Probabilistic forecasting tools for high-wind penetration areas: an Irish case study

Corinna Möhrlen and Ulrik Vestergaard
WEPROG, Assens, Denmark
Email: com@weprog.com, ulv@weprog.com

James Ryan and Kenneth Conway
EIRGRID Plc., Dublin, Ireland
Email: james.ryan@eirgrid.com, kenneth.conway@eirgrid.com

***Abstract*—An increasing penetration of Renewable Energy Sources (RES) in combination with climate change will make operating the electric grid in a secure and economically efficient way more challenging due to increasing amounts of extreme and highly variable weather conditions. With grid operations becoming more automated, reliable weather information is essential. New tools are required to handle the growing need for weather information, not only for the day-ahead planning, but also on the short-term horizon of hours and minutes, where local measurements can be assimilated to the forecasts in order to provide a more reliable picture of the current state of the electric system. In this paper, we will describe a new probabilistic tool for managing ramping reserves for the current and future high penetration of RES from a case study in Ireland.**

We will introduce a probabilistic ramping reserve forecasting tool required to handle extreme weather situations that have already started to occur regularly in Ireland. The tool is designed with planning analysts and control room staff in mind who need to take decisions that both optimise grid security and efficiency.

I. Introduction

Today's climate change debate reminds us to take action, while there is still control over our most important systems. One of them, the electricity system is one of the backbones of our society. Protecting the environment by increasing the percentage of energy coming from renewable energy sources (RES), while concurrently growing electricity demand are ambitious and challenging objectives. Under benign weather conditions, system operation with large amounts of RES on the grid has proven to be quite manageable with todays technology. However, when planning the grid and its future secure operation, the effects of extreme weather events need to be examined and ultimately mitigated.

Ireland is the first country in Europe to experience Atlantic storms propagating from west to east. Information on the track and intensity of low pressure systems in the Atlantic is sparse. Therefore, Irish weather forecast uncertainty is higher than other European countries and the growth rate of the uncertainty is often even higher during storm events. For this reason, growing wind power capacity on the island of Ireland is a perfect show case. EirGrid and SONI, the Irish Transmission System Operators, are implementing solutions to improve their use of forecasts during balancing, market scheduling and real-time decision making in the control room.

During storm events, it takes less than one hour from when the wind speeds pick up at the west coast of Ireland until a major fraction of the wind farms experience high winds which can result in high speed shutdown (HSSD). This ramping poses a significant risk of electricity supply demand imbalances during storm events. These risks are mitigated by limiting the wind generation in advance of the event and ensuring sufficient reserve is available to cover any reduction in output due to HSSD.

II. Background for Ramp and Reserve Forecasting in Ireland

The installed wind power and solar capacity in Ireland exceeded 5 GW in early 2019 and a further 1.5 GW of Variable Generation (VG) capacity is expected by 2020. At these capacity growth levels it is very likely that gross forecast errors will increase as forecast accuracy is unlikely to improve fast enough.

Heretofore, EirGrid and SONI have scheduled the system to meet the median expected forecast of variable generation. Forecast errors have been counteracted using the resources that have been available by default. As the installed capacity of variable generation continues to grow, the magnitude of forecast errors could begin to exceed the capability of resources that happen to be available. Therefore, as part of the "Delivering a Secure, Sustainable Electricity System" (DS3) program [1], EirGrid and SONI have introduced three Ramping Margin (RM) reserve products: RM1, RM3 and RM8. These three new reserve products will explicitly ensure sufficient resources are available to counteract probable forecast error events that evolve over time horizons of one, three and eight hours.

The RM products are designed to maintain security of supply in a system with very high levels of variable generation. However, only lower than expected production by variable generation, forecasts that transpire to be higher than realised generation, can increase the risk of supply/demand imbalances. Therefore, the RM products are one-sided, scheduling capability from resources that are able to increase their production. Taking a simplified view, the RM products determine which thermal resources can remain "cold" with startup times in excess of 8 hours and which resources need to be kept "hot" (immediate availability) or "warm" (available on short notice) in case a forecast error event occurs. Capacity qualifying for RM1 also qualifies for RM3 and RM8 unless they do not meet the duration requirements of the other two products. For example a storage unit may only qualify for RM1.

Figure 1 provides an overview of the three RM products, their time scales and indicative cost profiles. The red upper horizontal bar illustrates the RM1 product with its one hour ramp and two hour sustained horizon, the second, yellow bar, illustrates time scales for the 3-hour ramping margin product and the third and green horizontal bar illustrates the time scales for the 8-hour ramping margin product. Note, that the shaded areas at start-up of the products illustrate the partial non-availability of the resources and the mixed colours in the first 3-hours of the lower bar indicate the overlapping capabilities and corresponding cost profiles.

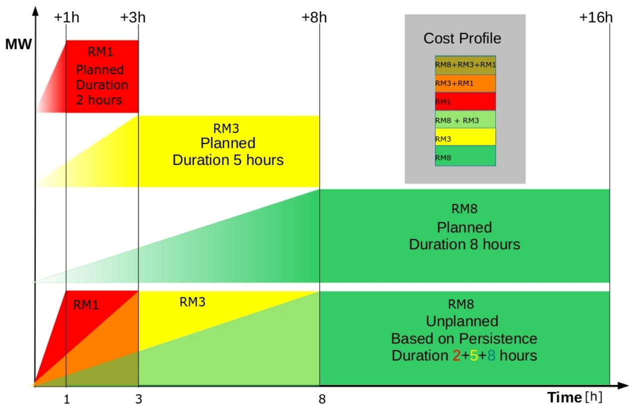

Fig. 1: Overview over the ramp products and time scales of the ramping reserve products RM1 (red) for the 1-hour ramping, RM3 (yellow) for the 3-hour ramping and RM8 (green) for the 8-hour ramping and duration. The first 3 horizontal bars show planned reserve, the lowest bar shows handling possibilities for unplanned ramping needs.

Although several other sources of generation supply and demand uncertainty contribute to the dynamic (changing between each scheduling interval) requirement for RM reserves, variable generation supply uncertainty is the source of the largest possible error. Weather-based uncertainty is the root cause for variable generation supply uncertainty. In order to translate the weather-based uncertainty into a reserve requirement, EirGrid and SONI have procured a new Variable Generation Ramping (VGR) forecast. The purpose of this study was to analyse VGR forecasts that can be utilised as a contribution to EirGrid and SONI's scheduling of RM products.

III. Definition of Ramp Forecasting

The objective of ramp forecasting is solely to forecast the change in power generation between two time stamps relative to the schedule, which is defined by a wind power forecast.

Using a deterministic forecast process for the VG schedule implies that the corresponding deterministic ramp forecast is constantly zero. A deterministic forecast would need to be complemented with past statistical results to produce some kind of a ramp result. Applying an ensemble forecasting process generates more degrees of freedom, or in other words, is capable of generating the uncertainty associated with the ramping events. The schedule from an ensemble is typically a soft curve derived from many ensemble members. This could be either a percentile P50, median or the average of the ensemble members. In this case each ensemble member provides a ramp result relative to the schedule. In a low uncertainty event the ramp values are close to zero and in a high uncertainty event the ramp values increase in magnitude in both up and down directions.

The longer the horizon of the ramp, the more accurate the ramp value is by each ensemble member, because the longer horizon filters out some phase error. This means that the uncertainty of an 8 hour ramp allocation is normally less than the 3 hour horizon, counted per hour of allocation.

IV. Ramp forecast objectives in Ireland

The purpose of ramp forecasts is to be able to schedule ramping products to fit the uncertainty of the weather and hence power generation.

Optimal scheduling during longer periods of varying forecast error implies the lowest possible spill of reserve, i.e. holding enough or too much reserve without risking not having enough. If an unnecessary amount of reserve is being carried then this is "spilled" (not utilised for power generation). If insufficient reserve is carried there is a risk of generation shortages. The intended usage of the ramping products is to cover the low side of the uncertainty band, where there is an increased risk of supply shortages in case of large forecast errors.

Alternatively, the ramping products could also be scheduled to cover all uncertainty. The benefit would be a simpler dispatch for the operatorand less dependency of the volume in the continuous intraday market. The downside is a significant higher cost than what the continuous intraday market can offer based on Short-Term (ST) forecasts on the 1–2h hour horizon. Allocation of reserve on 3–8 hours horizons based on less accurate forecasts will generate a spill of reserve with associated cost increases due to the forecast errors. Offers in the continuous intraday market, which match the ST forecast at 1–2 hours will reduce the imbalance, whereas reserve can eliminate imbalance. The ramping products seem to be required during hours, where the wind power forecast schedule is above the actual generation. For hours of excess wind power the ramping products could be in disfavour of capacity, which is more favourable in down regulation.

V. Forecast Methodology

For this study, we used forecasts from a multi-scheme ensemble approach [2]. The WEPROG multi-scheme ensemble prediction system (MSEPS) consists of 75 numerical weather prediction (NWP) models or "ensemble members" [3]. Every ensemble member is calibrated by VG unit for both solar and wind. Reserve forecasting has from the start of the design of the MSEPS system been one of the intended main applications and was introduced to the community in 2011 [4], 2014 [5] and latest in 2019 [6], but the commercial demand for reserve forecasting has been modest. In Europe focus has been largely on increased transmission capacity in expectation that large fractions of the produced energy will be exported and partially balanced remotely. This practise was adopted by Denmark/Germany, but large scale capacity increases makes this practise undesirable in the long run and forecasting reserve requirements is a suitable tool to mitigate the associated challenges, such as grid congestion and lack of transition capacity. In the study, real-time forecasts were used that are generated 4 times per day in a fixed 6-hourly schedule. Each ensemble member generates a forecast for each VG unit. The forecasts are for full availability, only limited by the static Maximum Export Capacity (MEC) value.

When wider probability bands are desired, e.g. in periods of high variability of the weather or the development of storms with uncertainty in the observational network for the NWP forecasts, time lagging is a straight forward method to achieve this. In this case up to 4 times the amount of ensemble members can be generated. When this is done, it should however be noted that the schedule is no longer centred in the probability interval and the number of members ramping down faster than the schedule may change from hour to hour in that case. In other words, by increasing the amount of ensemble members to e.g. 300, the probability space tends towards assymetry, because the 75 members from the latest forecast cycle are in a minority.
VG Ramp Rates (VGR) are defined as the difference forward in time of the VG Primary Power (VGP):

$$VGR = VGP_{end} - VGP_{begin} \qquad (1)$$

where VGP_{end} is the end interval and VGP_{begin} is the starting interval.

When we subtract two values from the same ensemble member we eliminate a major fraction of the power curve's bias relative to other ensemble members. The uncertainty becomes in that way more related to the wind speed.

The VGR probability term hides internal ramping to a large extend and hence there are inherent uncertainties that are not visible. The VGR probability band on the other hand visualises how forecasts cross each other due to periods of low predictability. Such periods differ between the members in amplitude and phase.

The generated production will have a flatter spectra with energy on a scale smaller than the ensemble members can resolve. Percentiles of VGR are suitable predictors of ramping reserve, because they are robust and smoother compared to the individual members. The ensemble average is a questionable predictor, because it is influenced by outlier members. For the VGP schedule the average is a conservative forecast, but for VGR there is no outlier with opposite sign. The outliers have therefore too much influence.

A. Blending Reserve forecasts

Forecasts are optimised by forecaster providers according to their usage. Reserve forecasts need to be optimised in a different way than forecasts for trading of generation from wind and solar in the market. For reserve, forecasts are nothing more or less than the computation of the inherent uncertainty of the forecast that derives the schedule of the expected generation in the market. The optimisation process will therefore be sensitive to this schedule. Blending of these forecasts however always has to be on absolute MW values.

B. Optimisation Criteria and Incentivisation

The word "incentivisation" may indicate that economic value counts more than system security. In this context, we use incentivisation therefore as a kind of optimisation criteria that could essentially be used for economic incentivisation. For example, if a cost function is chosen that penalises "misses" (failure to forecast a ramp event) and particularly large "misses", this is mostly to ensure system security and does not necessarily impact any economic value. In fact, this example illustrates that it is particularly important that there is sufficient penalty of large "misses" to ensure that the ramp forecast is sufficiently optimised towards grid security.

A cost function serving grid security need not be more complicated than:

$$cost = 4 * |MW_{missing}| + \frac{1}{2} * |MW_{spill}| \qquad (2)$$

The lowest cost is achieved if no spill and no missing MW are forecasted, but it is 8 times more expensive to miss than spill. Spill may or may not include the case where there is excess wind power. In other words, the coefficients can differ between RM products.

There is reason to believe that a simple cost function such as the above example will work well in extreme cases. HSSD should be covered, when misses are penalised sufficiently. If this is not the case, then it is because the forecaster does not try to predict the HSSD, as the forecaster considers HSSD a rare event with low likelihood of a successful forecast. If this is the case, then the forecaster's methodology is not suitable for the conditions in Ireland.

VI. Description of Ramp Forecasts and generated uncertainty forecasts

The aggregated wind generation on the Irish system continuously ramps up and down except when there is no wind. There are two main reasons for this situation on the Irish system: (1) the wind speed is seldom constant over large areas and (2) the capacity density varies across the Irish system.

Large scale wind speeds with variations in the range of 16 to 22 m/s mostly occur during periods from the end of October to the end of February. The remainder of the year is characterised by smaller weather systems with less constant wind.

The forecaster's task is then to assist the system operator in optimising the ramp products required to balance these error waves. This is a a paradigm change in wind power forecasting and it is justified by the fact that wind power capacity is increasing and there is little prospect in the near future to determine the weather in a detail needed to generate a "perfect" weather forecast on the day-ahead horizon.

For a very large power system inside a large continent of uniform installed capacity, it is possible to determine the generation with higher accuracy. Ireland's direct exposure to the Atlantic and its island grid means that system balance with a high penetration level of wind can only be maintained economically with a combination of (a) ST forecasts on the 0-2 hour horizon and (b) dynamic allocation of reserve to balance wind power forecast errors.

The predictability of the atmospheric waves increases with their wave length due to the amount of energy involved, the lifetime of the wave and the amount of measurements indicating the size of the waves. Also, approximations in weather forecast models are most valid for the longer atmospheric waves. Thus, the ratio of predictable energy versus total energy of a given wavelength and time scale is best for long waves and decays with the length of the wave. Waves exist on all scales and different wave lengths interact in the atmosphere. Therefore, the ideal reserve related to the forecast uncertainty should be computed by taking wave amplitude, wave length and the uncertainty of these two into account.

VII. The Relationship between Ramp and Reserve Forecasts

The following is a decomposition analysis of the primary drivers of RM product requirements. The purpose is to illustrate how VG forecasts and their uncertainty contribute to the need for RM products and how the RM scheduling process can use this to allocate optimised amounts of reserve.

The function of a RM product is to ensure that balance can be maintained between Load, VG, Scheduled Primary Power Generation (SPPG) and Import/Export over the product's deployment horizon.

$$LOAD = VG + SPPG + Import - Export + RM_{max} * D \tag{3}$$

where $RM_{max} * D$ is the deployed reserve and SPPG is the synchronous primary power generation. During the ex-ante scheduling process, as in equation 4, the equation is balanced when D has a value of zero.

$$scLOAD = scVG + SPPG + Import - Export \tag{4}$$

where scLOAD is the scheduled load, scVG is the scheduled variable generation VG.

By the time of the next scheduling run or on realisation of the schedule, the load is expected to have incremented by scLOADr and the VG is expected to have incremented by scVGr. For simplicity, the 3 last RHS terms are here assumed to remain constant. In order for equation 5 to remain balanced, the load and VG increments must be offset by the previously scheduled RM capability scRM.

$$\begin{aligned} scLoad + scLOADr &= \\ SPPG + scVG + scVGr + Import - Export + scRM \end{aligned} \tag{5}$$

Solving equations 4 and 5 for the scheduled reserve marginal product (scRM) we find

$$scRM = scLOADr - scVGr \tag{6}$$

scRM is the load (and VG) following capability that is known to be required due to the load and VG varying throughout every scheduling interval. To avoid allocating new RM every hour, scRM will be computed as the maximum required value between several auctions.

We assume that scRM will be deployed at some stage. Therefore, RM_{max} must cover the sum of the scheduled and non-scheduled part.

$$RMmax = scRM + nsRM \tag{7}$$

The non-scheduled reserve marginal product nsRM part contains uncertainty around the schedule of all terms in equation 3, thus

$$\begin{aligned} nsRM &= \\ max(\textit{Largest-Single-Point-of-Failure}&, \\ LOAD - VG - scLOAD + scVG&) \end{aligned} \tag{8}$$

Where all terms in the second part are weather dependent. In a power system in a continental climate with several large load centres, the weather dependent load can be considered as part of the VG, because extreme low temperatures and low VG values can occur simultaneously. More conveniently, we shorten the uncertain VG part to:

$$nsVG = VG - scVG \tag{9}$$

With nsVG there exists a risk of missing reserve capacity in cases where the capacity available on short notice is limited and large amounts required. If risk based forecasting is used for forecasting nsVG it is possible to order RM products in advance.

In the Irish system there exists the choice between allocation of RM8, RM3 and RM1 for nsRM. For very large volumes RM8 is required and would be activated 8 hours before the maximum nsRM value is expected to occur. The ramping products are building blocks, which can build up a too long lasting reserve or temporary peaks.

In case RM3 is preferable, it is the $nsVG_{t+3h} - nsVG_t$ value which determines the need of RM3. Only nsVG down-ramps can increase nsRM. This allows us for convenience reasons to define a non scheduled ramp down over 3 hours nsVGrd3 to:

$$nsVGrd3 = max(nsVG_t - nsVG_{t-3h}, 0) \tag{10}$$

The optimal allocation of RM products can be determined from a cost function or from a profile of the uncertainty. When using an ensemble of forecasts, equation 10 provides us with a probability distribution.

The above decomposition and the required allocation time of RM products explain why VG forecast uncertainty turns into ramp forecasting, although scheduled changes impact RM allocation as well.

There are essentially two methods to generate alternative scenarios to scheduled forecasts and from them compute the non-scheduled VG ramp down (nsVGrd) values for RM allocation:

1) Short-Term (ST) forecast updates in between Long-Term (LT) forecast delivery result in new deterministic nsVG profiles. However, the difference between LT and ST in the intraday can have many sign shifts, which would make allocation of nsVGrd very volatile. To prevent this, maximum shifts from one 15min interval to the next need to be defined.
2) Ensemble forecasts can be used to calculate probabilities of nsVGrd1, nsVGrd3, nsVGrd8 at any forecast horizon. The challenge here is to apply economic and grid security optimisation to the result, because grid

security increases cost and cost reduction increases risk of insufficient non-scheduled reserves (nsRM). Probabilistic forecasts are easily derived, but cost optimisation using the individual ensemble members will increase forecast value, if a precise cost function can be formulated.

A. Grid security and economic value from reserve forecasts

Forecast verification over a year has shown (see VIII) that the average mean absolute error (MAE) is of the order of 4.5% of capacity or 225MW at a wind operating capacity of 5 GW and over an 8-hour forecast horizon. However, the wind power forecast error can peak at 6 times this amount even on the 8-hour horizon.

From the distribution of the forecast errors for 2018 we found that a P50 (percentile 50) is approximately 130MW and a P90 (percentile 90) is 400MW, which is still within the traditional reserve amount. The more challenging values occur only in approximately one out of 10 hours with a range of 400–1500 MW (see details in section VIII and Figures 2 to 3). Such high error periods often last over several hours. Hence, there are long periods of low error in between peak error periods.

1) Motivation for using dynamic reserve allocation: Being able to improve scheduling at 10% or more of the time is the motivation for using dynamic reserve. It is not necessarily the peak error times, where this target may be achieved, but rather in times, where the forecasts are to be trusted.

There were several High Speed Shut Down (HSSD) events in the second half of 2018. In each HSSD event there is a risk of a sudden drop in the wind generation. These events occur within the periods of highest wind penetration. Additionally, there is increased risk of wind farm hardware faults. These events cause particular concern in system operation. Prior to a potential HSSD event wind farms can be pre-curtailed to zero output. If there are many events of this kind, pre-curtailment to zero may be a costly solution. So far the timing of HSSD events have been accurate, but the strength of each event is very uncertain, because the forecast uncertainty is limited to around 2m/s in the wind range around 25m/s. For this reason, it seems more economic to allocate reserves to cover for HSSD than to curtail wind for a longer period around the HSSD event.

2) Current Situation: The wind power capacity on the Irish system is primarily distributed in North, South West and South East regions. Each of these regions can be looked upon as a wind farm cluster, where the generation is correlated. Between the regions there is rather sparse capacity. The non-uniform capacity distribution contributes to waves seen in the aggregated generation, because the geographical extent of each high wind speed area varies and is often smaller than Ireland itself. Each wave causes forecast errors, unless the forecast is exact on the phase of the wind extremes.

3) Impact of dynamic reserve allocation: Both with respect to waves and the HSSD events we can directly argue that the difference between WEPROG's MSEPS ensemble members will in advance objectively indicate, whether there is likelihood of a significant forecast error. The MSEPS ensemble members are designed to provide a physically realistic difference between the ensemble members. Whenever certain changes in the power generation are more or less uncertain, the MSEPS members auto-adapt to the uncertainty of the weather situation.

A sudden event, which is uncertain, will be simulated with a longer wave of a small amplitude to avoid two peak errors, which otherwise would result in a wrong phase for the forecast wave and the measured wave respectively. If the timing of the event is certain, then the ensemble members are aligned with each other. The automatic dampening of uncertain events cause fewer peak errors and therefore a small reserve volume. At the same time the individual ensemble member's oscillations justify that there will be waves of low predictability, which will need to be balanced by reserve.

The expected impact of such dynamic reserve allocation is therefore both on grid security and a less costly grid balancing at a higher wind power penetration level.

VIII. Results

The validation period of this study included 13 months forward from January 2018 to January 2019, which was a difficult month with many ramps and extreme gradients in wind speeds. Each wind peak in January 2019 had a rather short duration, except for the weekend $12^{th}/13^{th}$ of January with up to 700MW wind power being curtailed. There was also one HSSD (high-speed shutdown) event encountered. A traditional January contains long periods of steady wind generation. In January 2019, the jet stream never turned stable. Also, the autumn months of 2018 were characterised by many up-ramps, followed by immediate down-ramps. The match between load and wind generation has been poor and the need of reserve high. The expected usage of RM products for this type of weather is therefore relatively high. It should be noted, that the results presented in this study have been derived directly from the raw ensemble forecast output. There has not been carried out any calibration of the ensemble data, which is a benefit of the physically based multi-scheme ensemble method, where the uncertainty is generated in every time step rather than for specific time horizons.

The following is a summary of the characteristics of the forecast method used in this study:

- raw ramp forecast output from each of the 75 ensemble members
- MW-difference forward in time per ensemble member
- no statistical methods used to tune the output data
- the positive wind power ramps are set to zero
- all values are computed for the potential generation
- the true potential generation used is a composite of the SCADA MW and an *Available Active Power* signal from the wind farms
- available active power values from wind farms were used whenever the MW values were exceeded with 5% of the installed capacity
- the schedule was set to the total MW value of the latest long-term (LT) forecast

The verification of the forecasts contains the following error measures:

- BIAS: mean error of forecasted ramps [MW]
- CORR: correlation of forecasted and measured ramps
- HITS:count of correct forecasted versus measured ramps in [hours]
- MISSES:count of missed forecasted versus measured ramps in [hours]
- FALSE ALARM: count of forecasted, but not measured ramps in [hours][1]
- MaxMISS: maximum *misses* values in [MW]
- MaxSPILL: maximum *false alarm* value in [MW]

The statistical results are shown for the persistence forecasts and 7 percentiles, where the persistence forecast is a backward ramp of the same length, but of opposite sign, carried out at the time of the LT forecast's generation time stamp. The measurements are named "OBS" for observations in the tables and figures. The percentiles are computed from 5 x 75 ensemble forecasts for the day-ahead horizon and 9 x 75 ensemble forecasts for the intraday horizon. The percentiles are derived from these ramp ensemble values with a numerical sorting algorithm.

Figures 2 and 3 and tables I, II show the results of the RM1 1hour and RM8 8hour ramp verification for the intraday. The RM3 3hour ramp verification and the verification of the day-ahead followed the same pattern and are not shown here.

TYPE	BIAS	CORR	HITS	MISSES	FALSE ALARM	MAX MISSES	MAX SPILL
max	251.38	0.23	9447	2	2076	316.6	3253.1
P90	45.71	0.33	8204	44	24	424.6	2636.9
P80	21.16	0.37	7467	64	5	481.3	553.5
P70	6.85	0.37	6836	83	1	523.9	510.9
P60	-3.47	0.36	6246	96	1	566.5	468.3
P50	-11.18	0.34	5716	112	1	594.9	425.8
P40	-17.19	0.31	5303	127	0	609.1	383.2
P30	-21.89	0.27	4957	142	0	609.1	326.4

TABLE I: *Statistical results for EirGrid on Intraday for RM1 containing 9516 events. The numerical distribution of data is shown on figure 2.*

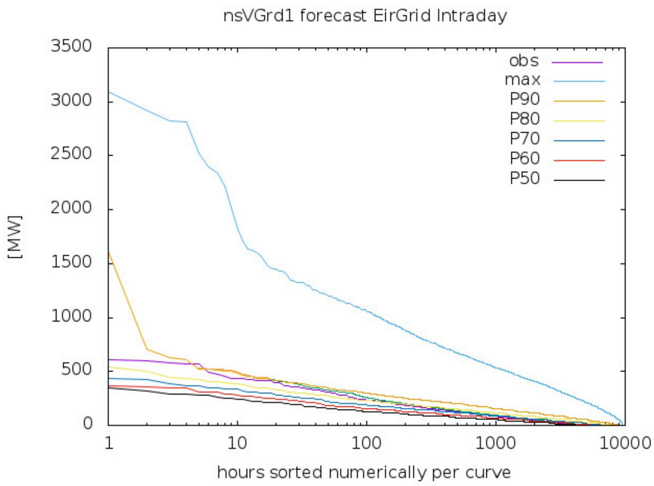

Fig. 2: *Numerically sorted time series of observed, persistence and percentiles for EirGrid at Intraday for nsVGrd1. The statistical results are found in table I.*

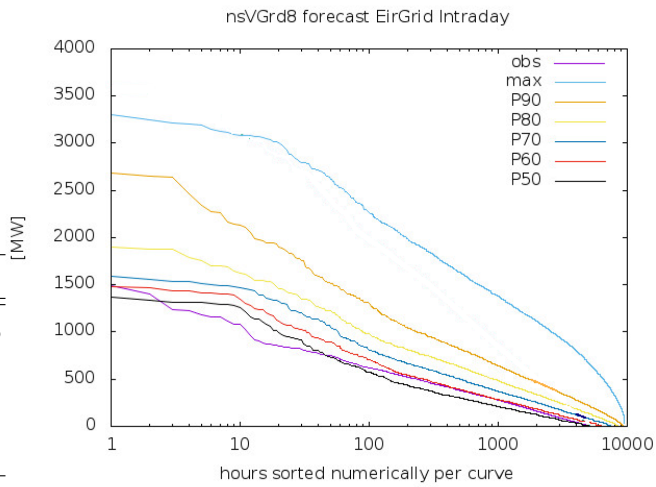

Fig. 3: *Numerically sorted time series of observed, persistence and percentiles for EirGrid at Intraday for nsVGrd8. The statistical results are found in table II.*

TYPE	BIAS	CORR	HITS	MISSES	FALSE ALARM	MAX MISSES	MAX SPILL
max	693	0.35	9462	6	7289	335	3253
P90	216.66	0.37	8564	182	2284	1010	2723
P80	121.06	0.39	7834	335	1189	1225	1787
P70	61.08	0.39	7135	496	611	1371	1626
P60	16.59	0.38	6427	643	335	1435	1509
P50	-17.93	0.36	5714	853	170	1538	1333
P40	-45.19	0.34	5149	1084	79	1611	1126
P30	-66.07	0.30	4742	1312	37	1611	932

TABLE II: *Statistical results for EirGrid on Intraday for RM8 containing 9516 events. The numerical distribution of data is shown on figure 3.*

What is characteristic for the entire verification of the three ramping products RM1, RM3 and RM8 over a period of 13 months and roughly 9500 events, is that the maximum forecast has quite strong BIAS, which means that following this forecast implies too much reserve allocation. Nevertheless, the forecast maximum showed some relevance during events, where P90 is close to zero. The same is true for persistence on correlation. The forecast is not off track, but when evaluated on cases we found that persistence overpredicts especially in the strong ramp events.

The range of forecasts between P30 and P50 is hardly relevant, because they suppress ramping in most cases. These forecasts are actually performing well on correlation, while the spill drops and the misses increase when looking at the results downwards from P50 to P30.

The range from P60 to P90 provides useful forecasts. These forecasts have the highest correlation values and show most reliable scores on all measures. There is very little difference from table to table in this pattern. This indicates that a dynamic selection of percentiles that is temporally dependent on grid conditions and may not relate directly to wind and solar power generation itself, will for this reason create higher value than a fixed choice of percentiles.

[1]Here, only values of 400MW above actual generation has been considered a *false alarm*.

The results also reveal that the ramps are over predicted due to the use of the applied fixed schedule that was not centred in the uncertainty range due to lagging of several 75 member ensembles.

IX. Lessons learned: ramp examples from October 2018 to January 2019

As mentioned before, the jet stream over the Northern Atlantic area didn't become stable for almost 6 months during winter. The result was that there were almost no periods of stable westerly flow with high predictability. Most strong wind events had wind speeds close to the HSSD point (25 m/s). So the strong wind areas were smaller areas of very strong wind instead of larger areas and less strong wind.

Furthermore, forecast errors that vary in time due to the varying predictability of the weather changes and the installed capacity gradients across the island contributed to the complexity and hence forecast error level. This pattern can be explained by the weather being a result of atmospheric waves on different scales. The majority of these waves have phase errors between minutes and up to 2-3 hours. Two error peaks of opposite sign occur at each passage of a wave across a wind capacity maximum. The forecasters should ideally keep phase errors below 15 minutes. Meteorologically, this is however not realistic on forecast horizons, where the wind speed is solely forecasted by weather forecasts and the short waves are the least predictable in the system.

Therefore it is interesting to study the four month ramp verification from October 2018 to end of January 2019 and compare these results to the annual results.

TYPE	BIAS	CORR	HITS	MISS	FALSE ALARM	MAX MISS	MAX SPILL
max	-515.14	0.34	2947	0	1934	186.3	-2420.9
P90	-130.38	0.43	2652	40	249	570.8	-1347.4
P80	-66.95	0.44	2410	66	108	647.4	-1186.3
P70	-27.72	0.44	2172	89	48	764.6	-1054.5
P60	0.74	0.43	1953	122	21	852.6	-937.4
P50	22.88	0.42	1758	169	12	953.8	-790.9
P40	39.84	0.38	1596	210	7	1056.4	-629.8
P30	52.46	0.32	1479	245	2	1159.0	-424.7

TABLE III: *Statistical results for EirGrid on Intraday for RM3 for 4 months (2018-10 to 2019-01, 2964 counted values). The results are comparable to table II*

The statistics in table III illustrate that the ramp forecast performed better in difficult weather than in less difficult weather. The reason is that the lagging of many ensemble members causes a smoothing effect of the percentiles, being less dependent on the recent measurements. The lagged ensemble is more robust both with respect to mean and spread. This is of benefit when the weather has erratic characteristics. The initial conditions are less accurate and in that case blending more results is better than trusting in only a few forecasts.

X. Discussion of results

The results have demonstrated that the uncertainty of the ramping around the wind power forecast schedule varies. There is a pattern of mostly small ramps with the potential of one peak per day in average.

A. Characteristics of Ramping Uncertainty

The results highlight how often ensemble members cross each other within the primary power uncertainty band. The ramp uncertainty looks very different from primary power uncertainty. The timing of extremes varies between the ensemble members and the ramp computations often indicate maximum ramp uncertainty around the local extremes in the wind power generation.

B. Time scales of Ramping Uncertainty

The different ramp horizons show increases in the uncertainty on different time scales. They also highlight how uncertainty is built up prior to extremes. Such uncertainty can be due to a small phase error of the large-scale weather, which may increase linearly with time during the forecast from the time when the low pressure system forms. The differences to the schedule are estimated and contribute mainly on the RM8 horizon. Although the weather uncertainty is far out in the Atlantic, it is visible as slower or faster increases in the wind speeds over Ireland. The difference can be on the 1^{st} or even 2^{nd} decimal of the wind speed, but still amounts to many MW.

C. Physically based variability of Uncertainty

The study's results confirm that uncertainty is variable, but it has a physical nature. We can demonstrate with correlations of 0.41 that the MSEPS ensemble does simulate a good portion of these uncertainty factors even though they are extremely complicated and located in a part of the world, where there is limited accurate and detailed measurements. This is not like a chess game of complicated, but well defined rules. Weather forecasting over the Atlantic area is a highly under-determined and non-linear problem, where we have to trust the forecasts to a great extent.

D. Extreme Event Example on 6^{th} February 2019

If we consider the intense low arriving on the south coast of the island on the 6^{th} of February 2019 in the afternoon, it nicely illustrates the complexity of the forecasting problem. The low was just visible as a wave on the isobars above New Foundland at 6UTC on the 5^{th} of February. It was isolated south of the main flow, but managed to get on a track across the Atlantic and reached Ireland 33 hours later. In that process, it had undergone significant development by interacting with the warm ocean and another low pressure system in higher altitudes. It continued the evolution and got 4hPa deeper between the west and east coast of Ireland. The intensity increase of the low pressure system implied stronger wind speeds and in that case strong ramp rates up to 700MW/h in the forecast.

Looking at 300 ensemble forecasts for that event we found differences in the structure of the low pressure system between all members. Most of the members had the low pressure centred within an area corresponding to half of Ireland, but shifted a bit southward. However, the shape of the low differed between all forecasts and thereby also the location of the extreme wind speeds.

The uncertainty of the low pressure system's position is very small compared to the 33 hour travel distance. The system produced an excellent meteorological forecast for the event. Nevertheless, the uncertainty in power is huge due to the non-uniform wind power capacity that amplifies this uncertainty in terms of MW.

From a single or two forecasts one could not be reasonably sure of the path of the low across the Atlantic. It would take very little difference to make the low go south of Ireland and it did in many forecasts with no impact in Ireland but instead in the UK.

E. State of the art approach to capture extremes

Primitive spatial shifts of the weather forecasts to align wind extremes in different ways are not physically consistent neither across a coast line nor in the proximity of roughness gradients, which is typical for the wind farm regions in Ireland. Movement of the wind farms would be a more correct approach, but then the trained power curves are not valid any more given the change in local roughness. Therefore, massive numbers of ensemble simulation is the only state of the art approach to simulate the sensitivity to the location of the extreme winds and thereby predict objective ramp rates in extreme weather conditions.

The performance of the physically based MSEPS ensemble justifies in this case that it has forecast capabilities far beyond human subjective forecasting and also deterministic forecasting. The event demonstrates uncertainty, but also that uncertainty is limited in space and time.

F. Recommendation for decision making

To justify economic decisions to curtail wind and to allocate ramping resources requires some objective forecast methodologies. Our results indicate that the P70 and P80 forecasts are about equally good choices for all ramping horizons. Capacity in markets, the effective penetration as a function of load and the conventional power plant mixture have influence on what is the preferred ramp forecast percentile. Basically, the cost versus grid security weight is not a constant function, which the forecaster should adapt to.

It is better that a more robust informative set of data is delivered and available to EirGrid. In that way, internal rules for decision making can be developed and the level of grid security can be evaluated versus economic aspects. A typical example that illustrates the complexity of the decision making is a weekend night, where a HSSD event is forecasted and additionally there exist concerns due to maintenance schedules of some other power plant. In such cases, it goes beyond the skill and responsibility of the forecaster to evaluate the system security versus cost ratio.

Therefore, EirGrid will be provided with percentiles of the 3 ramp horizons instead of only one value. Such data will allow EirGrid to choose the ”optimal” percentile for the grid situation at hand. Ongoing and future work suggests that EirGrid are now looking into strategies for the choice of percentile for various grid situations. Further research work is however required to assess how cost versus reliability trade offs should be handled.

XI. Conclusion

EirGrid’s initiative to conduct this study is an outstanding showcase for the use of probabilistic forecasting products in grid operation. The study has revealed a number of new lessons in the use of wind power forecasting and is a milestone in the use of forecasting for high-penetration areas.

The Irish ramping reserve (RM) products complement each other and it was found that the forecasts for ramping reserves are feasible and cost efficient to conduct from an ensemble system that generates physically based uncertainty in every time step like the MSEPS system.

Nevertheless, the successful application of such products for wind power uncertainty does not only rely on the appropriate forecasts, but also on competitive reserve prices and optimal scheduling, which again depends on their cost profile and grid state. In fact, it is an optimisation question on how cost versus reliability trade offs should be handled in various situations, especially in extreme situations.

In that context, the study also revealed that lagging of ensembles is a core element in increased efficiency, because it is not always the most recent forecast that describes a weather situation with extreme elements best. This is a well known phenomena in meteorological data assimilation. More effort must be applied in finding previously generated ensemble members that match the current measurements. Geographical dispersion of the measurements and a high-level of quality are important tools in that context.

References

[1] EirGrid and SONI, *DS3: System services consultation new products and contractual arrangements*, Tech. Rep., 8th June 2012, EirGrid, SONI, Dublin, IRL.

[2] Möhrlen, C., Uncertainty in Wind Energy Forecasting, PhD Thesis, pp19-21, 2004. Online: http://library.ucc.ie/record=b1501384~S0 full text: http://hdl.handle.net/10468/193

[3] Pahlow, M., Möhrlen, C., Jørgensen, J.U., *Application of cost functions for large-scale integration of wind power using a multi-scheme ensemble prediction technique*, in *Optimization Advances in Electric Power Systems*, Editor Edgardo D. Castronuovo, NOVA Publisher NY, ISBN: 978-1-60692-613-0, (2008). http://download.weprog.com/pahlow_et_al_2008.pdf

[4] Jørgensen, J.U., Möhrlen, C., *Increasing the Competition on Reserve for Balancing Wind Power with the help of Ensemble Forecasts*, Proc. 10th Int. Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Farms, Aarhus, 25 - 26 October, 2011. Online: http://download.weprog.com/public_paper_WIW11_032_joergensen_et_al.pdf

[5] Möhrlen, C., Jørgensen, J.U., *Reserve forecasting for enhanced Renewable Energy management*, Proc. 12th Int. Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Farms, Berlin, 11 - 12 November, 2014. Online:http://download.weprog.com/Paper_WIW14-1035_moehrlen_joergensen_online.pdf

[6] Würth, I., and Valldecabres, L. and Simon, E. and Möhrlen, C. and Uzunoğlu, B. and Gilbert, C. and Giebel, G. and Schlipf, D. and Kaifel, A., Minute-Scale Forecasting of Wind Power - Results from the collaborative workshop of IEA Wind Task 32 and 36, Energies 12, no. 4: 712, 2019.