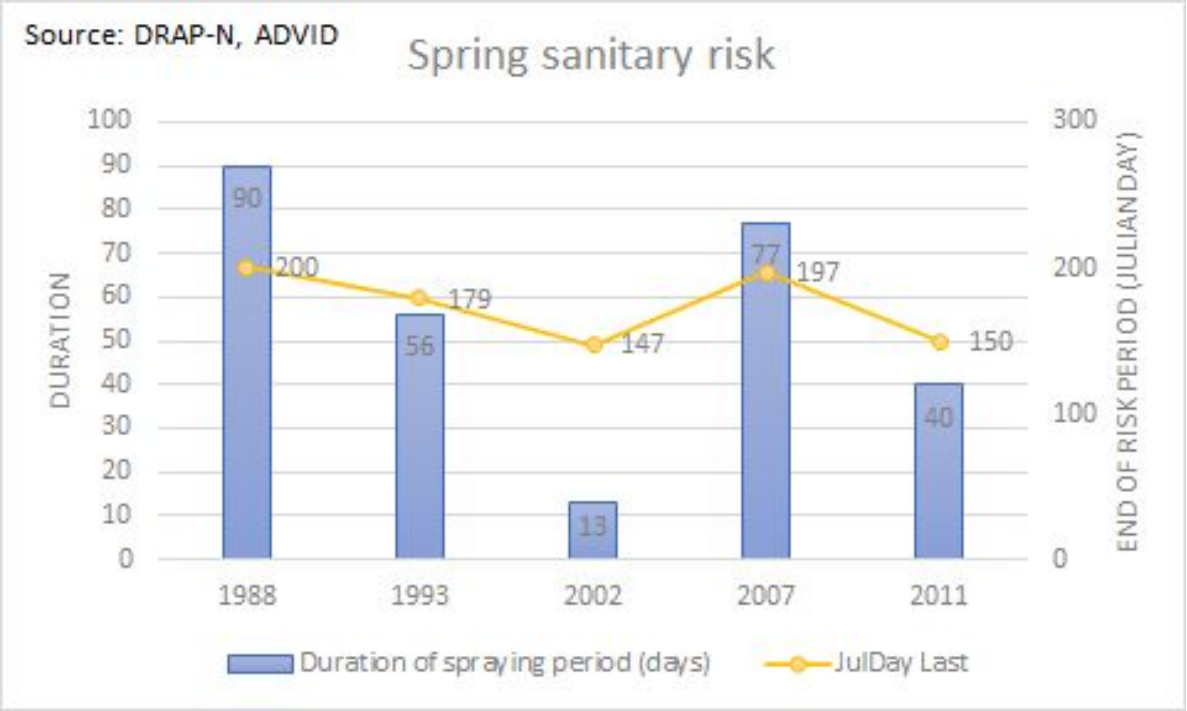


Figure 4-2 Total Douro declared wine production in chosen years.

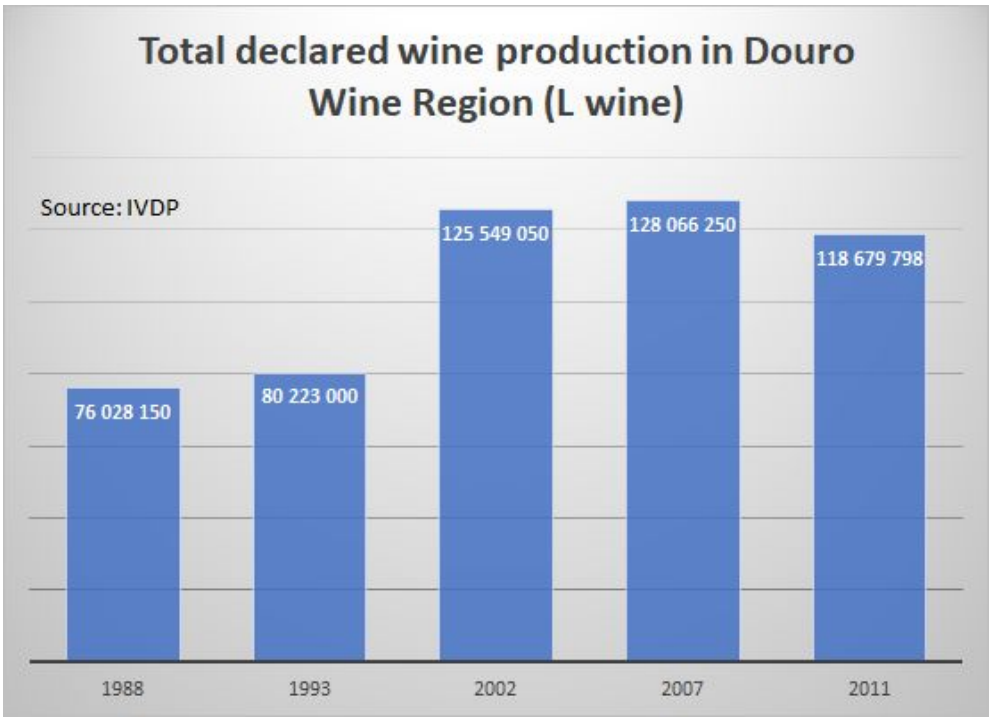
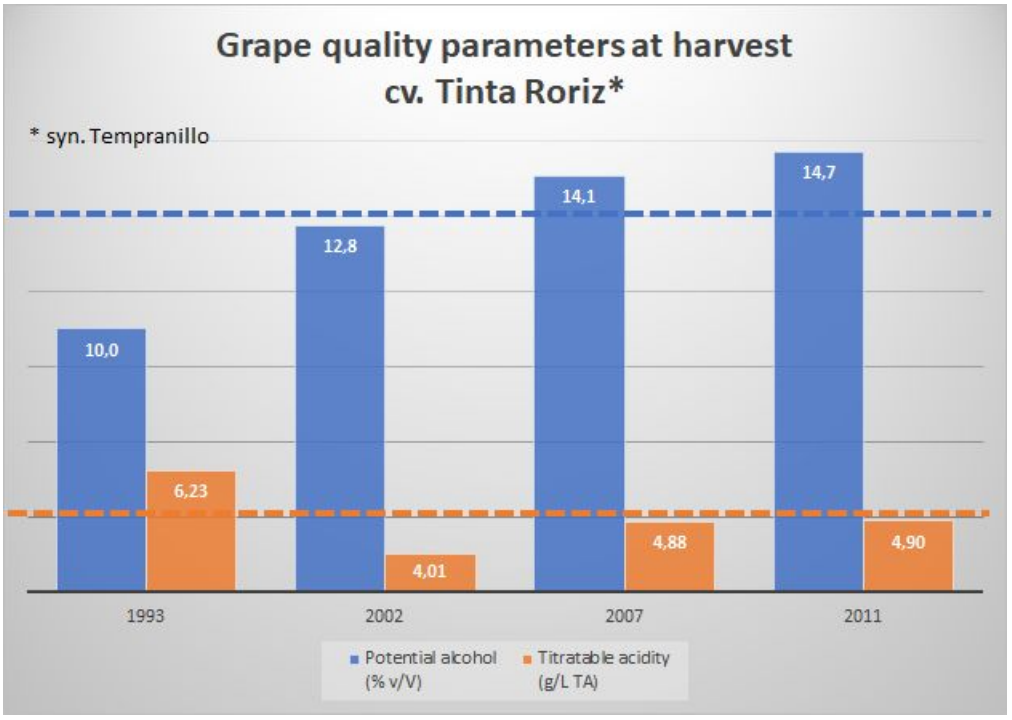


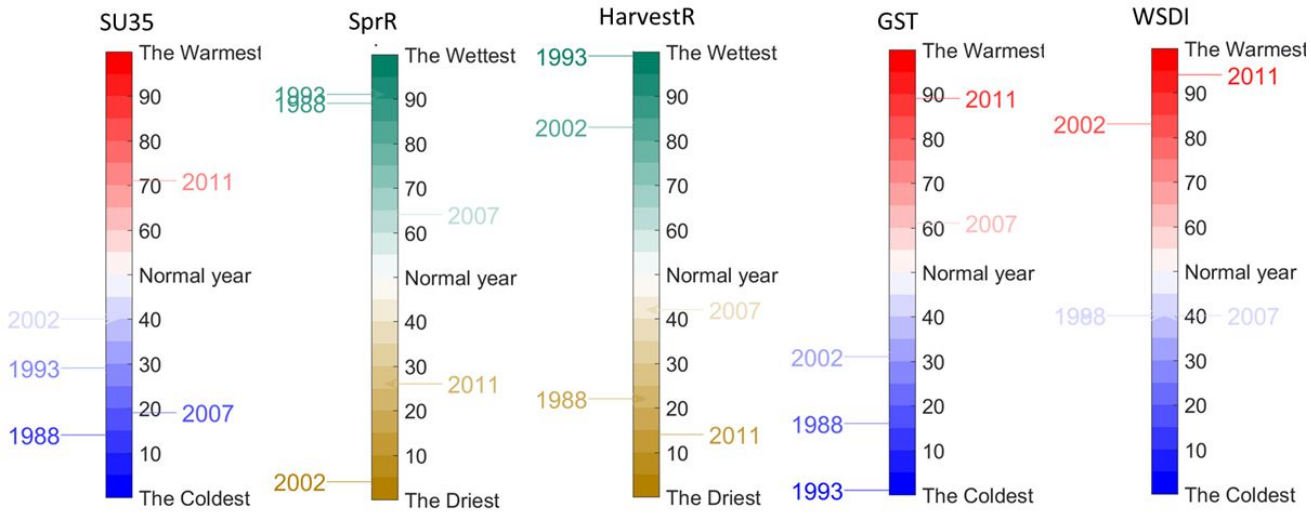
Figure 4-3 Grape quality parameters for a Douro wine region representative grape variety in chosen years as recorded in SOGRAPE vineyards (dashed lines indicate target values).



4.3 'Good' and 'Bad' Years in Terms of Bioclimatic Indicators

The characterization of 'good' and 'bad' years in terms of the selected bioclimatic indicators described above from the IPMA Weather station of Santa Barbara [RD.1] is reported only as an example in figure 4-4. The weather station of Santa Barbara (Lat=41.173; Lon=-7.549) was chosen as a representative sample of IPMA weather stations with a long time series and very little data missing, and because it is very close to one of the SOGRAPE vineyards in the Douro region. All the chosen bioclimatic indicators are reported here in terms of percentile, ranked over the reference period 1979-2018. On the left side of each colorbar, the ranking of the 'bad' years is reported, while on the right side are reported the 'good' ones. Similar results have been obtained by considering the closest grid-point to the IPMA weather station of Santa Barbara from several available gridded observational datasets (ERA5, E-OBS v19, PTHRES,...) and are not reported here. More details will be reported in the forthcoming deliverable 3.3 "Report on the climatic, bioclimatic and extreme climate indices developed in the wine pilot services".

Figure 4-4 Ranking of 'bad' (on the left side of each bars) and 'good' years (on the right side of each bars) in terms of percentiles of the selected bioclimatic indicators for the representative IPMA weather station of Santa Barbara.



The color bars shown in figure 4-4 are quite consistent with the narrative description of the case years just above reported. The SprR is also quite consistent with the sanitary periods represented in figure 4-1. The representation of 'good'/'bad' years in terms of percentiles of bioclimatic indicators was presented to SOGRAPE users on the occasion of the Workshop of the 27th of May 2019 at SOGRAPE's facilities in Avintes, Vila Nova de Gaia, fully described in [RD.3]. Summary of some of the main features shown in figure 4-4.

- The 'bad' years 1988 and 1993 are characterised by a very high value of SprR and concomitantly low GST values (below the 30th percentile).
- 1993 exhibits the highest HarvestR of the entire period.
- The 'good' years 2007 and 2011 show above-normal GST values (above the 60th percentile and near the 90th percentile, respectively), while all 'bad' years exhibit GST below-normal.

Important details on the capability of the bioclimatic indicators of interest for the wine pilot service will be reported in the forthcoming MED-GOLD Deliverable 3.3 "Report on the climatic, bioclimatic and extreme climate indices developed in the wine pilot services". In addition, in [RD.1] the capability of several gridded observational datasets to reproduce the selected bioclimatic indicators has already been widely reported and analysed for the region of main interest for the MED-GOLD wine pilot service, the Douro region in Portugal.

4.2 Development of the MED-GOLD Compound Wine Risk Indices

Looking at the results presented in figure 4-4, it is rather difficult to identify a 'good' or a 'bad' year considering only the ranking of a single bioclimatic indicator (with a partial exception for GST): for instance, two years with a completely opposite outcome for the wine sector in the Douro region (namely 1988 and 2011) exhibit quite similar values for HarvestR. On the other hand, three 'bad' years are at the two opposite ends of the SprR ranking (1988 and 1993 around the 90th percentile, while 2002 is below the 10th percentile). The reason is that the final behaviour of a specific season is generally due to the concomitance of several factors that could increase the risk of, for instance, having a higher number of infestations, or a quality/quantity loss due to heat stress.

To better take this into account, ENEA and SOGRAPE have co-designed and co-developed two new compound risk indices that can integrate into a single measure the different sources of risk, during the growing season, for two relevant events that may have a relevant impact for the wine company:

- Pressure of fungal disease in grapes (Sanitary Risk Index)
- Heat stress (Heat Risk Index)

These two new compound risk indices were developed using the ranking of the bioclimatic indicators above selected weights to resemble the high-level narrative description of 'good' and 'bad' years. The overall methodological strategy is to use some new ad-hoc implemented integrated indices (based on and in addition to existing bioclimatic indicators) that can give an idea of the main risk factors that can affect wine production (in terms of quality, quantity and derived inherent value) in the Douro region. First, we had to test and tune the methodology on the past seasons, using observational datasets to assess whether this strategy can correctly reproduce what happened in the recent past from a wine production perspective (favorable or unfavorable conditions for obtaining value from wine production). Subsequently, we applied the same methodology to the output of available climate predictions (from seasonal to longer time scales).

4.2.1 MED-GOLD compound Sanitary Risk Index

The MED-GOLD compound Sanitary Risk Index integrates in a single measure the possible sources of fungal disease (mildew and rot) risk in the Douro region during a growing season (this is linked to critical decisions to be made during the season, such as the number of sanitary treatments to be applied). From the above description of 'good'/'bad' years and the judgement of SOGRAPE experts, possible sources of risk of infection by fungus have been identified:

1. High/Low SprR (with very high SprR, the overall sanitary risk is even higher; with high GST, the risk related to SprR is increased)
2. High HarvestR (especially in case of low GST)
3. Low GST

Starting from these considerations, a first version of a normalized compound Sanitary Risk Index was introduced based on the percentile values of GST, SprR, HarvestR, properly combined in order to adjust to the behaviour observed in the test-case years. The weights of the single coefficients and the rules reported below were finalised after several fine-tuning and double-checking interactions between ENEA and SOGRAPE.

$$Sanitary\ Risk\ Index = off_{gts} * off_{sp} * (percentile(SprR)) + off_{hart} * percentile(HarvestR) + (100 - percentile(GST);$$

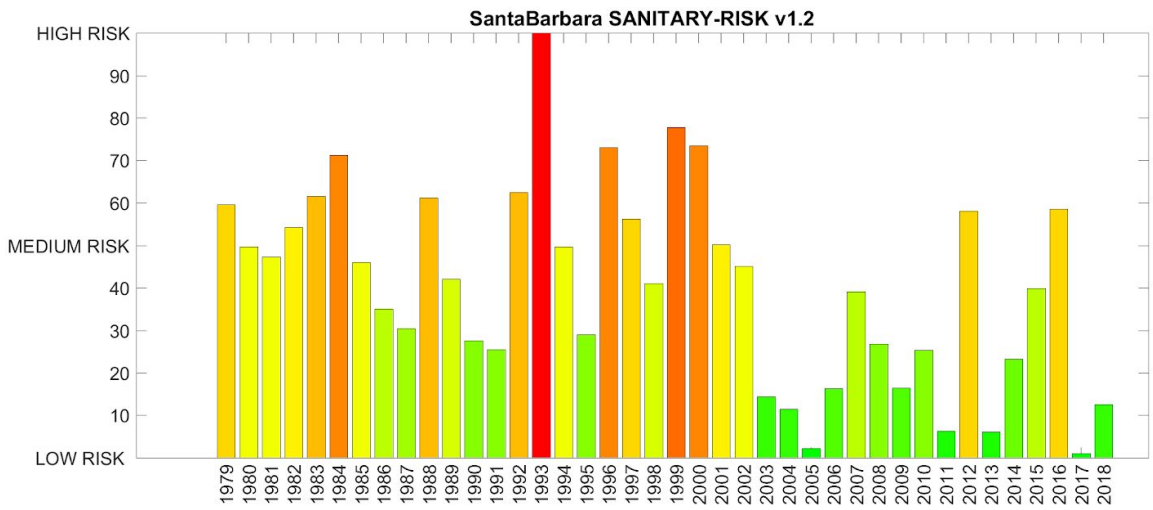
with: $off_{hart}=1.$; $off_{sp}=1$; $off_{gts}=1.$;

- ➔ *if percentile (SprR)>= 60;* $off_{sp}=1.5$;
- ➔ *if percentile (GST)<= 40;* $off_{hart}=1.5$;
- ➔ *if percentile (GST)>=70;* $off_{gts}=1.5$;

where off_{hart}, off_{sp}, off_{gts} are the coefficients for weighing the different sources of risk: harvest precipitation, spring precipitation and growing season temperature, respectively. For a season where the SprR exceeds the 60th percentile, we'll set off_{sp}=1.5: in such a way the weight of the risk associated with SprR is increased further by 50%. The risk associated with SprR is also increased if the GST is very high(i.e. if GST is above the 70th percentile we'll put off_{gts} *=1.5*). On the other hand, if the GST is below the 40th percentile, the grapes may ripen more slowly and the risk associated with high HarvestR is increased by setting off_{hart}*=1.5.*

The index is then reported to normalized units between 1 and 100 over a common reference period, the same chosen for computing the percentile distribution of the bioclimatic indicators. For instance, for the IPMA weather station of Santa Barbara, the reference period is 1979-2018 as shown in figure 4-5. When considering different datasets, the options for the reference period may be different.

Figure 4-5 Values of Sanitary Risk Index for the reference period 1979-2018 obtained from the representative IPMA Santa Barbara weather station data.



4.2.2 MED-GOLD Compound Heat Risk Index

The MED-GOLD compound Heat Risk Index integrates in a single measure the possible sources of heat stress risk in the Douro region during a growing season. From the above description of 'good'/'bad' years and the judgement of SOGRAPE experts, possible sources of heat stress risk for grapes have been identified:

1. High GST
2. High SU35
3. High WSDI

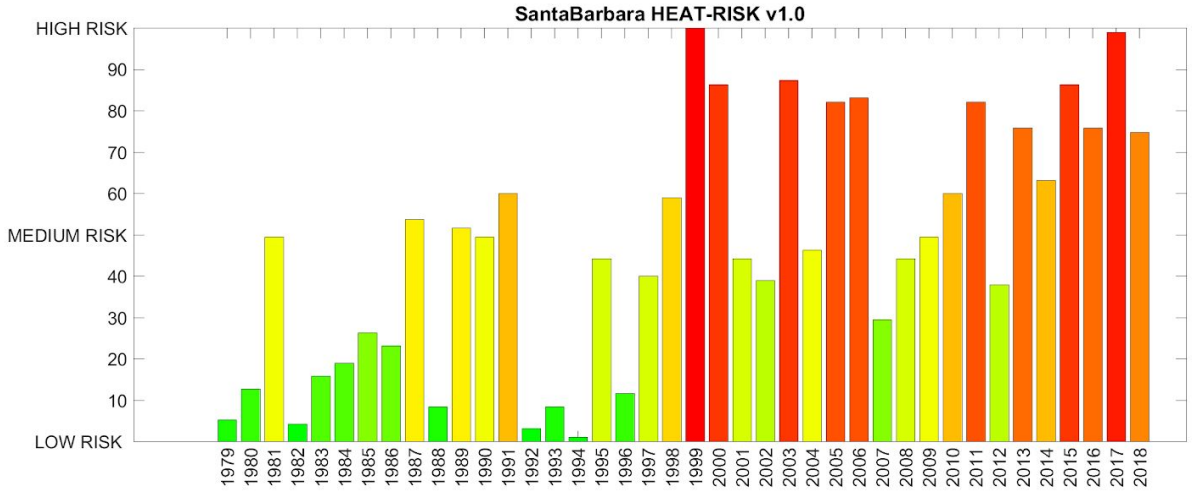
A first hypothesis of a heat risk index based on the percentile values of GST, SU35 and WSDI includes the same weight for all three factors:

$$Heat\ Risk\ Index\ =\ percentile(GST) + percentile(SU35\_AMJJASO)\ +\ percentile^{**}(WSDI\_AMJJASO).$$

***percentiles of WSDI have been readjusted putting equal to zero the value of percentile corresponding to seasons with WSDI=0 and then normalized to not overweight the seasons with 0 events.*

As for the Sanitary Risk Index, also the Heat Risk Index is then reported to normalized units between 1 and 100 over a common reference period 1979-2018 (for the IPMA Weather Station of Santa Barbara) as shown in figure 4-6 below.

Figure 4-6 Values of Heat Risk Index for the reference period 1979-2018 obtained from the representative IPMA Santa Barbara weather station data.



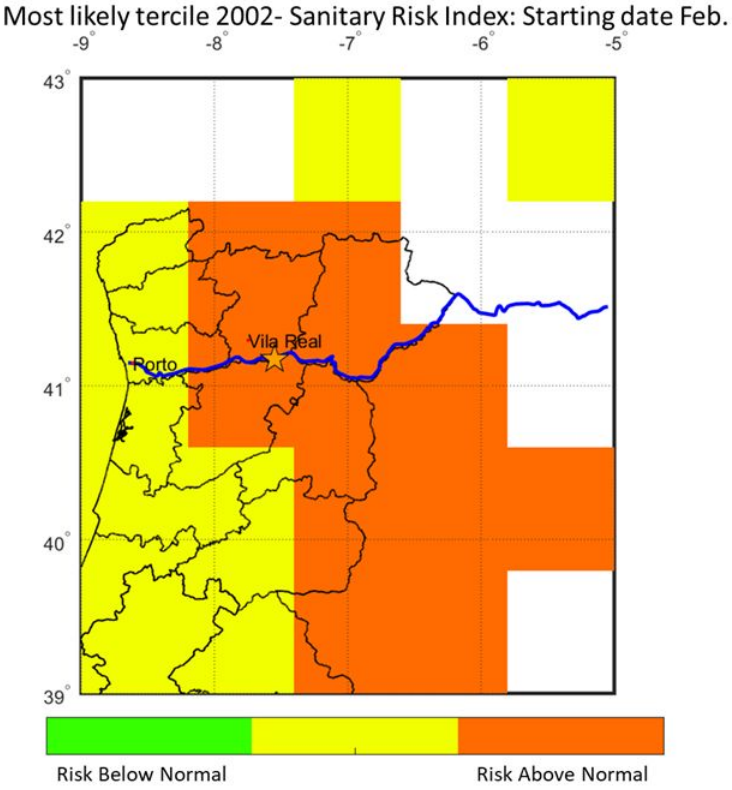
The formulas of the MED-GOLD compound wine risk indices can be further tuned and refined in the testing phase of the tool. Major details and outcomes will be reported in MED-GOLD Deliverable 3.3 "Report on the climatic, bioclimatic and extreme climate indices developed in the wine pilot services".

4.4 Methods for Development of the Climate Service based on Seasonal Predictions

As explained before, one of the components of the wine climate service is based on seasonal predictions. BSC developed this component of the climate service based on seasonal predictions obtained from CDS-C3S, in particular with the ECMWF SEAS5 dataset described in section 3. Data. The products obtained from this component of the climate service are based on seasonal predictions of essential climate variables (ECVs; i.e. mean temperature, maximum temperature, minimum temperature and precipitation), bioclimatic indicators (described in subsection 4.1) and risk indices (described in subsections 4.2). These products are displayed through the MED-GOLD Dashboard, as agreed from feedback gathered from users at the workshop held in May 2019 at SOGRAPE's facilities in Vila Nova de Gaia [RD.4].

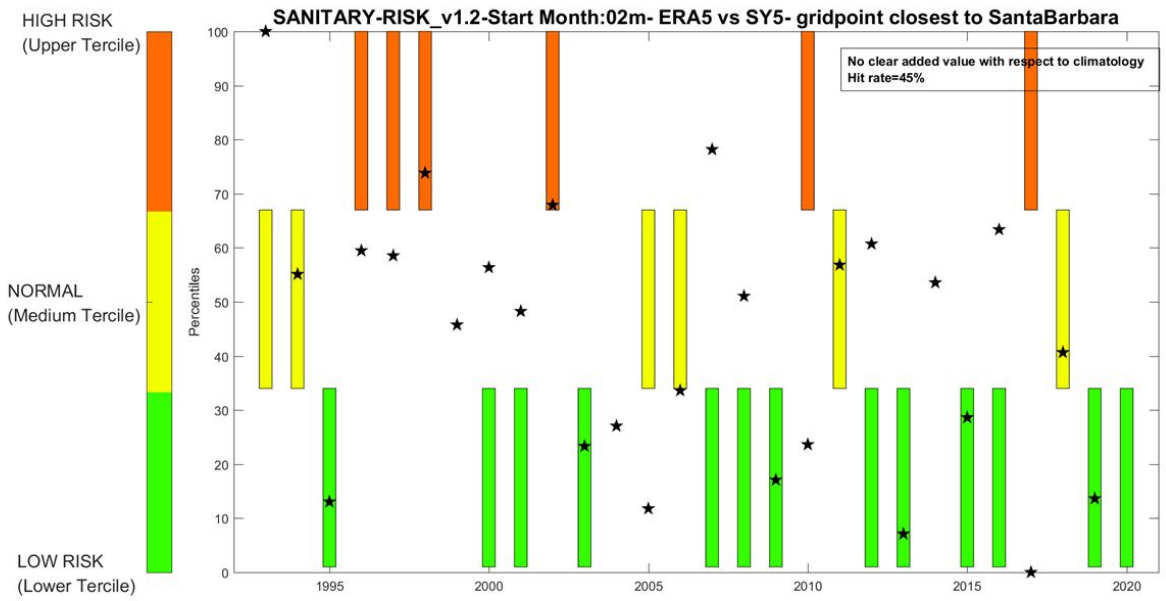
An illustration of the products that have been used as starting reference for the graphical solutions provided through the MED-GOLD Dashboard V1.0 is shown in figures 4-7, 4-8 and 4-9 for the Sanitary Risk Index. Figure 4-7 shows a map with the seasonal prediction of the most likely tercile category of the Sanitary Risk Index foreseen for year 2002 for Northern Portugal, including the Douro Valley. It should be noted here that in this first version of the Dashboard, the spatial resolution of the maps (1 degree) still needs to be updated to cover the Douro region with a higher resolution.

Figure 4-7 Most likely tercile category map for Sanitary Risk Index foreseen for year 2002. The dataset used to compute this map corresponds with the seasonal predictions issued in February coming from CDS-C3S ECMWF SEAS5.



In addition to the most likely tercile category map shown above, the end-user can select a particular grid point on the map to obtain information from the historical seasonal predictions for the period 1993-2020 (note that this period includes information from both CDS-C3S hindcast, 1993-2016, and forecasts, 2017-2020). Figure 4-8 shows the temporal evolution of the most likely tercile category of the Sanitary Risk Index obtained with seasonal predictions (coloured bars) in comparison with the observed values (black stars) for an example February start date. The skill of seasonal predictions (RPSS) is reported in the box at the top right of the figure (explained as the added value of using the predictions of the variable with respect to the use of climatology, i.e. RPSS >0). The percentage of times observation falls into the most likely tercile is also reported in the box. In principle this approach could inform the user whether the forecasts bring any additional value with respect to having climatological ones. The box information (score, hit rate) have not been included yet in the first release of Dashboard. Note that when it is not possible to determine the most likely tercile category (either the probability of the tercile categories does not exceed 40% or the probability of two tercile categories being equal) the information for this particular year is not shown.

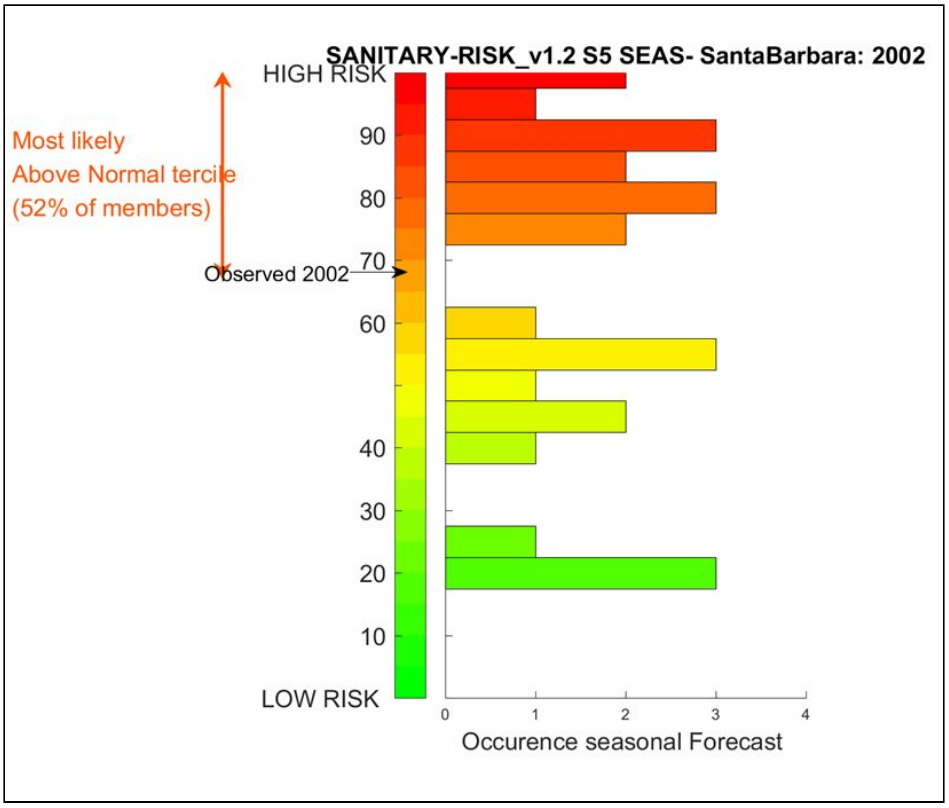
Figure 4-8 Temporal evolution of the most likely tercile category of the Sanitary Risk Index obtained from the seasonal predictions issued in February from the CDS-C3S ECMWF SEAS5 (coloured bars) to the closest grid point of the weather station of Santa Barbara (Douro Valley, Portugal). The stars show the observed value (in percentile) of the Sanitary Risk Index obtained with the ERA5 reanalysis dataset.



End-users can also obtain the probabilistic distribution of the seasonal prediction for a particular grid point and year in comparison with observation, as shown in figure 4-9 for the Sanitary Risk Index over the Santa Barbara weather station.

Even though the figures selected here for illustration purposes refer to products based on the Sanitary Risk Index, similar products based on the rest of the risk indices, bioclimatic indicators and ECVs are available to the end-user from the MED-GOLD Dashboard.

Figure 4-9 Probabilistic distribution of Sanitary Risk Index for the year 2002 obtained from the seasonal predictions issued in February from the CDS-C3S ECMWF SEAS5 to the closest grid point of the weather station Santa Barbara (Douro Valley, Portugal) in comparison with observed value in the ERA5 reanalysis dataset.



Nevertheless, before obtaining the final products of the service, described above, several post-processing techniques have been applied to the seasonal predictions to ensure they are tailored to the end-users' needs. They consist of: (i) bias adjusting the predictions taking both reanalysis ERA5 and observational PTHRES dataset as reference (described in section 3. Data), (ii) interpolating the predictions for the Douro region and (iii) assessing the quality of the predictions through a verification process. A complete description of the methods used by BSC in each case is described in the following subsections. In addition, to choose the best workflow of the seasonal service, BSC has carried out a study on how the order of sequential application of the interpolation and bias correction methods affects the quality of the predictions. The results and main conclusions drawn from this study guided the final workflow of the service and are described in subsection 4.3.3.

In addition to the methods mentioned above, BSC has respectively developed two methodologies for the provision of the climate service with real-time predictions and for the economic assessment of using retrospective seasonal predictions in a decision-making context. These two methodologies are described in the subsections 4.3.5 and 4.3.6.

4.3.1 Simple Bias Correction (SBC)

The state-of-the-art forecast systems produce climate estimates with systematic errors due to model imperfections, as well as other sources of uncertainty, such as initial and boundary conditions. Thus, the bias in predictions needs to be adjusted before the provision of the forecast to the end-user. After testing different bias adjustment techniques, the simple bias correction approach has been chosen for implementing the pilot wine service workflow.

Simple bias correction generates an ensemble of predicted datasets in which the mean and standard deviation are the same as the observations. Note that it assumes that the distributions of both the observational and predicted datasets are close to the Gaussian (normal) distribution. Most of the time, this assumption is valid for monthly/seasonal mean data. This method has been widely used to correct temperature and precipitation [RD.11]. The simple bias correction can be formulated as follows,

$$y_{j,i} = \left(x_{ij} - \bar{x}\right) \frac{\sigma_{obs}}{\sigma_{mod}} + \bar{o}$$

where $y_{j,i}$ is the bias corrected forecast, computed by adjusting the anomalies with the ratio of standard deviation of observation to hindcast and adding the climatological observation (denoted as $\bar{o}$). The daily anomaly is calculated by subtracting the ensemble mean x ($\bar{x}$) of the hindcast data set from the daily value (x_{ij}) for each member *i* and each year *j*. σ_{obs} (σ_{mod}) is the standard deviation of observation (hindcast).

This bias correction is performed for each grid point separately, resulting in a new forecast ensemble that has the identical ensemble mean and standard deviation as the observation. The underlying assumption of the simple bias correction is that the forecast distributions and the observation data are expected to follow the Gaussian distribution. Therefore, a 15-day moving window is further applied to make the distribution of datasets closer to the Gaussian distribution. The formula is given below.

$$y_{j,i} = \left(x_{ij} - \overline{x_{15-MW}}\right) \frac{\sigma_{obs,\,15-MW}}{\sigma_{mod,\,15-MW}} + \overline{o_{15-MW}}$$

The equation is almost identical to the simple bias correction shown previously. The key difference is that we take the 7 days before and after the computing day (15 days in total) when computing the two standard deviations ($\sigma_{obs,\,15-MW}$ and $\sigma_{mod,\,15-MW}$) and the two climatologies ($\overline{x_{15-MW}}$ and $\overline{o_{15-MW}}$). For instance, to correct the bias of the 10th day of the second year of the series, hindcast

values from the third to the 17th day of all the years (except for the second year) for all ensemble members are used to compute the standard deviation and the ensemble-mean climatology. Similarly, the observational values from the same 15 days of all years (except for that second year) are taken for the calculation of observational average and standard deviation.

Please note that the two climatologies and the two standard deviations remain the same for the first seven days to simplify the need of extension of time series. This applies to the last seven days as well.

4.3.2 Data Interpolation

Since the spatial resolutions of SEAS5 and PTHRES datasets are 1 degree and 1 km, respectively, the Inverse Distance Weighting (IDW) method is applied to interpolate SEAS5 data set to the PTHRES resolution as a downscaling method.

IDW takes the values of a set of known points nearby the target location and deterministically computes a weighted average taking into account the distance between the target location and each known point selected. The formula is given as below.

$$y_p = \frac{\sum_{i=1}^{n} \left(\frac{x_i}{{d_i}^p}\right)}{\sum_{i=1}^{n} \left(\frac{1}{{d_i}^p}\right)}$$

where y is the estimate at the unknown location. *n* is the number of known locations which are used for estimation. d_i is the Euclidean distance between the unknown location and each of the *ith* known locations with its value x_i. *p* is the power to the d_i. The IDW interpolation is applied with the power of one in preliminary experiments.

To better estimate the values over the spatial domain of interest at the finer resolution, only the two grid points of the CDS-C3S ECMWF SEAS5 forecast (i.e., 325°E & 353°E; 41°N) are selected as these points cover most of the Douro Valley.

4.3.3 Combination of Bias Correction and Interpolation

IDW interpolation together with the ordinary SBC are combined for adjusting biases and increasing the spatial resolution of forecasts. Otherwise, the forecasts cannot be used in decision-making due to the inherent biases of the forecasts system and its coarser resolution.

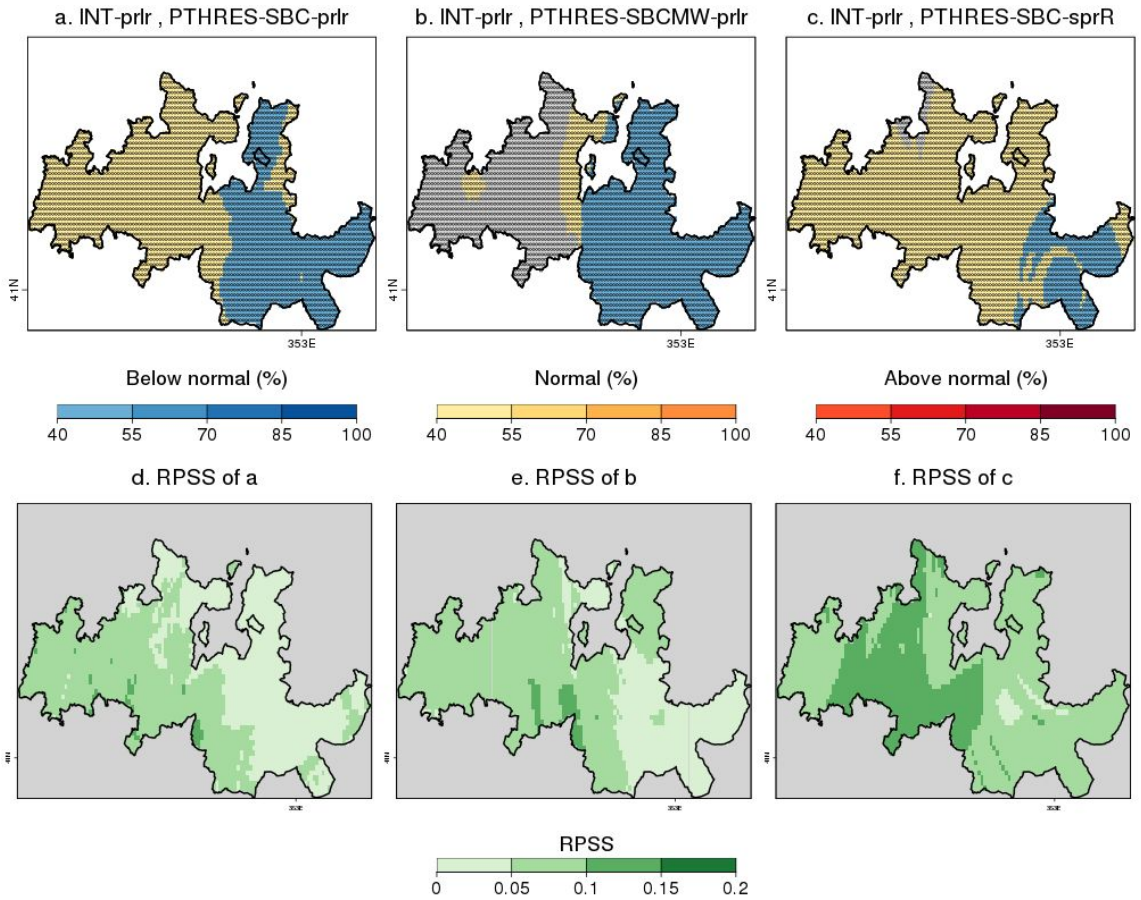
However, the aforementioned methods can be applied to either ECVs and bioclimatic indicators with a number of combinations. For example, one possible combination could be to interpolate the ECV

first, compute the indicator and, finally, bias correct it. Another would be to bias correct the ECV, compute the indicator and interpolate it to a finer resolution, but there are more possible combinations. The effect on the predictions' performance of the order in which these post-processing methods are applied is, to our knowledge, not well established in current literature (e.g., on the most likely tercile prediction).

Therefore, in order to clarify the impact of the order of application of the IDW and the SBC methods, BSC conducted a number of experiments in which all possible combinations for obtaining the SprR indicator from precipitation were assessed. The experiments were conducted using the ECMWF SEAS5 and PTHRES datasets for the Douro Valley and for the period 1993-2015 (with the starting date of April). The main findings were that changes in the position in which the interpolation (IDW) was applied didn't affect the predictions performance, but it did affect the position where bias adjustment was applied. Consequently, it was decided to firstly interpolate the seasonal predictions of precipitation to PTHRES resolution. In the second stage, in order to test how the position of bias adjustment affects the prediction of SprR indicator, three different configurations were analysed: (*i*) SBC was applied to the interpolated precipitation for obtaining the SprR; (*ii*) SBC with a 15-day moving window (hereafter SBCMW) was applied to the interpolated precipitation for calculating the SprR; (*iii)* the interpolated precipitation was aggregated to compute SprR before the application of SBC. Finally, the most likely tercile category and the associated FRPSS were calculated following the equations introduced in the previous subsections.

The results of the experiment with the three configurations described above are shown in figure 4-10 (taking the year 1993 as an example). The seasonal predictions of the most likely tercile category foresee normal conditions on western DV and below normal conditions on eastern DV for configurations *i* and *iii (*figure 4-10, panels *a* and *c* respectively*)* but for configuration *ii* western DV was not categorized (figure 4-10, panel *b*). When comparing FRPSS values (figure 4-10, panels d, e, f), the configuration with the highest values is *iii* (figure 4-10, panel f)*,* which corresponds to interpolation of the precipitation first, then calculating the SprR indicator and finally simple bias correcting the SprR indicator. The above experiment had been tested by using a temperature indicator and a similar result was found as well (results not shown). Thus, for the seasonal component of the wine climate service developed in WP3, the sequence chosen for the application of the IDW and SBC methods is the one described under configuration iii.

Fig 4-10 Seasonal predictions of the most likely tercile category of SprR (first row, panels a, b and c) and map of FRPSS values (second row, panels d, e and f) over the Douro Valley. (a) and (d) correspond to configuration i: SBC was applied to the interpolated precipitation for obtaining later the SprR, (b) and (e) corresponds to configuration ii: SBC with a 15-day moving window (SBCMW) was applied to the interpolated precipitation for calculating later the SprR; (c) and (f) corresponds to configuration iii: the interpolated precipitation was aggregated to compute SprR before the application of the SBC. The grid points cross represent FRPSS > 0 and the grey grid cells mean the probability of each category < 40%.



4.3.4 Quality assessment of seasonal predictions

To evaluate the quality of seasonal predictions for both non bias-corrected and bias-corrected predictions, Pearson's correlation and Fair Ranked Probability Skill Score are applied as verification metrics for the indices. The methods are described as follows.

Pearson's correlation

Pearson correlation coefficient [RD.12] between the ensemble-mean predicted and the observed is used as a measure of the linear correspondence between the retrospective prediction and the reference. This can be defined as:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}$$

where x_i and y_i are the observed and the ensemble-mean predicted values in each season over the *i* = 1, 2, … *n* years. The $\bar{x}$ and $\bar{y}$ are the average of the ensemble-mean predictions and the observations over the *n* years.

The correlation coefficient ranges between -1 and 1. If r_{xy} = 1 there is a perfect association between the ensemble-mean of predictions and observations. When r_{xy} = 0 indicates that there is no association between the ensemble-mean of the predictions and the reference dataset, which in turn shows that the ensemble-mean of the predictions does not provide any added value with respect to retrospective climatology. A negative correlation coefficient indicates that the observed climatology should be used instead of the predictions. In addition to its predictability, only positive values of correlation coefficient indicate that the seasonal predictions are able to provide added values because of the need for accurate data distribution [RD.13].

Fair Ranked Probability Skill Score (FRPSS)

For the evaluation of categorical events (e.g., below normal, normal and above normal here) obtained from probabilistic predictions, the Ranked Probability Score (RPS; [RD.12]) is one of the comprehensive verification metrics that has been frequently used in the context of seasonal predictions.

RPS is the sum of the squared distance between the cumulative probabilities of the n predictions-reference pairs (for the entire inter-annual time series) for the *k* equiprobable forecast categories (e.g., tercile when *k* = 3). The formula is given as follows:

$$RPS = \frac{1}{n}\sum_{i=1}^{n} \sum_{k=1}^{K} \left[\left(\sum_{j=1}^{k} y_{i,j}\right) - \left(\sum_{j=1}^{k} x_{i,j}\right)\right]^2$$

where $y_{i,j}$ and $x_{i,j}$ are the cumulative predicted and observed probabilities, respectively, assigned by the ith forecast (*i* = 1, … *n*) to the kth category (*i* = 1, … *k*). The $x_{i,j}$ = 1 indicates that the observation is in category *k*, and $x_{i,j}$ = 0 otherwise.

To better infer the predictive skill of forecast using RPS, it is often expressed as a skill score (i.e., Ranked Probability Skill Score, RPSS) which represents the added value of the prediction relative to the observations (usually referring to the climatology). The RPSS is given by:

$$RPSS = 1 - \frac{RPS}{RPS_{clim}}$$

As shown in the equation, RPSS ranges from −∞ to one. The prediction is considered as unskillful when a negative RPSS returns. In other words, climatology would be preferable in this case, because the predictions do not provide additional value. In the case of a positive RPSS, the higher the RPSS, the better the predictions are than the climatology. RPSS = 1 corresponds to a 'perfect' prediction.

The RPSSs shown in this deliverable have been calculated for the verification of terciles (three equiprobable categories which are related to the two thresholds of the climatological distribution of the reference data set). The probabilities of forecast have been computed as the fraction of ensemble members (leave-one-out cross-validation) in the corresponding category.

Fair (Skill) Scores to ensemble predictions ([RD.14-15]) favors the prediction with its ensemble members being performed as if they had been sampled from the same distribution than the reference data set. The fair version of the RPSS (i.e., FRPSS) estimates the skill of prediction when an infinite ensemble size is used (a measure of potential skill).

4.3.5 Real-Time Seasonal Predictions for the Year 2020

The methodologies reported in the development of the seasonal component of the wine pilot service were first implemented considering the seasonal hindcast period available through CDS-C3S, 1993 to 2016. Afterwards, in the GA 2019, the wine Champion SOGRAPE pointed out the importance of having real-time examples of this bias corrected climate information for a better assessment of its potential added value. In this framework, BSC assembled a real-time workflow example considering

the predictions issued in February 2020 for the six bioclimatic indicators developed and the ECVs. In spring 2020, real-time predictions of temperature, precipitation, GST, SprR and HarvestR as well as the sanitary and heat stress indices were displayed in the Dashboard by ENEA and BeetoBit. Next year it is planned that all six bioclimatic indicators together with the ECVs and the two risk indices will be available through the Dashboard.

At the time of the calculations of the bioclimatic indicators, the available seasonal predictions were those issued in February 2020 and covering the period February to August 2020. This temporal coverage of the seasonal predictions for the year 2020 was made in order to include some adaptations to the data to allow the calculation of the indicators. For instance, taking the GST as an example, we issued the forecast in February. Since the GST covers the April-October months, this means that the SEAS5 prediction only covered the April-August period. The remaining two months, September and October, were included in the GST through their SEAS5 hindcast climatological values (from 1993 to 2019).

For SprR, on the other hand, the prediction issued in February covered all the index aggregation months and, thus, no additional step was needed to complete it. Finally, for HarvestR, since it ranges from 21st August to 21st October, only the last days of August came from the February forecast. The rest was taken from the SEAS5 hindcast climatology.

4.3.6 Methodology for Economic Assessment of Using Retrospective Seasonal Predictions in a Decision-Making Context

During the first interactions with SOGRAPE (MED-GOLD GA 2018 in Porto), a gap was identified between the understanding of seasonal forecast information and its actual use in decision-making. In fact, end-users identified the difficulty of using climate predictions from past years (also known as hindcast) to test the payoff for seasonal forecasts against their customary methods. As a result, several interactions between the seasonal climate predictions provider (BSC) and the end-user (SOGRAPE) have been taken to co-develop an illustrative methodological example (case study) of the economic impact of using seasonal predictions for a specific decision-making using a long retrospective period (1993-2015).

This case study was carefully selected together with SOGRAPE and involved the use of seasonal forecasts of Spring Rain (SprR) to make decisions regarding *plant protection* and *canopy management* to compare the economic outcome with the customary use of climatology. The underlying question needing answers from a vineyard decision-maker perspective were: when is the signal offered by a forecast useful? What are the risks that I am willing to bear? How and when can forecasts influence these choices? And, perhaps more importantly, how does the use of seasonal forecast information compare with our current methods?

The approach is based on a methodology widely used in climate services evaluation: decision theory. BSC applies a retrospective analysis that relies on the comparison of hindcasts and observations to simulate changes in decisions triggered by information provided by seasonal forecasts (the service). Ultimately, the value of the service is equivalent to the value of the information conveyed by the service to the final user [RD.6].

The decision theory involves a single agent, the climate service user in our case, who has to make a decision in order to maximise (or minimise) an objective. In climate services evaluation, it is normally assumed that the agent has some prior knowledge corresponding to the climatological probabilities of the distribution of a climate variable. In the absence of climate forecasts, the agent makes the decision that maximises her/his expected payoffs based on climatological probabilities. Instead, when the agent is using a climate service s/he updates expectations based on the forecasts' probability distribution and makes the decision accordingly. The value of the forecasts is the difference in payoffs between making the decision with climate service and without [RD.7].

The theory works only under the assumptions that the agent's decisions are independent from other agents' choices and institutional setting. This implies that the agent's decisions are based only on her/his own payoffs and, moreover, that the value is user and decision specific. However, even if the results cannot be generalised to any decision or decision-maker, the protocol does so.

We simulate the farmer's decisions each year based on the probabilities associated with the 3 scenarios provided by the seasonal forecast (A_{sf}). The farmer has to make a decision on which action (A) at time t will have different outcomes depending on climate conditions at time $t+1$. The climatic conditions are summarised in 3 possible scenarios, in line with forecasts terciles:

- Above Normal (*An*)
- Normal (*N*)
- Below Normal (*Bn*)

The "normal" level is defined by the climatological average for the variable, location and period under study. The same decision will lead to different outcomes depending on the verified climate condition. The outcome is measured in yields (Y). Moreover, each A has different costs (c) associated. The farmer aims to achieve the highest payoff ($\prod$), meaning the highest yield at the lowest possible cost in the occurring scenario.

The farmer can choose A with or without seasonal forecast. The seasonal forecast provides the agent with a probability p associated with each of the 3 possible scenarios. On the other hand, without seasonal forecast, the farmer's actions are based on prior knowledge, which in decision theory studies is often associated with climatology (33% probability is associated with the occurrence of each scenario). The agent chooses A that maximises the expected payoff, given that each scenario has equal probabilities.

However, in reality, many decisions in farming are based on climate memory (meaning personal memory of the climate conditions of recent years) or do not take into account climate risk. Farmers are increasingly aware of climate risks, but changing strategy often entails additional costs and may entail other risks. Therefore, it might happen that the farmer repeats the same strategy every year (defined as Business-as-Usual, BaU) regardless of possible climate anomalies. If extreme weather occurs, the farmer will take last minute measures to preserve crops from damage. Clearly, with late adjustments to the strategy, the farmer often incurs higher costs and a higher risk of yield loss compared to timely intervention. To better capture reality, for the decision where this happens, our model will take *BaU strategy* as previous knowledge.

We then do a retrospective analysis using hindcasts and comparing them with observations. We simulate the farmer's decisions each year based on the probabilities associated with the 3 scenarios provided by the seasonal forecast (A_{sf}). We then compare it with the corresponding observation and derive the payoff of *A*. The payoff is compared with the payoff of the same decision made based on previous knowledge or BaU (A_{BaU}). The difference in payoffs of the actions performed with and without seasonal forecasts represents the value of the climate service for the farmer (V) in the year *x*.

$$V_x = \prod_x(A_{sf}) - \prod_x(A_{BaU})$$

We then repeat the exercise for many years (as many as the hindcasts available, 1993-2015 in our case) to infer the value of the forecast over a period of time.

From Forecast to Action

The seasonal forecast provides the farmer with the probability associated with the 3 possible scenarios (p_{An}, p_N, p_{Bn}). How are these probabilities translated into action? It depends on the decision strategy of the farmer. We tested different decision strategies:

- *Most likely tercile*. The farmer chooses the action that maximises the payoffs in the scenario where the highest probability occurs. This seems a logical choice, but it may not always be the case. In fact, depending on the costs of *A*, potential benefits associated, and alternative options, the farmer may give different weights to the probabilities of the scenarios.

- *Expected payoffs.* Normally, the farmer should choose the action that maximises her/his expected payoffs given p_{An}, p_N, $p_{Bn.}$ However, when calculating $\prod$ we do not account for other important factors such as risk aversion, competitor's choices or the institutional setting. Therefore, this does not necessarily correspond to the action they would actually do given p_{An}, p_N, $p_{Bn.}$ It remains an important strategy to assess and easily applicable in an ex-ante analysis as well.

- *Probability thresholds*. The farmer takes action when a certain probability associated with a scenario is reached. Sometimes a small probability is enough to make the farmer react because the impact associated with the occurrence of such a scenario is extremely high. On the other hand, it may take a very strong signal of an event to occur to make the farmer act to protect the crops from the impacts of the scenario. We test the payoffs of taking an action when different probabilities are associated with the scenario. Farmers are particularly interested in setting minimum thresholds to rely on a forecast. This is, of course, user and decision-specific. Repeating the exercise for different thresholds will allow us to discuss with farmers the optimal probability level based on observed payoffs, their yearly volatility and any other relevant factor. Ultimately, this process aims to support the farmer in integrating seasonal forecasts in her/his decision-making process.

From Action to Value

Regardless of the decision strategy we are using, the forecasts are assigning a probability to each scenario, while in reality only one scenario will occur. Once the farmer has taken action using the seasonal forecast, there are 3 possible consequences: Hit, Miss, 1-level False Alarm and 2-level False Alarm.

Table 4-1 Possible combinations of tercile forecasts and observations

	Observation		
Forecast	*An*	*N*	*Bn*
An	Hit	1-level False Alarm	2-level False Alarm
N	Miss	Hit	Miss
Bn	2-level False Alarm	1-level False Alarm	Hit

A hit allows the user to achieve the highest yields at the lowest cost. A miss can negatively affect yield and/or increase costs compared to a hit. With a 1-level False Alarm, the farmer incurs in higher and unnecessary costs, but generally without yield losses. The 2-level False Alarm indicates extreme cases when the farmer makes a decision expecting an anomaly (for instance a rainier than normal Spring) and the reality turns out to be the opposite (dry Spring). This situation is usually a high-damage situation, characterised by heavy yield losses and extra costs. As previously mentioned, the value of forecasts corresponds to the difference in payoffs of the decision made with and without forecast. In other words, it corresponds to the value of the information conveyed to the farmer. This implies that although the forecast allows the farmer to choose the correct *A,* if the farmer would have taken the same decision with prior knowledge (or BaU), the forecast delivers no value. By the same token, even if *A* chosen according to the forecast is not the optimal decision, if this leads to a higher payoff than the one without forecast, the value is positive because it has

delivered an improvement in decision-making. The results obtained from this study will be reported in D3.4 “Assessment of the added value for the decision-making process for the wine sector”.

4.5 Methods for Development of the Climate Service Based on Climate Change Projections

4.5.1 Regional Climate Models Evaluation

The RCM simulations used in the project have an horizontal resolution of 0.11° (about 12 km), and are compared with the E-OBSv17 dataset (resolution 0.22°; approximately 25 km) for the period 1971-2000. The E-OBSv17 data have a lower resolution than the RCMs, therefore the RCM simulations were aggregated in the 0.22° grid used by E-OBSv17. More specifically, the aggregation was achieved by calculating an area-weighted average of all grid cells of the RCM grid that overlapped each of the E-OBS grid boxes [RD.24].

The evaluation analysis regarding the models performance against E-OBSv17 was performed for a selected number among the 27 ETCCDI (Expert Team on Climate Change Detection and Indices) core climate change indices for both the Mediterranean and the Douro Valley, with the analysis revealing considerable biases for the examined temperatures and precipitation indices respectively (RD.2). Therefore, we opted to perform bias adjustment to each model’s output.

4.5.2 Testing Bias Adjustment Methods

The following bias adjustment methods were considered as candidates to bias adjust the climate projections in MED-GOLD:

- **Variance Scaling of Temperature**: the method corrects the mean and variance of temperature time series. It is only applicable for temperature.

- **Empirical Quantile Mapping (EQM):** this is a very versatile bias correction method which consists on calibrating the simulated Cumulative Distribution Function (CDF), by adding to the observed quantiles both the mean delta change and the individual delta changes in the corresponding quantiles. This method is applicable to any kind of variable.

- **Cumulative Distribution Function-transformation (CDFt):** The method aims to link the cumulative distribution function (CDF) of a large-scale variable with the CDF of the same variable at a much smaller scale, and to downscale and correct CDFs from which local-scale data can be generated. When observations are of similar resolution as the climate model, CDFt can be viewed as a bias-adjustment method. This method is applicable to any kind of variable.

- **Parametric Quantile Mapping (PQM)**. It is based on the initial assumption that both observed and simulated intensity distributions are well approximated by a given distribution, therefore is a parametric q-q map that uses the theoretical instead of the empirical distribution. For instance, the gamma distribution is described in [RD.25] and is applicable to precipitation.

- **Generalized Quantile Mapping (GPQM)** is also a parametric quantile mapping method but uses two theoretical distributions, the gamma distribution and Generalized Pareto Distribution (GPD). By default, it applies a gamma distribution to values under the threshold given by the 95th percentile and a GPD to values above the threshold. The threshold above which the GPD is fitted is the 95th percentile of the observed and the predicted wet-day distribution, respectively. For variables other than precipitation, values below the 5th percentile of the observed and the predicted distributions are additionally fitted using GPD and the rest of the values of the distributions are fitted using a normal distribution.

All methods except CDFt are part of the climate4R statistical toolbox ([RD.19]; [RD.26]) whereas CDFt [RD.27] can also be found as a separate package in the R statistical computing project. In order to assess and compare the performance of the different bias adjusting methods the cross validation framework was used. More specifically, cross validation tests whether the relationship established between the predictor (RCMs) and predictand (E-OBSv17) remains valid outside the training period [RD.28]. In the framework of MED-GOLD, the available data (n = 30 years) were partitioned into k-non-overlapping “folds” or subsets, each containing n/k elements (k=5 1971:1976, 1977:1982, 1983:1988, 1989:1994,1995:2000). The bias adjusted methods were then calibrated and validated k times, considering each of the folds as a test set and training the method with the remaining (k−1) sets. The resulting k-test series are typically joined and validated together in a single series spanning the whole analysis period. The major findings of the comparison can be summarized as follows. 1) Regarding the annual cycle, all methods capture the observed annual cycle of the essential climate variables in the Douro Valley, with no significant differences among the methods. 2) Among the methods a better performance was found for the EQM method for threshold based indices, such as the number of days with daily maximum temperatures above 35°C. Based on these findings, NOA opted to use the EQM method for bias adjusting of RCM simulations in MED-GOLD. For more details on the EQM method and its implementation, the reader is referred to the studies of [RD.19] and [RD.20].

4.5.3 PTHRES Integration

During the MED-GOLD progress, PTHRES was adapted for the needs of the WP3. In order to integrate PTHRES in the WP3 workflow, the RCMs daily output were remapped onto the PTHRES grid using bilinear interpolation and the EQM bias adjustment method was then performed. Consequently, the results were obtained for the selected bioclimatic indicators as well as for the ECVs. For each one of the indices and for each grid point over the Douro Valley, the differences between each one of the the future periods and the reference one are considered robust; the

changes in at least three out of five models are found statistically significant and the change in the same models is of the same sign. The first criterion is examined by using the 95th percentile confidence intervals as derived by bootstrap [RD.21-22]. If only one of the criteria is met, the change at the specific grid point is not considered significant. This simple and transparent method, which was proposed by [RD.23], summarizes multi-model projections and clearly separates lack of climate change signal from lack of model agreement by assessing the degree of consensus on the significance of the change, as well as the sign of the change. The main idea is that if multiple models agree on a result, there is a higher confidence than if the result is based on a single model, or if models disagree on the result. All results produced have been uploaded to the ICT platform and consequently presented on the WP3 dashboard.

4.6 Development of MED-GOLD Dashboard Interactive Tool

The overall process followed to co-design and co-develop the first release of MED-GOLD Dashboard (https://dashboard.med-gold.eu) has already been described in [RD.5], as it is a direct consequence of the interactions and feedback collected from the users. MED-GOLD Dashboard is a web-based application designed to provide a visualization tool for stakeholders, allowing them to browse and view, in a user-friendly way and without any programming knowledge or the need to manage climatic data files (typically in NetCDF or GRIB formats).

Users can interact with the dashboard via a graphical user interfaces, subdivided in three different sections (historical climate, seasonal forecast and long term projections) and organized in a step-by-step process (see figure 4-11 as example) that guides them through the selection of all the relevant parameters (such as time range, geographic location between Douro Valley and Iberian Peninsula, scenario type/forecast starting month, …) for the given section; the selected query is executed real time on the ICT platform and its results are dynamically plotted on a geographical map, using mapbox as a backend service, as a heatmap, whose specific colour palettes were chosen according to users' feedbacks. Clicking on the map shows more detailed information about the selected grid point according to the specific query. The data available since the first release (May 2020, see [RD.5]) are reported in table 4-2.

Figure 4-11 Step-by-step parameters selection control for Long Term Projections.

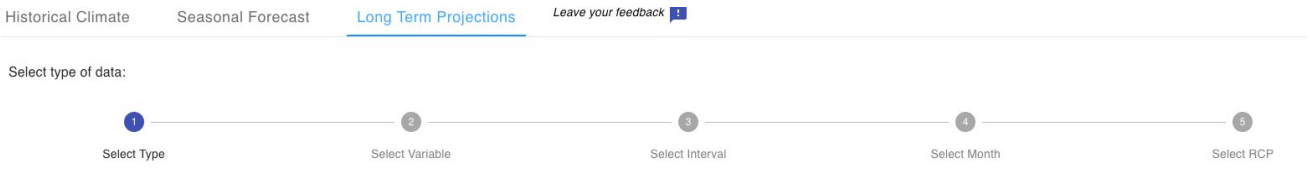
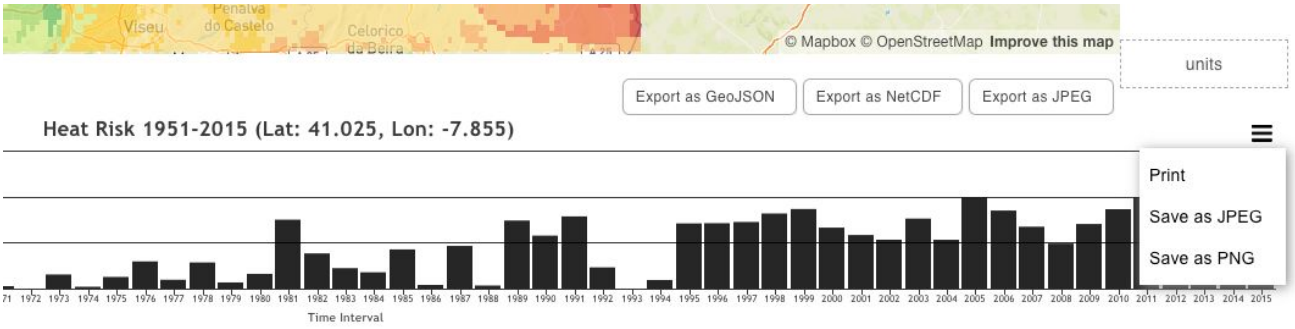


Table 4-2 Data availability on the first release of the MED-GOLD Dashboard (May 2020).

	Historical Climate	Seasonal Forecast	Long Term Projections
CLIMATE			
Precipitation monthly	V	V	V
Tmax monthly	V	V	V
Tmin monthly	V	V	V
Taverage monthly	V	V	V
BIOCLIMATIC			
GST	V	V	V
HarvestR	V	V	V
SprR	V	V	V
SU35	V	V	V
WSDI	V	X	V
GDD	X	X	V
WINE risk indexS			
Sanitary Risk	V	V	X
Heat Risk	V	X	X

Since the MED-GOLD Dashboard is meant to be a comprehensive environment for the stakeholders, it also provides simple data export options, allowing users to quickly export the data being visualized in any given screen for further off-line processing. Data can be exported (see figure 4-12) as NetCDF (standard format for array-oriented scientific data), GeoJSON (an open standard format designed for representing geographical features, useful for importing MED-GOLD data into mapping and GIS applications); plotted maps and graphs can be exported as image files (PNG or JPEG). As suggested by users' feedbacks, a CSV export option will be added as well during the next development iteration.

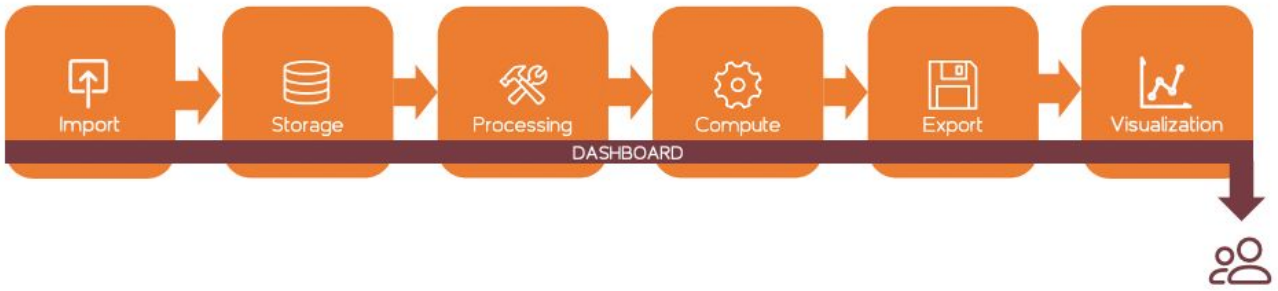
Figure 4-12 MED-GOLD Dashboard's data and image export options.



Furthermore, we also provide a Dashboard REST API that allows automatic and remote access to all the data and query capabilities. The REST API, described in more detail in ANNEX A, is, in fact, the technological foundation upon which the Dashboard itself is built and relies, in turn, on the data pipeline implemented by the ICT platform, that takes care of data import, storage, processing and format conversions. The availability of a REST API that provides the same functionalities of the Dashboard application, enabling further machine-to-machine integrations between external systems and the MED-GOLD platform, has been identified as an important point by users' feedback.

The diagram shown in figure 4-13 highlights the relationship between the Dashboard and the ICT platform in a concise way.

Figure 4-13 As shown in the diagram below, MED-GOLD Dashboard (dark red path) utilizes all sub-modules of the ICT Platform's data pipeline (in orange) in order to provide value to users.



Some screenshots for illustrative purposes of the Dashboard front-end for the three different sections (historical climate, seasonal forecast and long term projections) are reported in figures 4-14, 4-15 and 4.16, respectively.

Figure 4-14 Sanitary Risk Index from high-resolution observational dataset PTHRES for 2015 as reported in the MED-GOLD Dashboard front-end with the time series for a specific grid-point.

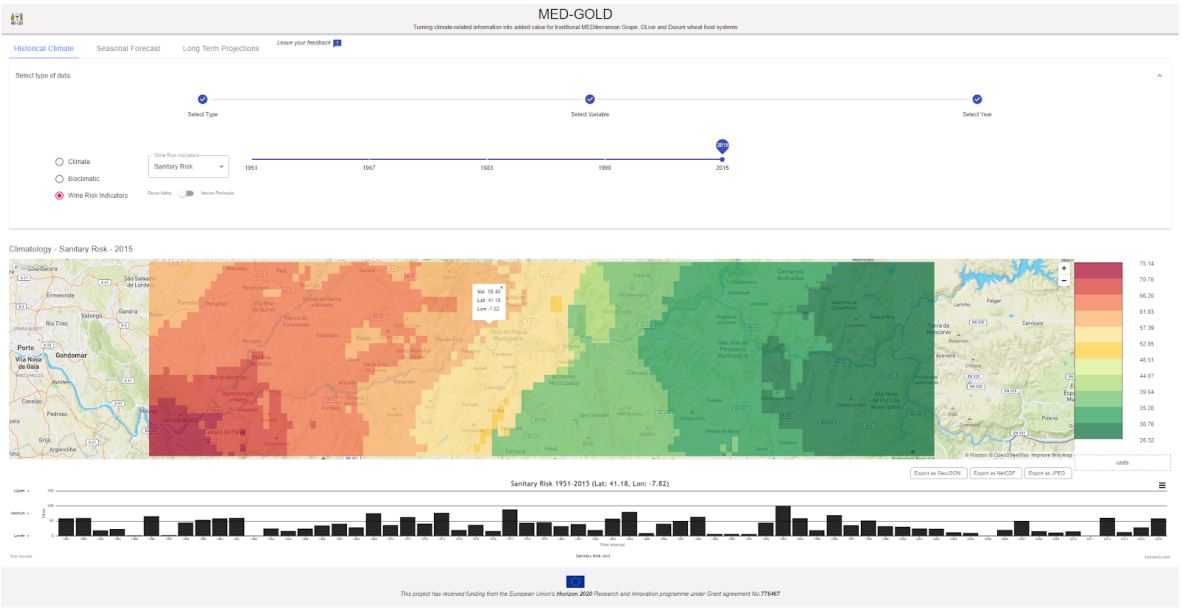


Figure 4-15 Seasonal forecast of Sanitary Risk Index for 2020: Map of the most likely tercile as reported in the MED-GOLD Dashboard front-end with the time series for a specific grid-point. The blue-dots are the reference values from ERA5, the squares in the time series are for the most likely tercile (if possible to assign) for the entire time series 1993-present.

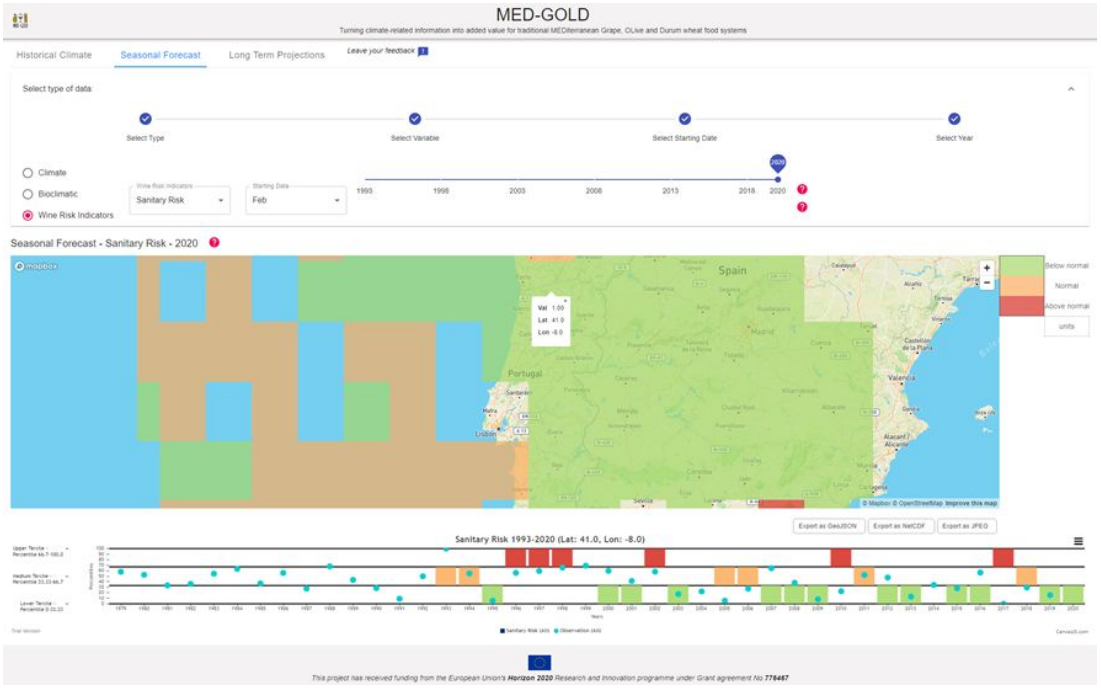
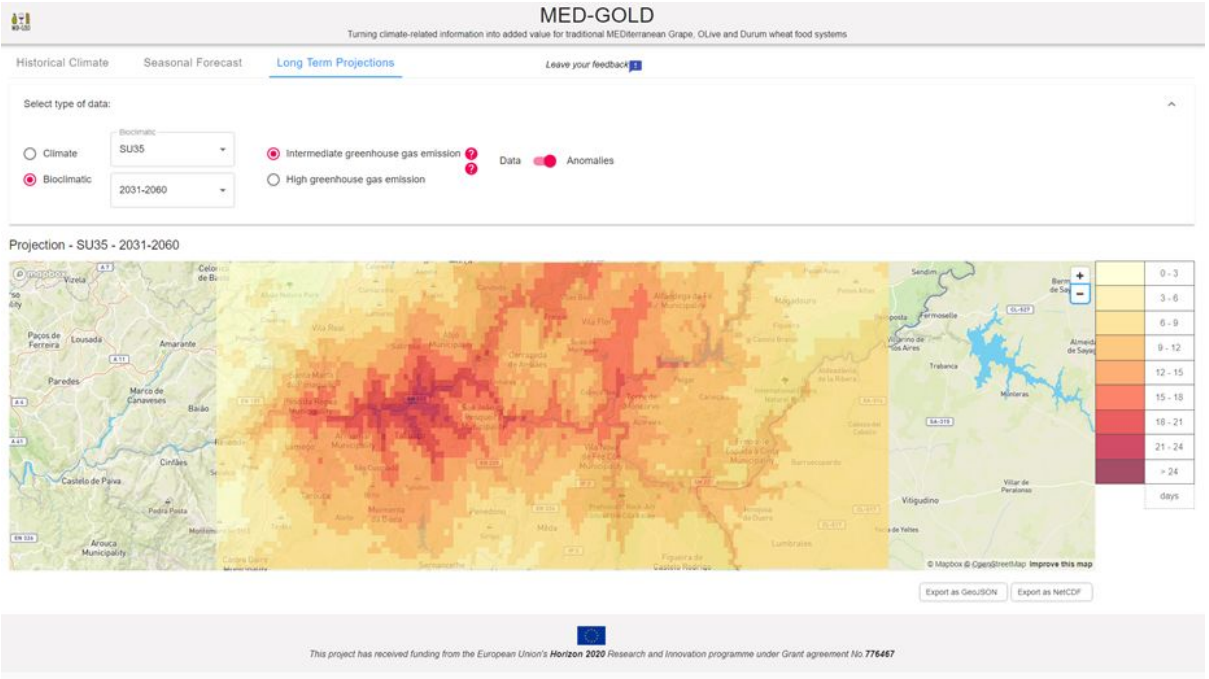


Figure 4-16. SU35 over the Douro Valley: Ensemble mean change in the number of days with TX > 35ºC during AMJJASO between 2031-2060 and 1971-2000 under RCP4.5.



4.7 Development of the Climate Service Based on Satellite Data

Earth observation remote sensing missions acquire information on the state of the planet without making physical contact with the target. Space-based remote sensing compliments in-situ measurements, taken at the target location. The remote-sensing observations made from orbit can be directly validated in the terrestrial environment that is being investigated. [RD.5] Regarding the sensors, the first remote sensing instruments were radiometers in the visible, infrared and microwave spectrum. Passive remote sensing was a precursor to active remote sensing given that the system is simpler (no transmission), the processing easier, and the requirements on power on the satellite are lower.

The evolution and impact of space-based remote sensing is accelerating as the technologies develop in the domain of electronics, ground processing, communications, and the inclusion of the private sector. Earth observation missions provide a wealth of information on the planet. Forty years of space-based remote sensing has transformed our knowledge on the planet and how it is changing. Remote sensing has an immeasurable effect on day-to-day applications.

Operational missions are launched to fulfil specific goals, to provide users with information, images and data products they need. Examples of operational missions are the Copernicus Sentinel

missions. Operational missions have strong requirements on reliability (low risk, redundancy) and longer lifetime to provide data continuity. Many operational missions span decades.

Research missions are intended for a one period study of a certain area. New technologies are tested in research missions, anything from guidance, navigation and control techniques, to calibration, data processing to non-space tested instrumentation certification.

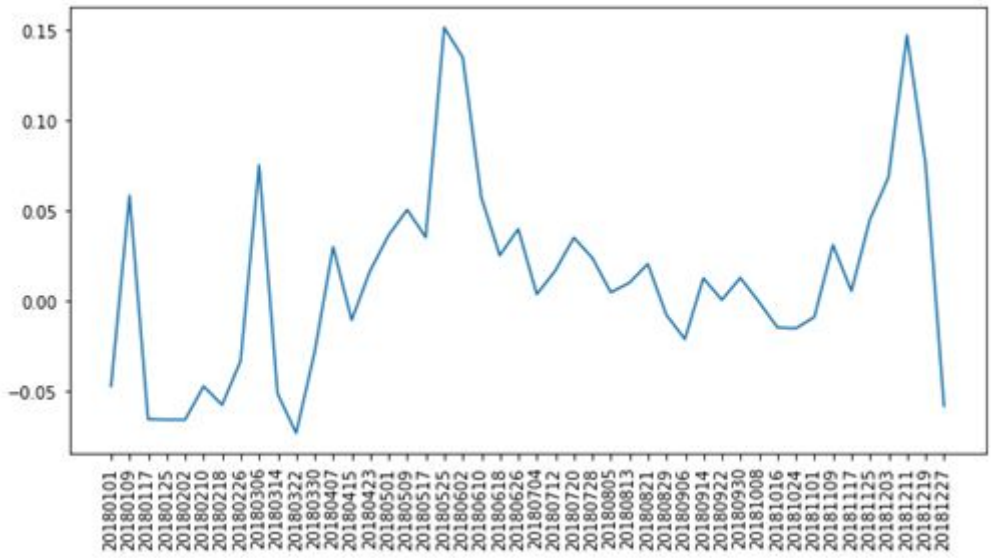
In the framework of the MED-GOLD project, some satellite missions were identified as interesting sources of information. In this point it is important to highlight that all indirect measures could accumulate errors (random or systematic) derived from several sources. Some factors such as clouds, sensor problems, transmission error, etc. can appear in the scenes affecting derived products like indices or mapping.

As was indicated above, the use of MSI imagery in MED-GOLD is focused to calculate some indices and measures as LST which could indirectly show climate variations through the comparison between EO results and climate behaviour (from reference or modeled datasets). The figure 4-17 shows an example of the NDWI behaviour during 2018 on a specific location in the Douro Valley. Some studies suggest [RD.8] the relation between climatic factors and the Vegetation Index. The calculations of different indices depends on factors associated with satellite to satellite specifications. The table 4-3 shows the indices calculated for the MED-GOLD project.

Table 4-3 indices calculation for WP3.

Satellite	Index	Spatial Resolution	Time Coverage
S2	NDVI	20m	2017
S2	NDWI	20m	2017
L8	NDVI	30m	2013
L8	NDWI	30m	2013
L8	LST	30m	2013
L8	TVX	30m	2013
MODIS	NDVI	460m	2000 to 2018
MODIS	NDWI	460m	2000 to 2018
MODIS	NMDI	460m	2000 to 2018

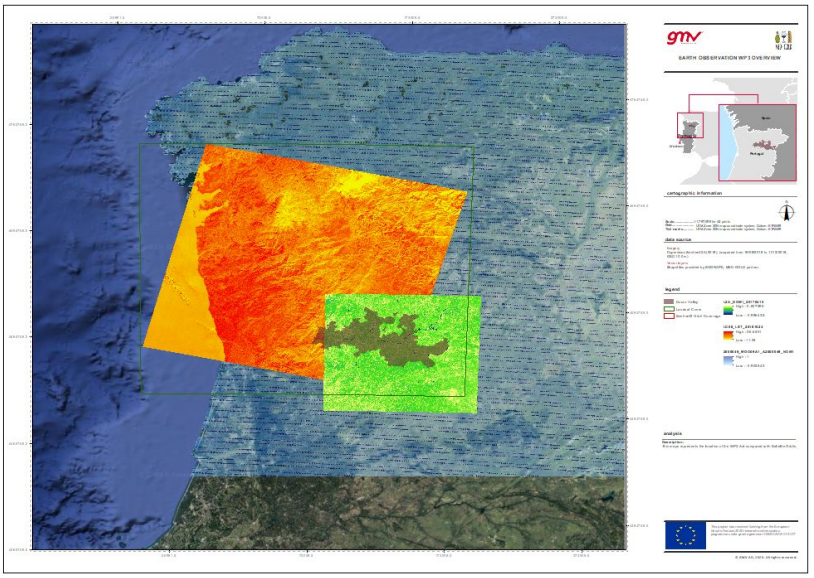
Figure 4-17 Example of the NDWI index variations (calculated from MODIS) during 2018 over the Douro Valley.



In the context of spatial coverage, the MED-GOLD WP3 grape/wine is focused in the Douro Valley located at the northeast of Portugal. Despite the spatial coverage of the satellite missions described above guarantee the covering of the MED-GOLD areas of study, its usefulness will depend on technical factors like the spectral covering (Bands), the covering Area (maximum extension covered) the spatial Resolution and the temporal covering (when the mission started and frequency of data acquisition).

Sentinel2 (S2) covers an area of 12000km^2. However, the orbit described by the satellite divides the Area of Interest (AoI, Douro Valley). Since the purpose to use EO imagery is mainly to have an ancillary measure to validate the MED-GOLD outputs, the scene selected to calculate the indices is 289TPF which covers approximately 92% of the Douro Valley. Figure 4-18 shows the location of the AoI compared with S2 extension. The Landsat8 scene, approximately covers 34200km^2, with a 75% coverage of the AoI. Finally, in the case of MODIS Imagery the total extension of the AoI is covered. The figure 4-18 shows the area covered by Sentinel2, Landsat8 and MODIS imagery.

Figure 4-18 EO imagery covered area. Green area corresponds to Sentinel 2; the orange, to Landsat 8; and the blue toned, to MODIS.

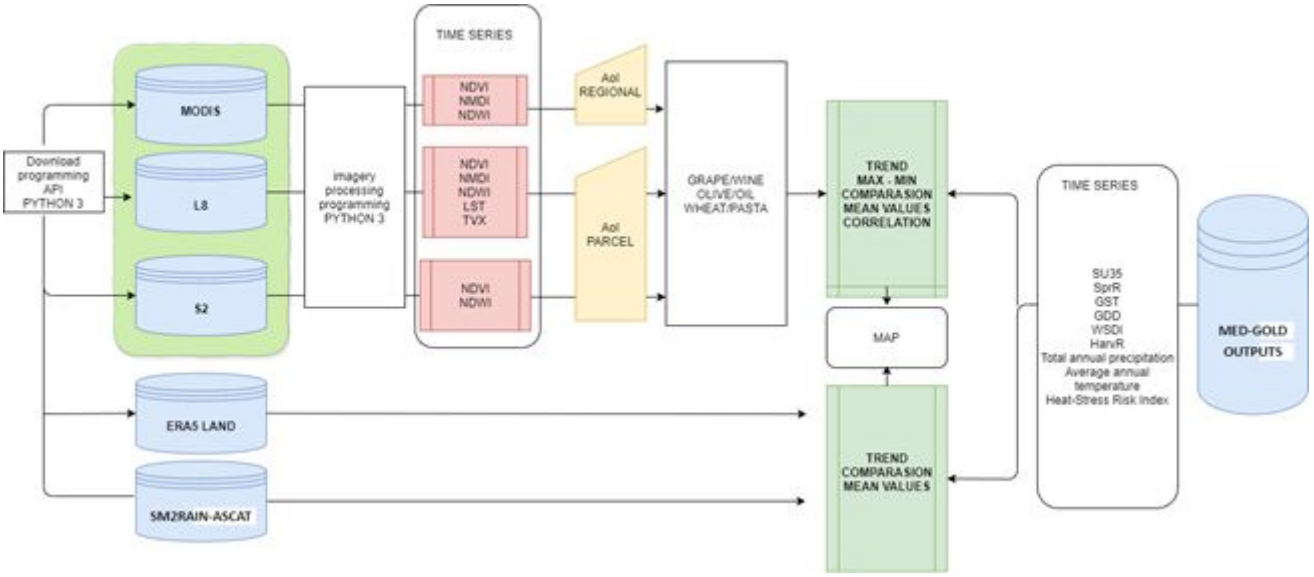


This section describes briefly how the EO imagery will be analysed as a way to verify results. EO will be used to verify indirectly, when the physical conditions allow it, some results obtained by MED-GOLD. In the same way, this information will be available to be shared as geospatial information linked to MED-GOLD through the dashboard, if supported. A total of 866 MODIS (2000-2018), 44 L8 (2014-2018), 33 S2 (2017-2018) to WP3 has currently processed and uploaded to MED-GOLD ITC focused on the second objective of this EO datasets. The availability of the imagery depends on external providers, atmospheric conditions (cloud covering), sensor transmission, etc. All subproducts (indices) calculated in the MED-GOLD framework have a standardized geographic metadata based on ISO TC 211 (https://committee.iso.org/home/tc211). This metadata contain relevant information about the dataset and is available as *xml* file:

- fileIdentifier
- language
- contact
- dateStamp
- metadataStandardName
- metadataStandardVersion
- spatialRepresentationInfo
- referenceSystemInfo
- identificationInfo
- distributionInfo
- applicationSchemaInfo

The imagery analysis in the context of MED-GOLD aims to verify some of the results obtained by the project. This process will be done through comparison based on the behaviour (trend, correlations, standard deviation) of the indices curve with MED-GOLD results as well as statistical analysis. Figure 4-19 shows a diagram to summarize the methodology that will be applied for validation purposes.

Figure 4-19 Methodology proposal for MED-GOLD validation using EO dataset.

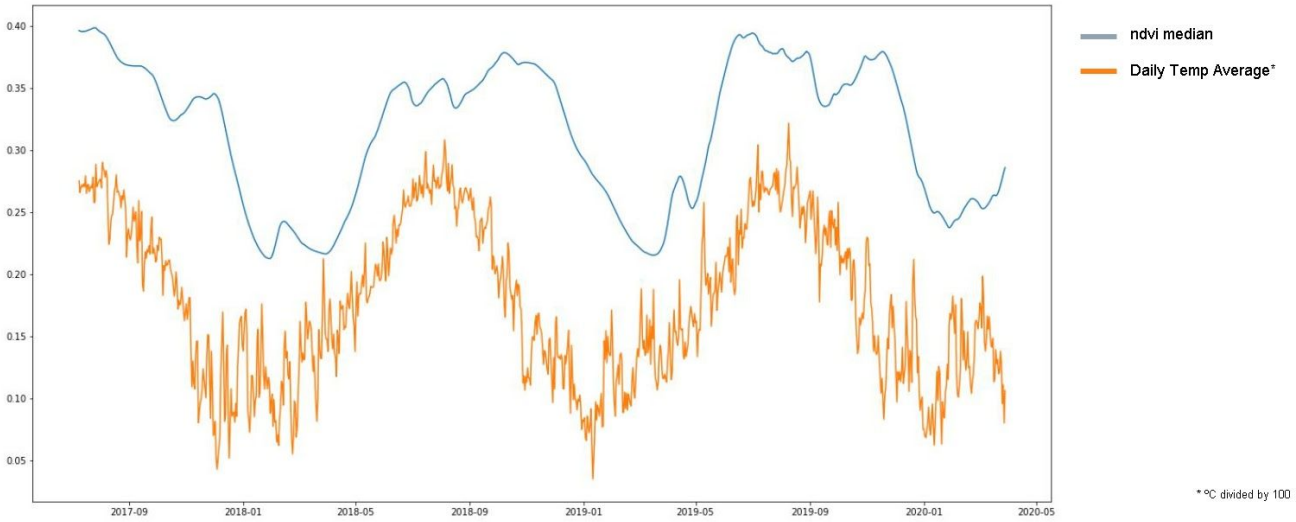


The first part of this proposal implied to develop python scripts to download dataset using API's, direct download or download under request approval. Then another set of scripts were developed to calculate the indices based on time series considering the spectral availability according to each sensor. Some of the spatial analysis libraries used were geopandas, gdal, shapely, pandas.

Considering the AoI (region or parcel) some aleatory areas within the Douro Valley will be used to compare the results obtained by MED-GOLD. Depending on the imagery spatial resolution, some parcels will be used to analyse the average value (derived from L8 and S2). These parcels correspond to vineyards provided by the MED-GOLD partner SOGRAPE. indices calculated from MODIS imagery will be used to compare largest extensions (maintaining homogeneity based on land cover) considering the spatial resolution of the MED-GOLD results.

This method is based on the assumption that different stages of the soil or vegetation are correlated with climate variations [RD.8]. The figure 4-20 shows an example of the correlation between NDVI average (blue) and temperature average (orange) in the vineyard region in Spain (2017-2020). The results of this validation applied to WP3 (grape/wine sector) will be described in the deliverable specified for this purpose.

Figure 4-20 NDVI average (blue line) and temperature average (orange line) comparison.



5. Conclusions

During the period between M4 and M30, the scientific and technological partners (Beetobit, BSC, ENEA, Met Office, NOA and UTH) and end-users (SOGRAPE) have intensively worked in the co-development of an end-to-end wine pilot climate service. This deliverable D3.2 reports the methodological strategies followed for the implementation and testing of each of the features included in the service. As a result of this intensive work a pilot wine climate service has been co-created following the end-users needs and requirements. The service provides climate information to end-users at different time horizons: historical climate, seasonal predictions and long-term projections through the MED-GOLD Dashboard interactive tool. The products resulting from this service are based on essential climate variables (i.e. mean temperature, maximum temperature, minimum temperature and precipitation), bioclimatic indicators (growing season average temperature, spring rain, harvest rain, number of heat stress days, warm spell duration index) and risk indices (sanitary and heat). Additionally, a first methodological approach to assess the economic impact of the use of retrospective seasonal predictions in a decision making framework has also been co-developed.

The main conclusions achieved during this wine climate service co-development are as follows.

- The MED-GOLD wine climate service is able to provide climate information at different time scales that respond to specific questions of the wine sector related to their current needs for plan control on fungal diseases, management of protective treatments, forecast quality of berries and wine, plan for harvest operations or evaluate suitability of varieties in the forthcoming years.

- The closer interaction between scientific partners and the wine champion has allowed the identification and calculation of specific bio-climatic indicators of relevance for the wine sector, as well as, the co-design and co-development of two new compound risk indices that can integrate into a single measure the main risk factors that can affect wine production (in terms of quality, quantity and derived inherent value) in the Douro region.

- The MED-GOLD wine climate service has been already effective in interacting with users to interpret the probabilistic information delivered by seasonal predictions, highlighting the importance of the skill and reliability associated with this type of predictions.

- As pioneer work, the climate service has gone a step further for understanding the use of seasonal forecasts in decision-making, co-developing a methodology to estimate the economic impact of using retrospective seasonal forecasts of Spring Rain (SprR) to make

decisions regarding plant protection and canopy management compared to their customary use of climate memory.

- Even though the final aim of MED-GOLD project is not to provide a real-time operational climate service, the products obtained from the wine climate service have been provided to the end-users using real-time seasonal predictions for the spring season 2020 and it is expected to have another round of real-time predictions for spring 2021, which will eventually result in improved feedback loops for evaluation, validation and streamlining of the tool resulting from the project..

- The wine climate service has been materialised in the MED-GOLD Dashboard interactive tool. This web-based application is meant to provide a visualization tool for stakeholders, allowing them to browse and display climate data, in a user-friendly way and without any programming knowledge. The information provided through this tool will be refined in the forthcoming months to achieve the needs stated by the wine end-users.

- The code workflow followed for the implementation of the MED-GOLD wine climate service is documented and accessible via the MED-GOLD ICT Platform and it is conceived for future application of the service in a real-time operational context.

- The MED-GOLD wine climate service is underpinned by the most advanced global climate models, satellite products and observational datasets, taking advantage of Copernicus products available through the Climate Data Store. This enables the replicability of the work done for the wine climate sector to other areas and agriculture sectors in a seamless and straightforward way.

Following all the conclusions stated above, it is worth to say that MED-GOLD project is reaching its overall objective of co-create an end-to-end climate service for the wine sector that strengthens the efficiency and sustainability of the Mediterranean wine industry in their seasonal- and long-term business strategies, with a special focus on the Portugal's Douro Valley, home of Port and Douro wine appellations and their iconic brands.

Annex A. Dashboard Rest API Reference

API Version: 1.0.0

How to use the API

Click on Authentication section and then on authentication endpoint

Click on Try it out
Insert your Med-gold Dashboard username and password
Copy token value, without double quotes
Click on Authorize Button
Paste token value and click on Authorize

You are now authorized to use the Med-Gold Dashboard API!

INDEX

1. AUTHENTICATION 4

1.1 GET /authentication 4

2. ES DATA API 5

2.1 GET /index/match-all 5

2.2 GET /index/match-location 6

2.3 GET /index/match-location/seasonal_forecast 7

3. GENERAL API 8

3.1 GET /all 8

3.2 GET /getlink 8

3.3 GET /ids 9

3.4 GET /{id}/info 10

Security and Authentication

SECURITY SCHEMES

KEY	TYPE	DESCRIPTION
dev-mgd-dashboard	apiKey	

API

1. AUTHENTICATION

Access Token endpoint

1.1 GET /authentication

The endpoint returns the authentication token needed to use API

REQUEST

QUERY PARAMETERS

NAME	TYPE	DESCRIPTION
*password	string	Password of Med-gold Dashboard account
*username	string	Username of Med-gold Dashboard account

RESPONSE

STATUS CODE - 200: Access token

RESPONSE MODEL - */*

```
{
    token  string  access token API
}
```

STATUS CODE - 401: Unhautorized

RESPONSE MODEL - */*

{

MESSAGE STRING UNHAUTORIZED ERROR MESSAGE

}

STATUS CODE - 404: NOT FOUND

RESPONSE MODEL - */*

{

MESSAGE STRING NOT FOUND ERROR MESSAGE

}

2. ES DATA API

ELASTICSEARCH'S DATA ENDPOINTS

2.1 GET /INDEX/MATCH-ALL

THE ENDPOINT RETURNS ALL THE INFORMATION LOADED ON THE MAP BASED ON INFORMATION PASSED

REQUEST

QUERY PARAMETERS

NAME	TYPE	DESCRIPTION
*TAB	ENUM ALLOWED: CLIMATOLOGY, SEASONAL_FORECAST, ERA5, PROJECTION	TAB NAME BETWEEN: CLIMATOLOGY, SEASONAL_FORECAST, ERA5 AND PROJECTION

*VALUE	ENUM ALLOWED: PRECIPITATION_MONTHLY, TMAX_MONTHLY, TMIN_MONTHLY, TAVERAGE_MONTHLY, GST, HEAT_RISK, HARVESTR, SU35, SPRR, WSDI, SANITARY_RISK, GDD, SPRTX	NAME OF THE INDEX OR VARIABLE SELECTED
*STYPE	ENUM ALLOWED: ECV, INDEX, ECVAVG, INDEXAVG, ECVANOM, INDEXANOM, ECVP, INDEXP	

TYPE OF INFORMATION CAN BE: INDEX OR ECV.

*YEAR	STRING	VALUE WITHIN THE RANGE DEFNED ACCORDING TO THE CHOSEN TAB
RCP	ENUM ALLOWED: 5, 45	RCP VALUE: 5 OR 45. REQUIRED IF TAB IS PROJECTION AND TIME INTERVAL IS NOT 1 1-2000
MONTH	STRING	VALUE WITHIN THE RANGE DEFNED ACCORDING TO THE CHOSEN TAB, STYPE AND VALUE. REQUIRED IF STYPE IS ECV.
*TOP_LEFT	STRING	COORDINATES IN LON, LAT FORMAT OF THE

		UPPER LEFT GRID		
*BOTTOM_RIGHT	*STRING*	*COORDINATES IN BOTTOM RIGHT GRID*	*LON, LAT*	*FORMAT OF THE*

RESPONSE

STATUS CODE - 200: K

RESPONSE MODEL - */*

```
{
    TYPE STRING LOCATION STRING
    VALUE        STRING
    RPSS         STRING
}
```

TYPE OF VARIABLE, CAN BE ECV OR INDEX

LAT,LON COORDINATES OF DATASET LOCATION

OAT VALUE OF THE REQUIRED INFORMATION

RPSS MASK VALUE, RETURNED ONLY IF TAB IS SEASONAL_FORECAST

STATUS CODE - 401: UNHAUTORIZED

RESPONSE MODEL - */*

```
{
MESSAGE STRING UNHAUTORIZED ERROR
                MESSAGE
}
```

5 OF 11

STATUS CODE - 404: NOT FOUND

*RESPONSE MODEL - */**

```
{
    MESSAGE  STRING  NOT FOUND ERROR MESSAGE
}
```

2.2 GET /INDEX/MATCH-LOCATION

THE ENDPOINT RETURNS CHART INFORMATION FOR HISTORICAL CLIMATE AND PROJECTION TAB BASED ON INFORMATION PASSED

REQUEST

QUERY PARAMETERS

NAME	TYPE	DESCRIPTION
*TAB	ENUM ALLOWED: CLIMATOLOGY, ERA5	TAB NAME BETWEEN: CLIMATOLOGY, ERA5 AND PROJECTION
*STYPE	ENUM ALLOWED: ECV, INDEX	TYPE OF INFORMATION CAN BE: INDEX OR ECV
*VALUE	ENUM ALLOWED: PRECIPITATION_MONTHLY, TMAX_MONTHLY, TMIN_MONTHLY, TAVERAGE_MONTHLY, GST, HEAT_RISK, HARVESTR, SU35, SPRR, WSDI, SANITARY_RISK, GDD, SPRTX	NAME OF THE INDEX OR VARIABLE SELECTED

LOCATION	STRING	LOCATION SELECTED IN FORMAT LAT,LON
MONTH	STRING	VALUE WITHIN THE RANGE DEFNED ACCORDING TO THE CHOSEN TAB, STYPE AND VALUE

RESPONSE

STATUS CODE - 200: IST OF VALUES FOR SELECTED LOCATION

RESPONSE MODEL - */*

{

VALUE	STRING OAT VALUE OF THE REQUIRED INFORMATION
YEAR	STRING PERIOD OF THE REQUIRED INFORMATION
MONTH	MONTH VALUE OF THE REQUIRED INFORMATION
COLOR	COLOR USED FOR CHART REPRESENTATION

}

STATUS CODE - 401: UNHAUTORIZED

RESPONSE MODEL - */*

{

MESSAGE STRING UNHAUTORIZED ERROR MESSAGE

}

STATUS CODE - 404: NOT FOUND

RESPONSE MODEL - */*

{

MESSAGE STRING NOT FOUND ERROR MESSAGE

}

2.3 GET /INDEX/MATCH-LOCATION/SEASONAL_FORECAST

THE ENDPOINT RETURNS CHART INFORMATION FOR SEASONAL FORECAST TAB BASED ON INFORMATION PASSED

REQUEST

QUERY PARAMETERS

NAME	TYPE	DESCRIPTION
*STYPE	ENUM ALLOWED: ECV, INDEX	TYPE OF INFORMATION CAN BE: INDEX OR ECV
*VALUE	ENUM ALLOWED: PRECIPITATION_MONTHLY, TMAX_MONTHLY,	NAME OF THE INDEX OR VARIABLE SELECTED

	TMIN_MONTHLY, TAVERAGE_MONTHLY, GST, HARVESTR, SPRR, SANITARY_RISK, SPRTX	
*LEADT	ENUM ALLOWED: 1, 2, 3, 4	TARTING DATE VALUE IN NUMBER FORMAT
MONTH	STRING	VALUE WITHIN THE RANGE DEFNED ACCORDING TO THE CHOSEN TAB, STYPE AND VALUE
*LOCATION	STRING	LOCATION SELECTED IN FORMAT LAT,LON

RESPONSE

STATUS CODE - 200: IST OF VALUES FOR SELECTED LOCATION

RESPONSE MODEL - */*

{

VALUE	STRING OAT VALUE OF THE REQUIRED INFORMATION
YEAR	STRING PERIOD OF THE REQUIRED INFORMATION
MONTH	STRING MONTH VALUE OF THE REQUIRED INFORMATION
COLOR	STRING COLOR USED FOR CHART REPRESENTATION
LEADT	LEADTIME STARTING DATE VALUE OF THE REQUIRED INFORMATION

NAME	*NAME OF TERCILE CAN BE UPPER TERCILE, MEDIUM TERCILE AND OWER TERCILE*

}

STATUS CODE - 401: UNHAUTORIZED

RESPONSE MODEL - */*

{

MESSAGE STRING UNHAUTORIZED ERROR MESSAGE

}

STATUS CODE - 404: NOT FOUND

RESPONSE MODEL - */*

{

MESSAGE STRING NOT FOUND ERROR MESSAGE

}

3. GENERAL API

GENERAL INFORMATION ENDPOINTS

3.1 GET /ALL

THE ENDPOINT RETURNS ALL INFORMATION OF THE ELEMENTS IN THE DASHBOARD

REQUEST

NO REQUEST PARAMETERS

RESPONSE

STATUS CODE - 200: IST OF DICTIONARY WITH ALL INFORMATION OF DASHBOARD ELEMENTS

RESPONSE MODEL - */*

{

}

STATUS CODE - 401: UNHAUTORIZED

*RESPONSE MODEL - */**

```
{
    MESSAGE  STRING  UNHAUTORIZED ERROR MESSAGE
}
```

STATUS CODE - 404: NOT FOUND

*RESPONSE MODEL - */**

```
{
    MESSAGE  STRING  NOT FOUND ERROR MESSAGE
}
```

3.2 GET /GETLINK

THE ENDPOINT RETURNS A LINK OF THE FLE BASED ON INFORMATION PASSED

REQUEST

QUERY PARAMETERS

NAME	TYPE	DESCRIPTION
*TAB	ENUM ALLOWED: CLIMATOLOGY, SEASONAL_FORECAST, ERA5, PROJECTION	TAB NAME BETWEEN: CLIMATOLOGY, SEASONAL_FORECAST AND PROJECTION
*VALU	ENUM	NAME OF THE INDEX OR VARIABLE SELECTED

E	ALLOWED: PRECIPITATION_MONTHLY, TMAX_MONTHLY, TMIN_MONTHLY, TAVERAGE_MONTHLY, GST, HEAT_RISK, HARVESTR, SU35, SPRR, WSDI, SANITARY_RISK, GDD, SPRTX	
*STYP	ENUM	TYPE OF INFORMATION CAN BE: INDEX OR ECV
E	ALLOWED: ECV, INDEX, ECVAVG, INDEXAVG, ECVANOM, INDEXANOM, ECVP, INDEXP	
*YEAR	STRING	VALUE WITHIN THE RANGE DEFNED ACCORDING TO THE CHOSEN TAB

NAME	TYPE	DESCRIPTION
MONTH	STRING	VALUE WITHIN THE RANGE DEFNED ACCORDING TO THE CHOSEN TAB, STYPE AND VALUE. REQUIRED ONLY IF STYPE IS ECV.

LEAD T	*ENUM* *ALLOWED: 1, 2, 3, 4*	*TARTING DATE VALUE IN NUMBER FORMAT.* *REQUIRED ONLY IF TAB IS SEASONA_FORECAST.*

RESPONSE

STATUS CODE - 200: URL OF REQUIRED FLE

RESPONSE MODEL - */*

```
{
}
```

STATUS CODE - 401: UNHAUTORIZED

RESPONSE MODEL - */*

```
{
    MESSAGE  STRING  UNHAUTORIZED ERROR MESSAGE
}
```

STATUS CODE - 404: NOT FOUND

RESPONSE MODEL - */*

```
{
    MESSAGE  STRING  NOT FOUND ERROR MESSAGE
}
```

3.3 GET /IDS

THE ENDPOINT RETURNS THE IDS OF ELEMENTS IN THE DASHBOARD

REQUEST

NO REQUEST PARAMETERS

RESPONSE

STATUS CODE - 200: IST OF ID OF THE ELEMENTS OD THE DASHBOARD

RESPONSE MODEL - */*

{

}

STATUS CODE - 401: UNHAUTORIZED

RESPONSE MODEL - */*

{

MESSAGE STRING UNHAUTORIZED ERROR MESSAGE

}

STATUS CODE - 404: NOT FOUND

RESPONSE MODEL - */*

{

MESSAGE STRING NOT FOUND ERROR MESSAGE

}

1

3.4 GET /{ID}/INFO

THE ENDPOINT RETURNS ALL INFORMATION BASED ON ID PASSED

REQUEST

PATH PARAMETERS

NAME	TYPE	DESCRIPTION
*ID	ENUM ALLOWED: CLIMATOLOGY, EASONAL-ORECAST, PROJECTION	ID NAME BETWEEN: CLIMATOLOGY, EASONAL-ORECAST AND PROJECTION

-

RESPONSE

STATUS CODE - 200: SON OBJECT WITH INFORMATION OF SELECTED ID

RESPONSE MODEL - */*

```
{
}
```

STATUS CODE - 401: UNHAUTORIZED

RESPONSE MODEL - */*

```
{
    MESSAGE  STRING  UNHAUTORIZED ERROR MESSAGE
}
```

STATUS CODE - 404: NOT FOUND

RESPONSE MODEL - */*

```
{
    MESSAGE  STRING  NOT FOUND ERROR MESSAGE
}
```