An innovative algorithm for the power loads forecasting in Italian transmission grid: development and main results of the PREVEL software of Osmose project

Dario A. Ronzio
RSE - Sviluppo sostenibile e Fonti Energetiche
Milano, Italy
dario.ronzio@rse-web.it

Elena Collino
RSE - Sviluppo sostenibile e Fonti Energetiche
Milano, Italy
elena.collino@rse-web.it

Giuseppe Lisciandrello
Terna S.p.A., Innovation Factory System Operator
Rome, Italy
giuseppe.lisciandrello@terna.it

Luca Orrù
Terna S.p.A., Innovation Factory System Operator
Rome, Italy
luca.orrù@terna.it

***Abstract*—In the framework of the European OSMOSE project, the Zonal Energy Management System requires, every 15 minutes, the forecast of the loads at each node of the electricity transmission and sub-transmission network for the following three hours. The very short-term forecasts are provided using an Analog Ensemble scheme and an autoregressive method whose inputs consist of the last two months of short-term forecasts and load measurements. The required short-term forecasts result from a Random Forest algorithm trained using the last four months of meteorological data and load measurements. This article describes the first results of the ongoing trial, applied in Southern Italy.**

Keywords — weather forecasting, short-term forecasts, very short-term forecast, congestion resolutions, osmose project

I. NOMENCLATURE

DTR:	Dynamic Thermal Rating
DSR:	Demand Side Response
EMS:	Energy Management System
Z-EMS:	Zonal Energy Management System
NWP:	Numerical Weather Prediction model
PREVEL:	PREdictiVe system for Electrical Load
RAMS:	Regional Atmospheric Modeling System model
ST:	Short-Term forecast (up to 2 days ahead)
VST:	Very Short-Term forecast (up to 3 hours ahead)
WRF:	Weather Research and Forecasting model

II. INTRODUCTION

The ongoing global energy transition, driven by the Paris Treaty (COP21, 2015), is leading to the replacement of fossil-fuelled power plants with renewable energy plants, to mitigate the dramatic effects of climatic change. Both incentives for green energy and decreasing costs of solar and wind power technologies [1] facilitate this transition, posing several challenges in the operation of power systems, both for Transmission (TSO) and Distribution System Operators (DSO). Among others, balancing requirements increase, due to the variability of some renewable energy sources (wind and solar), while balancing capacity is reduced due to the phasing out of many fossil plants. The congestion may become a growing problem as renewable generation increases, as is the case in Germany at present [2]. One must also consider that, while in the past the increase renewable capacity increment was slow, so transmission capacity could be adjusted accordingly in time, at present, renewable plants are installed much faster and they are less coordinated. These problems lead to an increase of violations of transmission grid restrictions, with a consequent adoption of redispatch mechanisms [3]. In addition, the increase in Distributed Generation, characterized by many plants of much smaller size with respect to fossil plants, requires a more active coordination between TSOs and DSOs when flexibility resources are used to avoid congestions and for grid stability management.

Effects of congestions have been investigated mainly in relation to the prediction of wholesale electricity prices [4] and several congestion management procedures have been investigated [5]. In contrast, congestion forecasting is not so common. While in [6] the authors developed a transmission congestion forecasting mechanism based on neural networks, but testing their approach only on simulated data of an IEEE 14-bus system, in [7] the congestion forecast has been approached through a neural network tested on empirical data of the German transmission system.

The overall objective of the project Optimal System-Mix Of flexibility Solutions for European electricity (OSMOSE, [8]), funded under the European call H2020 "LCE-04-2017– Demonstration of system integration with smart transmission grid and storage technologies with increasing share of renewables", is to test and evaluate the reliability of different flexibility solutions for the grid in a real working environment, in order to avoid congestions and better manage grid stability. Congestions can be tackled with several techniques, such as Dynamic Thermal Rating (DTR) and Demand Side Response (DSR). The advantage of DTR is its ability to avoid congestions without changing any grid parameters. The project is organized into several work packages with different objectives. Terna, the Italian TSO, leads the Working Package 5 (WP5), devoted to evaluate and to improve the technical and economic efficiency of the provision of flexibility services through a smart Energy Management System (EMS), that will coordinate DTR sensors and algorithms, DSR by large industrial customers and renewable generation (wind, even with battery energy storage). WP5 involves several university

978-88-87237-50-4

departments, research centres, industrial partners and aggregators. When DTR is not sufficient to resolve a congestion, DSR could be applied, adjusting the amount of power flowing on certain lines by changing the load profile of some industrial customers. The Zonal Energy Management System (Z-EMS), developed by RSE and IBM for the sake of the project, manages short-term congestions up to 3 hours ahead through the coordination of the available resources (mainly DTR and DSR). The implemented mechanism starts with a snapshot every 15 minutes of the grid configuration and produces a forecast of the grid configuration for the next three hours, with a temporal resolution of 15 minutes. Thus, a forecast of the meteorological variables involved in DTR strategy together with a prediction of generation and loads in any subnet of the electric grid are required by Z-EMS. In any grid subnet, only the result of the operative run-time AC load flow result – referred to in this article as *exchange* – is available, therefore the forecast of this kind of quantity has to be evaluated instead of generation and load separately.

The new method of exchange forecasting, developed by RSE in the framework of the OSMOSE project, is described in the following sections. More precisely, in Section III the input data and the implemented methodologies are presented, some preliminary results are described in Section IV, while the main conclusions are illustrated in Section V.

III. Materials and Methods

A. Grid Data

Fig. 1 shows the locations of the network nodes (substations) used in the test. The Demo area is the 380-220-150 kV network of a portion of South Italy (red rectangle) that spans from 39.8N and 41.9N of latitude and from 15.1E to 18.4E of longitude. The forecast of the exchange is performed on all the nodes belonging to the larger area shown in the figure, that extends from 38 to 44 degrees North and from 11.6 to 18.4 degrees East. The substations are 1055, of which 203 belong to the demo area.

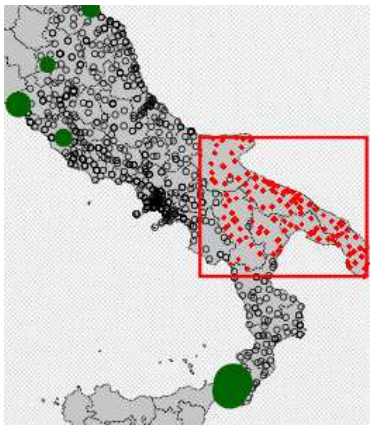

Fig. 1. Location of the nodes of the electric network at which the forecast of the exchange is made. The red rectangle contains the area of the demo HV network, the green solid circles the boundary lines. The size of the symbols is proportional to the energy flow in June 2021.

The need to produce forecasts for all loads in the extended area, not just for the ones included in the demo area, is due to the numerical requirements of the Z-EMS, as is the requirement for the prediction of the power flowing on some lines at the boundary of the integration domain (indicated by large solid green circles in Fig. 1).

Generations, loads or both can be connected to each substation. Furthermore, the number of loads is not constant, but can vary in number and time. At one moment, the snapshot takes a picture of the loads feeding the various substations, but 15 minutes later a fraction of these may change, e.g. the same substation may have different combinations of loads, new active loads may appear and others may no longer be active.

As the Z-EMS only needs information on the exchange at the substation level, it was decided to aggregate all generation and loads pertaining to the same substation together, and to produce the forecast on this basis. This choice seemed to be the most efficient and would have reduced the forecast errors (the performance of an aggregate of plants is usually better than the sum of the individuals). However, the frequent variations of load combinations in the same substation cause significant issues in the training sets needed by the forecasting algorithm, because a historical dataset of sufficient length with a stable configuration has to be built. The set of possible load configurations is shown in Table I.

TABLE I. Number of possible different loads referring to the same subnet, for the demo area and for the whole area.

Number of different loads	**1**	**2**	**3**	**4**	**5**	**6**	**7**
demo area	84	133	32	5	0	0	0
whole area	596	587	104	21	1	1	1

As said before, the available measurements at the substation level consist in the result of the AC load flow, performed at the current quarter-hour by Terna algorithms. From a forecasting point of view, this presents a twofold challenge:

- Generation and load forecasting techniques usually adopt different predictors, since – basically – the former depends on wind intensity / surface irradiation, according to the source, and the latter on calendar information and subjective temperature. As no information on generation is available in advance, the exchange predictors must be able to detect both of these features.

- The exchanges come from an AC load flow. This implies that somewhere and sometimes there may be weak correlation between meteorological data and exchanges due to the load flow optimization effect. In addition, the exchanges in some substations may suddenly vary from one baseline to another one, even very different from the previous one. This new baseline may last a few minutes or a few hours or many days, without any reasonable pattern (from a meteorological perspective). This is another issue in training and forecasting.

The measurements are available approximately four minutes after each nominal execution round.

B. Meteorologica Data

The forecasting system needs some meteorological variables to produce the short-term and very short-term forecasts. The variables are derived from two different numerical weather prediction (NWP) models: the Weather Research and Forecasting Model (WRF, [9]) and the Regional Atmospheric Modeling System (RAMS, [10]). They are limited area models and therefore require initial and boundary conditions to solve their differential equations. In this study these conditions are provided by two different global numerical models: the Integrated Forecast System, developed by the European Centre for Medium Range Weather Forecasts, UK, and the Global Forecast System, developed by

the National Oceanic and Atmospheric Administration, USA. The weather variables, used in input to the exchange forecast system, are calculated by averaging the outputs of these four NWP model combinations. The used weather variables are astronomical angles, global, diffuse and direct normal irradiances, temperature, relative humidity and subjective temperature at 2m agl, wind speed at different heights agl (10m, 50m, 75m, 100m, 150m), precipitation and mean sea level pressure. Both WRF and RAMS are run once a day, starting at 12 UTC, and each run supplies the weather forecast for the following two days, with a spatial resolution of 4 km and a time resolution of 15 minutes on the grey dotted area shown in Fig. 1, covering the territory of Central and Southern Italy. These runs are carried out by RSE's meteorological and climate group in Milan. As the geolocation of the substations is only known at municipal level, the outputs of the four NWP models are interpolated on the centroid of the municipalities to which the loads belong. These outputs are then transferred on the RSE's ftp server. These data are loaded into an Oracle database by Terna, and in this way they are made available to the *PREdictiVe system for Electrical Load* (PREVEL).

C. Methods

The PREVEL system is the exchange forecasting system. The only constraint imposed is that the grid state remains frozen for the next three hours, a condition that will not always be verified.

Due to the issues described above, the exchange forecasts are carried out in two stages:

- A short-term forecast (ST), with a temporal horizon of 60 hours and a time resolution of 15 minutes. This prediction is calculated for each substation and for any possible configuration of loads (Table I). It is delivered once a day, late in the evening, using meteorological and calendar information in input. This step is necessary to better describe the behaviour of the real-time series during the last four months, to smooth out the outliers and some short periods during which the baseline changes drastically.

- A very short-term forecast (VST), updated every 15 minutes, with a forecasting horizon of three hours, time resolution of 15 min. The VST is carried out using the ST forecast as main predictor together with the latest real-time exchange measurements. This forecast should only take into account the trend of the last measurements, which might be very different from the baseline predicted by the ST.

The ST is obtained by applying a Random Forest (RF, [11]) algorithm (using 350 trees, seven variables randomly sampled at each split, 80% of data for training and 20% for evaluation). The predictors are both calendar factors (day of the week, such as public holiday working day, bridge or holiday), and weather variables (air temperature, relative humidity, wind speed at 10 and 75 m, global, direct and diffuse irradiation, astronomical angles). The training is built using a sliding window on the last four months of measurements of exchanges and the corresponding aforementioned weather variables, and a further set of six previous weekends. The RF algorithm is applied to the time series of each substation, developing a specific model for each of them. This procedure is activated on Sunday morning and the same model is used throughout the following week. The ST forecasts for that week are obtained by feeding the built model with the current weather forecast provided daily. The RF has been implemented by means of the package *ranger* [12] of the statistical library R, and both model creation and prediction are parallelised to reduce CPU time.

For the VST forecast, two different approaches have been considered: an auto-regressive integrated moving average model with exogenous input (ARIMAX, [13]), and the Analog Ensemble method (AnEn, [14]). Both methods need of a historical dataset, in this case shorter than that one used for RF: three weeks for ARIMAX and two months for AnEn. ARIMAX was chosen because it is a fast algorithm, thus appropriate for online use, and should adapt to sudden changes in the exchange [15]. AnEn is a statistical technique that would require a fairly long training period. Its use in this context of the VST is not usual, but it has achieved rather encouraging results. The input variables of ARIMAX are the past ST, the solar zenith angle and the exchange measurements up to the start time of the new run. For AnEn, instead, the predictors are the ST, the air temperature at 2 m and the surface global irradiance. The AnEn method uses the meteorological forecast to rank the past events most similar to the current forecast, by ordering them with respect to a specific metric. Once the past situations with the most similar weather forecasts have been identified, the predicted exchange is calculated by averaging the past exchange measurements at those selected times. The distance metric adopted for evaluating the similarity between the current vector C of predictors and the past forecasts P is shown in (1):

$$d_{DM}(t,t') = \sum_{v=1}^{N_v} \left[\sum_{j=-\tilde{t}}^{\tilde{t}} \left(C_{v,t+\tilde{t}} - P_{v,t'+\tilde{t}} \right)^2 \right]^{1/2} / \sigma_v$$

where v represents each variable of the set of N_v predictors, $C_{,t}$ the *v-th* component of the array of the current forecast for the forecast horizon t, $P_{v,t'}$ the *v-th* component of the past forecasts relative to the same temporal horizon of $C_{v,t}$, and the time window $\tilde{t} = 1$. Each variable used to evaluate the distance is previously normalized using the standard variation σ_v of each variable inside the historical period considered.

Due to the presence of outliers (which could be introduced by the AC load flow procedure), training data must be pre-processed for cleaning purposes. Moreover, with real-time data it is not so easy to identify outliers when they occur at the last steps.

The CPU time required to calculate the VST predictions of about 1300 different load configurations on a Linux server with 12 intel Xeon E5-2660 processors is about 90 seconds for AnEn and 130 seconds for ARIMAX.

IV. Preliminary Results

A preliminary description of the first results of the application of the two-stages forecasting method is given here, and, in particular, only the Analog Ensemble implementation as a VST method is analysed, leaving the comparison between ARIMAX and AnEn to a later study. The preliminary analysis covers the period from 10th to 16th July 2021. Thus, 672 runs have been considered, involving approximately 850 substations in the entire area and 200 in the demo area.

It is necessary to consider at least three different groups of substations:

- Demo area: this contains about 200 substations.

- Entire area (LOADS): about 850 substations are located outside the demo area but inside the forecasting (and Z-EMS) domain.

- Boundaries: there are seven HV lines at the boundary of the forecasting domain (large solid green circles in Fig. 1). Since no information about these lines is available, the forecast is performed as if these lines were loads.

The capacity of these substations is highly variable and can range between a few kW and ±300 MW. The power flowing through the boundary lines can vary between -1400 MW and 1000 MW. So far, the same forecasting methodology has been applied to the three datasets, as no information characterising the different types of loads (or lines) was available. Fig. 2 shows the distributions of the maxima (in absolute value) of the exchange, on the left, and the distribution of the total power transited in the various substations during June 2021, on the right. It can be seen that the most frequent maximum value is between 10 MW and 50 MW.

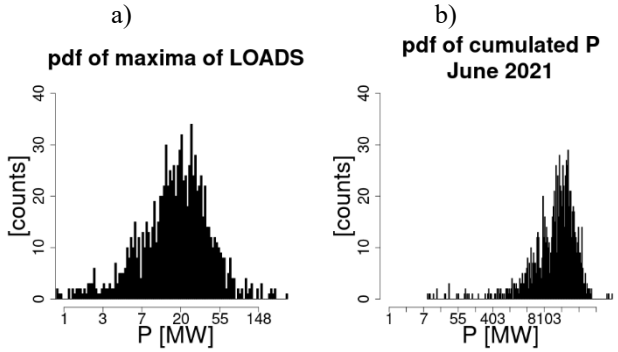

Fig. 2. Histograms of the a) maximum absolute value of the active power and b) of the cumulative (in absolute value) power transited through the substations of the entire domain in June 2021. The x-axis uses a logarithmic scale.

Identifying a key performance indicator for exchange forecasts is not so straightforward. The most used error indicators are the BIAS, defined as the average of the errors, the MAE, the average of the absolute errors, the RMSE, the square root of the average squared errors, and the MAPE, the average of the percentage errors. BIAS, MAE and RMSE can be divided for the average measurements obtaining the corresponding relative (percentage) errors rBIAS, rMAE and rRMSE. Defining the error as:

$$e_{sub}^{forec}(t) = P_{sub}^{forec}(d, t_{start}) - M_{sub}(d, t_{start} + forec)$$

where $\boldsymbol{P}$ is the prediction for each substation at starting time t_{start} of the day d and the leadtime $forec$, $\boldsymbol{M}$ the corresponding measurement, and the percentage error as:

$$p_{sub}^{forec}(t) = 100 \cdot e_{sub}^{forec}(t)/M_{sub}(d, t_{start} + forec)$$

the scatter plots of the instantaneous percentage error p_{sub}^{forec} against the measurements $\boldsymbol{M}$ for some leadtimes (+15, +90, +120 and +180 minutes) are shown in Fig. 3 for the entire area (small grey dots) and the demo area (large dark red circles), because an objective of the analysis was to determine how much the performance of the forecast would decay with increasing forecasting horizon. For this reason, the VST forecasts for different lead times have been compared with each other. The graphs in Fig3 show a general underestimation for positive loads and it is important to be aware that p_{sub}^{forec} is a poor-accuracy indicator when small loads are involved, because of the insignificant amount used in the denominator to calculate the percentage.

The interesting feature that emerges from Fig. 3 is the strong reduction of the percentage error with the size of the exchange for both positive and negative values, i.e. when load or generation are dominant respectively.

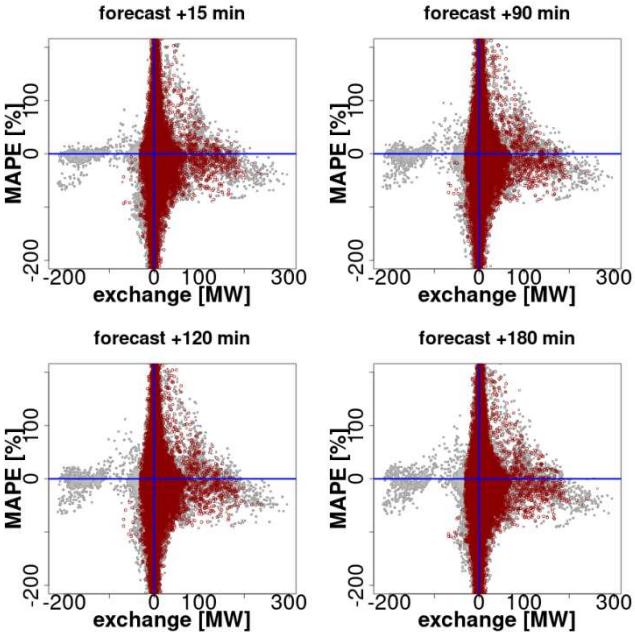

Fig. 3. Scatter plots for leadtimes +15, +90, +120 and +180 minutes. In abscissa the measured active loads, while in ordinate the instantaneous percentage error for the entire area (small grey dots) and for the demo (larger dark red circles) are shown. Boundary fluxes are not included.

Fig. 4 shows instead the errors for the demo and the whole area, varying the time horizon. In this case we find a worsening of 27% for rMAE and 9% for rRMSE, from 15 min to 180 min of forecasting horizon.

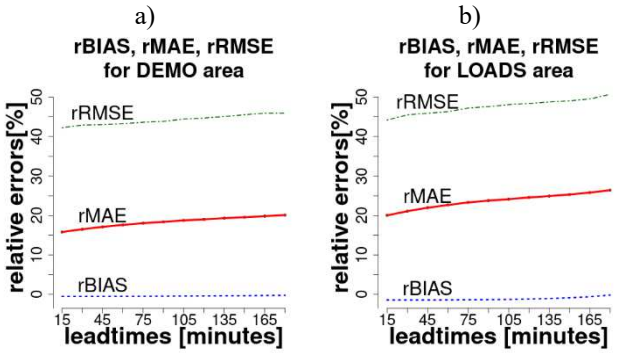

Fig. 4. Trend of relative errors of BIAS, MAE and RMSE for a) demo area and b) whole area as the time horizon changes.

In order to better understand the extent of errors with respect to the amount of involved energy, the relative differential MAE and BIAS have been calculated and reported in Fig. 5 for the demo area. The measurement series were divided into homogeneous intervals, and MAE and BIAS were evaluated for each interval, considering only the forecasts associated with the measurements belonging to that interval. Relative values were calculated by normalising with respect to the average value of the interval. Due to the fact that the population of each sub-interval considerably differs, especially when the (absolute) value of the measures increases, only the indices of samples made of at least 20 items are shown and a neighbourhood of the origin is masked, due to the presence of large errors associated to small

denominators, and therefore with little influence in terms of moved energy.

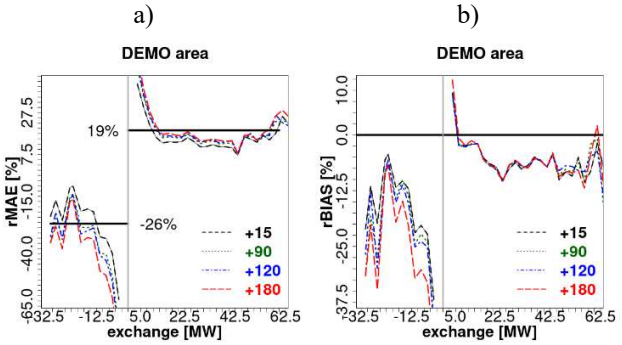

Fig. 5. Relative percentage a) rMAE and b) rBIAS for the demo area for specific ranges of measurements. The error trend against exchange measurement is shown for four leadtimes with different colours. Sampling was carried out at a rate of 2.5 MW.

As said before, large errors are present for small exchange values, just like for MAPE, while errors show a more stable trend for larger values, ranging around a rMAE of 19% for positive exchanges, and –24% for the negative ones. A quite similar trend is obtained in terms of relative BIAS. It is clear that the forecasting system based on Random Forest for ST and Analog Ensemble for VST is reasonably stable when varying the time horizon.

Another aspect to consider is the daily variability. For this purpose, for each leadtime the absolute errors $|e_{sub}^{forec}(t)|$ have been represented by means of whiskers plots, by grouping data with respect to the daily quarter-hours. Fig. 6 shows the boxplots for the first (+15') and the last (+180') time horizon. The worsening is more evident during the central hours of the typical day both in the interquartile range and in the 5th–95th quartiles, to which the forecast error of the distributed generation probably contributes.

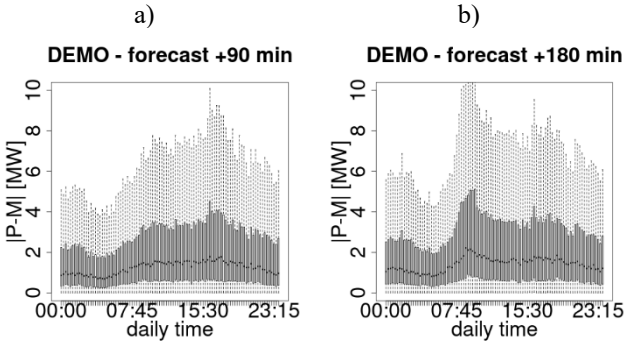

Fig. 6. Quarter-hourly boxplots for the DEMO area at leadtime a) +15 and b) +180 minute. There are shown the 5th and 95th quartile (longer dotted lines), the 25th and 75th (vertical rectangles), and the median values (black points). Outliers have been omitted.

Some insights into these issues can be gained by looking at the temporal evolution of the forecasts, to also understand the different situations that may arise. Figures 7 to 10 show some peculiarities of the time evolution of the ST and VST forecasts with respect to measurements. Moreover, the persistence is also shown, i.e. the time series obtained by keeping the last measure constant for the following three hours.

A case of baseline change is shown in Fig. 7. On Monday 12th July, the exchange jumps from 7 MW to 20 MW. This behaviour was not predicted by the ST methods (blue lines) on the basis of the past behaviour of the loads at that station, while the VST required many days to match the measurements again.

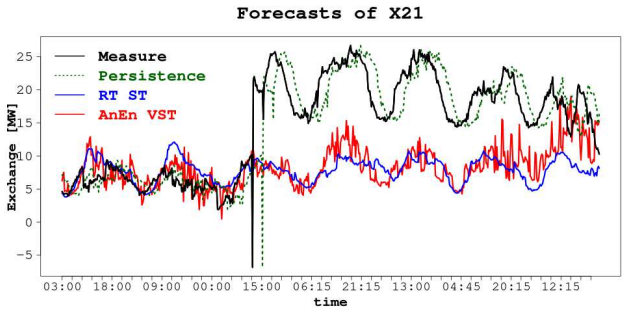

Fig. 7. Demo case. Temporal evolution of ST forecast (RF, blue), VST (AnEn, red), measures (black) and persistence (green dotted line). The period is from 10th to 16th July 2021.

Fig. 8 depicts a case of very fast oscillations, with cycles several times a day. The ST can describe the average trend, and the VST can follow (at +180 minutes) measurements quite well.

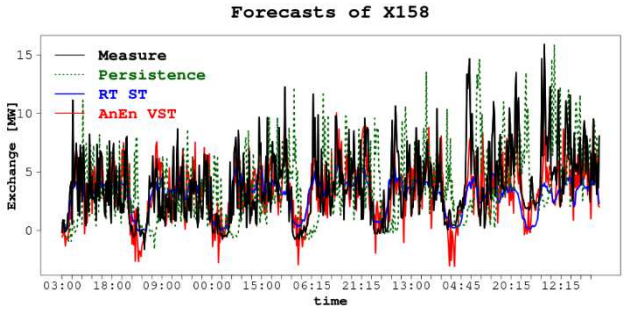

Fig. 8. Demo case. Temporal evolution of ST forecast (RF, blue), VST (AnEn, red), measures (black) and persistence (green dotted line).

Fig. 9 shows a case for which the solar generation forces the time evolution. Here the ST performs very well, and the VST gains a R=0.956, compared to an R=0.97 of the ST.

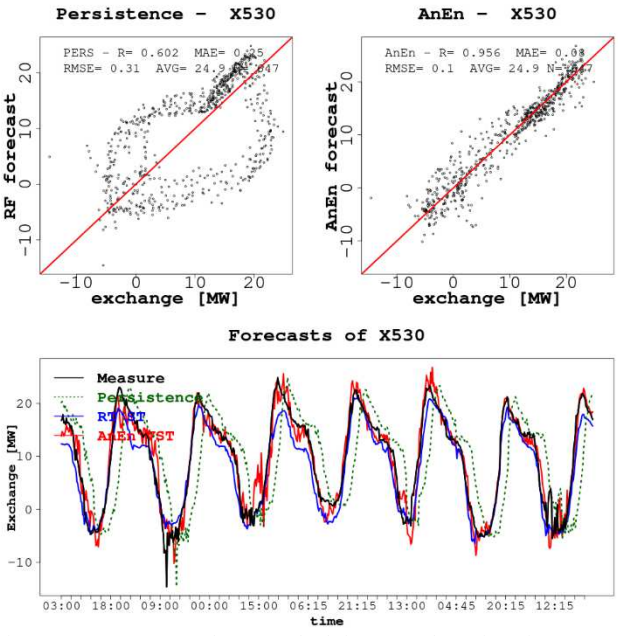

Fig. 9. Demo case. Above: on the left scatter plots of persistence vs. measurements, on the right AnEn forecast vs. measurements. Below: temporal evolution of ST forecast (RF, blue), VST (AnEn, red), measures (black) and persistence (green dotted line).

The worsening of the performance of VST for (some) ST predictions is an example of the difficulties in applying the AnEn method, which is based on past measurements and bears little relation to the trend of the latest measurements, resulting in spurious fluctuations around the actual measurement. This effect is essentially independent on the time horizon.

Finally, Fig. 10 shows the time evolution at a line on the boundaries. The power involved in some of these boundaries is very large, as in the example, but for this case data can be modelled quite well, with a Pearson correlation factor of 0.926 at +15 to 0.697 at +180 minutes, much better than that of persistence (0.453).

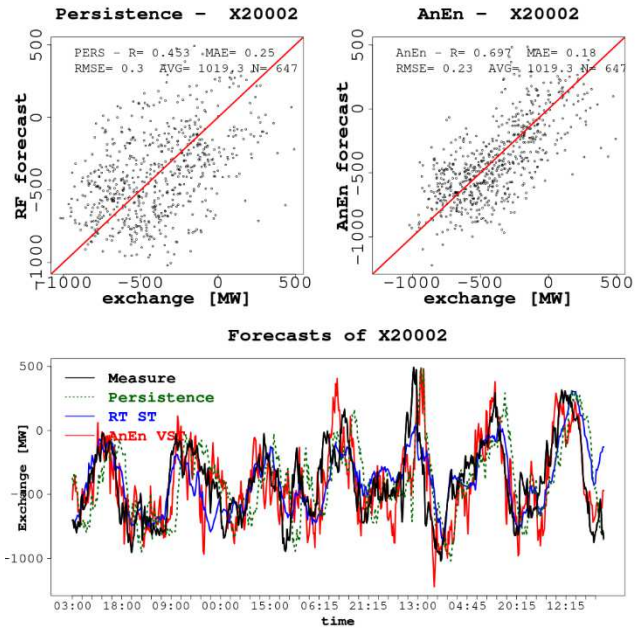

Fig. 10. Boundary case. Above: on the left, scatter plots of persistence vs. measurements, on the right AnEn forecast vs. measurements. Below: temporal evolution of ST forecast (RF, blue), VST (AnEn, red), measures (black) and persistence (green dotted line).

Fig. 7-10 makes it clear that there is a great variability between substations. As specified above, in addition to not knowing what type of load and/or generation is referred to the substations, ramps, outliers and baseline changes occur.

V. Conclusions and Future Works

In the framework of the European project OSMOSE a demonstrator has been implemented to test and assess the reliability of implementing flexibility strategies for congestion identification and resolution. These strategies, as DTR and DSR, are activated by a zonal EMS that analyses the grid status every 15 minutes and produces the optimal flow for three hours ahead, using as input the forecasts of the loads at every primary substation. The word *load* is not precise, as the available measurements correspond to the differences between load and generation, both of which are present in many primary stations, processed by an optimal power flow. Furthermore, the demonstration area is a smaller region than the forecasting one, due to the numerical stability requirements of the Z-EMS. In addition, some lines at the boundary of the domain, have to be forecasted to close the equations implemented in the Z-EMS.

In this paper, we have presented some preliminary results concerning the PREVEL software, the exchange forecasting system implemented in the WP5 demonstrator. While the short-term forecasts are developed by means of a Random Forest algorithm and produce exchange time series for two days ahead, the very short-term forecasts are calculated using two different approaches: an autoregressive method and one based on the Analog Ensemble scheme. The results obtained by applying the RF+AnEn method have been described here, while a comparison between the two methods is still in progress. The forecasting methods are continuously monitored and adjusted to face with/tackle the challenging issues which in a working environment can arise, such as ramps and baseline changes or the presence of some non-modelling time series of the exchange. In addition, the PREVEL software, which runs on an operational server within Terna's organisation, has to manage the synchronisation of a number of processes such as input data acquisition, DTR forecast systems and the production of weather forecasts developed externally at RSE.

Acknowledgment

D. Ronzio thanks F. Imer (Terna S.p.A.) for the IT support in the operational management of the software on the operating server.

References

[1] Renewables 2020 Global Status Report; REN21: Paris, France, 2020; ISBN 978-3-948393-00-7.

[2] H. Schermeyer, C. Vergara, and W. Fichtner, "Renewable energy curtailment: A case study on today's and tomorrow's congestion management," Energy Policy, vol. 112, pp. 427–436, 2018.

[3] K. Trepper, M. Bucksteeg, and C. Weber, "Market splitting in germany–new evidence from a three-stage numerical model of europe," Energy Policy, vol. 87, pp. 199–215, 2015.

[4] N. Amjady, "Day-ahead price forecasting of electricity markets by a new fuzzy neural network," IEEE Transactions on power systems, vol. 21, no. 2, pp. 887–896, 2006.

[5] Pillay, S. P. Karthikeyan, and D. Kothari, "Congestion management in power systems–a review," International Journal of Electrical Power & Energy Systems, vol. 70, pp. 83–90, 2015.

[6] S. Sharma and L. Srivastava, "Prediction of transmission line overloading using intelligent technique," Applied Soft Computing, vol. 8, no. 1, pp. 626–633, 2008.

[7] P. Staudt, B. Rausch, J. Gärttner and C. Weinhardt, "Predicting Transmission Line Congestion in Energy Systems with a High Share of Renewables," 2019 IEEE Milan PowerTech, 2019, pp. 1-6.

[8] https://www.osmose-h2020.eu/project-overview/

[9] C. Skamarock, B. Klemp, J. Dudhia, O. Gill, D. Barker, G. Duda, X. Huang, W. Wang, G.A. Powers: "Description of the Advanced Research WRF Version 3.", 2008, doi:10.5065/D68S4MVH.

[10] R.A. Pielke, W.R. Cotton, R.L. Walko, C.J. Tremback, W.A. Lyons, L.D. Grasso, et al.: "A Comprehensive Meteorological Modeling System—RAMS.", Meteorol. Atmospheric Phys. 1992, 49, 69–91, doi:10.1007/BF01025401.

[11] L. Breiman, "Random Forests, Machine Learning.", 2001, 45 (1), 5–32, https://doi.org/10.1023/A:1010933404324

[12] Y. Li, Y. Su, L. Shu: "An ARMAX Model for Forecasting the Power Output of a Grid Connected Photovoltaic System.", Renew. Energy 2014, 66, 78–89, doi:10.1016/j.renene.2013.11.067.

[13] M.N. Wright, A. Ziegler, "Ranger: A fast implementation of random forests for high dimensional data in C++ and R.", Journal of Statistical Software, 2017, 77(1). https://doi.org/10.18637/jss.v077.i01

[14] S. Alessandrini, L. Delle Monache, S. Sperati, G. Cervone: "An Analog Ensemble for Short-Term Probabilistic Solar Power Forecast.", Appl. Energy 2015, 157, 95–110.

[15] E. Collino, D.A. Ronzio, "Exploitation of a New Short-Term Multimodel Photovoltaic Power Forecasting Method in the Very Short-Term Horizon to Derive a Multi-Time Scale Forecasting System.", *Energies* 2021, *14*, 789. https://doi.org/10.3390/en14030789