VISCA

VINEYARDS INTEGRATED SMART CLIMATE APPLICATION

Deliverable 2.2: Report on weather forecast extreme events

WP	2	Climatic & Agricultural Data	
Task	2.2	Quality assurance and Risk management	
Dissemination level[1]	PU	**Due delivery date**	30-04-18
Nature[2]	R	**Actual delivery date**	18-05-18

Lead beneficiary	MET
Contributing beneficiaries	All partners

Version	1.2
Total number of pages	30

[1] Dissemination level: **PU** = Public, **PP** = Restricted to other programme participants (including the JU), **RE** = Restricted to a group specified by the consortium, **CO** = Confidential, only for members of the consortium

[2] Nature of the deliverable: **R** = Report, **P** = Prototype, **D** = Demonstrator, **O** = Other

Version history

Document Version	Date	Author	Comments[3]
1.0	20-09-17	Ignasi Porras (MET)	Main structure
1.1	10-05-18	Ignasi Porras (MET)	First draft of the deliverable
1.2	17-05-18	Raül Marcos (BSC)	Review
1.3	18-05-18	Ignasi Porras (MET)	Amendments
1.4	18-05-18	María Navarro (MET)	Final Review

Deliverable abstract

A detailed description of the weather models and processes applied to develop the weather forecast products for the VISCA project in the context of task 2.2 in work package 2 is presented in this document.

In the VISCA project, two kinds of products are delivered: short-term weather forecasts and mid-term weather forecast. While having in two days in advance the weather forecast product is useful for near future in-field activities, having weather forecast information ten days in advance could be useful to wine producers to minimize risks related to coming extreme events. For instance, if a heat wave is forecasted in 6 days, wine producers could act in advance irrigating the field more than usual before the event.

Meteosim (MET) has developed and supplied weather forecasts services, which consisted in deliver the best prediction of high impact weather variables at forecast time scales from hours up to 10 days (240 hours). For the short-term forecast (up to 48 hours), a regional weather forecast model is used. On the other hand, a probabilistic model is used for mid-term forecast (up to 10 days).

Finally, all weather forecast products are sent to the VDI (VISCA Data Interface), where other modellers and end-users can download these information for their use in the VISCA platform.

[3] Creation, modification, final version for evaluation, revised version following evaluation, final.

Contents

List of Tables

List of Figures

Acronyms list

AEMet	Meteorological agency from Spain
CNTL	Control weather forecast
CMS	Cumulus parametrization
ECA&D	European Climate Assessment & Dataset
GEFS	Global Ensemble Forecast System
GFS	Global Forecast System
GRIB	Gridded Information in Binary
HRES	Deterministic NWP models
MPH	Microphysics parametrization
NCEP	National Centers for Environmental Prediction
NetCDF	Network Common Data Form
NWP	Numerical Weather Prediction
SWR	Short-wave radiation
SYNOP	Surface synoptic observations
VDI	VISCA Data Interface
WRF_ARW	Weather Research & Forecast - Advanced Research WRF

1. Introduction

Knowing weather conditions in advance it is extremely important for the society, the economy on which it depends, and the environment to avoid certain risks in advance. Moreover, extreme weather conditions can cause substantial disruptions in daily life and incur monetary costs and even cause deaths. Therefore, having the best prediction of high impact weather variables at different time scales could be useful to wine producers to minimize risks related to coming extreme weather events.

1.1 Roles and Responsibilities

The main goal of this report is to provide a detailed description of the weather models and processes applied to develop the weather forecast products for the VISCA project in the context of task 2.2 in work package 2. In this task, Meteosim (MET) developed and supplied weather forecasts services, which consisted in deliver the best prediction of high impact weather variables at forecast time scales from hours up to 10 days (240 hours). This information could be useful to wine producers to minimize risks related to coming extreme events (for instance, if a heat wave is forecasted in three days, wine producers could act in advance irrigating the field two days before the event).

1.2 Structure of the document

This report is organised in the following manner:

- **Chapter 1** includes this introduction and a description of the document;
- **Chapter 2** introduces the schematic data flow from input to output of the implemented weather forecast models.
- **Chapter 3** provides information about the applied numerical weather prediction models.
- **Chapter 4** describes a description of the output data

2. System architecture

This chapter describes the schematic data flow of the weather forecast production from input to output within the context of task 2.2. It includes the weather forecasts up to 2 days which is developed by Meteosim (MET) and up to 10 days, also post-processed by MET.

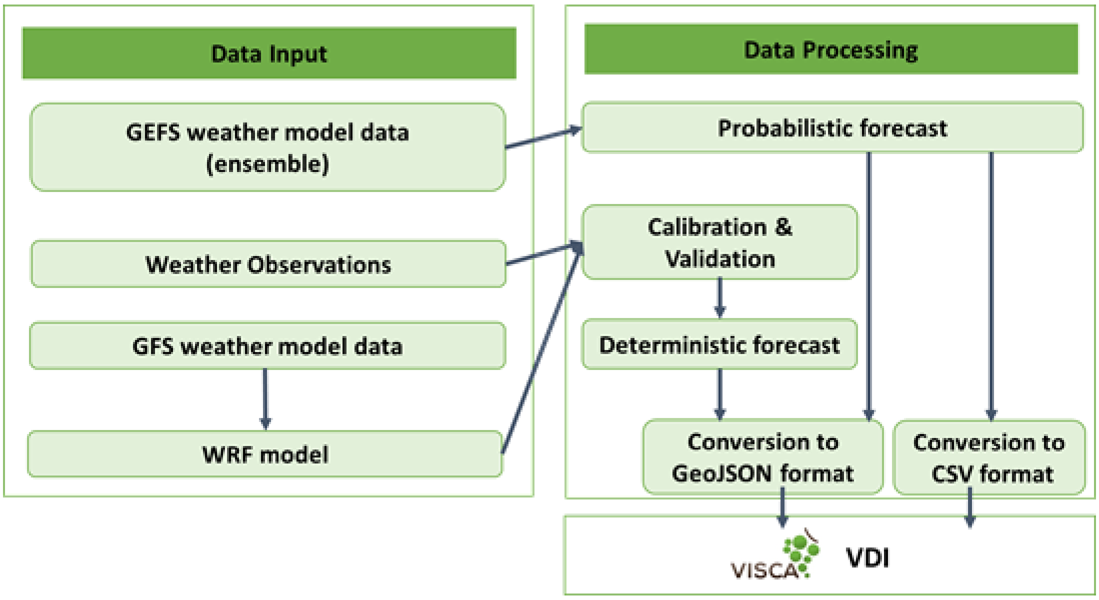

Figure 1. MET data processing flow chart regarding VISCA data.

Figure 1 shows the data processing flow chart, which consists of two main blocks; data input and data processing. As an input for weather forecast production, the weather observations are taken from surface synoptic observations (SYNOP), and initial weather model data from two numerical weather prediction models. For the short-term forecast (up to 48 hours), a regional weather forecast model is used. On the other hand, a probabilistic model is used for mid-term forecast (up to 10 days).

For the short-term product, the SYNOP data is provided in the ASCII text format, and the weather model data in the GRIB (Gridded Binary) format. The first phase in the data processing block is weather forecast validation. For validation purposes, different weather forecast experiments are done to select the better physical parameters of the model to improve weather forecast. When the different forecasts are validated, the final physics options are set to provide more reliable and accurate weather forecasts. Then, each day the model is run twice per day producing NetCDF (Network Common Data Form) format data and converted to GeoJSON format.

The mid-term is produced by a probabilistic global forecast model where the data obtained in GRIB format and post-processed. Then, the data is formatted into a CSV format per each demo-site and in GeoJSON format per each demo-area.

Finally, all weather forecast products are sent to the VDI (VISCA Data Interface).

3. Numerical weather prediction models

Weather forecasting is mainly based on the use of Numerical Weather Prediction (NWP) models. NWP focuses on taking observations of current weather and processing these data with computer models to forecast the future state of the weather. Current weather observations serve as input to the numerical computer model through a process known as data assimilation to produce outputs of different weather variables. These observations are gathered from specific sites at ground level and from soundings where radiosonde measures the vertical structure of the atmosphere. Satellite Earth observation data is also essential part of data input for models.

Weather models consists of mathematical equations that are based on the laws of physics, fluid motion and chemistry. The models use a coordinate system which divides the forecast domain into a three-dimensional (3D) grid (Figure 2), where every grid cell represents the average value within the cell. The horizontal domain of a domain of a model is either global or regional. Regional models can use a global model to obtain lateral boundary conditions to allow systems from outside the regional model domain to move into its area. Compared to global models, regional models (limited are models) can produce finer spatial resolution output because computational resources are focus on a specific area instead of covering whole globe. In the VISCA project, we utilize the limited area model WRF-ARW (Weather Research & Forecast - Advanced Research WRF) for short-term forecasts that range from a few hours to a couple of days, and the global model GEFS (Global Ensemble Forecast System) for longer-term forecasts that range from a few days to ten days.

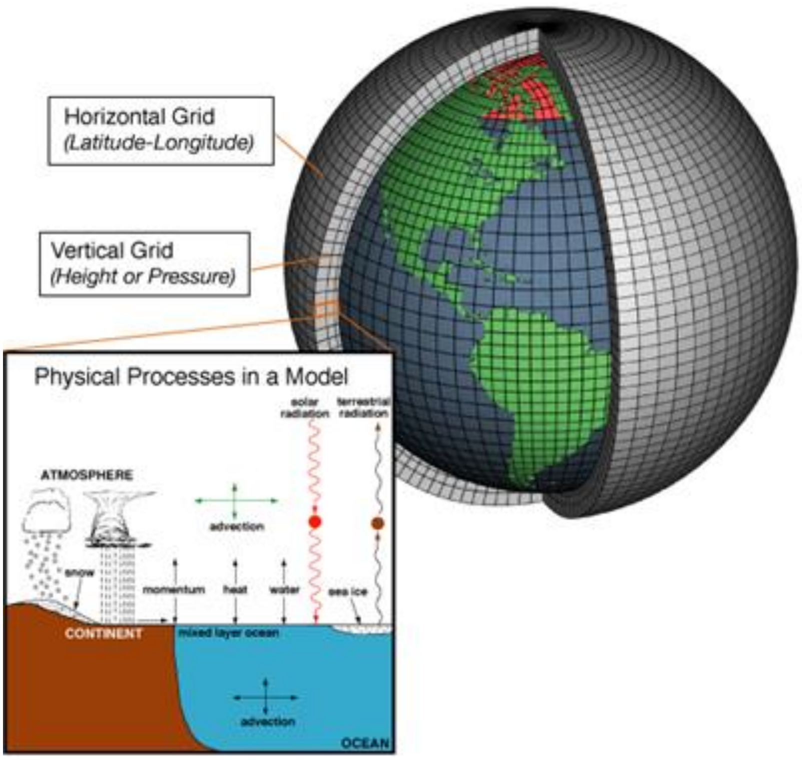

Figure 2. An illustration of the 3D numerical weather prediction model principle.

Weather forecasts are either deterministic or probabilistic. As we go from the short-term to the long-term, the nature of the forecasts also changes. Short-term forecasts are deterministic, in what we state what is going to happen, when and where. In general, as we go to longer times scales the proportion of the forecast which must be probabilistic increases.[4]

3.1 Deterministic forecasting

During the last years, the use of deterministic NWP models (HRES) has increased significantly due to: higher accuracy of the models, easier access to community models, computational advances, etc. Despite the inherently uncertain that they can have due to the initial conditions of the atmosphere, several calibrations can be done to minimize the discrepancies against the real weather conditions. Due to NWP models have a wide range of options to set up: physical options, dynamical options, horizontal model resolution, number of vertical layers and density, etc., it is crucial configurate the model with the properly parametrizations and model options.

3.1.1 WRF-ARW

The meteorological model applied for the short-term weather operational system is the Weather Research and Forecasting model. WRF-ARW is a new generation, non-hydrostatic and modular structure meteorological model designed to execute multi-task operational weather forecasting. The WRF-ARW modeling system, a complex mesoscale model, is capable of providing reliable and precise hourly weather parameters. The WRF-ARW system has been developed for diagnostic and forecasting purposes, which makes it the most optimal modeling system for all the activities to be performed for this project.

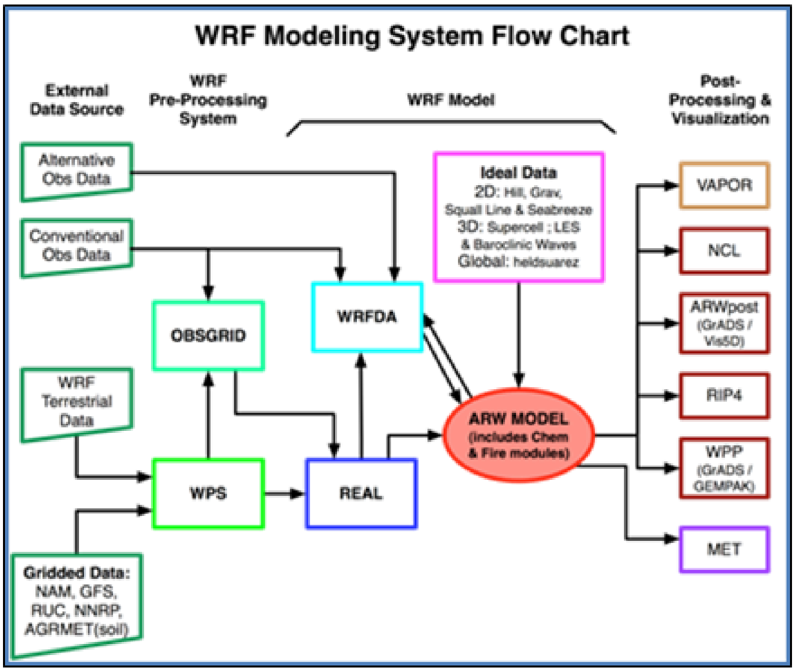

Figure 3. Modular Scheme of WRF-ARW modeling system.

[4] https://www.unc.edu/courses/2008ss2/geog/111/001/ForecastTypes/ForecastTypes.htm

WRF-ARW model have a wide range of options to set up: physical options, dynamical options, horizontal model resolution, number of vertical layers and density, domains architecture and nest down options, assimilation data, time-step, spin-up time, etc. It is a fundamental factor when configuring a model, the selection of the parameterizations and options that are used. And the best combination for one region is not necessarily applicable to another.

The preferred initial and boundary conditions for the operational configuration over the coarsest regional domain is supplied by the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) with horizontal resolution 0.25°.

For this project, the definition of an optimal configuration testing methodology of a specific WRF-ARW is considered to improve meteorological forecasting for operational purposes. The proposed methodology is based on previously configured forecasting systems and performance achievements reached by the modelers for this project (Arasa *et al.*, 2016)[5].

3.1.2 Model configuration

The meteorological model is configured and executed daily across nested grid domains for each demo-area as was defined in the proposal.

GFS global model domain (red squares from Figure 4 to Figure 6) feeds the regional model WRF-ARW. Within the proposes of this project, 3 different model simulations are being done daily to provide the most reliable and accurate weather forecast. In this way, WRF-ARW model is built over a mother domain with 12 km spatial resolution (green squares). The first nested domain, with a spatial resolution of 3 km (yellow squares) includes all the three countries involved in each demo area; Italy, Portugal and Spain. In addition, to reproduce the local scale, nested domains with 1km of horizontal resolution (lilac squares) are developed in the meteorological forecasting system for each demo-area; Mastroberardino, Symington and Codorniu demo-areas.

Figure 4. WRF-ARW domains structure for Mastroberardino demo-area.

[5] Arasa, R., Porras, I., Domingo-Dalmau, A., Picanyol, M., Codina, B., González, M.A., Piñón, J., 2016. Defining a standard methodology to obtain optimum WRF configuration for operational forecast: application over the Port of Huelva (Southern Spain). Atmospheric and Climate Sciences, DOI: 10.4236/acs.2016.62028.

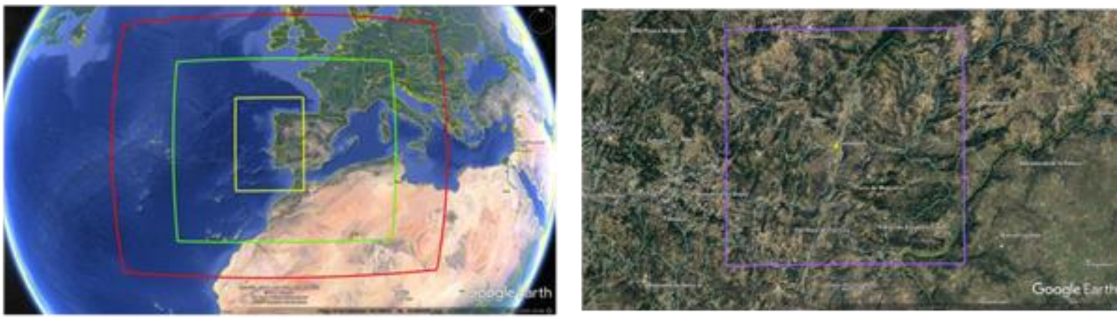

Figure 5. WRF-ARW domains structure for Symington demo-area.

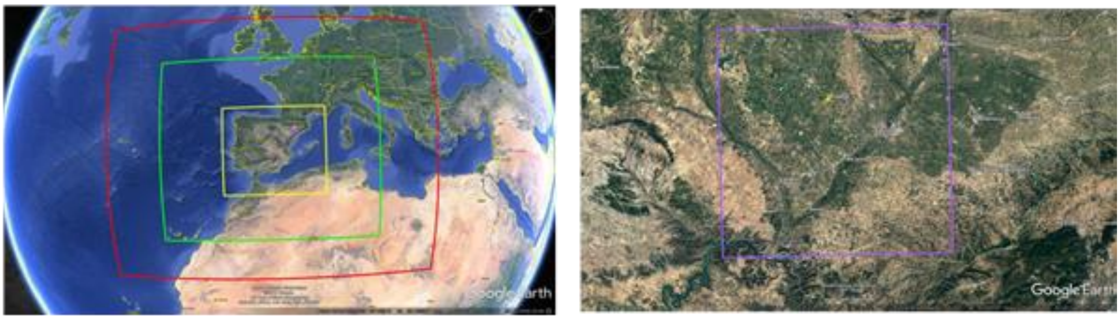

Figure 6. WRF-ARW domains structure for Codorniu demo-area.

More information about the features of the domains in Table 1.

Table 1. Domain's modelling features.

Domain	Horizontal resolution	Region analysis	Vertical levels	Points domain	Size domain (km)
Green	12 km	South – West Europe	30	220 x 220	2640 x 2640
Yellow	3 km	Italy, Portugal, Spain	30	461 x 393	1383 x 1179
Lilac	1 km	Demo areas	30	103 x 103	103 x 103

The temporal resolution of the forecasts is 1 hour for lead times up to 48 hours. The weather parameters that are used from the WRF-ARW data output stream are shown in Table 2.

Table 2: Parameter details of the WRF-ARW forecasts.

Short name	Long name	Unit	Temporal resolution (hours)	Lead times (hours)
t2m	2 metre temperature	K	1	0…48
rh2	2 metre relative humidity	%	1	0…48
wspd	10 metre wind speed	m/s	1	0…48
swrad	Downward short-wave flux at ground surface	W/m^2	1	0…48

3.1.3 Calibration and verification of the deterministic forecast

To calibrate the model, we will consider a past period and we will analyze different options. Modelling results will be validated using a representative past period. At this point of the report, the calibration and validation of the different experiments shown in Table 3 are being carried out. Therefore, the results of these are not shown in this report.

Simulations will be conducted in different periods of the year 2017 (selected because is the most recent year). Numerical simulations will be executed for 30 hours corresponding on every day (hereinafter referred to as daily simulations) included in the period compressed between 01/01/2017 and 12/31/2017, taking the first 12 hours as spin-up time to minimize the effects of initial conditions. To calibrate the model, we have considered the months of January, April, July and October for the year 2017. These months represents the climate variability of the region, being the coldest, driest, warmest, wettest month respectively. A total of 5040 WRF daily simulations will be conducted, corresponding to 120 days, 11 different experiments and 3 different regions.

3.1.3.1 Numerical experiments definition

We have defined up to 11 numerical experiments based on the use of different CMS (cumulus parametrization), SWR (short-wave parametrization) and MPH (microphysics parametrization) schemes. We have focus our experiments on MPH and CMS. The reason to focus on these experiments and not on others is related with the important effect of these schemes over the precipitation and temperatures fields (which they are the most important meteorological variables for agricultural issues).

In this line we have designed: four experiments to analyze optimum microphysics scheme; four experiments to analyze cumulus; one experiment to analyze a different shortwave radiation scheme. Furthermore, we have defined 1 experiment modifying land use and topography databases.

The process followed for the development of the experiments consists in:

- to analyze physical options: modifying the MPH scheme carrying out up to 4 experiments (MPH1-4), comparing against the observed data and selecting the MPH scheme that minimizes the uncertainty; later modifying the SWR scheme and keeping the selected MPH and the rest of schemes defined initially, and comparing again against the observed data and selecting the MPH scheme that minimizes the uncertainty; and finally doing the same for the different CMS experiments (CMS1-4).

- once selected best physical options, experiment with different physiographical database information will be realized.

All experiments will be executed using Noah LSM as land surface model (Chen and Dudhia, 2001)[6].

In all WRF-ARW sensitivity experiments we will use a *ceteris paribus* experimental approach (Campra and Millstein, 2013)[7]. This approach is based into modify only one configuration option and holding all else constant. This approach has always been followed except when exist some WRF-ARW restrictions. In Table 3 the different experiments designed are showed.

[6] Chen, F., Dudhia, J., 2001. Coupling an advanced land surface hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. Monthly Weather Review, 129, 569-585.

[7] Campra, P. and Millstein, D., 2013. Mesoscale climatic simulation of surface air temperature cooling by highly reflective greenhouses in SE Spain. Environmental Science and Technology, 47, 12284-12290.

Table 3: Numerical experiments designed corresponding to physics options and physiographical model database.

Exp.	PBL	SUR	CMS	SWR	LWR	MPH	VER	PHI
INI	*YSU (Hong et al., 2006)*	*MM5 similarity*	*Kain-Fritsch (Kain, 2004)*	*Dudhia (Dudhia, 1989)*	*RRTMG* (Iacono et al., 2008)	*WMS3 (Hong et al., 2004)*	*30*	GTOPO30 and GLCC
MPH1	YSU	MM5 similarity	Kain-Fritsch	Dudhia	RRTMG	WDM6 (Hong et al., 2010)	30	GTOPO30 and GLCC
MPH2	YSU	MM5 similarity	Kain-Fritsch	Dudhia	RRTMG	SBU-Lin (Lin et al., 2011)	30	GTOPO30 and GLCC
MPH3	YSU	MM5 similarity	Kain-Fritsch	Dudhia	RRTMG	NSSL 2-mom w/o hail (Mansell et al., 2010)	30	GTOPO30 and GLCC
MPH4	YSU	MM5 similarity	Kain-Fritsch	Dudhia	RRTMG	HUJI BSM 'fast' (Khain et al., 2010)	30	GTOPO30 and GLCC
SWR1	YSU	MM5 similarity	Kain-Fritsch	RRTMG (Iacono et al., 2008)	RRTMG	Best MPH	30	GTOPO30 and GLCC
CMS1	YSU	MM5 similarity	Grell-Freitas (Grell et al., 2014)	Best SWR scheme	RRTMG	Best MPH	30	GTOPO30 and GLCC
CMS2	YSU	MM5 similarity	KF-CuP (Berg et al., 2013)	Best SWR scheme	RRTMG	Best MPH	30	GTOPO30 and GLCC
CMS3	YSU	MM5 similarity	Multi-scale KF (Zheng et al., 2016)	Best SWR scheme	RRTMG	Best MPH	30	GTOPO30 and GLCC
CMS4	YSU	MM5 similarity	New Tiedtke (Zhang and Wang, 2011)	Best SWR scheme	RRTMG	Best MPH	30	GTOPO30 and GLCC
HRP1	Best Physical Options						30	ASTER (Abrams et al., 2003) and CLC2006

3.1.3.2 Methodology to evaluate

To calibrate and validate the WRF-ARW model data for the different periods, we are using observational data from the European Climate Assessment & Dataset (ECA&D) and the Meteorological agency from Spain (AEMet). Modelled WRF-ARW values from the domain of 3 km horizontal resolution and observed values from local meteorological stations are being compared using different statistics. Values from all the local meteorological stations have been considered as observed values.

To evaluate the model performance of the meteorological variables temperature, relative humidity, wind velocity and wind direction we have use a deterministic numerical evaluation considering four statistics the large amount of methodologies that can be applied[8] [9]: the Mean Bias (MB), the Mean Absolute Gross Error (MAGE), the Root-Mean-Square Error (RMSE) and the Index of Agreement (IOA). These statistics provide information on how uncertain a model is, regarding to the observations[10] and according to them, a benchmark is given[11] [12]. In the next lines are showed the different mathematical expressions of every statistical.

$$MB = \frac{1}{N}\sum_{i=1}^{N}(M_i - O_i) \qquad MAGE = \frac{1}{N}\sum_{i=1}^{N}|M_i - O_i|$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)^2} \qquad IOA = \frac{\sum_{i=1}^{N}(M_i - O_i)^2}{\sum_{i=1}^{N}\left[\left|M_i - \acute{O}\right| - \left|O_i - \acute{O}\right|\right]^2}$$

where O_i and M_i corresponds to the observed and modelled values respectively, is the $\acute{O}$ observed mean value and N corresponds to the product of the number of hours and stations used in the evaluation.

MB is an evaluation of the data tendency; positive values mean that the simulated values are overestimating the observed ones. Similarly, negative values indicated an underestimation of the simulated values over the real ones. MAGE is used to measure the closeness of the modeled and observed values. IOA provides a measure of the match between the departure of each prediction from the observed mean and the departure of each observation from the observed

[8] Carvalho, A.C., Carvalho, A., Gelpi, I., Barreiro, M., Borrego, C., Miranda, A.I. and Pérez-Muñuzuri, V., 2006. Influence of topography and land-use on pollutants dispersion in the Atlantic coast of Iberian Peninsula. Atmospheric Environment, 40, 3969–3982.

[9] Pielke, R.A., 1984. Mesoscale meteorological modelling. Academic Press, London.

[10] Denby, B., Larssen, S., Guerreiro, C., Douros, J., Moussiopoulos, N., Fragkou, L., Gauss, M., Olesen, H. and Miranda, A.I., 2008. Guidance on the use of models for the European Air Quality Directive. ETC/ACC Report.

[11] Emery, C., Tai, E., 2001. Enhanced Meteorological Modeling and Performance Evaluation for Two Texas Ozone Episodes. Final report submitted to Texas Natural Resources Conservation Commission, prepared by ENVIRON, International Corp, Novato, CA

[12] Tesche, T.W., McNally, D.E. and Tremback, C., 2002. Operational Evaluation of the MM5 Meteorological Model Over the Continental United States: Protocol for Annual and Episodic Evaluation. Prepared for US EPA by Alpine Geophysics, LLC, Ft. Wright, KY, and ATMET, Inc., Boulder, CO. http://www.epa.gov/scram001/reports/tesche_2002_evaluation_protocol.pdf

mean. This is a statistic to evaluate with a single value the goodness of fit of a modeling system with respect to the observations[13]. IOA has theoretical range of 0 to 1, with a value of 1 suggesting perfect agreement. RMSE is calculated as the square root of the mean squared difference in modeled and observed values. It is commonly used as a measure of the overall model performance.

Regarding wind direction, statistics have to be considered very carefully due to circular nature of wind direction. For this reason, in the case of the wind direction a modification of the traditional formula of MB and MAGE has been applied[14 15 16].

The evaluation performed is focused on the inner domains, d03 and d04, since the final aim of this study is to find the best model setup for high resolution domains. Statistical evaluation of the meteorological data is achieved by comparing the modeled parameters to the meteorological station observations of mean sea level pressure, temperature at 2 m, wind speed at 10 m, wind direction at 10 m and relative humidity at 2 m. The statistics used for each meteorological parameter and its benchmarks are shown in next table. Wind speed and wind direction are calculated considering calms below 1 ms^{-1}, as wind direction is not reliable for lower speeds[5,16]. The statistics have been calculated from hourly data of the model and observations.

Table 4. Statistics used for model evaluation and benchmark values in case of temperature, relative humidity, wind velocity and wind direction.

Meteorological parameter (reference height)	Statistic	Benchmark
Temperature (2 m)	MB	< ±0.50 K
	MAGE	< 2.00 K
	IOA	≥ 0.80
Wind speed (10 m)	MB	±0.50 ms-1
	RMSE	< 2.00 ms-1
Wind direction (10 m)	MB	< ±10.00°
	MAGE	< 30.00° *
Relative humidity (2 m)	MB	< 10.00%
	MAGE	< 20.00%
	IOA	≥ 0.60

*The benchmark value of 30º is a valid reference value for meteorologically simple areas (locations with low topography complexity and/or land use variance and which meteorology depends basically on the synoptic scale). Typically for the meteorologically complex areas this reference value is significantly higher.

[13] Pérez, V.A., Arasa, R., Codina, B. and Piñón, J., 2015. Enhancing air quality forecasts over Catalonia (Spain) using Model Output Statistics. Journal of Geoscience and Environment Protection, 3, 9-22.
[14] Jiménez-Guerrero, P., Jorba, O., Baldasano, J.M. and Gassó, S., 2008. The use of a modelling system as a tool for air quality management: Annual high-resolution simulations and evaluation. Science of the Total Environment, 390, 323-340.
[15] Soler, M.R., Arasa, R., Merino, M., Olid, M. and Ortega, S., 2011. Modelling Local Seabreeze Flow and Associated Dispersion Patterns over a Coastal Area in North-East Spain: a case study. Boundary-Layer Meteorology, 140, 37-56.
[16] Jiménez, P.A., González-Rouco, J. F., García-Bustamante, E., Navarro, J., Montávez, J.P., Vilà-Guerau de Arellano, J., Dudhia, J. and Roldán, A., 2010. Surface wind regionalization over complex terrain: Evaluation and analysis of a high-resolution WRF numerical simulation. Journal of Applied Meteorology and Climatology, 49, 268-287

Statistical showed in the previous table are complemented with the Directional Accuracy (DACC) statistic[17] [18]. This statistic is a parameter that quantifies the percentage of occasions that the atmospheric model uncertainty for this variable is less than 30º (Eq. 1) being D_i the difference between modelled and observed values.

$$DACC = \frac{100}{N}\sum_{i=1}^{N}\begin{cases}1 & si\ 0° \leq D_i \leq 30° \\ 0\ in\ other\ case\end{cases} \tag{1}$$

For the model complete evaluation also has been defined parallel statistics for the rest of variables: Temperature Accuracy (TACC, Eq. 2), Wind Speed Accuracy (WACC, Eq. 3) and Relative Humidity Accuracy (RHACC, Eq. 4). Values obtained for these statistics have been considered by the authors as the numerical representation of the accuracy, confidence level or reliability of the meteorological modeling system.

$$TACC = \frac{100}{N}\sum_{i=1}^{N}\begin{cases}1 & si\ 0° \leq D_i \leq 2K \\ 0\ \ in\ other\ case\end{cases} \tag{2}$$

$$WACC = \frac{100}{N}\sum_{i=1}^{N}\begin{cases}1 & si\ 0° \leq D_i \leq 2ms^{-1} \\ 0\ \ in\ other\ case\end{cases} \tag{3}$$

$$HRACC = \frac{100}{N}\sum_{i=1}^{N}\begin{cases}1 & si\ 0° \leq D_i \leq 20\% \\ 0\ \ in\ other\ case\end{cases} \tag{4}$$

And finally, to evaluate the accuracy of the precipitation, we have used a categorical evaluation based on the use of the statistics POD, CSI, FAR and SR. The definition of these statistics is showed in the next lines.

$POD = \frac{A}{A+C}$	Probability of Detection. Gives the rate of we—predicted locations among the locations where the reference is over the threshold; a perfect forecast has POD=1.
$CSI = \frac{1}{\frac{1}{(1-FAR)}+\frac{1}{POD}-1}$	Critic Successful Index. Indicates how well exceedances of a threshold were predicted considering false alarms and exceedances not forecasted. A perfect forecast has a CSI=1.
$FAR = \frac{B}{A+B}$	False Alarm Rate. Gives the rate of ill-predicted locations among the locations where the forecast is over the threshold; a perfect forecast has FAR=0.
$SR = 1 - FAR$	Successful Rate. Corresponds to the complementary value of FAR. A value of 100% means that false alarms are not reproduced. A perfect forecast has SR=1.

Where A, B and C are parameters in relation with a comparison between if the observed/modelled value exceed or not exceed a threshold value. The threshold value

[17] Santos-Alamillo, F.J., Pozo-Vázquez, D., Ruiz-Arias, J.A., Lara-Fanego, V. and Tovar-Pescador, J., 2013. Analysis of WRF Model Wind Estimate Sensitivity to Physics Parameterization Choice and Terrain Representation in Andalusia (Southern-Spain). Journal of Applied Meteorology and Climatology, 52, 1592-1608.

[18] Santos-Alamillo, F.J., Pozo-Vázquez, D., Ruiz-Arias, J.A. and Tovar-Pescador, J., 2015. Influence of land-use misrepresentation on the accuracy of WRF wind estimates: Evaluation of GLCC and CORINE land-use maps in southern Spain. Atmospheric Research, 157, 17-28.

corresponds to 0.2 mm/3h. In the next table is showed the description of A, B and C values. This table is known as contingency table.

Table 5. Contingency table defining the parameters A, B, C and D used to calculate POD, CSI, FAR and SR statistics.

Forecasted	Observed	
	Yes	No
Yes	A (successfully)	B (false alarms)
No	C (exceedances not forecasted)	D (exceedances not forecasted and not occurred)

3.2 Ensemble forecasting

Weather forecasts are inherently uncertain because the initial state of the atmosphere can never be known perfectly, and the model equations must be expressed through approximations and simplification in the model system. Moreover, even the smallest uncertainties in the initial conditions of the forecast model tend to rapidly increase over time because of the chaotic nature of the atmosphere. Therefore, rather than integrating a single forecast from a supposedly best guess of the initial state (as was done in short-term forecasting), it has been shown that a better approach would be to start the forecast from several slightly different initial conditions, e.g. 20, and then derive as many, presumably somewhat different, outcomes from these differing initial conditions.[19] This approach is called ensemble forecasting and as outcome, produces forecasts that are given as probability distribution. From these distributions it is possible to calculate local probabilities for different weather events b using thresholds (Figure 7). This information about forecast uncertainty is relevant to provide to all forecast users, i.e. provide weather forecasts un probabilistic terms. Especially when forecasting rare and high-impact weather events.

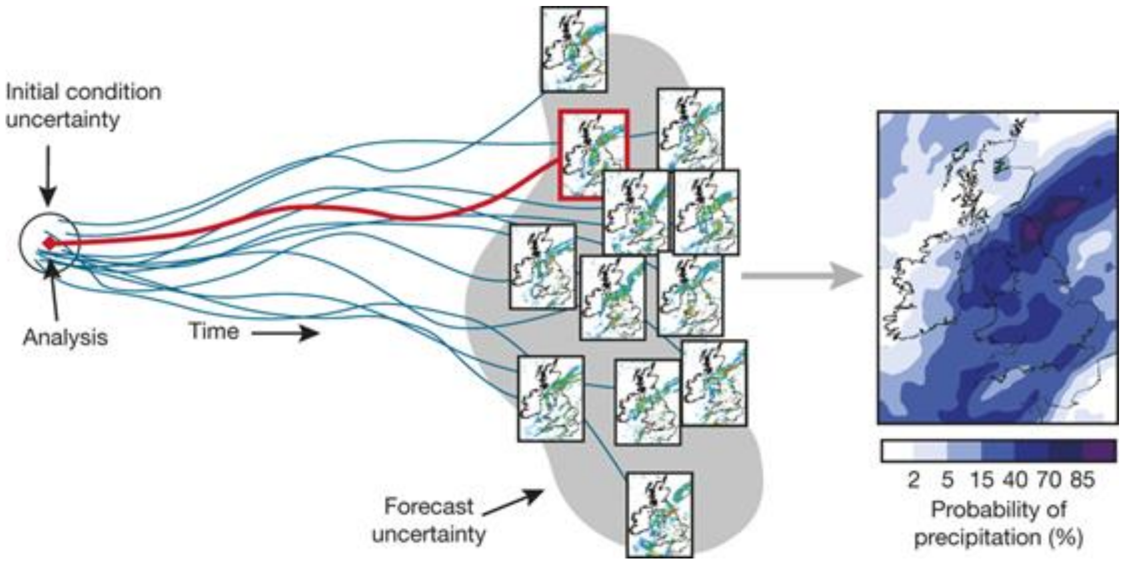

Figure 7. The principle of ensemble-based probabilistic forecasting.

[19] https://www.ecmwf.int/en/about/media-centre/fact-sheet-ensemble-weather-forecasting

3.2.1 GEFS

The National Centers for Environmental Prediction (NCEP) is the leading center that produces global forecasts. The Global Ensemble Forecast System (GEFS) forecast ensemble Is based upon the notion that erroneous forecasts result from a combination of initial analysis errors and model deficiencies, the former dominating during the first five days or so. Analysis errors amplify mostly in the sensitive parts of the atmosphere where strong low-pressure areas develop. These errors then move downstream and amplify and thereby affect the large-scale flow. To estimate the effect of possible initial analysis errors and the consequent uncertainty of the forecasts, small changes to the analysis are made creating an ensemble of many different “perturbed” initial states. In order to save computational time, the ensemble members are run at a lower (~40 kilometres) spatial resolution.

The GEFS consists of 21 individual ensemble members that are created for lead times of up to 16 days (384 hours) four times a day. However, within the VISCA proposes, GEFS products are considered up to 10 days forecast. The temporal resolution of the ensemble forecasts is 3 hours for lead times up to 192 hours, and 6 hours for lead times from 198 to 384 hours. The first member of the GEFS is called the control forecast (CNTL). It utilizes the same current condition and description of model physics but at a coarser spatial resolution. Its significance for the ensemble is that it provides the unperturbed member to which the perturbations for the remainder of the ensemble members are applied. The 21 perturbed members are similar to the CNTL but their initial states and model physics have been perturbed to explore the currently understood range of uncertainty in the observations and the model. They provide a range of possible future weather states. When averaged over many forecasts (although not necessarily for any particular forecast), these have lower skill than either the HRES or the CNTL. However, they do provide an estimate of the forecast uncertainty or confidence. More importantly, the ensemble provides information from which the probability of alternative developments is calculated, in particular those related to risk of extreme or high-impact weather. The weather parameters that are used from the GEFS data output stream are shown in Table 6.

Table 6. Parameter details of the GEFS forecasts.

Short name	Long name	Unit	Temporal resolution	Lead times
t2m	2 metre temperature	K	3 h	0…240 h
tmax	2 metre maximum temperature	K	3 h	0…240 h
tmin	2 metre minimum temperature	K	3 h	0…240 h
rh2	2 metre relative humidity	%	3 h	0…240 h
wspd	10 metre wind speed	m/s	3 h	0…240 h
swrad	Downward short-wave flux at ground surface	W/m2	3 h	0…240 h

4. Weather forecast products

In the VISCA project, two kinds of products are delivered as explained in previous sections; short-term weather forecasts and mid-term weather forecast. While having in two days in advance the weather forecast product is useful for near future in-field activities, having weather forecast information ten days in advance could be useful to wine producers to minimize risks related to coming extreme events. For instance, if a heat wave is forecasted in 6 days, wine producers could act in advance irrigating the field more than usual before the event.

Weather forecast data is distributed in the GRIB file format which is used by the operational meteorological centres for storage and the exchange of gridded fields. Meteosim will convert the processed forecasts products to the GeoJSON format for upload to the VISCA system. Ensemble forecasts for each demo-site will be computed from GEFS data and converted to the TXT file format for the computation of the irrigation products.

4.1 Short-term

Meteosim provides deterministic forecast for the variables identified in Table 7. The deterministic forecasts are calculated by using WRF-ARW output data.

Table 7. The deterministic weather forecast variables provided to VISCA.

Variable (unit)
Temperature (°C)
Wind speed (m/s)
Accumulated precipitation (mm/h)
Relative humidity (%)
Downward short-wave flux (W/m^2)

Figure 8 to Figure 12 below show examples of the deterministic weather forecasts obtained with WRF-ARW model. In these examples, the forecast valid time is the 27th of March 2018, 15:00 UTC. The figures are presented in GeoJSON format as example sent to the VDI.

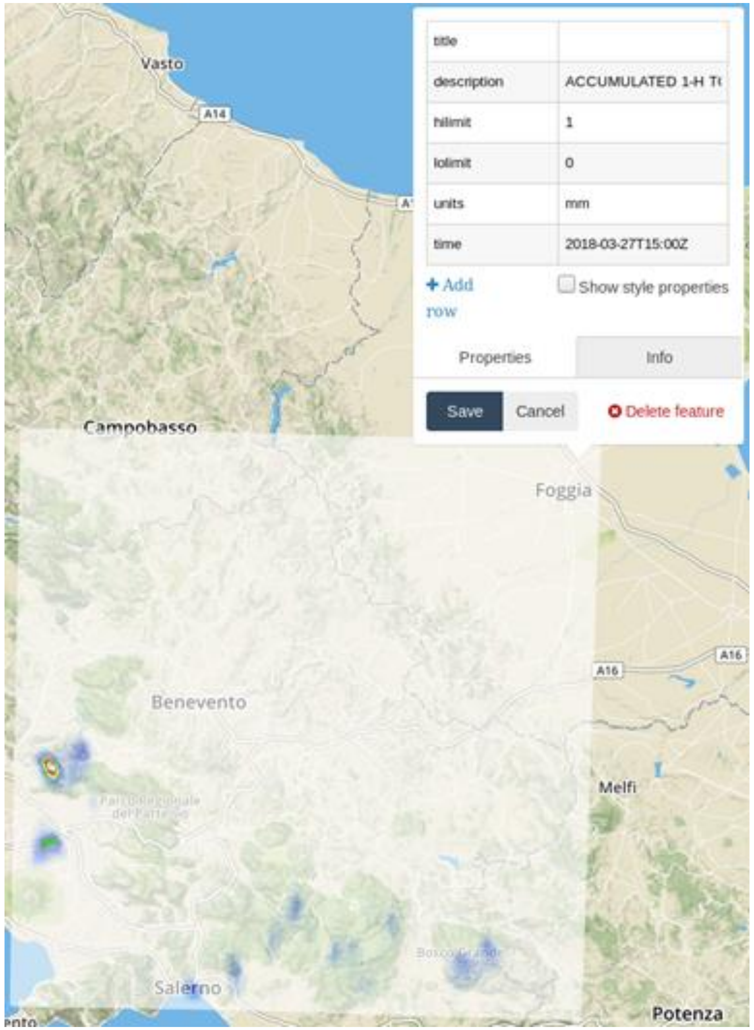

Figure 8. Deterministic forecast for the accumulated precipitation (mm/h).

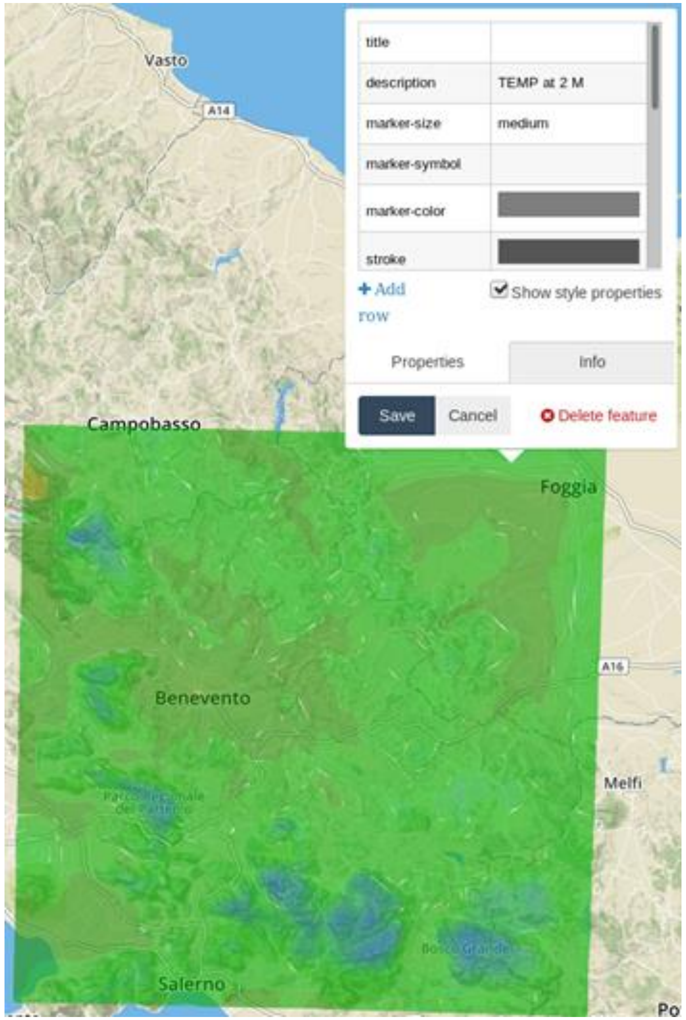

Figure 9. Deterministic forecast for the temperature (°C).

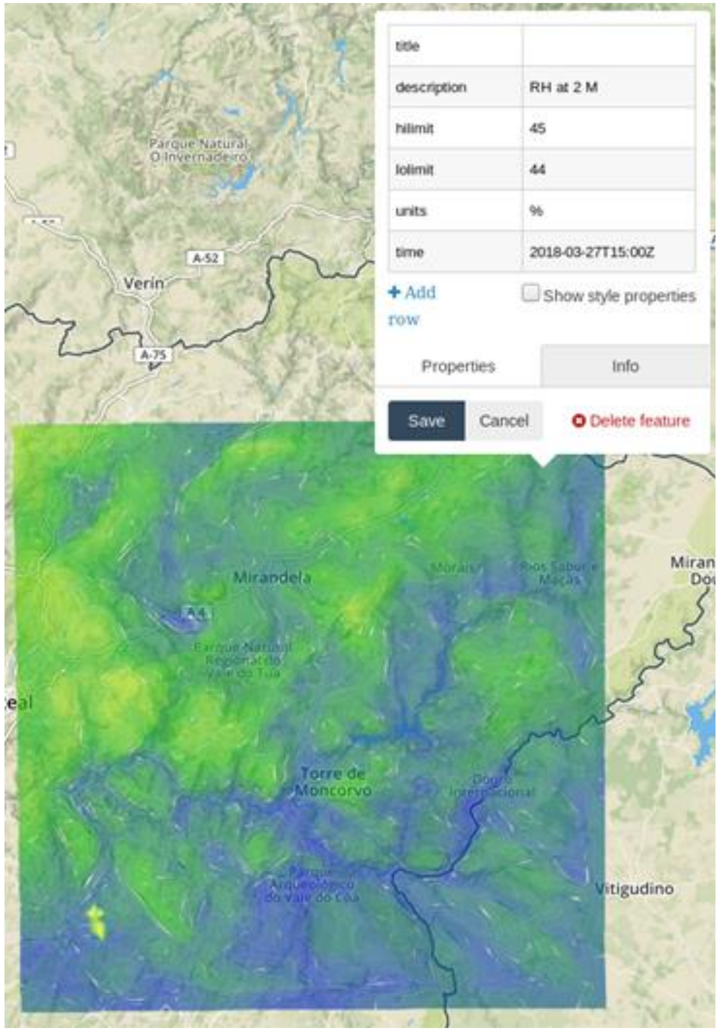

Figure 10. Deterministic forecast for the relative humidity (%).

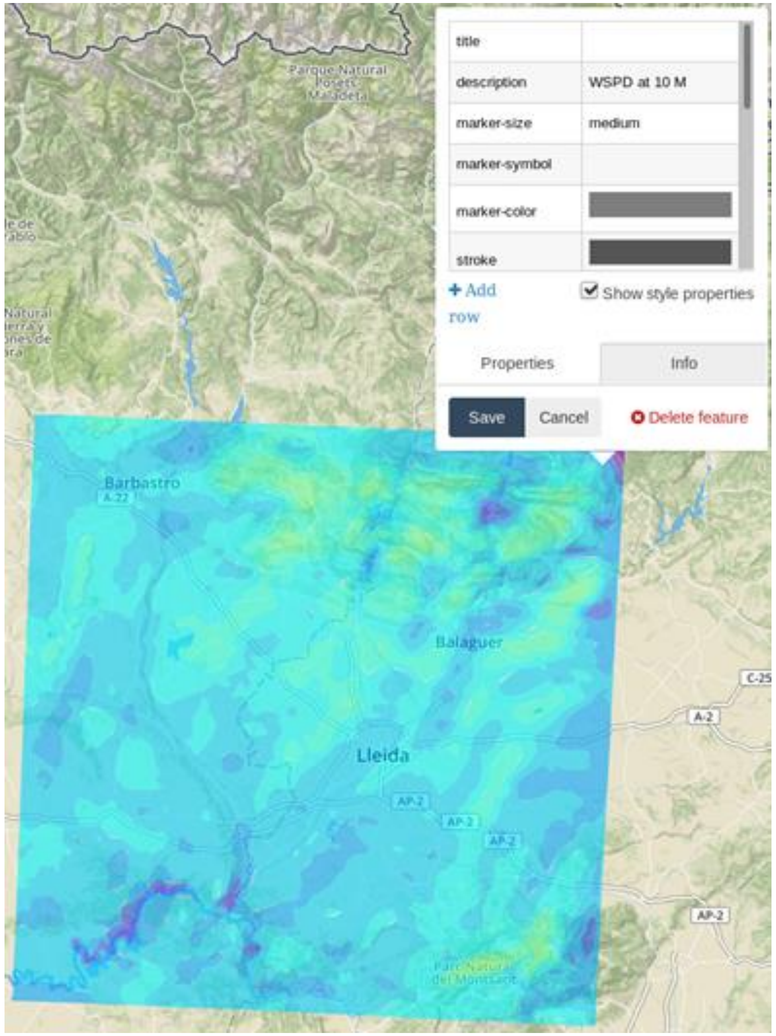

Figure 11. Deterministic forecast for the wind speed (m/s).

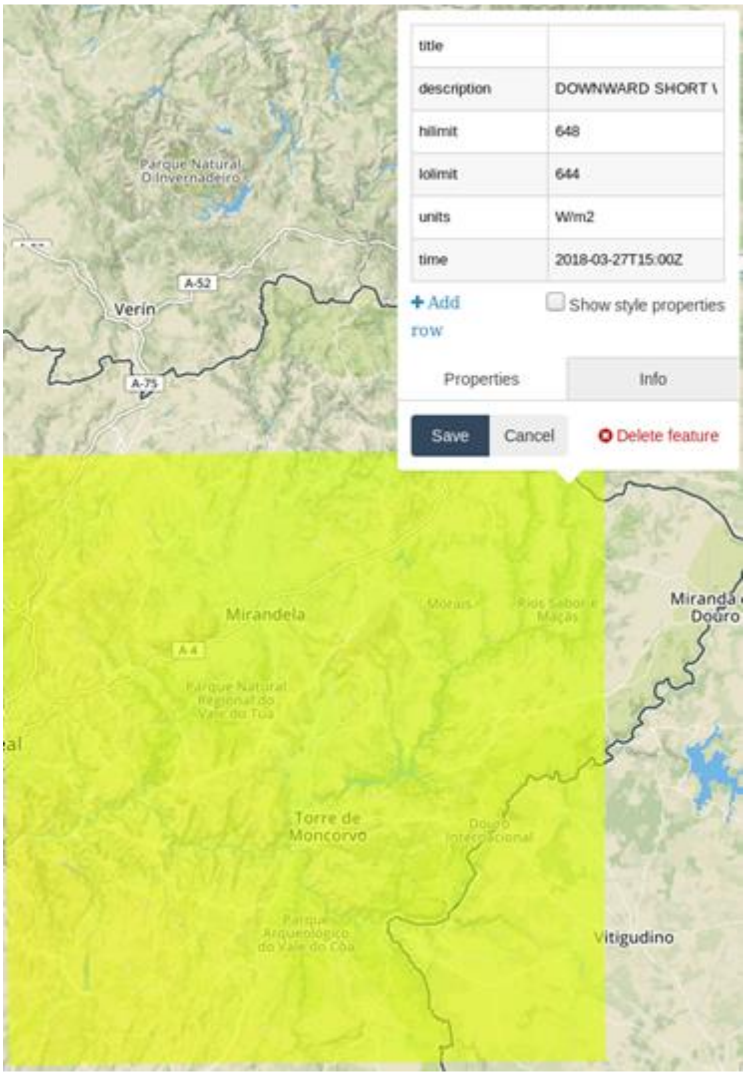

Figure 12. Deterministic forecast for the downward short-wave flux (W/m^2)

4.2 Mid-term

Meteosim also provides probabilistic forecast for some of the variables identified in Table 9. The probabilistic forecasts are computed by using GEFS data. While grid output forecast data will be computed, also point-series output forecast are going to be extracted from the grid for each demo-site.

4.2.1 Grid forecast

A European grid with the probabilistic weather forecast is going to be delivered to the VDI. In this way, for the above variables, a set of bins were defined (Table 8). GeoJSON files contain for each lead time forecast and meteorological field the probability of members within each bin. From Figure 13 to Figure 17 examples about mid-term forecast product are shown. Table 8 shows the variables computed and the associated bins considered to calculate the probabilistic forecast.

Table 8. Bins for each meteorological field.

Variable (unit)	Bins
Mean temperature (°C)	-10 – 40 in °C
Maximum temperature (°C)	0 – 50 in °C
Minimum temperature (°C)	-20 – 30 in °C
Wind speed (m/s)	0, 0 – 2, >2 in m/s
Accumulated precipitation (mm/h)	0, 0 – 10, >10 in mm/day
Relative humidity (%)	0 – 30, 30 – 60, 60 – 90, 90 – 100 in %

4.2.2 Point-series forecast

Point-series forecast for each demo-site are also being extracted from the GEFS data. This data is going to be used by IRTA in order to run the irrigation model. To do that, a TXT files are going to be computed and delivered to the VDI. The meteorological fields considered are the ones showed in Table 9. Moreover, Table 10 shows an example of a file where weather forecast data for each ensemble member (MEMBER), valid forecasting date (valid_date), mean temperature (T2M), relative humidity (RH2M), wind speed (WSPD10M), precipitation (PREC), downward short-wave flux (DSWRF), longitude, (LON), latitude (LAT) and the end-user (COMPANY). In the below example some rows are shown for the demo-site of Codorniu. By the way, the weather forecast data were also extracted for demo-sites of Symington and Mastroberardino.

Table 9. The probabilistic weather forecast variables provided to VISCA.

Variable (unit)
Mean temperature (°C)
Wind speed (m/s)
Accumulated precipitation (mm/h)
Relative humidity (%)
Downward short-wave flux (W/m^2)

Table 10. Mid-term forecast example TXT file.

MEMBER	valid_date	T2M	RH2M	WSPD10M	PREC	DSWRF	LON	LAT	COMPANY
1	2018032700	6.2	66	4.1	0.0	0	0.50742	416.625	Codorniu
1	2018032703	4.8	74	4.0	0.0	0	0.50742	416.625	Codorniu
1	2018032706	4.8	80	4.4	0.0	0	0.50742	416.625	Codorniu
1	2018032709	11.9	63	8.2	0.0	260	0.50742	416.625	Codorniu
1	2018032712	18.7	50	10.7	0.0	480	0.50742	416.625	Codorniu
1	2018032715	20.7	46	12.3	0.0	710	0.50742	416.625	Codorniu
1	2018032718	17.1	56	10.7	0.0	493	0.50742	416.625	Codorniu
1	2018032721	13.4	69	7.9	0.0	0	0.50742	416.625	Codorniu
1	2018032800	11.0	78	5.4	0.0	0	0.50742	416.625	Codorniu
1	2018032803	9.4	84	4.0	0.0	0	0.50742	416.625	Codorniu
1	2018032806	8.0	85	2.7	0.0	0	0.50742	416.625	Codorniu
1	2018032809	15.2	55	4.3	0.0	270	0.50742	416.625	Codorniu
1	2018032812	20.9	37	5.9	0.0	498	0.50742	416.625	Codorniu
1	2018032815	22.8	25	6.5	0.0	740	0.50742	416.625	Codorniu
1	2018032818	18.4	34	3.3	0.0	518	0.50742	416.625	Codorniu
1	2018032821	13.9	39	3.2	0.0	0	0.50742	416.625	Codorniu
1	2018032900	11.9	45	3.0	0.0	1	0.50742	416.625	Codorniu
1	2018032903	10.6	56	3.4	0.0	0	0.50742	416.625	Codorniu
1	2018032906	9.4	76	4.5	0.0	0	0.50742	416.625	Codorniu

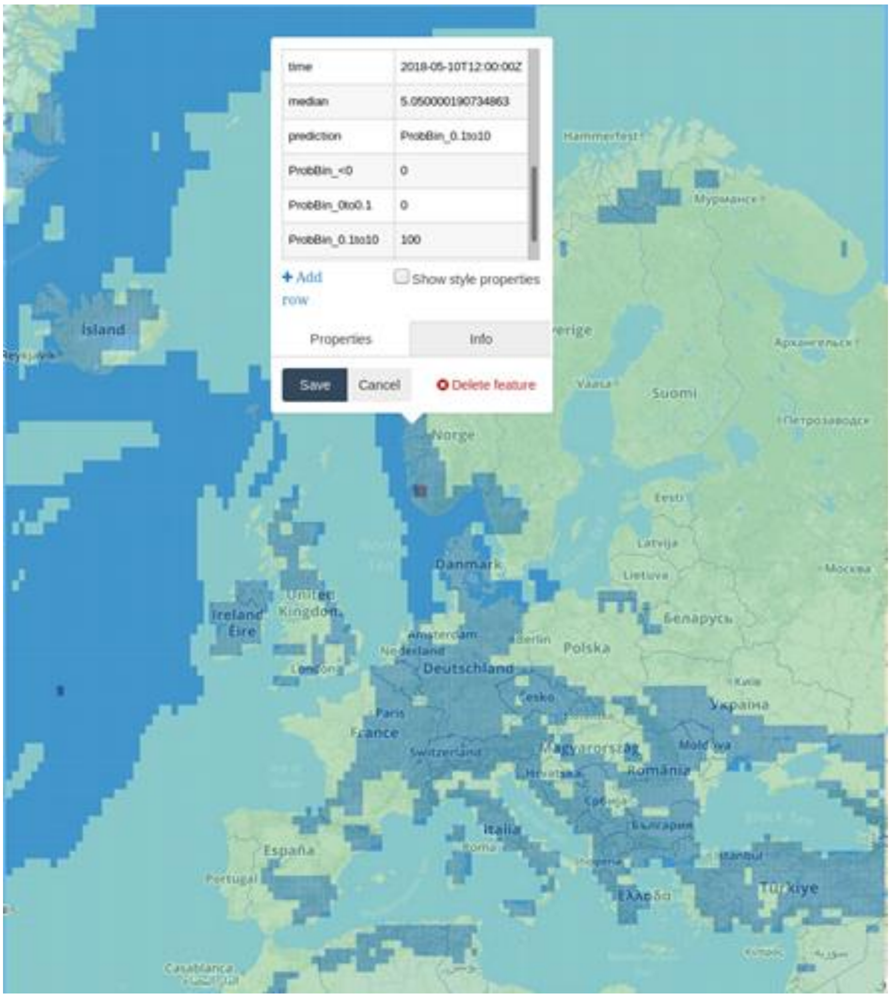

Figure 13. Probabilistic forecast for the accumulated precipitation (mm/h).

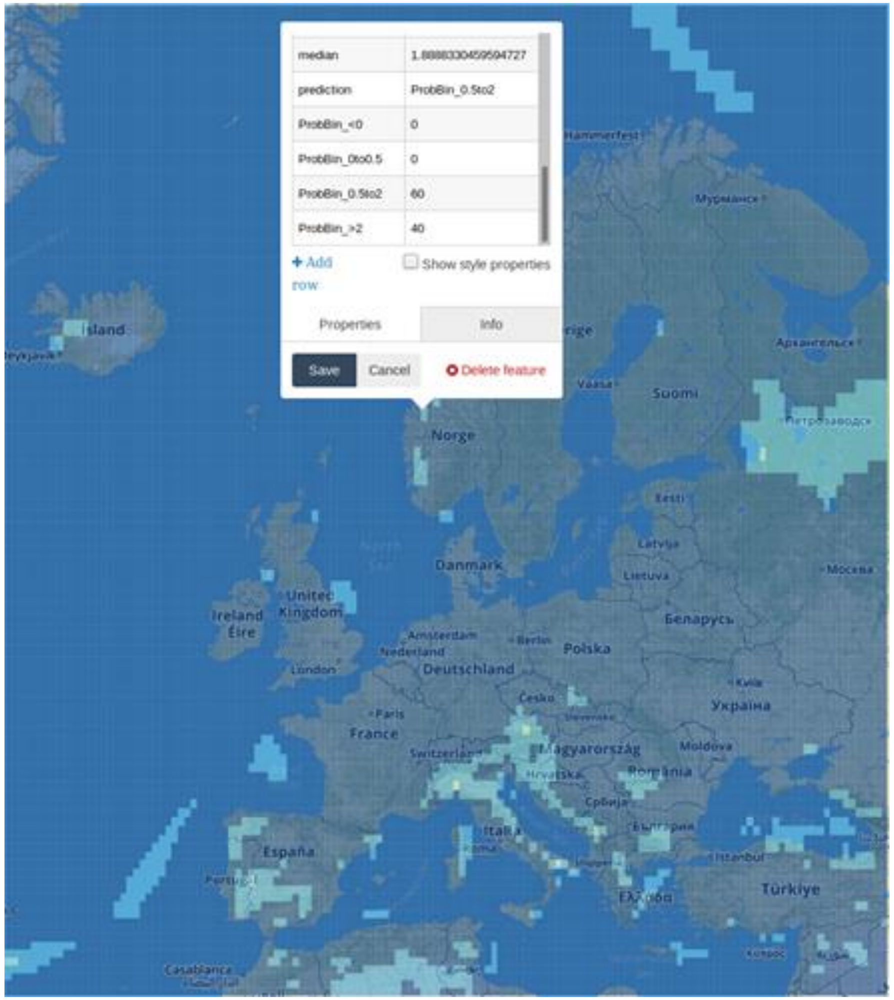

Figure 14. Probabilistic forecast for the wind speed (m/s).

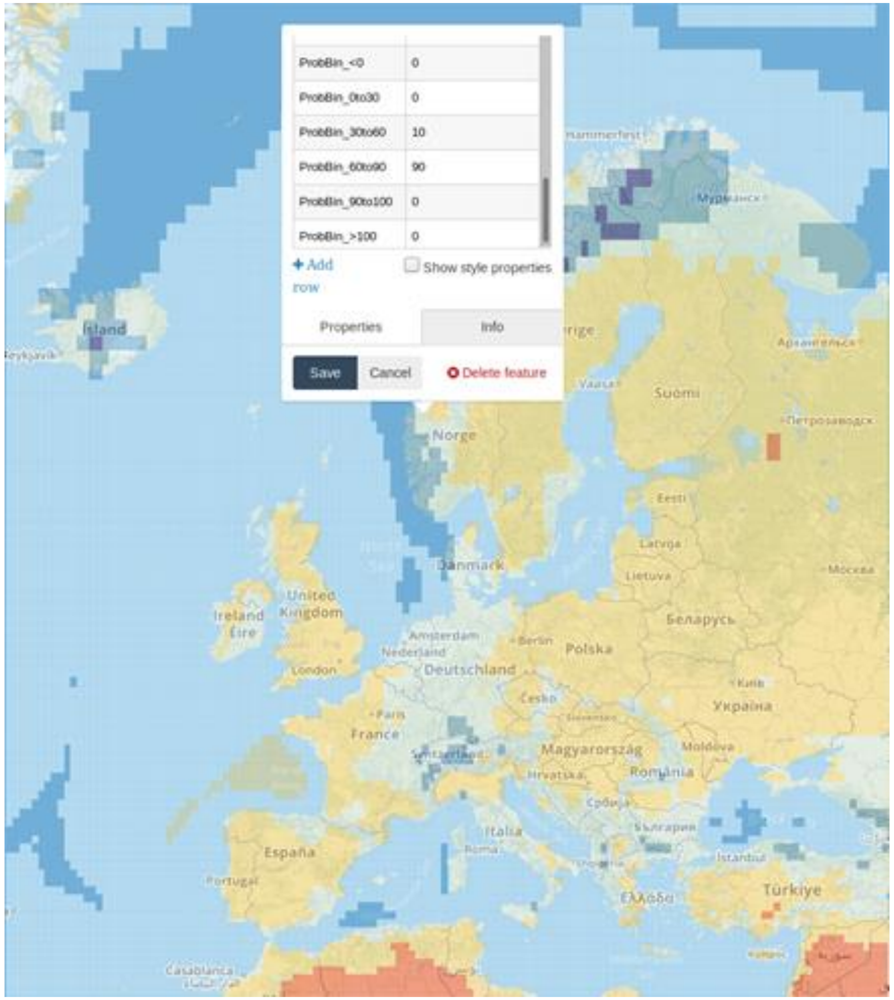

Figure 15. Probabilistic forecast for the relative humidity (%).

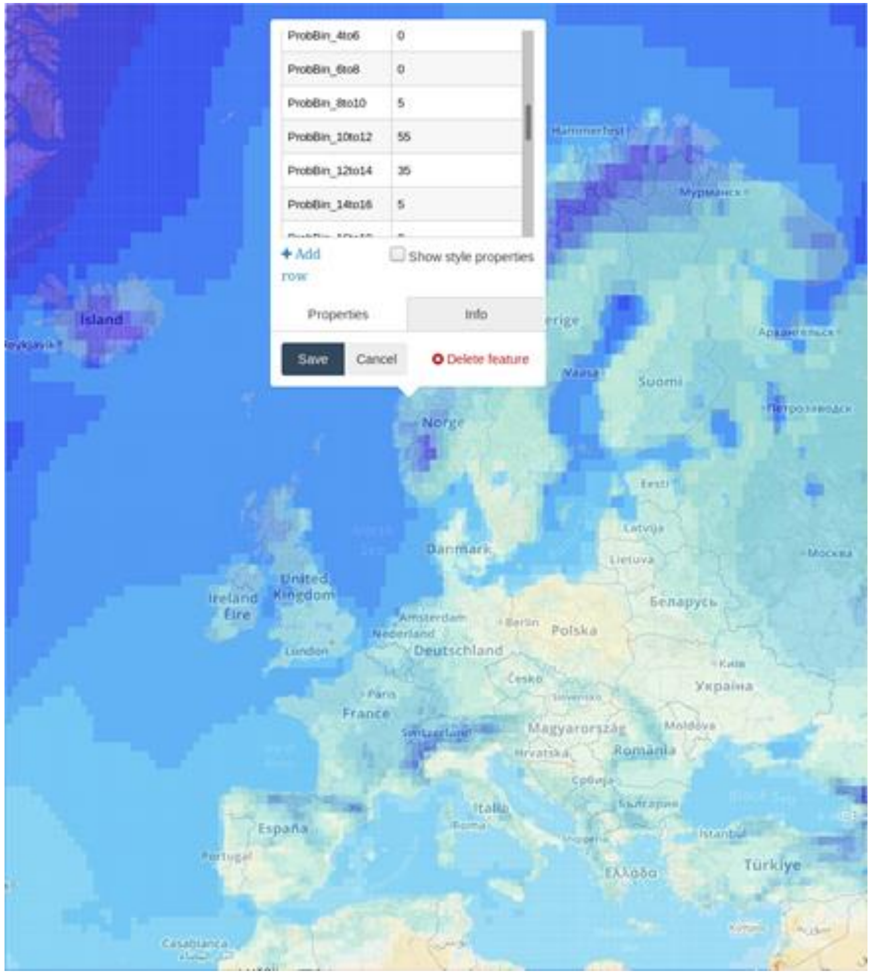

Figure 16. Probabilistic forecast for the maximum temperature (°C).

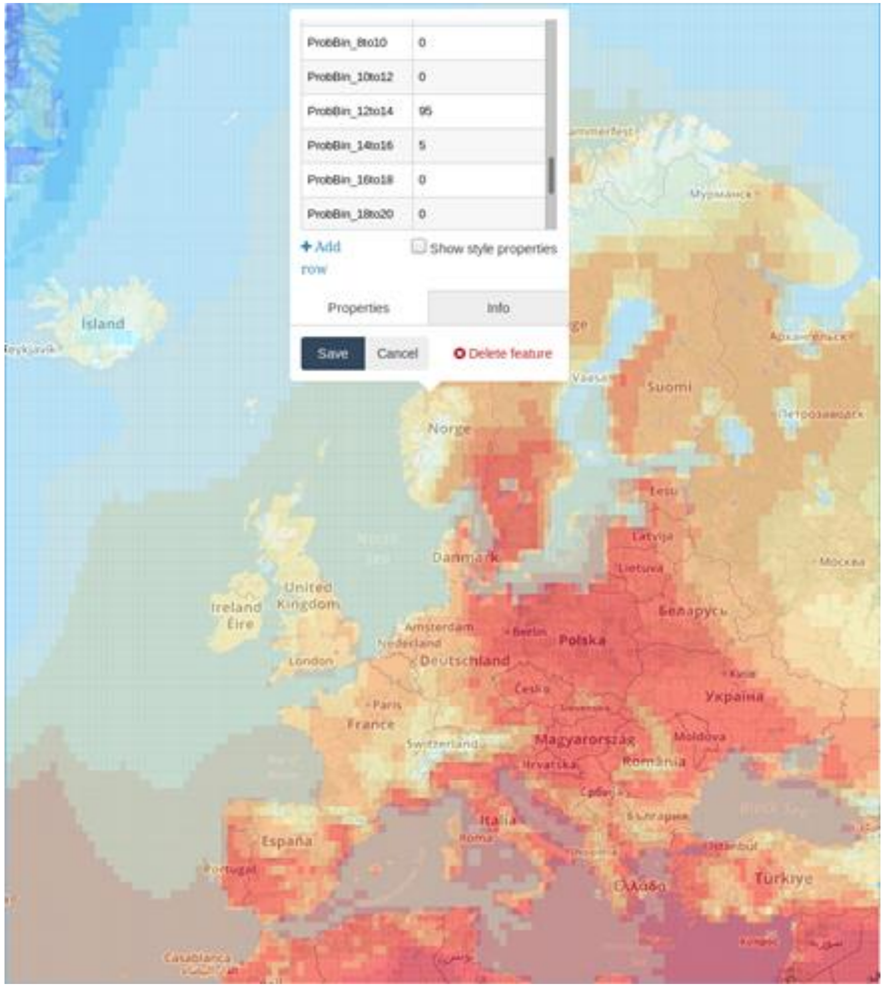

Figure 17. Probabilistic forecast for the minimum temperature (°C)

4.1 Conclusions

In this report, the process of weather forecast development and production from input to output data was described.

As was shown in section 3.1.3, some experiments were defined to improve deterministic forecast. The experiments are being carried out, so the sensitivity analysis results will be shown when the calibration and validation will finish for the three demo-areas. The calibration should improve the overall short-term forecast that is already running without the best model configuration; experiments are still running. The existing of ensemble members allow us to compute probabilistic forecast that can be useful to obtain the uncertainty of the model.

The purpose of these developed weather forecast products is to provide information of the weather conditions, so that necessary preparedness and mitigation actions can be made by wine producers.