This project has received funding from the European Union's Horizon 2020
research and innovation programme under grant agreement No 864337

Smart4RES

Strategies for RES-oriented NWP models' enhancement

D2.1 Strategies for RES-oriented NWP models' enhancement

WP2, T2.1

Version V2.0

Authors: Ivana Aleksovska, Bastien Alonzo, Marie Cassas, Laure Raynaud, Quentin Libois, *Météo-France*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

Disclaimer

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

Technical references

Project Acronym	Smart4RES
Project Title	Next Generation Modelling and Forecasting of Variable Renewable Generation for Large-scale Integration in Energy Systems and Markets
Project Coordinator	ARMINES – MINES ParisTech
Project Duration	November 2019 – April 2023
Deliverable	D2.1: Strategies for RES-oriented NWP models' enhancement
Dissemination level [1]	PU
Nature [2]	R
Work Package	WP 2 – Next generation of weather forecasting models for RES purpose
Task	T 2.1 – RES-oriented NWP
Lead beneficiary	Météo-France (MF)
Contributing beneficiary(ies)	ALL
Reviewers	Tuhfe Göçmen (DTU), Björn Witha (EMSYS)
Due date of deliverable	31 January 2022

[1] PU = Public
PP = Restricted to other program 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)

[2] R = Report, P = Prototype, D = Demonstrator, O = Other

Document history		
V	Date	Description
0.1	15/12/21	First draft submitted by Laure Raynaud
1.0	22/12/21	Review by Quentin Libois (WP leader)
2.0	17/01/22	Review by Tuhfe Göçmen and Björn Witha
2.1	27/01/22	Review by George Kariniotakis (coordinator)
3.0	01/02/22	Final public version released by Coordinator

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

Executive summary

Smart4RES is a H2020 research project that aims to develop and validate next-generation tools enabling to:

- Increase the performance of Renewable Energy Sources (RES) production forecasts by at least 15%,
- leverage the economic value of end-use applications under RES uncertainty.

This document reviews the strategies considered within the Smart4RES project to improve the prediction of atmospheric variables relevant to the RES using Numerical Weather Prediction (NWP) models, with a focus on surface solar irradiance and wind at hub height. Weather forecasts errors are indeed the main source of uncertainty in RES forecasts, and RES intermittency is the main obstacle to the penetration of RES in the electricity grid. Better predicting the evolution of the atmosphere from a few hours to a few days is thus at the core of the energetic transition supported by Smart4RES.

The document first discusses how the representation of aerosols and cloud optical properties in NWP models could be updated to better simulate solar irradiance. In addition it is shown that valuable additional information could be extracted from such models in complement to the variables currently used by most end-users. In particular, outputting cloud optical thickness and the spectral distribution of solar irradiance could be beneficial in many cases. Then a variety of post-processing tools are presented, that take advantage of the wealth of information contained in ensemble simulations, which tend to become the standard of NWP. These include building pseudo-deterministic forecasts (single time series built from ensemble simulations), which outperform deterministic or ensemble-mean forecasts. Seamless forecasts (forecasts that link in time different models, generally having different resolutions and leadtimes) are also presented, which take advantage of a high resolution (short leadtime) model for the first part of the forecast, and minimize the discontinuity when moving to a lower resolution (longer leadtime) model. In addition, probabilistic products targeting the occurrence of relatively rare but critical events, such as wind ramps and cut-out, are derived from the unique high temporal resolution of Smart4RES dedicated simulations. All these tools aim at providing to the final user relevant information in an accessible format to maximize the information passed from the NWP model to the end-user. They participate to reducing by approximately 10% the RMSE for forecasts of the relevant variables. Finally, we present paths to dedicated RES forecasts, which includes the optimization of NWP models for RES purposes and training meteorologists to become experts in RES forecasting.

This work is carried out in the framework of Smart4RES WP2 entitled *Next generation of weather forecasting models for RES purpose*. This report will be later on complemented by deliverable D2.2 that will specifically focus on the added value of increasing the spatial and temporal resolution of NWP models for RES forecasting. Deliverable D2.3 will introduce alternative sources of weather information and will present innovative tools to merge NWP forecasts with these complementary products. Finally deliverable D2.4 will investigate the potential of very high-resolution simulations including assimilation of very fine-scale observations.

<u>Key messages:</u>

- ***Physical variables relevant to the RES and internally used in NWP models could be outputted at very low cost, with significant gain for the users***
- ***Ensemble simulations contain a wealth of information, enabling to build deterministic-like forecasts that outperform deterministic forecasts, and to derive probabilistic decision-aid tools relevant to the RES sector***
- ***Enhanced communication between the atmospheric modeling and energy communities is key to improve RES forecasts***

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

Table of contents

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

List of tables

List of figures

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

Introduction

Numerical Weather Prediction (NWP) models have been developed for nearly a century, primarily to help predicting the occurrence of extreme meteorological events and protecting people and goods. Nowadays, the primary mission of national weather services is to protect people and goods from natural hazards related to atmospheric processes (heavy rain, storms, heatwaves, avalanches, etc.). However the number of economic activities relying on NWP has greatly increased in the last decades, now including aeronautics, agriculture and more recently the energy sector. Despite the constant evolution of the expectations of the final users, only limited changes have been implemented in the way NWP models are designed, optimized, evaluated and run. As a consequence, many users simply take what is available from operational forecasts and have developed sophisticated post-processing tools to adapt the forecasts to their actual needs. As a consequence, interactions between the atmospheric sciences community and the final users of weather forecasts have remained limited, which has in particular precluded a faster improvement of RES forecasts.

In this report we highlight strategies that should be followed to adapt NWP models to the needs of the RES sector. We address this issue by presenting ongoing research carried out at Météo-France and elsewhere in that direction, and try to quantify the added value of implementing such strategies, bearing in mind the objectives fixed in the Smart4RES project in terms of improvement of the weather forecasts for RES-relevant variables, namely that we target a 10-15% improvement.

1. Adapting NWP models to RES

1.1. Some background

NWP models predict the evolution of the atmosphere based on initial (and also boundary in the case of limited area models) conditions. To do so they solve the equations of motion for the air on a discretized grid, accounting for the conservation of mass and energy. NWP models are generally split into two main components: the dynamical core which solves the equations of motion for structures larger than a few grid cells, and the physical parametrizations that account for the impact of all subgrid unresolved processes. The latter mostly include turbulence, convection, microphysics and radiation. Because parametrizations involve many poorly constrained parameters, the latter are practically tuned to ensure that the model overall behaves correctly. The performance of NWP models is evaluated using scores, that routinely compare model forecasts with standard reference observations of wind gusts and surface precipitation. Other variables such as near-surface temperature, humidity, wind and sea-level pressure are commonly evaluated although they are not explicitly included in the scores. The tuning of the models is a complex and key process in atmospheric modeling, that aims at maximizing the scores (Hourdin et al., 2017). As a consequence all the variables that do not contribute to the scores are not expected to be as well predicted as those directly included in them. In particular, wind in altitude and solar irradiance are not specifically evaluated.

A strong constrain for NWP modeling is also the capability to perform forecasts more rapidly than the real time. In practice it takes less than one hour for a supercomputer to make a forecast for around 48 hours ahead. It means that any improvement to the model that has a negative impact on the total computation time cannot be implemented operationally. Improvements can come from a refinement of the physical

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

parametrizations and/or from an increase in spatial and temporal resolutions. Fortunately, NWP models rely on an ever-increasing computational performance which allows them to become increasingly more time-consuming.

Although a significant part of the research in atmospheric modeling aims at improving the physical parametrizations by refining our understanding of the physical processes themselves or by constraining the parameters based on observations and high-resolution models (Couvreux et al., 2021), this aspect will not be elaborated further in this report because this is a continuous, slow process that is not specific to the RES sector. Although improvements in the physical parametrizations are expected to increase the performance of the models, including for RES purpose, in this section we focus on actions that can be taken right away with minor effort, with a particular focus on solar irradiance.

1.2. Extracting new variables from NWP models

The radiative variables output from NWP models directly come from the radiative scheme of the model. This parametrization primarily aims at computing the net radiative flux at the surface (i.e. the sum of the net longwave and shortwave fluxes) from which the surface energy balance is computed, and the vertical profiles of radiative heating (or cooling) rates which describe how absorbed (emitted) radiation locally warms up (cools down) the atmosphere. To achieve this, the radiative scheme needs to account for the numerous properties of the atmosphere that drive the interactions with solar radiation. As the radiative properties of the atmosphere greatly vary across the solar spectrum, radiative schemes generally perform the calculations over several contiguous spectral bands. The resulting spectral fluxes, though, are not routinely outputted. Likewise, radiative schemes generally distinguish diffuse and direct radiation, the latter being loosely defined as will be explained further. Finally some internal variables such as cloud optical thickness are computed in the radiative scheme but not outputted either, while this is a valuable information for the final user. Below we detail which quantities could advantageously be provided by NWP models without any extra computational cost, the only drawback being the need to save and store more data.

1.2.1. Spectral fluxes

The response of solar panels and other solar devices greatly depends on the spectral distribution of solar irradiance (Fernandez et al., 2014). When only broadband fluxes are provided by NWP models, the user has to make some assumption on the spectral distribution. Under cloudy conditions this approximation can result in errors in PV production forecast larger than 10% (Lindsay et al., 2020). Using the spectral information contained in the radiative scheme would avoid making such an assumption, and it simply implies saving variables (namely spectral surface fluxes) that are already computed. For instance in the operational version of AROME (the operational limited-area model of Météo-France), the fluxes are computed over 6 bands, and in Integrated Forecasting System (the model of the European Center for Medium Range Weather Forecast) they are computed over 14 bands (Figure 1). We show in Figure 2 an example of direct solar fluxes computed with AROME. Assuming a simple scaling between spectral and broadband fluxes can result in errors up to 2% in the useful fraction of irradiance, which directly translate into errors in PV forecasts. Accessing this information would avoid a loss of information that a posteriori needs to be estimated, for instance by trying to compute the Average Photon Energy (APE, Norton et al., 2015).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

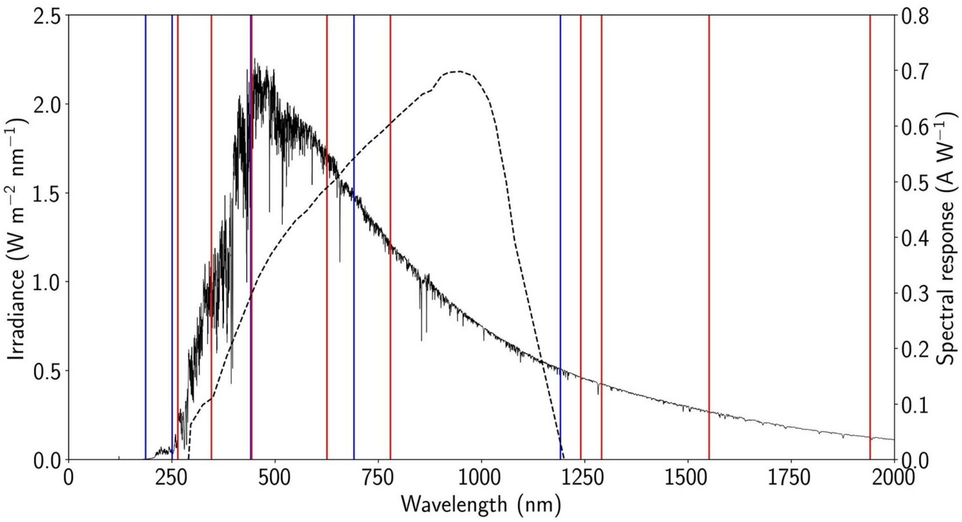


Figure 1: The dark line shows the spectral solar irradiance at top-of-atmosphere (Kurucz, 1995) and the vertical lines indicate the spectral bands from AROME (blue) and IFS (red) models. The dashed line shows the typical spectral response of a mono-crystalline silicium solar cell (https://pvpmc.sandia.gov/modeling-steps/2-dc-module-iv/effective-irradiance/spectral-response/).

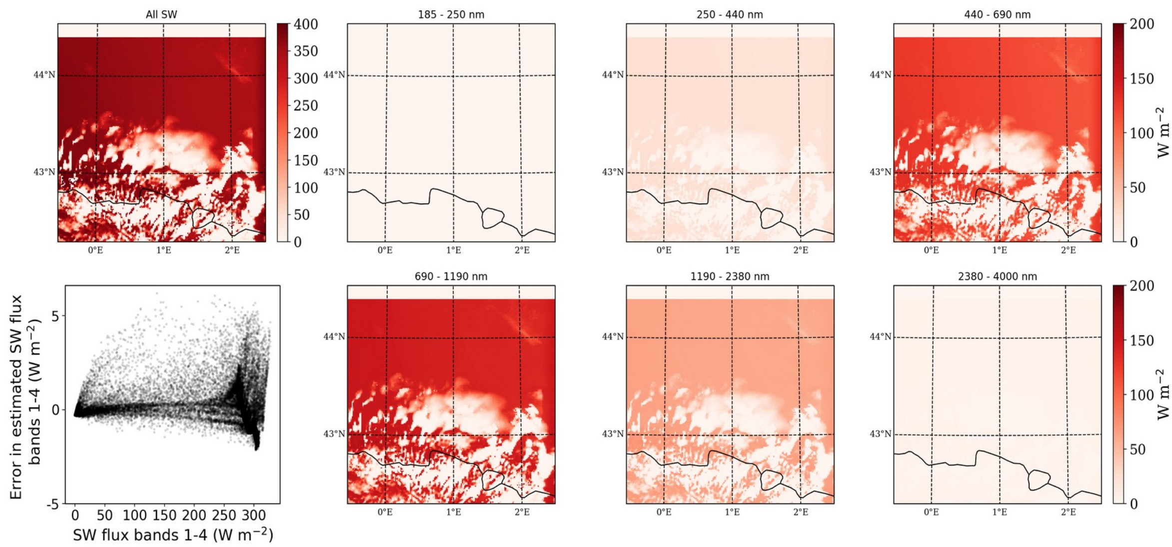


Figure 2: Instantaneous direct fluxes simulated by AROME in 6 spectral bands over a spatial domain centered around Toulouse (South-Western France). The top left panel shows the broadband SW flux. The bottom left panel shows the error in the estimated cumulated flux up to 1190 nm (roughly the cutoff wavelength of solar panels) assuming proportionality with broadband flux, as a function of the actual flux in this spectral range.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

1.2.2. Cloud fraction seen by the radiative scheme

Although algorithms designed to estimate solar energy production from NWP models forecasts do not use spectral outputs, they may rely on the cloud fraction predicted by the model. Usually, only the total cloud fraction is provided as output, which is the cloudy fraction of the grid point as seen from the surface or from satellite. Practically this total cloud fraction is diagnosed within the NWP model from the vertical profile of cloud fraction for each grid cell. The combination of these successive partial cloud fractions into a total cloud fraction is not straightforward, and requires assumptions on the overlapping of cloudy layers (Räisänen et al., 2004). These assumptions can be different in the radiative scheme (where it actually impacts the solar fluxes) and in the cloud scheme (where it is only a diagnostic). This can result in inconsistencies between the information outputted by the model (the cloud fraction from the cloud scheme) and that actually used to compute the solar fluxes. Hence it may be useful either to provide only the total cloud fraction seen by the radiative scheme, or to provide both variables if it occurs that both provide complementary information.

1.2.3. Cloud optical thickness

Clouds are responsible for the largest and fastest variations of surface solar irradiance, as they can reduce by up to 95% the irradiance compared to clear sky conditions. The primary quantity that drives the cloud radiative impact is the cloud optical thickness (COT). COT increases linearly with the total amount of cloud condensate (liquid water path (LWP) in the case of liquid clouds) and with the inverse of effective radius r_{eff} of cloud particles. The amount of condensate is a standard prognostic variable in NWP models, but r_{eff} is not. Generally it is estimated in the radiative scheme from the liquid water content (LWC), using empirical relations based on observations. Although the COT is the key quantity to describe the radiative impact of a cloud this quantity is typically not an output. As a consequence, the user only has access to insufficient cloud information to predict solar energy resource: cloud fraction, cloud altitude, LWP. As clouds modify the spectral distribution of light via preferential absorption and scattering, their optical thickness drives the spectral distribution of radiation. This information can to some extent complete or replace the information on the spectral fluxes.

1.2.4. Direct fluxes

It is worth noting that the definition of direct radiation can vary from one community to another, depending on the final use of this information (Xie et al., 2022). From the point of view of NWP models, any radiation that does not deviate much from the direction of the Sun is considered direct, because in the end the main difference between direct and diffuse radiation will occur at the surface, when the surface reflectance could differ between direct and diffuse radiation or when surface features can create shadows. A widely used approximation in radiative schemes is the two-stream approximation, which greatly simplifies the description of radiative transfer in a scattering atmosphere by considering only fluxes going upward or downward. However this approximation is more accurate when scattering is isotropic. Because scattering by large particles such as aerosols and clouds is characterized by a strong forward peak (typical of Mie scattering), this approximation is not very accurate. To circumvent this it is common to consider that any radiation slightly scattered can be treated as unscattered. The δ-Eddington approximation replaces the original phase function by the sum of a Dirac and a contribution which is much more isotropic (Joseph et al., 1976). Doing so means that the direct radiation computed by the model actually includes slightly scattered radiation. It is clear from the loose definition of forward scattering that the amount of scattered radiation that is assumed unscattered depends on the threshold chosen to distinguish

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

scattered and unscattered radiation. The choice made in NWP models primarily aims at getting the most accurate global horizontal irradiance. When direct irradiance is critical (in particular for solar concentration devices) this definition might not be satisfying. In this case it is either possible to adapt the delta-scaling (Villefranque and Hogan, 2021), or to scale the truly unscattered radiation (which only depends on total optical thickness via the Beer-Lambert law) to include slightly scattered radiation (Sun et al., 2016, Räisänen and Lindfors, 2019). In any case it should be clear to the user that the direct flux outputted by a NWP model can be a quite different quantity than the one of interest.

1.2.5. Extra time steps

Finally it's worth bearing in mind that the internal time step of NWP models is much shorter than the standard time step of the outputs. For instance the internal time step of AROME is 50 s, compared to the standard 1 hour output. The internal time step is directly related to horizontal resolution to satisfy the Courant-Friedrich-Levy condition which states that wind should not transport variables more than one grid away in one time step. This roughly means that Δx should be less than $V \cdot \Delta t$, where V is the typical wind speed. With a velocity of 20 m s^{-1} and a spatial resolution of 1.3 km this gives Δt less than 65 s. The reason why not all time steps are stored is mostly an issue of data storage and time of writing. Although it is probably redundant to extract all time steps, storing outputs more frequently could definitely be useful for a large variety of end-users. This will be explored in Section 2.4.

1.3. Refining the radiative scheme

The main issues in solar irradiance forecasts are related to the correct prediction of cloud occurrence (in particular fog, stratus and cirrus which are very difficult to simulate, Köhler et al., 2017), a shortcoming of NWP models not restricted to RES forecasts. Yet several improvements can be achieved by refining the representation of cloud-radiation and aerosol-radiation interactions, without directly tackling the issue of cloud modeling which combines turbulence and microphysical issues. Some paths to improvement are discussed below, which constitute active research actions at Météo-France initiated in the context of Smart4RES.

1.3.1. Near-real time representation of aerosols

Although clouds are responsible for the largest forecasting errors of solar irradiance, in many areas largely relying on solar energy such as Mediterranean countries, clear-sky conditions are dominating and the impact of aerosols is non-negligible (Gutiérrez et al., 2018). In most NWP models aerosols are accounted for via monthly climatologies, meaning that the amount of aerosols considered in the model at one location will be the same every year for a given calendar date. Although such climatology may capture the average annual aerosol load and the main seasonal variations, it does not capture aerosols events such as dust outbreaks (Córdoba-Jabonero et al., 2021) which can have a significant impact on the power production (Rieger et al., 2017). There are several options to tackle this issue. The first is the explicit simulation of the transport and physical evolution of aerosols as prognostic variables, as do chemical transport models (CTM) like MOCAGE (Josse et al., 2004). However it requires a huge amount of computation time and currently appears unrealistic for operational forecasts. An alternative is to use forecasts from a CTM to force aerosols concentrations in a NWP model. Such strategy is currently being explored at Météo-France for AROME. To quantify the potential added value of such an upgrade, we have used one year of AROME atmospheric outputs to simulate solar irradiance and PV production using the radiative

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

scheme ecRad (Hogan et al., 2018) and the PV code WO2PV (Lindsay et al., 2020). To this end we combined one year of atmospheric variables predicted by AROME with various sources of aerosols including Copernicus Atmosphere Monitoring Service (CAMS) monthly climatology (Bozzo et al., 2020) and CAMS near-real-time products. The comparison of both sets of simulations gives a hint to the errors encountered when using a monthly climatology, and to the potential improvement that could result from using near-real-time aerosols forecasts for operational weather prediction. Practically, instantaneous and local errors up to 100 W m^{-2} could be avoided. The annual RMSD of solar irradiance between both sets of simulations reaches 30 W m^{-2} in very polluted areas for a mean RMSD of around 8 W m^{-2} over AROME domain (Figure 3). This suggests that using aerosols forecasts in NWP models could already decrease the RMSE of solar irradiance forecast by about 3% in all conditions.

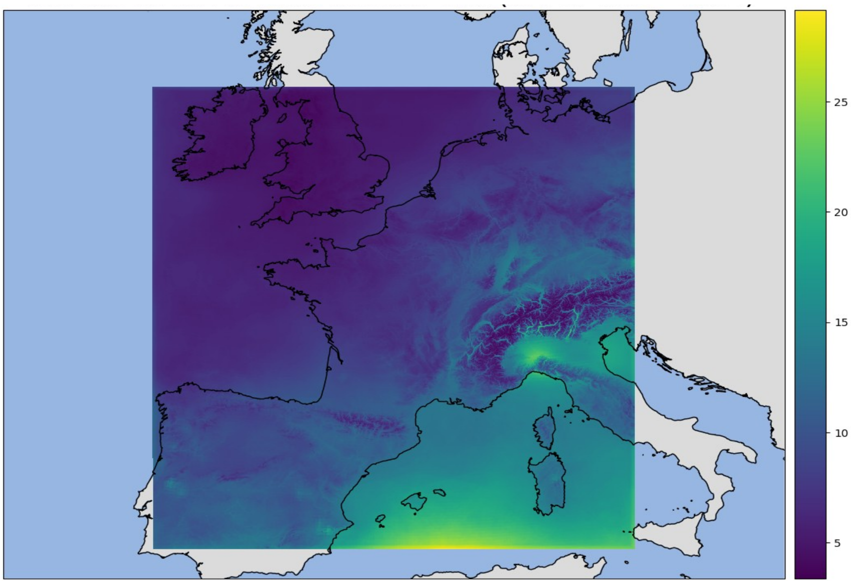


Figure 3: Root mean square deviation (RMSD, in W m^{-2}) between simulated solar irradiances using CAMS aerosols climatology and CAMS near-real-time products for the period August 2019 –July 2020, for the AROME operational domain.

To further illustrate the potential improvement of using near-real-time aerosols, we show in Figure 4 the scatterplot of irradiance forecast errors (with respect to observations from Météo-France's pyranometers network) as a function of the difference between the aerosol optical depth (AOD, the primary quantity that characterizes the radiative impact of aerosols) in the near-real-time products and in the operational climatology. As expected, when the operational climatology overestimates the aerosol load, the solar irradiance is underestimated. The linear regression of both quantities can be used to correct the irradiance forecasts by adding a linear contribution to the irradiance forecast. Doing so allowed to reduce the RMSE from 21.3 to 16.2 W m^{-2}, suggesting that nearly 25% of the RMSE in clear-sky conditions could be removed by appropriately accounting for the high-frequency variations of aerosols. Such a simple strategy could be utilized by end-users in post-treatment of NWP forecast at a very limited computational cost.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

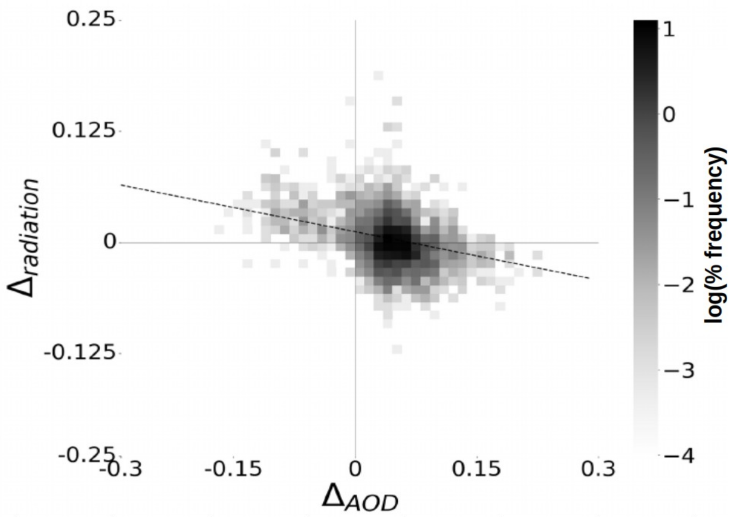


Figure 4: Scatterplot of the relative errors in solar irradiance forecasts for AROME in clear-sky conditions (with respect to ground measurements across France in August 2020) as a function of the difference in aerosol optical depth (AOD) between AROME climatology and CAMS near-real-time product. The dashed line indicates the linear regression used to correct solar irradiance forecasts.

1.3.2. Cloud optical properties

The modulation of surface irradiance by clouds depends on their detailed microphysical properties (concentration of droplets or ice crystals, particle shape, particles size distribution (PSD)), which are poorly simulated by NWP models. As a consequence the cloud optical properties are generally estimated from bulk cloud properties (mostly from water contents) based on empirical relationships, often derived from observations. In practice, such relationships were obtained for particular cloud conditions, but tend to be used universally in NWP models. It means for instance that from an optical point of view, cumulus clouds and fog are treated exactly the same way, although their microphysical properties are obviously different. It is possible, though, to improve this representation. First, more detailed microphysical schemes (the parameterization that describes the formation and evolution of cloud particles) are now implemented in NWP models (Jouan et al., 2020). These new generation models do provide quantitative information about particle size that can be advantageously used in the radiative scheme. In addition it is possible to use the same assumptions on the shape of the PSD in the microphysical and radiative schemes, ensuring consistency throughout the cloud physics. In this context we have developed a set of parameterizations of optical properties for liquid clouds

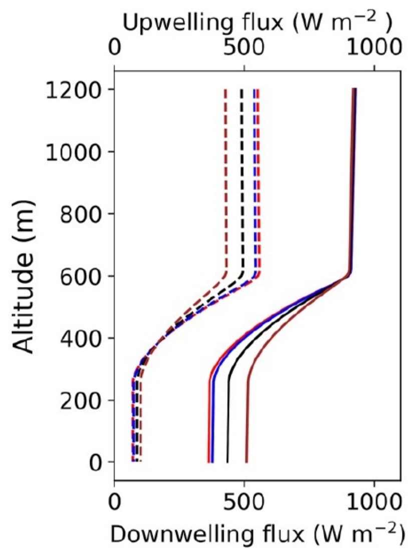


Figure 5: Simulated downward (solid) and upward (dashed) solar fluxes with a stratocumulus cloud. The colors correspond to various shapes od the PSD, all other things being equal. Taken from Jahangir et al. (2021).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

(tables that provide the fundamental optical properties of a cloud as a function of cloud droplets effective radius and shape of the size distribution). Our focus was on the shape of the PSD, because it both affects the estimation of the effective radius (Martin et al., 1994) (when this quantity is not provided by the microphysical scheme) and the estimation of the optical properties (Slingo, 1989). To illustrate how the assumed shape impacts the overall simulation of solar irradiance we show in Figure 5 the vertical profiles of downward irradiance under a stratocumulus cloud, for various PSD shapes. It shows that the surface irradiance can vary up to 20% depending on the assumed shape, highlighting the large, and up to now unquantified, impact of the PSD shape. The new parameterizations have been implemented in the research model Meso-NH (Lac et al., 2018) and will be implemented soon in the AROME model to further asses the impact in an operational context.

More generally developers of NWP models are more and more trying to build consistent physical parametrizations that communicate with each other and use as much information as possible coming from other parametrizations. This is a major change compared to the times when physical processes were split among distinct communities that hardly knew about the developments of the others.

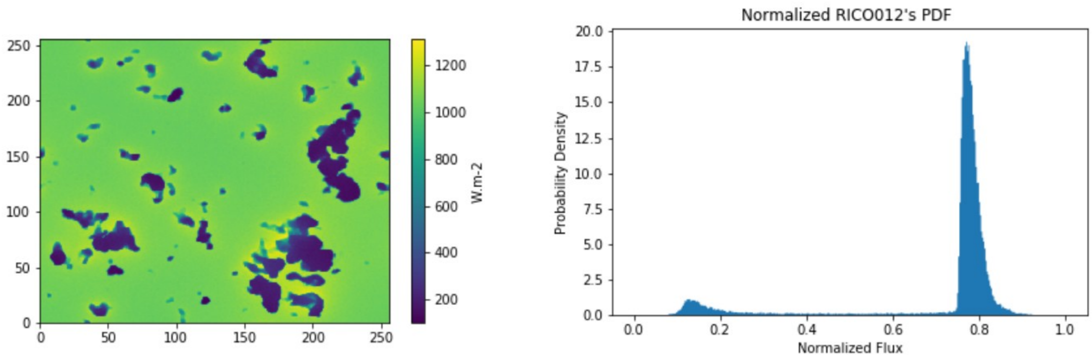
Another shortcoming of cloud-radiation interactions in NWP models is the fact that precipitating particles (rain, snow, graupel) are often not accounted for in the radiative scheme. It means that from a radiative point of view, falling snow is treated exactly as clear-sky, while obviously snow does reduce visibility, hence solar irradiance. It results from the assumption that the radiative impact of a few large particles is much less than that of numerous small suspended cloud particles. Although this might be true on average, Hill et al. (2018) pointed out that singular events do show a significant impact of precipitating particles. Given that a large fraction (about 75%) of condensed water in NWP models corresponds to precipitating particles, their radiative impact could be significant. Accounting for such particles would require the computation of optical properties distinct from those of the clouds, which are valid for particles up to 100 μm only, while precipitating particles can be one millimeter or more. Interestingly, accounting for snow when simulating lidar backscattering from outputs of GCMs, Cesana et al. (2021) showed an improvement between simulated and observed backscattered signals, suggesting that the radiative impact of snow is not negligible. Analysis of singular cases when solar irradiance was poorly predicted by AROME also pointed to situations with significant amounts of snow or rain. This overall suggests that accounting for the radiative impact of precipitating particles would improve the prediction of solar irradiance in a number of situations, although the overall impact needs further quantification.

1.3.3. 3D radiative transfer

In NWP models, the simulation of radiative transfer relies on the plane parallel hypothesis. It means that each atmospheric column is considered as a stack of infinitely extended layers. So-called 3D effects, which can be described as the interactions between adjacent atmospheric layers, are completely ignored, which results in inaccurate irradiance forecasts when clouds are present. A very simple illustration is the fact that in a model, the cloud shadow always lies below the cloud, whatever the position of the Sun in the sky. Likewise, a clear-sky column will feature clear-sky irradiance, even though a neighboring column may contain a cloud hiding the Sun. These shortcomings are mostly due to the computational efficiency required by radiative schemes, which cannot handle the whole 3D field of atmospheric variables and need to be as parallelized as possible, meaning that computations in columns should be independent of each other. This approximation has been considered acceptable in large-scale models where columns are sufficiently large, but it becomes a much more questionable assumption in high-resolution models at the kilometer-scale or less. A costly alternative would be to use 3D

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

radiative transfer schemes in NWP models, but this is still far from being an operational solution. Meanwhile, recent studies have focused on the simulation of solar irradiance fields based on NWP models, but using offline 3D radiative transfer simulations (Gristey et al., 2020). For instance the numerical tool htrdr (Villefranque et al., 2019) was developed to simulate irradiance fields from outputs of atmospheric models. Applied to large-eddy simulations (LES) where clouds are resolved, such simulations allow to study the detailed impact of individual clouds at the surface. We show for instance in Figure 6 a map of solar irradiance obtained under a field of cumulus (RICO case study, Van Zanten et al., 2011). Such a map can be used to further investigate the impact of clouds on the solar energy availability at the scale of a PV plant, with a focus on the high temporal and spatial resolution variations. Such simulations can also be used as references to assess the performance of radiative parametrizations and improve them. Note that the added value of LES for weather forecasts is explored in Smart4RES, and will be detailed in deliverable D2.2 and D2.4.



(a) Flux map under a RICO cloud field

(b) Flux pdf under a RICO cloud field

Figure 6: Simulation of solar irradiance under a cumulus field simulated with the model Meso-NH at 25 m resolution. The simulation is performed with the tool htrdr (Villefranque et al., 2019) to provide a surface map (a) from which the probability distribution of solar irradiance is computed (b).

2. Using ensemble simulations

2.1. Ensemble prediction

The intrinsic predictability of atmospheric dynamics, and in particular of small-scale phenomena such as those relevant for RES forecasting, is limited. In order to account for the uncertainty of weather forecasts it is now of common practice to use probabilistic forecasting, that aims at predicting the probability distribution of future atmospheric states, instead of a single deterministic weather forecast.

Probabilistic forecasting is currently implemented with Ensemble Prediction Systems (EPSs), that run in parallel multiple perturbed weather forecasts (also called "members"). Each forecast uses slightly different initial conditions, boundary conditions and model formulations in order to account for the different sources of uncertainty (Figure 7).

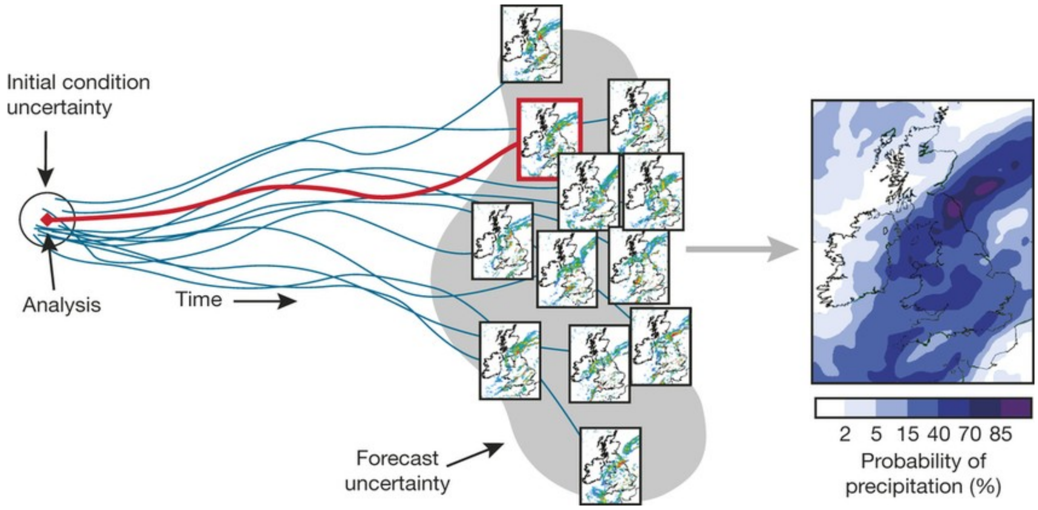


Figure 7: Principle of Ensemble Prediction. Different NWP forecasts are issued from slightly perturbed initial conditions. The resulting ensemble of forecasts samples the distribution of the future atmospheric states.

Different EPSs have been developed by several National Weather Services worldwide. At Météo-France, two EPSs are used for operational forecasting, the ARPEGE-EPS and the AROME-EPS. The ARPEGE-EPS is based on the global ARPEGE NWP model and provides forecasts up to 96 hours. Each member of ARPEGE-EPS starts from different initial conditions, designed with state-of-the-art methods including Singular Vectors (SV) and Ensemble Data Assimilation (EDA). In order to account for the uncertainty of subgrid scale processes each member uses a different package of physical parametrizations (Descamps et al., 2015). The AROME-EPS is based on the regional high-resolution AROME model, that runs over a Western Europe domain centered over France, and provides forecasts up to 51h. For each AROME member, perturbed initial conditions are derived from a specific AROME EDA, lateral boundary conditions (LBC) are provided by a selected ARPEGE-EPS member, and the model uncertainty is represented with stochastic perturbations of physics tendencies (Bouttier et al., 2016). The characteristics of the operational ARPEGE-EPS and AROME-EPS are given in Table 1.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

Table 1: Characteristics of AROME-EPS and ARPEGE-EPS for the currently operational configuration (denoted as Ref) and for the Smart4RES configuration.

	AROME-EPS		**ARPEGE-EPS**	
	Ref	***Smart4RES***	***Ref***	***Smart4RES***
Horizontal resolution	2.5 km	1.3 km	7.5 km	5 km
Output frequency	1 h	5 min	1 h	4 min
Size	16	25	35	35
Lead time	51 h	51 h	96 h	96 h
LBCs	ARPEGE-EPS ref	ARPEGE-EPS Smart	-	-

For the purpose of the Smart4RES project, enhanced configurations of ARPEGE-EPS and AROME-EPS have been developed to improve the prediction of RES variables. These enhanced configurations include a significant increase of horizontal resolutions: from 7.5 km to 5 km for ARPEGE-EPS and from 2.5 km to 1.3 km for AROME-EPS. In addition the frequency of forecast outputs is greatly increased: from hourly outputs to 4 minutes outputs (ARPEGE) and 5 minutes outputs (AROME). The size of the AROME-EPS is also increased from 16 to 25 members. Finally, the relevant output weather variables and time periods to run have been defined with Smart4RES partners. Ensemble forecasts over four periods are made available to the project: October 2018, August 2019, and February-March 2020.

The practical use of EPS outputs in RES applications is not straightforward and requires further investigation. One can consider different levels of integration, in particular:

- A direct use of ensemble members in RES prediction models: this allows the full EPS information to be propagated in the applications, however the cost of storing and handling ensemble RES simulations can be prohibitive for some users.
- To reduce the amount of information, ensemble outputs can be summarized before being used in RES applications: different solutions are possible, including for instance the selection of a smaller set of informative members, the use of ensemble statistics (mean, percentiles), the identification of the most likely member.
- Raw ensemble weather outputs can also be post-processed to derive RES-oriented Decision Support Tools (DSTs), highlighting specific risks for the management of RES production (for instance production intermittency due to wind ramp events).
- Finally, how to combine forecasts from the AROME-EPS and ARPEGE-EPS to provide seamless predictions from nowcast to 96h can be important for medium-range planning.

These questions have been examined at Météo-France, and innovative strategies have been developed to provide ensemble-based deterministic forecasts, seamless ensembles and RES-oriented DSTs. The proposed methods are described in the next sections.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

2.2. Dataset

For the need of the development and assessment of the three post-treatments proposed, the wind speed is extracted from the AROME and ARPEGE EPSs described previously. The AROME deterministic forecast used operationally at Météo-France (spatial resolution of 1.3 km, hourly outputs and maximum leadtime of 42 hours) is extracted as well. Evaluation of post-treated forecasts are performed against observed wind speed measurements from anemometers installed on the nacelles of 121 Vestas wind turbines with hub-heights between 95 m and 105 m. The locations of the turbines are displayed in Figure 8. Table 2 summarizes the dataset used in the study.

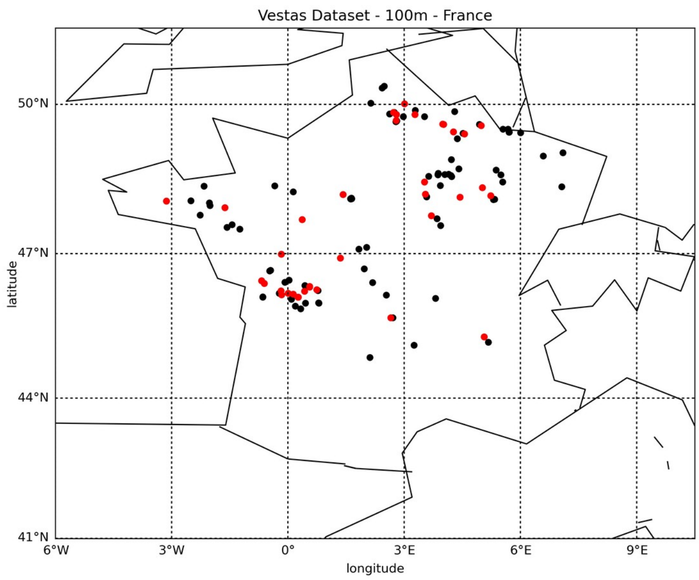


Figure 8: Locations of wind turbines at level around 100 m from which observations are available. The red points show the locations where turbine height is exactly 100 m.

Table 2: Details of the dataset used in this report. HR (High Resolution) refers to Smart4RES simulations.

	Observations	***AROME deterministic***	***AROME EPS HR***	***ARËGE-EPS HR***
Variable	~ 100 m 10-min mean wind speed	100 m wind speed	100 m wind speed	100 m wind speed
Locations	121 wind turbines (see Figure 8)		Closest gridpoints to observations	
Nature	Nacelle anemometers on wind turbines		NWP outputs	
Periods		Summer 2019 and Winter 2020		Winter 2020
Frequency	10 min	1 h	5 min	4 min

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

2.3. Pseudo-deterministic forecasts

In order to summarize the ensemble outputs, we propose different methods to extract a single deterministic trajectory which we call a pseudo-deterministic (PD) forecast. The methods are presented and evaluated in the next subsections.

In the following, we denote N the size of the ensemble associated with the index m, D the number of forecasted days associated with the index d, K the number of locations associated with the index k, T the forecast horizon associated with the index t. x and y respectively correspond to observations and forecasts at 10 min resolution.

2.3.1. Approach and methods

Three different approaches to extract a PD forecast from the AROME-EPS have been investigated, that are detailed below.

Optimal percentile

This method is based on the choice of the ensemble percentile that minimizes a given forecast score over a training period D_{tr}. Four different ways to optimize the percentiles are considered :

- A **constant optimized percentile**: The percentile minimizes the forecast score averaged spatially over France, temporally over the full forecast leadtime and over the training period.
- A **leadtime dependent optimized percentile**: The percentile minimizes the forecast score averaged spatially over France, and temporally over a given time window $W_q < T$ and over the training period.
- A **location dependent optimized percentile**: The percentile minimizes the forecast score for each location, averaged temporally over the full leadtime and over the training period.
- A **leadtime/location dependent optimized percentile**: The percentile minimizes the forecast score for each location, averaged temporally over a given time window $W_q < T$ and over the training period.

Statistics derived from the ensemble such as percentiles are generally smoother than a raw ensemble member. In particular, the high frequency variations are attenuated. For users interested in high frequency variations, a solution is to take the ensemble member closest to the optimal percentile. The closest member is defined by the minimal euclidean distance (squared deviation), averaged over a given time window $W_{closest} < T$:

$$y_{m_q} = \min_{m=1:N} \frac{1}{W_{closest}} \sum_{t=t_0}^{t_0+W_{closest}} |y_q - y_m|,$$

with y_q the optimized percentile found with methods explained previously, y_m a given member of the ensemble, and t_0 the starting timestep of the given time window.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

Weigthed mean method

Stanger et al. (2019) propose a method to weight ensemble members based on the rank histogram of past observations. The rank histogram measures whether the probability distribution of observations is well represented by the ensemble. To build a rank histogram, we first sort all members $y_{t,0}, y_{t,1}, \ldots, y_{t,N}$ together with the observation x_t. We find the rank r_t of x_t within this sorted list of $N+1$ values. As for the optimized percentile method, we build such rank histogram over the training period D_{tr}, for each location k and/or over a given time window $W_r < T$.

The height of the bins obtained from the normalized histogram correspond to weights applied to members which rank is contained within the bins boundary : the rank r_m of the forecast member y_m is computed in order to attribute to y_m the weight w_m corresponding to the bin which contains r_m. Then the forecasted variable is given by :

$$\overline{Y_{w_m}} = \sum_{m=1}^{N} w_m \times y_m$$

Preliminary classification step and optimal percentiles

Instead of computing the optimal percentile or mean from the full ensemble, another approach is to reduce the ensemble to some members that are representative of the main forecast scenario. This scenario is derived over a given time window $W_c < T$. Indeed, it is not realistic to define a single preferential trajectory constituted of the same members over the full forecast leadtime. For the purpose of identifying this scenario, different clustering techniques can be applied:

- Unsupervised standard clustering algorithms such as partition clustering (k-means and k-medoid), hierarchical clustering (ascendant HAC), and density based clustering (Dbscan).

- A classification method designed for wind power application. This alternative method is based on fixed categories derived from a wind turbine power curve: as presented in Figure 9, four main categories are identified regarding the typical wind speed U_{in}, U_{nom}, and U_{out}.

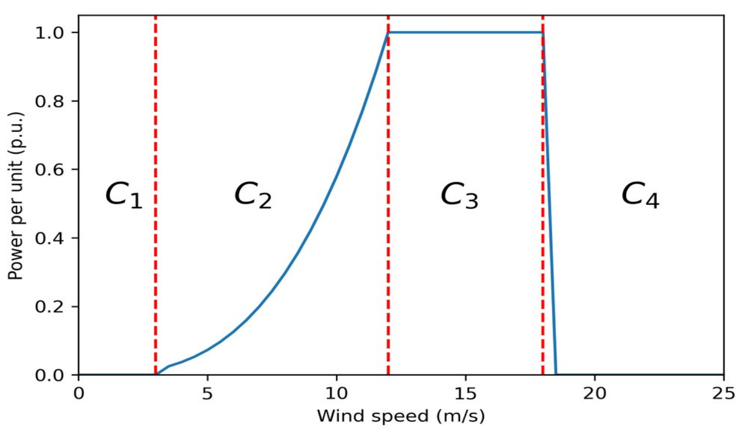


Figure 9: Fixed categories derived from a typical wind turbine power curve.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

For each clustering technique the largest cluster provides the preferential trajectory. A PD forecast is then built from this reduced set of members, using the optimized percentile method described previously.

The clustering time window W_c is set to 1 hour, 3 hours and 6 hours. Each clustering technique requires to set some parameters, which sensitivity has been tested as well. The three approaches require an optimization step, based on the training period D_{tr} which aims at finding the optimal percentiles, or at building a rank histogram and computing weights. For the percentile optimization, the forecast score minimized is the mean absolute error (MAE) which generic formula is given by :

$$MAE = \frac{1}{DKT} \sum_{d=1}^{D} \sum_{k=1}^{K} \sum_{t=0}^{T} |y_{d,k,t} - x_{d,k,t}|$$

Practically, we compute the MAE for each percentile from the 5th to 95th percentile (interpolated from the full cumulated probability distribution), with increments of 5% and keep the percentile which gives the minimum of the MAE obtained. The time window W_q is set to 1 hour. The time window $W_{closest}$, used to find the ensemble member closest to the optimal percentile, varies from 1 hour to 6 hours. Using larger window size like 6 hours improves the forecast temporal continuity, with the risk to deviate further from the optimal solution. As for the leadtime dependent percentile optimization method, the time window W_r to derive rank histograms for the weighted mean method is set to 1 hour.

The methods for deriving a PD forecast have been applied to the 100-meter wind speed forecasts provided by the high resolution AROME-EPS. The derived PD forecasts are compared to the operational AROME deterministic forecast. The evaluation of PD forecasts is performed against observed wind speed measurements described previously. The periods considered for the training and evaluation range from the 01/08/2019 to the 31/08/2019 and from the 02/02/2020 to the 16/03/2020, hereafter named Summer 2019 and Winter 2020, respectively (Table 2).

2.3.2. Results

The forecasts are assessed regarding the RMSE and the bias defined as follows :

$bias_t = (y_t - x_t)$, then averaged over D, and K

$$RMSE_t = \frac{1}{D} \sum_{d=1}^{D} \sqrt{\frac{1}{K} \sum_{k=1}^{K} (y_{d,k,t} - x_{d,k,t})^2}$$

The evaluation of the different methods shows that:

- The optimized percentile is generally around 30-40%, especially over Winter 2020. This relatively low value is related to the ensemble's slight over-estimation of the wind speed values.
- Using a single member closest to the optimized percentile highly degrades the RMSE of the forecast, compared to using the optimized percentile directly. However, it shows a good ability to represent the high frequency variability of the wind speed.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

- The clustering step, regardless of the method, generally degrades the RMSE, compared to direct estimation of the optimized percentile.
- The weighted mean has a better performance (as measured by the RMSE) than the standard (equally weighted) mean forecast.
- The best method in terms of RMSE improvement is the location-dependent optimized percentile method. This method is also better than the ensemble mean in representing the high frequency variability of the wind speed.

Figure 10 shows the bias of the ensemble mean forecast (black), the AROME deterministic forecast (red) and the PD forecast (green, location-dependent optimized percentile), for both Summer 2019 and Winter 2020 periods. Over both periods, the PD forecast allows to reduce the average bias compared to both the ensemble mean and AROME deterministic forecast. However, over Summer 2019, the bias shows a large diurnal cycle. The PD forecast only reduces the average bias, without reducing the amplitude of the diurnal cycle. Using the leadtime-dependent optimized percentile allows to better correct the bias over Summer 2019. However, it increases the RMSE over Winter 2020.

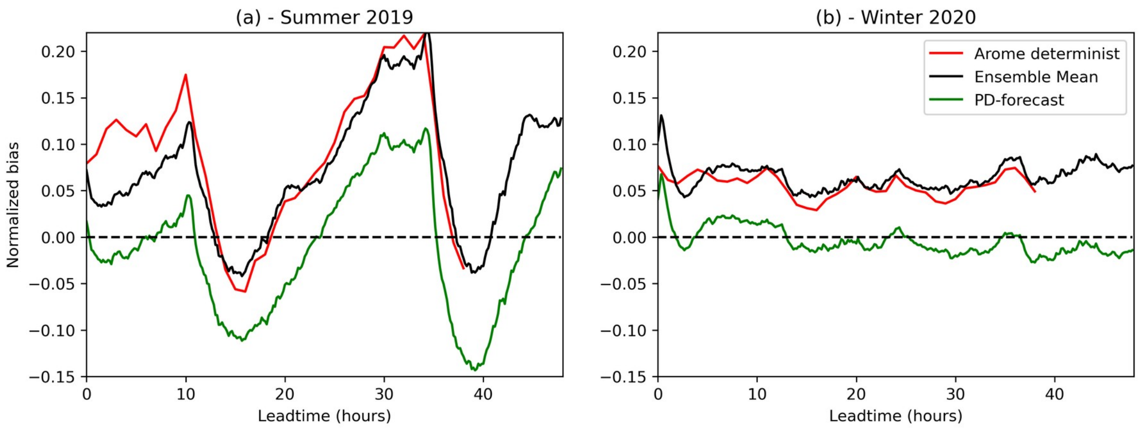


Figure 10: Normalized bias obtained from the AROME deterministic forecast (red), the ensemble mean forecast (black) and the PD forecast (green), as a function of the leadtime of the forecast : (a) for Summer 2019 period (b) for Winter 2020 period.

The improvement of a forecast for a given score S regarding a reference forecast is given by the following formula:

$$I_S = (S_{ref} - S_{for}) / S_{ref},$$

with S_{ref} the score of the reference forecast (here the ensemble mean or the AROME deterministic forecasts), and S_{for} the score of the forecast considered (here the PD forecast).

Figure 11 displays the improvement in terms of RMSE obtained with the PD forecast compared to the ensemble mean and AROME deterministic forecasts. Dashed red lines highlight the KPIs of 10% and 15% improvement aimed by the Smart4Res project (KPIs 1.1.b and 1.1.c: improvement of weather forecasting performance at next-hour and next-days horizons, cf. Smart4RES Deliverable D1.1 for more details). Compared to the AROME deterministic forecast, the PD forecast allows to reach the project's objectives and even

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

overshoot them for some leadtimes. Compared to the ensemble mean forecast, the PD forecast allows to almost reach the 10% objective over the Winter period 2020. On average, over both periods, the improvement with respect to the ensemble mean forecast amounts to 4%.

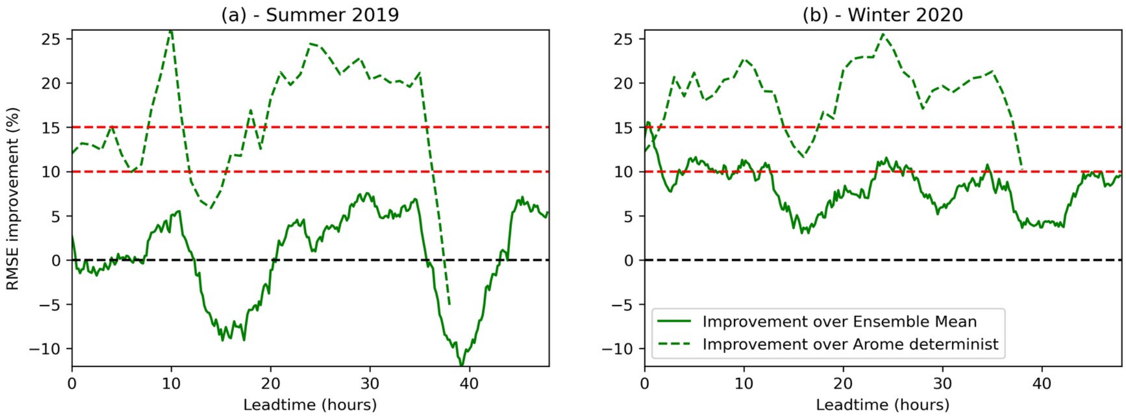


Figure 11: RMSE improvement with respect to the AROME deterministic forecast (dashed green curve) and the ensemble mean forecast (solid green curve) as a function of the leadtime of the forecast : (a) for Summer 2019 period (b) for Winter 2020 period. The dashed red lines displays the KPIs targeted by the Smart4RES project.

The same approaches have been tested for deriving a PD forecast of the irradiance. A dataset of 72 meteorological station of Météo-France have been considered. The result of this study shows that the optimal method for irradiance is the leadtime dependent weighted mean method. The improvement in terms of RMSE over Summer 2019 and Winter 2020 for day 1 and day 2 of forecast are given in Table 3. The improvement over the ensemble mean reaches 12% over Summer 2019 for day 1, and 15% over the AROME deterministic forecast. Over Winter 2020, the improvement remains lower.

Table 3: RMSE improvement obtained for irradiance PD forecast using leadtime dependent weighted mean method, for Summer 2019 and Winter 2020.

RMSE improvement (in %)	Summer 2019		Winter 2020	
	Day 1	Day 2	Day 1	Day2
Over ensemble mean forecast	12.1	11.1	4.0	3.3
Over AROME deterministic forecast	15.6	12.9	9.3	12.1

2.4. RES-oriented decision support tools

Another way to summarize the large ensemble information is to design decision support tools (DSTs) customized to end-users needs. Four DSTs are proposed and detailed in this section. They include two products giving access to the uncertainty and variability of the wind speed, which should help market trading strategies by better addressing risks of

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

extreme wind variations. Two other products address two specific events, namely the risk of wind speed exceeding the cut-out, and ramping events. Both are of high importance for wind farm operation and power system balancing.

2.4.1. Uncertainty

The uncertainty of the forecast is represented by the dispersion of the members in the ensemble. The dispersion of the ensemble generally increases with the forecast leadtime. The uncertainty of a forecast can be measured by the confidence interval range between centered percentiles, e.g. the interquartile range (IQR) gives the range between the 25th and 75th percentiles (the interval is centered around the median). It is often more useful to give a measure of a larger interval, for instance the 90% confidence interval range, because it addresses the extremes of the distribution (Figure 12). The coverage of a confidence interval is the percentage of observed values contained in the interval. If the ensemble is calibrated and of large enough size, the coverage should converge to the interval confidence, e.g. the coverage of the 90% confidence interval should tend to 90%. The IQR and the 90% confidence interval coverages over both Winter 2020 and Summer 2019, over all locations, and all leadtimes, are given in Table 4. It shows that the ensemble is overconfident as the intervals are undercovered. This is due to the underdispersive behaviour of the ensemble.

Table 4: Interquartile range and 90% confidence interval coverages

	Summer 2019	Winter 2020
IQR coverage	0.46	0.42
90% CI coverage	0.82	0.78

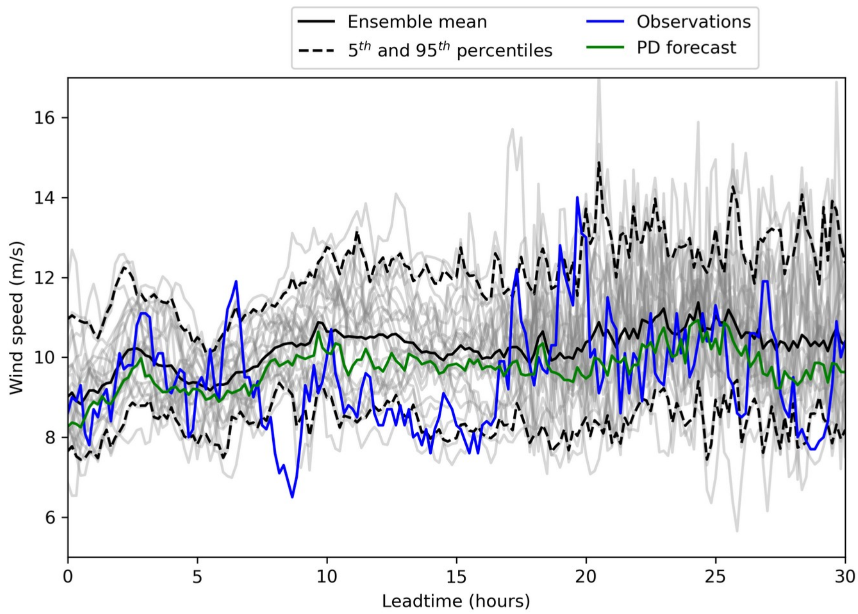


Figure 12: Example of 90% confidence interval given by the ensemble (dashed lines). The observed wind speed is shown in blue, the ensemble mean is given by the black solid line, and the PD forecast by the green solid line.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

2.4.2. Sub-hourly variability

As explained in the previous section, the PD forecast extracted from the ensemble does not properly represent the high frequency variations provided by the 5 min outputs. In addition to the PD forecast, it is proposed to provide time series of sub-hourly variance of the wind speed forecasts computed as follows.

Each member y_m of the ensemble is decomposed into a slowly varying signal $\overline{y_m}$ and a high frequency signal ϵ_m using a rolling mean filtering over a one-hour time window. The sub-hourly variability is given by the standard deviation of ϵ over the considered time window W and among the members :

$$\sigma_{sh}(h) = \frac{1}{NW} \sum_{m=1}^{N} \sum_{t=h}^{h+W} \left(\epsilon_{m,t} - \overline{\epsilon}\right)^2,$$

with $\overline{\epsilon}$ the average value of the high frequency signal over W and the ensemble. The time window W is set to 1 hour, a key timescale for trading strategy. The final DST product provided to the end-user is the hourly (and 30 minutes) timeseries of $2\sigma_{sh}$. The value of $2\sigma_{sh}$ represents the variability associated with the smooth PD forecast; it is a forecast of the variability, while the confidence interval is a measure of uncertainty. Figure 13 displays an example of the interval +/-$2\sigma_{sh}$ provided by the ensemble.

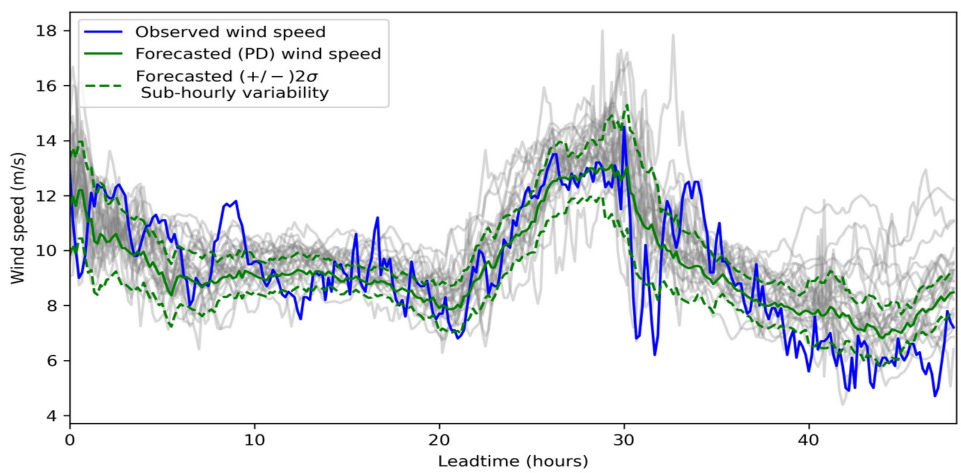


Figure 13: Example of the 2σ sub-hourly variability (W=1hour) forecasted by the ensemble (dashed green lines). The observed wind speed is shown in blue and the PD forecast in green.

2.4.3. Probability to exceed the cut-out wind speed

Several extreme events are of concern for wind energy producers. Among them, the probability to exceed the cut-out wind speed is of high importance as it results in a very fast decrease of the wind production, and/or put at risk the turbine itself. We define the following event: "The wind speed exceeds the cut-out wind speed (W_{out}) within a given hour". The idea is to use the high-resolution AROME-EPS ensemble, and in particular the

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

high frequency wind variations, to give the probability of this event to the user. We compute for each member the probability p_m of the event: the probability equals 1 if the W_{out} is reached for at least one time step during the hour. We then compute the probability p_e of the event in ensemble : $p_e = \frac{1}{N} \sum_{m=1}^{N} p_m$.

Figure 14 shows an example of such probability forecasted for the 29/02/2020. In this example, such an event is observed around leadtime 15h. No deterministic forecast (AROME, ensemble mean, and PD forecast) does predict the event, while in the ensemble several members capture this extreme event.

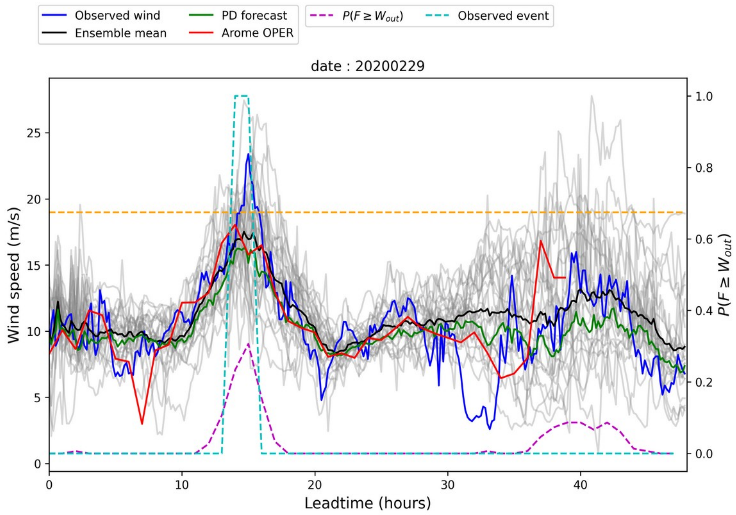


Figure 14: Example of forecasted p_e (purple). The observed event is displayed in light blue. The observed wind speed is displayed in blue, the AROME deterministic forecast in red, the ensemble mean forecast in black and the PD forecast in green.

A more comprehensive comparison of the deterministic forecasts skill regarding the ensemble shows a clear advantage of using the ensemble information.

The final product provided in this cas is the probability p_e, from which the end user can easily build a binary event forecast according to its own requirements.

2.4.4. Probability of ramping events

Ramping events are defined by a large and sudden change in production due to abrupt variations of the resource (and/or cut-out exceeding). The proposed product is limited to the wind speed variations, and thus does not address non-linearities induced by transformation from wind to power. However, it still contains valuable information for the end user. Several ways to define and detect ramping events exist in the literature. We use the so called "fixed time interval" method for ramp detection (Bianco et al. 2016). This simple method is based on the derivative of the wind speed over a given time window $(t, t+\Delta_t)$, to measure the increase or decrease of the wind speed :

$$S(t) = y(t+\Delta_t) - y(t)$$

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

The function S_t is then tested regarding a threshold S_0 to decide whether a positive/negative ramp event occurs or not. We define the two timeseries I^+ and I^- of positive and negative ramps occurrence, respectively :

$$I^+_{t:t+\Delta_t} = \begin{cases} 1 \text{ if } S(t) \geq S_0 \\ 0 \;\; elsewhere \end{cases}$$

$$I^-_{t:t+\Delta_t} = \begin{cases} 1 \text{ if } S(t) \leq -S_0 \\ 0 \;\; elsewhere \end{cases}$$

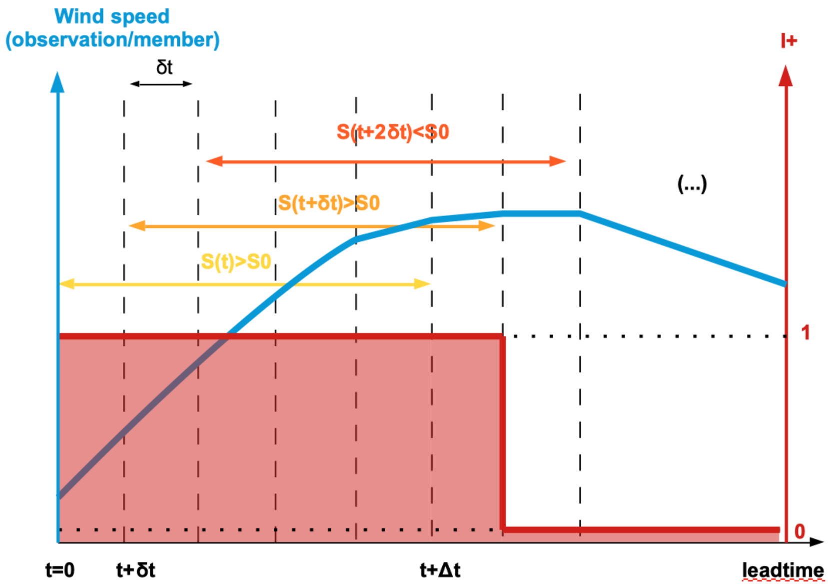


Figure 15: Schematic example of the filling of I+. The blue curve represents the wind speed (y-axis on the left), and the red line represents I+ (y-axis on the right). At the first iteration (in yellow), a ramp is found, so that I+ is filled by 1 over t:t+Δt ; at the second iteration, a ramp is still found, so that a 1 is added at timestep t+Δt+δt ; at the third iteration, no ramp is found : I+ is set to 0 for timestep t+Δt+2δt.

Note that if a I_t is filled by a 1 at an iteration step, the value 1 can not be changed afterwards. A schematic example of filling I^+ is shown in Figure 15. Each ramp is then characterized by its timing (ramp centre), its duration (window with consecutive 1 in the timeseries I^+ and I^-) and its amplitude (difference between wind values at the end and at the beginning of the ramp).

The parameters are here set to $\Delta_t = 3\,hours$ and $S_0 = 6\,m\,s^{-1}$ which correspond to a typical time window in the literature, and to about 50% of the rated wind speed (wind speed value for which rated power is obtained). Other settings or other ramp definition can be used depending on user requirements.

This ramp detection method is applied to observations and to each member in the ensemble. Thus, for positive ramps, we obtain for each member a timeseries I^+_m, as well

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

as for negative ramps I^-_m. Following the work of Bossavy et al. (2012), we compute the proportion of members forecasting a ramp at any leadtime of the forecast, to obtain a probability of ramp at each timestep of the forecast. For instance for positive ramps :

$$P_t(ramp+)=\frac{1}{N}\sum_{m=1}^{N} I^+_{t,m}$$

Several approaches to define a single ramp event from the ensemble forecast are proposed in Bossavy et al. (2012). Following one of them, a single event is considered forecasted by the ensemble when the probability of ramp $P(ramp+)$ or $P(ramp-)$ reaches local maxima. In order to define the characteristics of each single ramping event (i.e. its timing, duration, and amplitude), we average the characteristics of the ramp forecasted by all members which timeseries I_m equals 1 at time t_{max} of the local maxima. An example of positive and negative ramping event forecasts is shown in Figure 16.

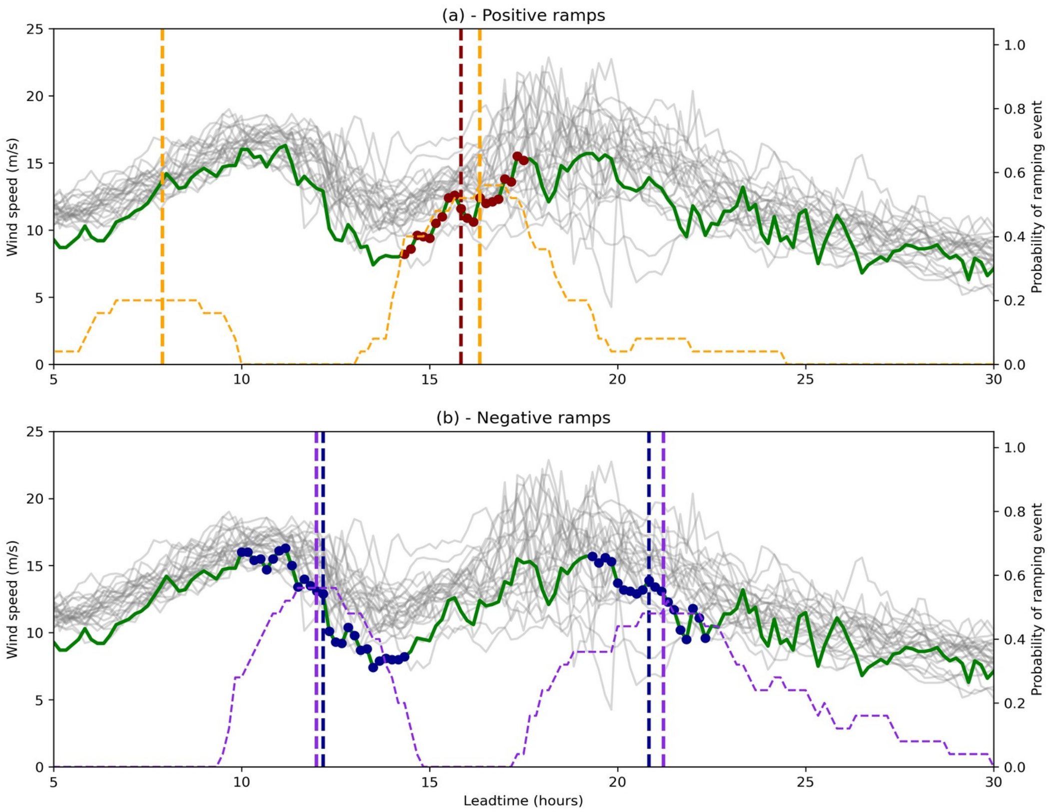


Figure 16: Observed and forecasted ramping events at one wind turbine in France, on the 12/02/2020 for (a) positive ramps (b) negative ramps. The red and blue dashed vertical lines indicate the observed ramp centers, and orange and purple dashed vertical lines, the forecasted ramp centers. The blue and red dots indicate an observed ramp. The orange and purple dashed lines represents the probability of ramp P(ramp+) and P(ramp-).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

2.5. Seamless ensemble forecasts

For stakeholders who are interested in using the full ensemble information from both AROME-EPS and ARPEGE-EPS, an approach is presented that combines both systems in order to provide seamless ensemble forecasts over the whole forecast range (4 days). The proposed method takes advantage of the enhanced performance of the high-resolution AROME-EPS for short leadtimes, while providing a smooth transition (with limited temporal discontinuities) to larger-scale ARPEGE-EPS for longer lead times.

2.5.1. Data set

This study focuses on wind turbines at 100m. Weather data are extracted from AROME-EPS and ARPEGE-EPS for February 2020. The forecasts from ARPEGE-EPS were linearly interpolated to a frequency of 5 min to match those of AROME-EPS.

The observations used come from 36 turbine measurements taken at exactly 100 m (red points in Figure 8).

2.5.2. Seamless design

The seamless design follows the method described in Aleksovska et al. (2021):

- over the forecast period 0-51h only the 25 members of the AROME-EPS are used,
- for longer ranges each member of the AROME-EPS is paired with a member of the ARPEGE-EPS according to some assignment rules,
- the selection of paired members, described below, is designed to minimize the discontinuities at the merging time (Figure 17).

This configuration is motivated by the higher performance of AROME-EPS over its forecast period (documented in Section 2.5.3).

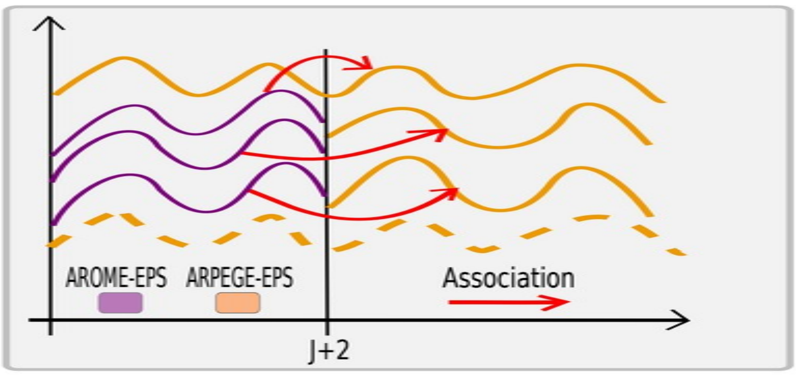


Figure 17: Schematic representation of the seamless junction between the AROME and ARPEGE ensemble forecasts.

The assignment is based on distance measurements between the AROME and ARPEGE members. For that purpose the Dynamic Time Warping (DTW, Berndt and Cliford, 1994) is used, and is computed over a period of length W before the merging time. In a second step an assignment algorithm selects the target members with the objective of minimizing the merging discontinuities, based on the computed DTW. This is an intuitive approach, which ensures a smooth transition between two predictions.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

The assignment method chosen uses the Kuhn-Munkres algorithm (Kuhn, 1955), also called the Hungarian method (HU), that finds the optimal bijective match between two samples in order to minimize the total distance of the assignment. Given a distance measure d, let d_{ij} be the distance between member i (from the AROME-EPS of size N) and member j (from the ARPEGE-EPS of size M). The Hungarian method finds, for each member i of the AROME-EPS, the corresponding member j_i of the ARPEGE-EPS as:

$$j_i = argmin_j \sum_{i=1}^{N} d_{ij}, j = 1, \ldots, M$$

2.5.3. Results

This section presents an evaluation of the seamless design for 100 m wind speed forecasts, following two criteria:

- The **probabilistic performance** of the seamless ensemble is assessed with the commonly-used CRPS score (Candille et Talagrand, 2005; Matheson et Winkler, 1976), that measures the distance between the forecast and observation distributions (Figure 18). Let $F = P\left[X \le x\right]$ be the distribution function associated with an ensemble forecast X and $F_0 = 1_{\left[X \le x_0\right]}$ be the distribution function associated with the reference x_0 (which is then a step function), the CRPS is defined by :

$$CRPS\left(F, F_0\right) = \int_R \left(F - F_0\right)^2 dx$$

  The CRPS is a measure of reliability of the forecast, it measures how close the forecast is to the observation at a given location and forecast time. In the remainder of this section the CRPS values are averaged spatially (for different wind farms) and temporally (for different forecast dates) in order to provide a more robust forecast evaluation.

  The lower the CRPS the better the ensemble. The statistical significance is calculated using the Wilcoxon score (Wilcoxon et al., 1963).

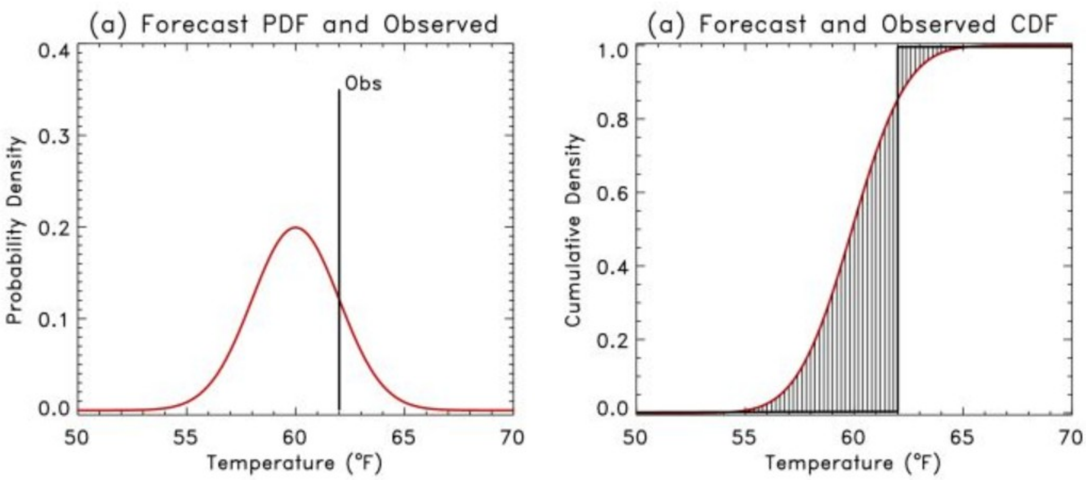


Figure 18: On the left the probability distribution of the forecast (red curve) and the reference (black vertical line), on the right the associated distribution functions. The CRPS corresponds to the area between the two distribution functions (shaded area). The horizontal axis represents the variable to be predicted. Source : (Herbasch.,2000).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

- The **temporal continuity** of the seamless ensemble is assessed by computing the average forecast difference at the merging time :

$$\Delta=\sum_{i=1}^{N}\left|\left(f_i(51\,h)-f_j(51\,h05)\right)\right|$$

where f_j is the prediction of the i^{th} member of the seamless ensemble after the merging time. The lower this difference the better the merging.

The seamless design will also be evaluated against a simple benchmark that randomly selects the ARPEGE-EPS merging members, without repetition (Wetterhall and Di Giuseppe, 2018). This method is referred to as random neighbor (RN).

Figure 19 presents the distribution of forecast differences at the merging time for the HU and RN strategies. The differences obtained with the ARPEGE-EPS members, which are by construction seamless scenarios, are taken as reference. As expected, the largest differences are obtained with the RN strategy, while the discontinuities observed with the HU assignment are much smaller and close to those observed in the ARPEGE-EPS forecasts. These results therefore indicate that the seamless design with the HU method is able to generate realistic forecasts with limited temporal inconsistencies.

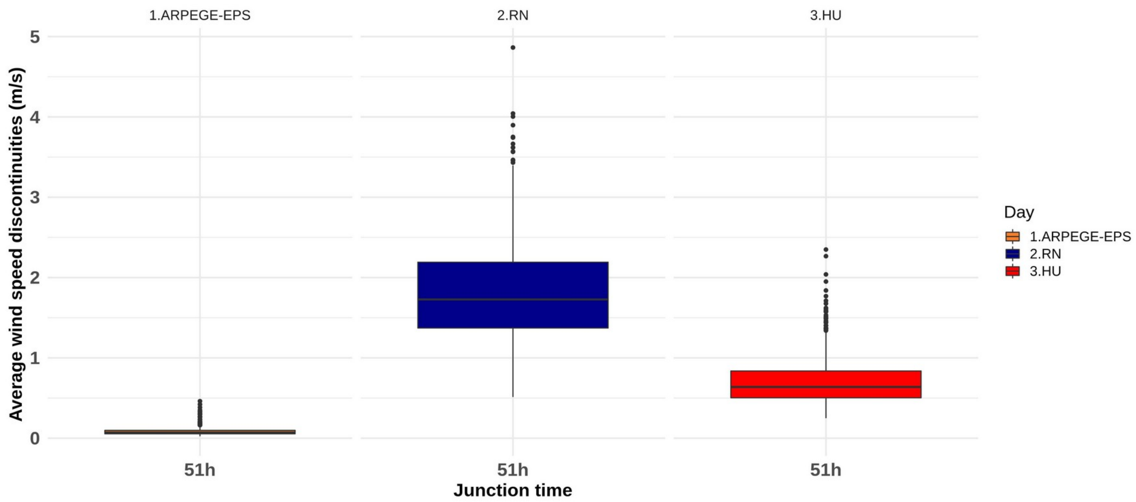


Figure 19: Distributions (in the form of box plots) of the absolute differences in predicted wind speed (averaged over all ensemble members) between the 51h and 51h05min, calculated for 36 wind turbine farms over the period 01/02/2020-29/02/2020. These distributions are presented, from left to right, for the ARPEGE-EPS members and for the members of the RN and HU (with W=3h) seamless ensembles.

The probabilistic performance of the ensembles obtained with the different connection strategies are evaluated against the performance of the raw ARPEGE-EPS. The CRPS presented in Figure 20 shows that over the first 48 hours of forecast, the use of the AROME-EPS forecasts leads to an average improvement of the CRPS. The HU strategy outperforms the RN strategy in the vicinity of the merging time (between 51h and 51h05 lead times). For longer leadtimes both strategies have similar performances.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

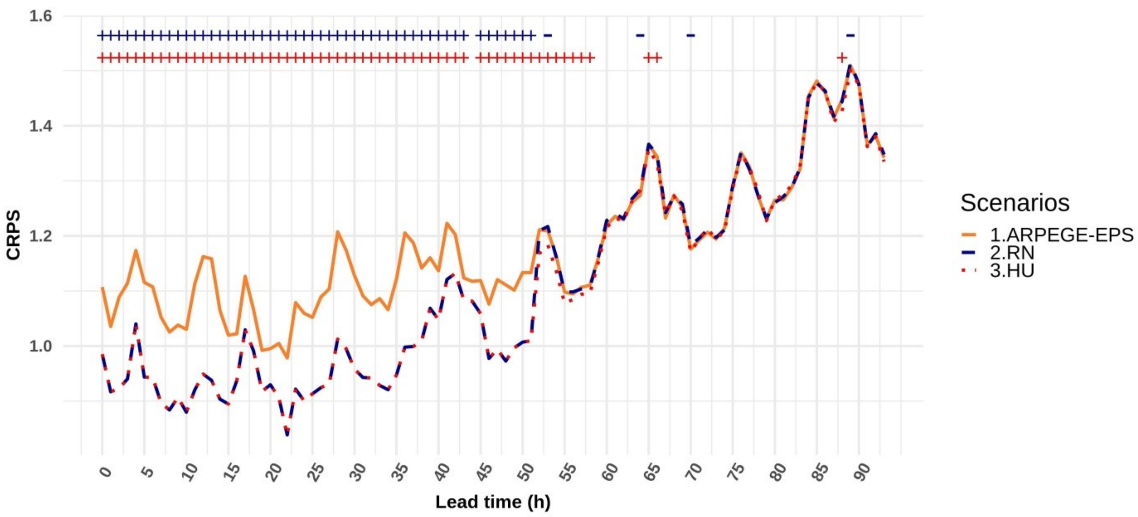
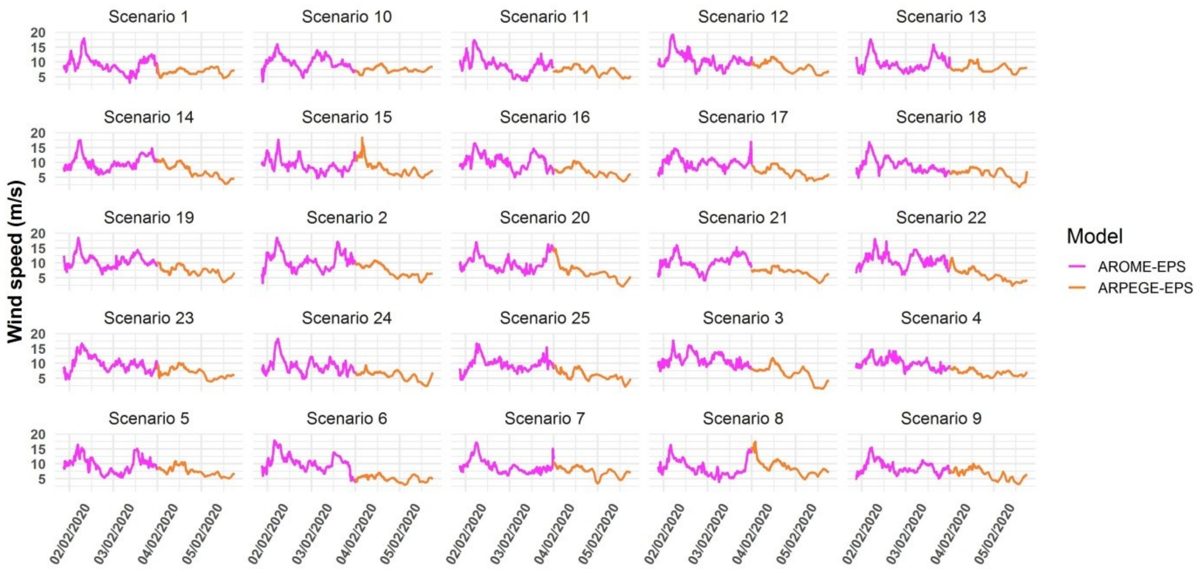


Figure 20: CRPS of 100-meter wind speed forecasts as a function of forecast leadtime . The scores are calculated for the 36 sites described in the experimental design over the period 01/02/2020-29/02/2020 for both strategies HU (with W=3h) and RN and the ARPEGE-EPS raw ensembles. The + (resp. -) signs indicate that the performance of AROME-EPS is statistically better (resp. worse) than the performance of ARPEGE-EPS, according to the Wilcoxon test.

A visualization of an ensemble of 100-m wind speed forecasts obtained with the HU and RN strategies is shown in Figure 21. These plots confirm the previous evaluation, indicating smooth transitions with the HU design compared to much larger discontinuities obtained with the RN merging.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

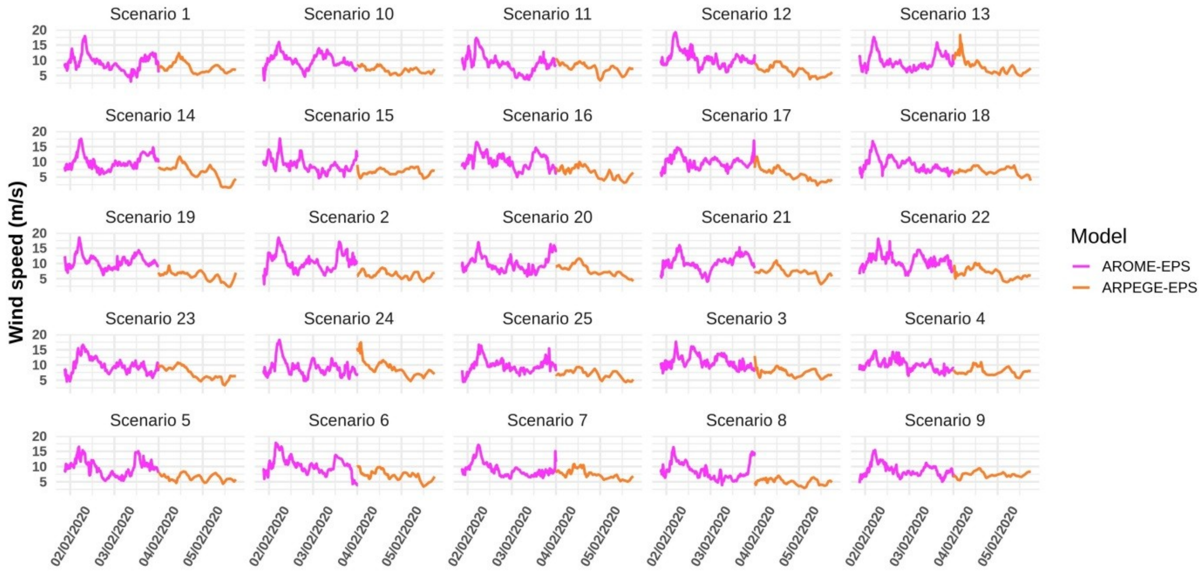


Figure 21: Ensemble of seamless forecasts constructed with (top) HU method, (bottom) RN method. Start date of the forecast is 01/02/2020, at 21:00 in Park 2040. Each seamless member is initialized by an AROME-EPS forecast (in pink) and then connected to an ARPEGE-EPS member (in orange) at 51h. The distances between members are computed over the last 3 hours of the common period (from 49h to 51h lead-time).

Finally, an analysis of the impact of W, the length of the period chosen for the assignment, is shown in Figures 22 and 23. It seems that the shorter W the smaller the merging discontinuity is. On the other hand, larger W (24h) seems to provide slightly better forecasts at longer lead times. An evaluation over a longer period and different seasons will be conducted to improve the robustness of these results. Taking both criteria into account it seems that an optimal value is reached for W=3h. However, the criteria must be evaluated over other time periods as well, in order to obtain more robust statistics to make the optimal choice for W.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

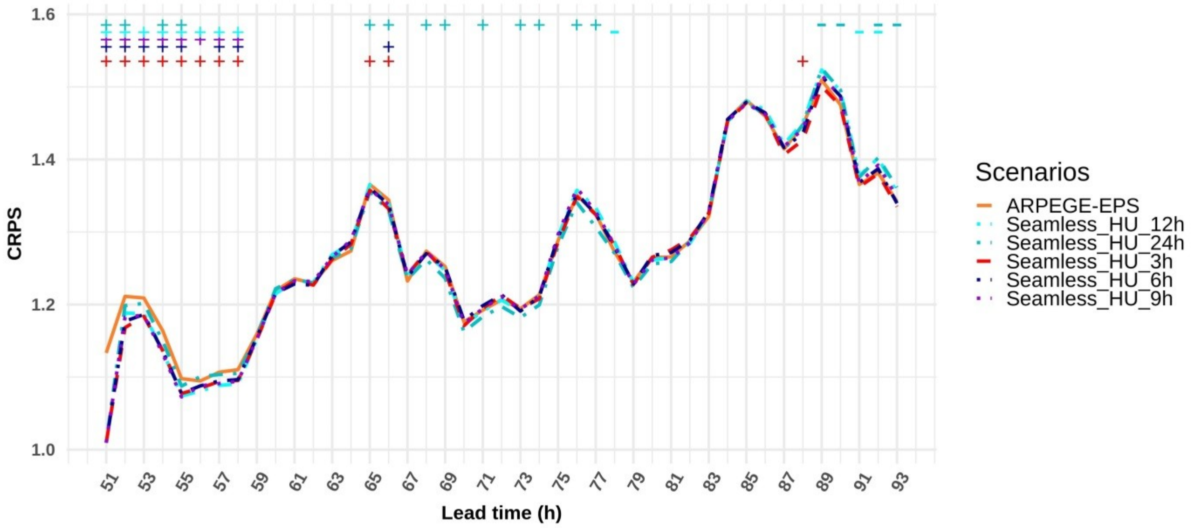


Figure 22: CRPS of 100-meter wind force forecasts as a function of forecast time frame. The scores are calculated for the 36 sites described in the experimental design over the period 01/02/2020-29/02/2020 for both strategies HU (W=3,6,9,12 and 24h) and the ARPEGE-EPS raw ensembles. The + (resp. -) signs indicate that the performance of AROME-EPS is statistically better (resp. worse) than the performance of ARPEGE-EPS, according to the Wilcoxon test.

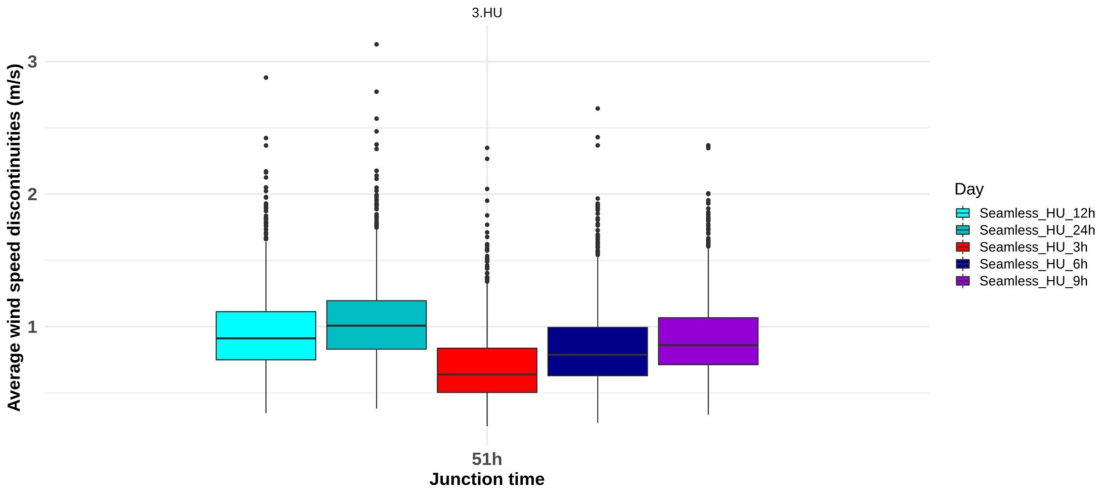


Figure 23: Distributions (in the form of box plots) of the absolute differences in predicted wind speed (averaged over all ensemble members) between the 51h and 51h05min, calculated for 36 wind turbine farms over the period 01/02/2020-29/02/2020. These distributions are presented for various time windows W.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

3. Towards RES-oriented weather services

As NWP models are initially designed to accurately forecast the variables of interest to ensure the safety of goods and people, national weather services have so far mainly focused on monitoring the quality of a reduced set of parameters. These variables are chosen either to check the overall quality of the weather processes modelled (for example low and mid troposphere air temperature and geopotential), or to check the quality of the main parameters at stake for the safety of goods and people (for example 2-meter air temperature, wind gusts, precipitation, mean sea level pressure, storm occurrence).

However, NWP models can predict many other variables and at nearly any vertical level either because these variables are used in the computations, or because they can be determined using diagnostic equations with some of the main NWP outputs (see Section 1.2). Even though the quality of these variables is not directly monitored, their validity is ensured by the physical approach of NWP modelling, which provides consistency among all atmospheric variables. This is why most NWP variables can be used as inputs for diverse applications such as renewable energy production forecasting (e.g. using 10 to 200-meter wind speed or solar irradiance).

General enhancements of the model usually are beneficial for all variables, or at least some of them without significantly deteriorating the others. However, since not every output variable is monitored, the impacts of such modifications are not always well identified. The lack of measurements for all variables with a sufficient geographical and temporal coverage makes the evaluation process even more difficult and uneven among variables. Moreover, because these variables are not sufficiently evaluated by national weather services, the improvement of their quality is not usually part of the goals targeted by further model developments. This leads to a situation of a vicious circle, “no evaluation / no improvement”, which is a major obstacle in the enhancement of weather prediction for RES generation in particular. In this section we discuss long-term efforts that could improve the quality of weather services for the RES sector.

3.1. NWP models evaluation and monitoring

A first step to specifically improve the quality of weather prediction for RES purposes could be to include more end-user oriented variables and quality indices in the NWP evaluation and near-real-time monitoring processes: global and direct or diffuse solar horizontal irradiance for solar power production and wind speed at levels between 50 and 200 meters for wind power production for example.

Power generation forecasters also seem interested in information about the quality of weather forecasts in real time to better monitor their installations or evaluate their own forecasts, and later on to work on their models improvement. Indeed, distinguishing the errors resulting from their modeling approach from those due to weather forecasts errors is key.

Monitoring more variables however raises the issue of the availability of measurements to be used as references for the computation of scores. Most national weather services operate a ground station network to get measurements of near surface parameters, but

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

the density of stations equipped with the adequate sensors differs depending on the parameter. In France for example, more than 2000 stations measure the 2-meter air temperature, but less than 200 measure global solar irradiance and less than 10 measure direct solar irradiance.

Solar irradiance measurements can relevantly be supplemented by satellite observations, however, this is not yet possible for other variables of interest such as wind speed at 50-200 meters. The density of weather stations measuring wind speed in the networks is usually satisfying for most use cases, but these measurements are only available at around 10 meters above the ground. The quality of 10-meter wind speed forecasts is not necessarily meaningful for the quality of the 100 or 200-meter wind speed forecasts, and meteorological masts or LiDARs capable of measuring the wind speed at such level are quite sparse and usually operated by researchers or end-users, rather than by national weather services themselves. However, national weather services have already decided to install dedicated sensors and weather stations at certain locations in order to meet one of their specific users' requirements, aeronautics. The measurement networks could thus also be adapted to meet other users' needs, such as the RES production industry, by installing and maintaining their own 100-meter wind masts all across the region for example.

Wind farms are generally equipped with anemometers on the turbines, and pyranometers are sometimes installed next to solar plants. Since the number of RES plants is growing quickly around the world, they might represent a crucial opportunity in RES oriented weather monitoring, if their measurements were to be shared with the weather community. More generally, collecting, filtering and adapting as much data as possible from all available sources (other scientific communities, end users, crowd sourcing, smart objects) to draw useful information to correct weather observations, as done by Mandement and Caumont (2020), or forecasts, is one of the biggest challenges and opportunites for the weather community for the next decade. Some data sharing and data market related issues are tackled in the WP 4 of Smart4RES.

An example of an operational application for RES-oriented NWP monitoring could be a platform where real time and climatological evaluation of NWP forecasts would be conducted for a certain set of variables at stake for the RES, including data from weather services as well as data from the end-users (for example anonymized or individualized). This kind of applications would also stimulate RES-oriented activities in the NWP modelling community and help to collect feedback from NWP output end-users through the shared data and experiences. Feedback from end-users is indeed of great value for NWP modelers who are not always the best informed regarding the qualities and limits of their models in operational use: the weather community has a lot to gain from a better collaboration and easier exchanges with other scientific communities and especially those using weather data in their applications.

3.2. RES-dedicated NWP models

Beyond evaluating models with respect to variables of particular interest for end users, the models themselves could be designed to outperform current operational models for such variables. However, apart from general improvements such as horizontal or vertical resolution increases, it is generally quite difficult to greatly improve the quality of some specific variables without deteriorating the quality of others, which is part of the reason why NWP improvements appear to be quite modest most of the time. This comes from

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

the tuning process that sometimes simply results in the best compensations ensuring balance between imperfect parametrizations. As a consequence a significant change in a physical parametrization could cause a rupture in this balance between parameters and therefore deteriorate the overall performance of the model, even though the change was initially proven to increase the performance of the targeted process.

Innovative methods to optimise parameterizations in an NWP model based on machine learning, such as the tools developed in the framework of the High Tune project (Couvreux et al., 2021), could support the joint improvement of the radiation, turbulence and microphysical schemes and their interactions, which are essential to accurately forecast solar irradiance and 100-meter wind speed.

Data assimilation also plays a crucial role in the quality of the forecasts, especially at short-term. There is room for improvement in the assimilation impacting the main variables of interest for the RES purposes, in particular for solar power. Indeed NWP models do not currently assimilate ground measurements of solar irradiance or cloud cover. Assimilation of data in cloudy conditions is still an active field of research and a major challenge for the NWP community. This challenge will be addressed in more details in the deliverable D2.4. Data from end-users of the energy sector could then not only be beneficial for the real-time monitoring of NWP models, but also to refine their operational initial states, with expected improvement in the forecasts.

Another approach to improve the quality of NWP models for the RES is to design and operate distinct dedicated models depending on end-users. This is why Jimenez et al. (2016) recently developed the WRF-Solar model and its ensemble version. This approach could be pushed further by optimizing and tuning the parameterizations of a specific NWP model version, only considering the variables of interest for a certain profile of users. For example, Météo-France could develop a "SolAROME" model, that is the AROME model tuned to forecast solar irradiance and other variables important for the solar power community as accurately as possible, even if this may imply that other variables are slightly deteriorated. This idea as presented here would however be quite costly in human and computing resources since it would require to develop, tune, maintain and run several versions of a NWP model at the same time for every type of user. This can be considered the future of NWP, which the ever growing computing capability could enable.

3.3. Weather services for the RES sector

Besides NWP, the weather forecast community can provide more innovative dedicated services to the RES end-users. A first way of improvement is the development of quicker and easier access to the weather forecasts, using APIs and data platforms capable of handling the very large amounts of data generated by NWP models and their ensemble versions in particular. These tools should facilitate the integration of weather data in the end-users' applications, help to reduce the transfer time of the data in real time and optimize their longer term storage and retrieval.

To partially avoid the problems caused by the large volumes of data and to help end-users exploit ensemble data more efficiently, post-processing techniques can be applied to summarise the information of many ensemble members into a reduced number of scenarios. This was investigated in Section 2.3 for instance.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

Post-processing methods are currently often applied to improve the overall quality of the forecasts, using machine learning techniques, and would still be relevant even if the quality of the NWP outputs was refined by some of the means described above. These techniques include ensemble calibration to correct biases and improve dispersion as presented in Taillardat and Mestre (2020), or seamless ensemble junction as described in Section 2.5 for example. More general post-processing methods, non-specific to ensemble forecasts, can also still be relevant to take into account local conditions around the end-users' site more precisely by using their data in the training, or to correct very short term forecasts by merging measured data with NWP forecasts for instance, which will be detailed in deliverable D2.3.

Finally, the quality of a weather forecast substantially relies on the "human-in-the-loop" or "human-over-the-loop" approach. Weather experts are able to:

- analyze the current atmospheric situation and possible scenarios predicted for the following hours and days,
- determine which scenario is the most likely to happen, or describe alternatives to an initially deterministic forecast,
- correct known errors in the models occurring in certain specific weather types and locations.

This human expertise could be extended to specific end-user's needs and stakes (as it is currently done for aeronautics) via:

- translating the forecasts and uncertainties into a more understandable message if the customer is not a weather expert (useful for grid management or solar power plant operating),
- sorting out and summarizing the information relevant for this user's specific field of application among different models' forecasts and available outputs.

Although this would require a specific training for these specialized forecasters it could bring an additional gain in quality of the forecasts and in their use along with the deeper NWP model improvements and post-processing methods described before.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

Conclusion and outlook

In this report we have presented the relevant information that NWP models currently provide for RES forecasts. Through a detailed description of how NWP models work, we have highlighted how the latter could be improved to better match the needs of the RES sector. This can be achieved in various ways. First, NWP models can be improved by refining the physical parametrizations that drive the key atmospheric variables, or by providing more physically-relevant diagnostics that could be useful to a variety of end-users. It is also possible to design new tools to extract the most critical information from the wealth of information contained in new ensemble simulations that tend to become the standard of NWP. These tools include building probabilistic products of high interest for the final users, or optimally condensing the information in a readily accessible format. Finally we have shown that NWP models dedicated to RES forecasts could be developed, and that forecasters could be trained to develop a specific expertise on that topic. This highlights the high potential of NWP models to tackle RES-related issues, with stimulating scientific challenges and beneficial exchanges between users who know the caveats of the models and model developers who can more efficiently refine the models when they know their defaults.

All these potential improvements rely on enhanced interactions between the atmospheric modeling community and the end-users. As a matter of fact, end-users should not be limited to using the commonly available information provided by NWP models, they should actively participate to the design of NWP models and post-processing tools. The aeronautics sector, one of the main customer of weather services, has for instance become fully involved in the development of NWP models, so that cutting-edge research on the physical parametrizations is now directly driven by this community, and a variety of diagnostics (near-surface visibility, icing in altitude etc.) have been developed for their needs. We believe that such enhanced communication between communities, along with shared experience, is key to an efficient improvement of NWP products for the RES sector, and unique to the Smart4RES consortium which gathers experts from the whole value chain of RES production.

The proposed strategies have already resulted in quantitative improvements of the performance of NWP models compared to currently available operational weather forecasts, with reductions of the RMSE in the range 10-20% for wind and solar irradiance forecasts, depending on the periods and locations. There are various paths to further improve these performances. First, physical parametrizations could be further refined in the light of RES variables. In particular, cloud physics could be revisited with the primary goal of correctly simulating the cloud transmittance while generally precipitation is the variable used to evaluate cloud parametrizations. Second, the evaluation of NWP models in terms of RES should become systematic, so that the model developers could get more insight into the defaults of their models that affect RES prediction. Finally, a big challenge of the upcoming years is to fully take advantage of the numerous power production measurements available throughout territories. Such measurements, which outnumber standard meteorological ground measurements, can be considered as proxys of weather variables, and as such could be assimilated in NWP models. These perspectives stress that much can be done to further improve RES forecasts.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

References:

Aleksovska, I., Raynaud, L., Faivre, R., Brun, F., & Raynal, M. (2021). Design and evaluation of calibrated and seamless ensemble weather forecasts for crop protection applications. *Weather and Forecasting*.

Bianco, L., Djalalova, I. V., Wilczak, J. M., Cline, J., Calvert, S., Konopleva-Akish, E., ... & Freedman, J. (2016). A wind energy ramp tool and metric for measuring the skill of numerical weather prediction models. *Weather and Forecasting*, *31*(4), 1137-1156.

Berndt, D. J., & Clifford, J. (1994, July). Using dynamic time warping to find patterns in time series. In *KDD workshop* (Vol. 10, No. 16, pp. 359-370).

Bossavy, A., Girard, R., & Kariniotakis, G. (2013). Forecasting ramps of wind power production with numerical weather prediction ensembles. *Wind Energy*, *16*(1), 51-63.

Bouttier, F., Raynaud, L., Nuissier, O., & Ménétrier, B. (2016). Sensitivity of the AROME ensemble to initial and surface perturbations during HyMeX. *Quarterly Journal of the Royal Meteorological Society*, *142*, 390-403.

Bozzo, A., Benedetti, A., Flemming, J., Kipling, Z., & Rémy, S. (2020). An aerosol climatology for global models based on the tropospheric aerosol scheme in the Integrated Forecasting System of ECMWF. *Geoscientific Model Development*, *13*(3), 1007-1034.

Candille, G., & Talagrand, O. (2005). Evaluation of probabilistic prediction systems for a scalar variable. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *131*(609), 2131-2150.

Cesana, G. V., Ackerman, A. S., Fridlind, A. M., Silber, I., & Kelley, M. (2021). Snow reconciles observed and simulated phase partitioning and increases cloud feedback. *Geophysical Research Letters*, *48*(20), e2021GL094876.

Córdoba-Jabonero, C., Sicard, M., López-Cayuela, M. Á., Ansmann, A., Comerón, A., Zorzano, M. P., ... & Muñoz-Porcar, C. (2021). Aerosol radiative impact during the summer 2019 heatwave produced partly by an inter-continental Saharan dust outbreak–Part 1: Short-wave dust direct radiative effect. *Atmospheric Chemistry and Physics*, *21*(8), 6455-6479.

Couvreux, F., Hourdin, F., Williamson, D., Roehrig, R., Volodina, V., Villefranque, N., ... & Xu, W. (2021). Process-based climate model development harnessing machine learning: I. A calibration tool for parameterization improvement. *Journal of Advances in Modeling Earth Systems*, *13*(3), e2020MS002217.

Descamps, L., Labadie, C., Joly, A., Bazile, E., Arbogast, P., & Cébron, P. (2015). PEARP, the Météo-France short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, *141*(690), 1671-1685.

Fernández, E. F., Almonacid, F., Ruiz-Arias, J. A., & Soria-Moya, A. (2014). Analysis of the spectral variations on the performance of high concentrator photovoltaic modules operating under different real climate conditions. *Solar Energy Materials and Solar Cells*, *127*, 179-187.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

Gristey, J. J., Feingold, G., Glenn, I. B., Schmidt, K. S., & Chen, H. (2020). Surface solar irradiance in continental shallow cumulus fields: Observations and large-eddy simulation. *Journal of the Atmospheric Sciences*, *77*(3), 1065-1080.

Gutiérrez, C., Somot, S., Nabat, P., Mallet, M., Gaertner, M. Á., & Perpiñán, O. (2018). Impact of aerosols on the spatiotemporal variability of photovoltaic energy production in the Euro-Mediterranean area. *Solar Energy*, *174*, 1142-1152.

Hill, P. G., Chiu, J. C., Allan, R. P., & Chern, J. D. (2018). Characterizing the radiative effect of rain using a global ensemble of cloud resolving simulations. *Journal of Advances in Modeling Earth Systems*, *10*(10), 2453-2470.

Hogan, R. J., & Bozzo, A. (2018). A flexible and efficient radiation scheme for the ECMWF model. *Journal of Advances in Modeling Earth Systems*, *10*(8), 1990-2008.

Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J. C., Balaji, V., Duan, Q., ... & Williamson, D. (2017). The art and science of climate model tuning. *Bulletin of the American Meteorological Society*, *98*(3), 589-602.

Jahangir, E., Libois, Q., Couvreux, F., Vié, B., & Saint-Martin, D. (2021). Uncertainty of SW cloud radiative effect in atmospheric models due to the parameterization of liquid cloud optical properties. *Journal of Advances in Modeling Earth Systems*, e2021MS002742.

Jimenez, P. A., Hacker, J. P., Dudhia, J., Haupt, S. E., Ruiz-Arias, J. A., Gueymard, C. A., ... & Deng, A. (2016). WRF-Solar: Description and clear-sky assessment of an augmented NWP model for solar power prediction. *Bulletin of the American Meteorological Society*, *97*(7), 1249-1264.

Joseph, J. H., Wiscombe, W. J., & Weinman, J. A. (1976). The delta-Eddington approximation for radiative flux transfer. *Journal of Atmospheric Sciences*, *33*(12), 2452-2459.

Josse, B., Simon, P., & Peuch, V. H. (2004). Radon global simulations with the multiscale chemistry and transport model MOCAGE. *Tellus B: Chemical and Physical Meteorology*, *56*(4), 339-356.

Jouan, C., Milbrandt, J. A., Vaillancourt, P. A., Chosson, F., & Morrison, H. (2020). Adaptation of the Predicted Particles Properties (P3) microphysics scheme for large-scale numerical weather prediction. *Weather and Forecasting*, *35*(6), 2541-2565.

Köhler, C., Steiner, A., Saint-Drenan, Y. M., Ernst, D., Bergmann-Dick, A., Zirkelbach, M., ... & Ritter, B. (2017). Critical weather situations for renewable energies–Part B: Low stratus risk for solar power. *Renewable Energy*, *101*, 794-803.

Kuhn, H. W. (1955). The Hungarian method for the assignment problem. *Naval research logistics quarterly*, *2*(1-2), 83-97.

Kurucz, R. L. (1995). The solar spectrum: atlases and line identifications. In *Laboratory and Astronomical High Resolution Spectra* (Vol. 81, p. 17).

Lac, C., Chaboureau, J. P., Masson, V., Pinty, J. P., Tulet, P., Escobar, J., ... & Wautelet, P. (2018). Overview of the Meso-NH model version 5.4 and its applications. *Geoscientific Model Development*, *11*(5), 1929-1969.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

Lindsay, N., Libois, Q., Badosa, J., Migan-Dubois, A., & Bourdin, V. (2020). Errors in PV power modelling due to the lack of spectral and angular details of solar irradiance inputs. *Solar Energy*, *197*, 266-278.

Mandement, M., & Caumont, O. (2020). Contribution of personal weather stations to the observation of deep-convection features near the ground. *Natural Hazards and Earth System Sciences*, *20*(1), 299-322.

Martin, G. M., Johnson, D. W., & Spice, A. (1994). The measurement and parameterization of effective radius of droplets in warm stratocumulus clouds. *Journal of Atmospheric Sciences*, *51*(13), 1823-1842.

Matheson, J. E., & Winkler, R. L. (1976). Scoring rules for continuous probability distributions. *Management science*, *22*(10), 1087-1096.

Norton, M., Amillo, A. G., & Galleano, R. (2015). Comparison of solar spectral irradiance measurements using the average photon energy parameter. *Solar Energy*, *120*, 337-344.

Räisänen, P., Barker, H. W., Khairoutdinov, M. F., Li, J., & Randall, D. A. (2004). Stochastic generation of subgrid-scale cloudy columns for large-scale models. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *130*(601), 2047-2067.

Räisänen, P., & Lindfors, A. V. (2019). On the computation of apparent direct solar radiation. *Journal of the Atmospheric Sciences*, *76*(9), 2761-2780.

Rieger, D., Steiner, A., Bachmann, V., Gasch, P., Förstner, J., Deetz, K., ... & Vogel, H. (2017). Impact of the 4 April 2014 Saharan dust outbreak on the photovoltaic power generation in Germany. *Atmospheric Chemistry and Physics*, *17*(21), 13391-13415.

Slingo, A. (1989). A GCM parameterization for the shortwave radiative properties of water clouds. *Journal of Atmospheric Sciences*, *46*(10), 1419-1427.

Stanger, J., Finney, I., Weisheimer, A., & Palmer, T. (2019). Optimising the use of ensemble information in numerical weather forecasts of wind power generation. *Environmental Research Letters*, *14*(12), 124086.

Sun, Z., Li, J., He, Y., Li, J., Liu, A., & Zhang, F. (2016). Determination of direct normal irradiance including circumsolar radiation in climate/NWP models. *Quarterly Journal of the Royal Meteorological Society*, *142*(700), 2591-2598.

Taillardat, M., & Mestre, O. (2020). From research to applications–examples of operational ensemble post-processing in France using machine learning. *Nonlinear Processes in Geophysics*, *27*(2), 329-347.

VanZanten, M. C., Stevens, B., Nuijens, L., Siebesma, A. P., Ackerman, A. S., Burnet, F., ... & Wyszogrodzki, A. (2011). Controls on precipitation and cloudiness in simulations of trade-wind cumulus as observed during RICO. *Journal of Advances in Modeling Earth Systems*, *3*(2).

Villefranque, N., Fournier, R., Couvreux, F., Blanco, S., Cornet, C., Eymet, V., ... & Tregan, J. M. (2019). A Path-Tracing Monte Carlo Library for 3-D Radiative Transfer in Highly

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

Resolved Cloudy Atmospheres. *Journal of Advances in Modeling Earth Systems*, *11*(8), 2449-2473.

Villefranque, N., & Hogan, R. J. (2021). Evidence for the 3D Radiative Effects of Boundary-Layer Clouds From Observations of Direct and Diffuse Surface Solar Fluxes. *Geophysical Research Letters*, *48*(14), e2021GL093369.

Wetterhall, F., & Giuseppe, F. D. (2018). The benefit of seamless forecasts for hydrological predictions over Europe. *Hydrology and Earth System Sciences*, *22*(6), 3409-3420.

Wilcoxon, F., Katti, S. K., & Wilcox, R. A. (1963). *Critical values and probability levels for the Wilcoxon rank sum test and the Wilcoxon signed rank test* (pp. 171-176). Pearl River (NY): American Cyanamid.

Xie, Y., Yang, J., Sengupta, M., Liu, Y., & Zhou, X. (2022). Improving the prediction of DNI with physics-based representation of all-sky circumsolar radiation. *Solar Energy*, *231*, 758-766.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337