Assessing the need for better forecasting and observability of PV

A white paper by the European Technology & Innovation Platform PV Working group on grid integration

www.etip-pv.eu

Authors

Active contributors to the white paper are (in alphabetical order):

- Pierre-Jean Alet
- Venizelos Efthymiou
- Giorgio Graditi
- Mari Juel
- David Moser
- Franco Nemac
- Marco Pierro
- Evangelos Rikos
- Stathis Tselepis
- Guangya Yang

Layout and Printing

Secretariat of the
European Technology & Innovation Platform PV
Tel: +49-89-720 12 722
Fax: +49-89-720 12 791
info@etip-pv.eu

Disclaimer

The opinions expressed in this document are the sole responsibility of the European Technology & Innovation Platform PV and do not necessarily represent the official position of the European Commission.

"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727272"

Table of Contents

1. INTRODUCTION

In the last decade, cumulative installed capacity of photovoltaic (PV) has grown at a compound average rate of 49% per year, reaching by the end of 2015 a worldwide installed capacity of 230 GW. In 19 countries the annual PV contribution to electricity demand was estimated to exceed the 1% mark, with Italy leading with at least 7.9% followed by Greece at 7.6% and Germany at 7%. Different IEA scenarios predict for 2050 a PV penetration between 6% and 16% of the world electric consumption [1], while other scenarios advocate for much more ambitious numbers in order to reach 100% renewable energy in the electricity sector in the same timeframe. For instance a roadmap from the Stanford University [2] envisions a total worldwide PV penetration of 47% in 2050, as shown in Figure 1.

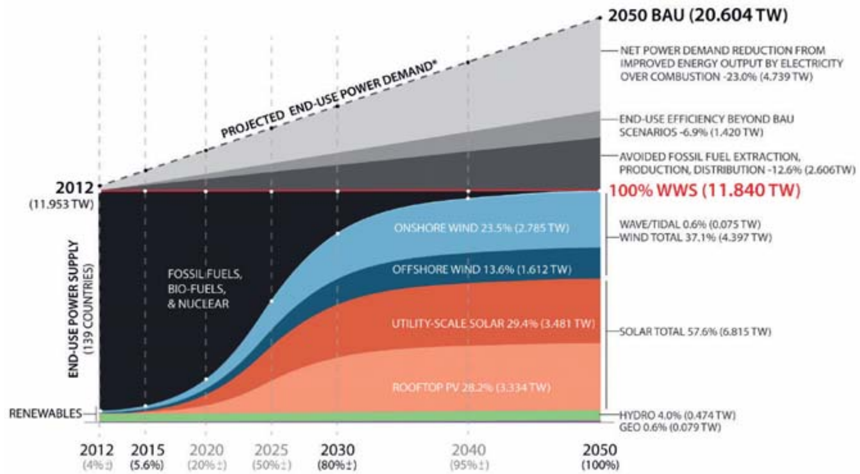

Figure 1: A vision of the structural change required in worldwide electricity generation and consumption in order to reach 100% coverage by wind, water and solar (WWS) by 2050. Note that besides PV, the 'Utility-scale solar' fraction in the figure includes about 10% CSP generation [2].

Electricity grids can be affected by high PV generation long before attaining such ambitious penetration levels, introducing a stochastic variability dependent on meteorological conditions. On the daily time scale in particular, PV production increases the rapidity of load ramps so that a greater secondary reserve and ready supply is needed. This is accentuated in the evenings when the rapid reduction of large amounts of PV power production combines with an increase in electricity demand, a phenomenon which has come to be known as a "Nessie" or "duck" curve.

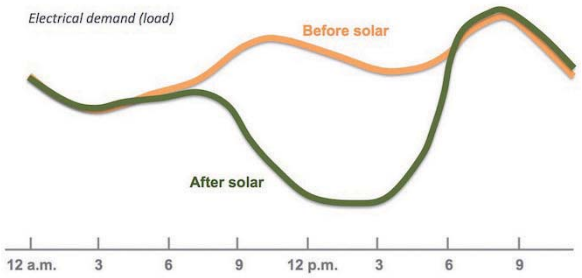

Figure 2: Example of a "duck" curve, showcasing typical grid load with and without high PV penetration over the course of a sunny day (from midnight to midnight) [3].

Thus, a large share of PV power introduces new challenges for the stability of the electrical grid, both at the local and national level, requiring the need of revised reserve policy (more distributed than central) and utilization complementing technology solutions (flexible generation, flexible demand response, storage, etc.) to ensure electrical balancing and overcome the unpredictability and variability of demand and intermittent generation. Moreover it implies an increase in costs related to transactions and dispatching/balancing operations on the Day-Ahead, Intraday and Real-Time Energy Markets. Despite the challenges, grids can sustain high penetration of distributed power generation provided that quality of supply is addressed at connection point through the capabilities of modern power electronics, distributed control, and the use of ancillary services.

PV power forecasts could mitigate the effects of high solar power injection into the electricity grid, both for grid management and on the energy market. Short-term forecasts (intra hours) could be used to predict power ramps and voltage flickers as well as to better control operations on the Real-Time market and dispatching management. Mid-term forecasts (intra-day and day-ahead) could be used, on one hand, for load following to control voltage and frequency instability and for transmission scheduling to reduce the secondary reserve. On the other hand, it could be employed for a better match between the intra-day and day-ahead market commitment and the real PV production, reducing the energy unbalancing costs. For these reasons, the site and regional day ahead forecast of the solar power generated by large PV producers and Distribution System Operators (DSOs) is now mandatory in many European and non-European countries (Italy, Germany, Spain, Romania, USA, Japan etc.) yet the required level of accuracy is still generally undefined.

In its review of the challenges and opportunities associated with massive deployment of solar PV generation [4], the Grid integration working group of the ETIP PV identified forecasting and observability as critical technologies for the planning and operation of the power system with large PV penetration. In this white paper we set out to spell out in more details what features are needed from these technologies and what is the state of the art.

Some very good reviews of forecasting techniques have been published in recent years [5, 6]. We have built on these by taking a step back and analysing the different use cases for forecasting in relation to PV, and by linking forecasting to the issue of observability i.e., the ability to evaluate at a given time the status of PV generation. Experts on power systems, PV technologies and forecasting contributed their knowledge of the field as well as first-hand results they have obtained and issues they have observed.

2. RATIONALE FOR PV POWER FORECASTING AND OBSERVABILITY

2.1. Dynamics of the power system

At any time in any power system, consumption (including losses and charging of storage systems) and production (including losses and discharging of storage systems) need to be equal. In a conventional power system operating in alternating current (AC) the frequency is a real-time indicator of this balance. However grid assets have finite dynamic characteristics so to ensure the balance any fluctuation of production or consumption needs to be anticipated as much as possible before it translates into frequency deviations.

Indeed, the characteristic time constants of power system components range from less than a second to ten years or more, as summarized in Table 1:

Table 1: Characteristic time constant of power system components

< 10 s	Inertia response Protection system operations Switching of power electronics Battery switching between charge and discharge
1 min	Fast start of pumped hydropower plants [7] Fast start of some combustion engines [8]
15 min	Gas power plant from 1/3 to full power [9]
1 h	Start-up and shutdown of most power plants
24 h	Commitment of generation units
1 year	Maintenance planning
10 years	Expanding transmission infrastructure
20+ years	Economic lifetime of PV systems Economic lifetime of grid assets

Prior to the introduction of variable renewable sources (wind and PV), power consumption was the only stochastically variable component in power system balance. Forecasting its variations was already introduced in the 1940s. It has since been refined to take into account "seasonal" variations (day of the year, day of the week, hour of the day) and the specific characteristics of different electricity uses (heating and cooling, cooking, industrial equipment, lighting, etc.). The focus has always been on regional or national aggregates [8].

The deployment of variable renewable generation is introducing new requirements on forecasting techniques. First of all PV and wind generators are much more sensitive to weather conditions. The main weather parameter with an influence on electricity demand is temperature, where heating or cooling is powered with electricity. This parameter varies relatively slowly in time and space. PV and wind on the other hand strongly depend on rapidly changing variables: as a first approximation, PV power is proportional to $G \cdot (1 + k \log G)$ where G is the global irradiance on the plane of the PV array, k is an installation-dependent parameter, and wind power varies with V^3 where V is the wind speed. As a result, the geographic distribution of the generators matters more for the aggregate variations than that of the loads. In addition, PV generation is highly distributed in terms of locations and ownership. It is therefore often necessary to forecast generation with a higher spatial resolution than demand. Indeed single MW-scale PV plants may be exposed to market trades, and microgrid operations with self-consumed PV electricity require forecasts at the building or district levels. Such granularity increases the forecasting difficulty: the standard deviation of PV power production is reduced as $1/\sqrt{S}$ and $1/\sqrt{N}$, where S is the surface area of a PV power plant and N is the number of aggregated plants [9, 10].

2.2. Drivers for PV forecasting

An important concept when dealing with forecasting in the power system is the balance group. Balance groups can include generation units, consumption units, or be "virtual" when operated by financial actors who only trade. Forming a balance group is a requirement to operate on wholesale electricity markets. All balance groups report to a balancing authority, which in Europe is generally the transmission system operator (TSO). This authority ensures that trades on the electricity market are balanced i.e., that contracted generation matches forecast consumption. Balance Group Managers (BGMs) are responsible to ensure that at each time step of market operations their contracted production and/or consumption matches the realised values. In case of mismatch between prediction and realisation, BGMs are penalised based on intraday market price; if the imbalance is in the same direction as the whole system (e.g., a producer under-delivering when there is a shortage in production), the penalty will be above the intraday market price and if the imbalance is in the opposite direction the penalty will be below.

PV generators were until recently shielded from this balancing responsibility. In Germany for example, transmission system operators carry the responsibility and operate a balance group for PV systems connected under the Renewable Energy Sources Act in their area [11]. Regulators are now pushing to increase exposure of PV generators to market conditions and increase their responsibility in the balancing mechanisms. A 2014 ruling by the Italian regulator introduced imbalance charges for renewable power generators of more than 1 MW in capacity; the mechanism is similar to that applied to conventional balance groups but the fees are modulated to take into account the inherent volatility of the different sources [12]. The resulting cost for PV generators is estimated around 5 €/MWh, which is still significantly lower than imbalance prices applied to regular balance groups in Europe [13, 14].

In addition, support mechanisms for large PV generators are evolving from feed-in tariffs to market premiums in France, Germany and the UK [15] under which these generators receive a regulated payment on top of market prices. As illustrated in Figure 3, these premiums can be floating i.e., cover the difference between the average market price over a certain period of time – generally one month – and a reference price set by the regulators, or fixed. In both cases generators have a direct interest in maximising the value on the market of the electricity they produce and the volumes they can effectively sell. Since a generator can only commit on the market amounts of power which it can confidently produce, accurate forecasts are essential to maximise these sold volumes.

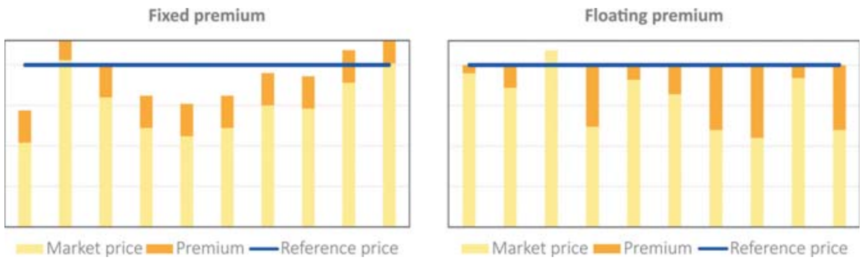

Figure 3: Working principle of market premiums; adapted from [15]

Finally, the development of micro-grids and of combined PV+storage systems requires local energy management which, for optimal operation, relies on predictive control. Single-system or neighbourhood-level power forecasts on timescales from a few minutes to 24 hours are therefore necessary.

Together with the dynamics of power system components described earlier, these drivers create a range of use cases for forecasts on time horizons ranging from 15 min or less to decades, and on geographical scales ranging from a single site to an entire region or country. These use cases are summarised in Table 2.

Table 2: Summary of use cases for PV power forecasting

	PV plant owners PV plant operators	DSOs Microgrid operators	TSOs Market operators
Scale / Time horizon	Single site (10 m – 100 m)	MV distribution grid (1 km – 10 km)	Transmission system (100 km – 1000 km)
15 min	Management of storage system	Management of active/reactive power	Activation of reserves
1 h	Management of storage system Intra-day trades	Storage and load management	Intra-day trades
24 h	Management of storage system Compliance with regulations Day-ahead trades	Storage and load planning	Booking of reserves Transmission scheduling Day-ahead trades
1 year	O&M contract	Planning of maintenance operations	Long-term trades
20+ years	Investment case	Infrastructure planning	Infrastructure planning

2.3. Current and future use cases for PV forecasting and observability

2.3.1. For investors in solar PV power plants

Long term assessments of energy production are essential to investors in PV plants. Investment decisions are made based on a comparison of the levelized cost of electricity for the PV generation with the benefits derivable from the sale of electricity to the grid or self-consumption, taking into account projections for future grid electricity prices. Such assessments are based on statistical weather data ("typical meteorological years") and physical modelling of PV power plants.

2.3.2. For operators of PV power plants

The operator of a PV power plant can be the plant owner himself, or a third party who is conferred operational authority by the owner. Regardless of its size, the primary operational goal of a PV plant is to achieve a profit or cost reduction.

Operators of small PV power plants have an incentive to receive short-term PV forecasting information if such information can be useful to determine their own operational strategies to yield lower energy bills and better utilization of solar PV generation.

For larger PV power plants, time-resolved forecasts of energy production are necessary to place bids on the wholesale electricity market including the fully dispatchable mode, or to honour other forms of contractual arrangements, e.g. with retailers.

2.3.3. For grid operators

The increasing penetration of intermittent energy resources requires grid operators (including Transmission System Operators – TSOs – and Distribution System Operators – DSOs) to pay attention to upcoming fluctuations in electricity generation. Forecasting of solar PV generation, as well as intermittent renewable energy in general, is critical for decisions on active/reactive power flow control, as well as for the operation and management of network components to avoid possible grid overload and to facilitate the economic operation of the system.

Grid operators increasingly need to consider the inclusion of intermittent renewable generation in their network planning models in order to achieve proper planning decisions. In distribution grids in particular, many of the planning models are still based on one-way power flow instead of bidirectional flow accounting for generation from distributed energy resources.

2.3.4. For electricity retailers/aggregators

An electricity retailer is an entity that purchases electricity from different channels, such as the wholesale market or power plants, and sells it directly to consumers with the goal of earning a profit. Retailers are not responsible for balancing the grid, but are however associated with a balancing group and they need to meet the collective profiles of their customers.

For electricity retailers, forecasting of the consumption from their consumers is directly linked to the energy they should procure from their suppliers. The profitability of retailers therefore largely depends on the accuracy of demand forecasting, as the cost of the purchased electricity increases with deviations from the actual consumption. Solar PV forecasting is essential to retailers with high PV installations in their customer group as it influences the consumption of other customers connected to the grid. Moreover, retailers / aggregators will utilise accurate forecasting techniques to plan their supply energy mix capable of meeting the forecasted demand profile of their customers including flexible demand response where provided.

2.3.5. For balance group managers

To formulate a schedule, a balance group manager (BGM) must collect the power generation and consumption information from all the generators and customers in the group. Negotiations may take place between BGMs in case the required schedule leads to any infringement of the grid security. In real time operation, deviations may occur from time to time due to uncertainty of energy generation and consumption in the period. Any deviations from the schedule will be handled at the time of market settlement, while the exact rules may differ depending on the system operator.

In general, short and near term PV forecasting is of great importance to the economic operation of balance groups. BGMs send in and/or modify the schedules in day-ahead and intraday markets. Since the forecasting accuracy generally degrades with the stretch of time horizon into the future, the effectiveness of forecasting results varies. Depending on the price scheme, the use of forecasting results by BGMs covers:

- Establishing the baseline generation/consumption of the Balance Group (BG)
- Generate possible scenarios for economic evaluation
- Trade with other BGs
- Optimize the energy schedule

Besides the above, the results can also be used to:

- Estimate possible operational security issues
- Monitor power quality issues

2.3.6. For balancing authorities

Balancing authorities are required to permanently maintain the power balance within the control area they are responsible for. For them, only aggregated PV forecasting results are of interest. The main application of forecasting for their operations is related to balancing service provision. There can be a number of benefits from PV forecasting for balancing authorities that can be identified and evaluated:

- Reduced cost for procurement of required upward and downward reserve control power and energy including the actual use
- Reduced amount and stress on the regulation units (wear and tear, efficiency), as well as their operational costs
- Reduced ramping capability requirements of the system. Solar irradiance can change drastically at the time scale of seconds, and fast and continuous ramps create problems for the plants and grid operation. Fast ramping capability relying on online capacity is precious and expensive
- Reduced requirements on grid inertia, including both from alternators and virtual
- Reduced amount of interruptible loads to be prepared
- Reduced energy storage requirements

2.4. Performance criteria

Because the use cases are so diverse, there is not a single metric which could characterise an absolutely "good" forecast. Instead, any of the three most commonly used metrics, which are listed in Table 3, can be preferred depending on the target application. These metrics are all based on the difference between the forecasted PV production $Y_{forecast}$ and actual yield $Y_{realised}$ over a given time period. They are generally reported in a normalised way; particular attention must be paid to the normalisation factor and to the integration period. It is in particular good practice to integrate the error only over day hours, since PV production is sure to be zero in the night. And while errors in irradiance forecasts are generally normalised by the average measured irradiance, those on power forecasts are often normalised by the nominal peak power of the system. This difference mechanically results in errors for power generation which are about three times lower than for irradiance.

Table 3: Main performance metrics used to assess forecasting methods

Metric	Formula	Application
Mean Bias Error	$MBE = \frac{1}{N}\sum_{i=1}^{N}(Y_{forecast} - Y_{realised})$	Investment decision
Mean Absolute Error	$MAE = \frac{1}{N}\sum_{i=1}^{N}\lvert Y_{forecast} - Y_{realised}\rvert$	Balance group management
Root-Mean-Square Error	$RMSE = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(Y_{forecast} - Y_{realised})^2}$	Optimisation of generation reserves
Skill Score	$SS = 100\left(\frac{RMSE_{reference} - RMSE_{forecast}}{RMSE_{reference}}\right)\%$	Comparison of the accuracy of forecasts in different locations or years

2.5. Quantifying the value of forecasting

In the electricity market deviations from declared profiles of supply and demand have a cost and this is bound to be the case for imbalances as a result of forecasting deviations of renewable energy sources (RES).

Imbalance prices are regularly published by balancing authorities, which in Europe are generally TSOs. For example, the half-hourly energy imbalance price (charged to suppliers and generators) in the UK on 7th January 2016 ranged from 20.1 GBP/MWh to 119.8 GBP/MWh (average: 39.0 GBP/MWh) [13]. Between 8th October 2015 and 6th January 2016, half-hourly imbalance settlement prices in France [14] ranged:

- From -3.6 €/MWh to 112.5 €/MWh for positive deviations (average: 31.8 €/MWh)
- From 0.3 €/MWh to 265.9 €/MWh for negative deviations (average: 44.4 €/MWh)

The cost which can be attributed to forecast errors is the difference between this imbalance price and market prices over the same time horizon.

For example, the relevant cost for hour-ahead forecast errors is the difference between imbalance prices and intraday spot prices. On average, this difference is typically 20 €/MWh. If a 1 MWp plant in the North of Italy were a balance group on its own it would then be charged this price. This situation is currently hypothetical but may soon become a reality, at least for large PV power plants. As shown in Table 3, the relevant metric for balance group management is the mean absolute error. Over four years for the above-mentioned 1MWp PV plant example it is 11.6% of nominal power with clear-sky persistence, and 7.1% with an advanced forecasting technique (numerical weather forecast plus support vector machine) [16]. Since only daytime is taken into account (yearly average duration of 12 hours), these errors translate into an annual imbalance of 0.50 MWh/kWp and 0.31 MWh/kWp, respectively. So the annual imbalance cost would be 10000 € and 6200 €, respectively. As a comparison, with power-purchase agreements at less than 80 €/MWh as are now contracted in Germany [17], annual income for this plant would be at most 80000 €. So two conclusions can be drawn:

- Forecasting errors can reduce the value of PV electricity by more than 12%
- Advanced forecasting techniques can generate a value of almost 4000 € per year for a 1 MWp plant.

2.6. Current forecasting techniques

The first approach in PV power forecasting relies first on the **prediction of relevant weather parameters** (at least temperature and irradiance), followed by a calculation of the corresponding power output. This approach can build on existing weather forecasting tools. The most appropriate tool to predict irradiance depends on the desired time horizon.

For resource assessment i.e. to predict patterns of energy generation over the lifetime of the system, statistically representative time series of weather parameters are generated based on interpolation of ground-level measurements (weather stations) or satellite images to produce “typical meteorological years”. Typical accuracies are of the order of 5% for satellite data, 3% for site adaptation techniques, and 2% for ground measurements.

For time horizons between six hours and three days, numerical weather prediction (NWP) is preferred. NWP data are generated by global or mesoscale simulation models which provide the numerical integration of the coupled differential equations describing the dynamics of the atmosphere and radiation transport mechanisms [18]. The initial conditions are given by satellite, radar, radiosonde and ground station measurements. NWP data are often corrected by post-processing algorithms called Model Output Statistics (MOS) which use historical ground measurements to partially remove systematic errors [19].

For time horizons between two hours and six hours, visible and/or infrared images are acquired by satellite-based sensors. The cloud index is computed by satellite reflectance measures and is typically used to derive ground-level global and direct irradiances through a model [20]. Since the measurements directly provide solar irradiances, as compared to NWP, only a few relatively simple modelling assumptions have to be applied to derive the solar resource. Persistence of cloud speed and direction (as derived from the two last images) is generally assumed. The dynamic nature of clouds challenges cloud-motion vector approaches as cloud distribution can change substantially within the 30 min horizon which is the typical rate of image refresh. Therefore, it is challenging to account for cloud convection, formation, dissipation, and deformation. However, since large-scale cloud systems (such as those associated with a cold weather front) are more persistent, satellite-based forecasts typically perform more accurately than NWP-based forecasting models up to 6 hours ahead, mostly because of ingestion, data assimilation, and latency of calculations required to “spin up” NWP-based forecasts. As classical satellite methods use only the visible channels (i.e., they work only in daytime), morning forecasts are less accurate than daytime ones because of a lack of time history; to overcome this issue, images from infrared channels (which work day and night) have to be taken into consideration [21].

For time horizons below 30 minutes, total sky imaging is the preferred method. It consists in four steps:

- Acquisition of the sky image from a ground-based, wide-angle camera
- Analysis of the sky image to identify clouds
- Estimation of cloud motion vectors
- Prediction of future cloud cover and ground irradiance.

The maximum accuracy with this method is generally obtained between 5 min and 20 min; with low and fast-moving clouds it can be reduced to 3 min and for high and slow-moving clouds it can be extended to 30 min.

The state-of-the-art accuracy for these methods is summarised in Figure 4.

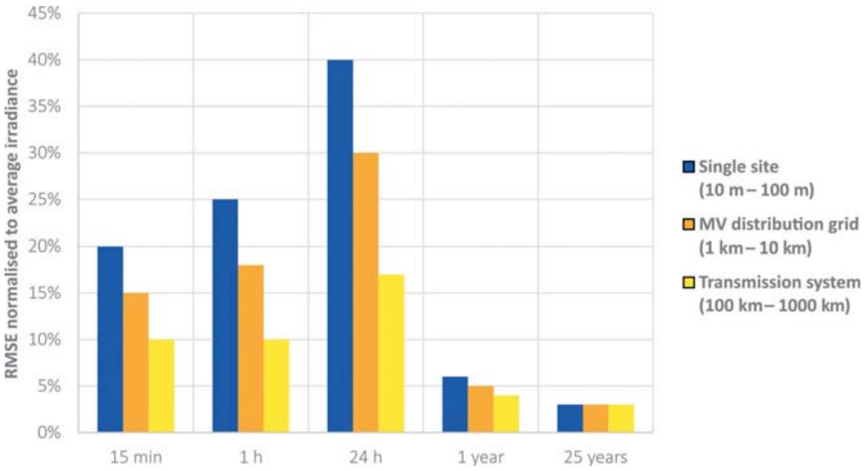

Figure 4: Error obtained with state-of-the-art physical forecasting methods for irradiance (source: ETIP PV)

The uncertainty on PV power modelling from irradiance and weather data comes then on top. In a review of major modelling tools, the hourly RMSE on AC power output was found to be below 7% in all situations [22]. To avoid this addition of errors and to deal with time horizons between 30 min and 2 h where there is no satisfactory physical forecasting technique for irradiance, stochastic learning techniques are used. These methods can be separated between [16]:

- Univariate methods i.e., methods where only time series of the target variable (here, PV power) are fed into the model. These include:

- STL: seasonal decomposition of time series by Loess
- Holt-Winters seasonal method
- TSLM: linear model fit with time series components
- ARIMA: autoregressive integrated moving average
- BATS: exponential smoothing state-space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components
- Nnetar: Feed-forward neural networks with a single hidden layer and lagged inputs for forecasting univariate time series

- Multivariate methods i.e., methods where exogenous variables such as measurements of ground irradiance, temperature or humidity levels are fed into the model in addition the the target variable. These include:

- MLR: Multi-Linear Regression Model
- SVM: Support Vector Machine
- ANN: Artificial Neural Network
- Regression Tree

3. CASE STUDIES

3.1. Single-plant power estimation based on an artificial neural network model (Italy)

In literature, several models have been developed to simulate the different components of PV power systems based on analytical or numerical approaches. In some of these approaches PV systems are simulated through mathematical equations for each component of the system (i.e., PV arrays, batteries, regulators, etc.). Generally, these systems are considered non-linear, thereby requiring complex modeling difficult to define through classical approaches.

In this section, an artificial neural network (ANN) approach is proposed to estimate the AC power production of a 1 kWp experimental micromorph silicon PV plant located at the ENEA Portici Research Center. The multi-layer-perceptron (MLP) is used for solar radiation and PV power production estimation. Feed-forward multilayer perceptron networks consist of units arranged in layers with only forward connections to units in subsequent layers.

The connections have weights associated with them. Each signal traveling along the link is multiplied by a connection weight. The first layer is the input layer; the input units distribute the inputs to units in subsequent layers. In subsequent layers, each unit sums its inputs, adds a bias or threshold term to the sum and nonlinearly transforms the sum to produce an output. This nonlinear transformation is called the activation function of the unit. The output layer units often have linear activations.

The goal is to estimate the PV plant AC power production (P_{ac}), as a function of two input parameters, i.e. the ambient temperature (T_{amb}) and solar global irradiance (GHI). Since the output-input relation used is non-linear the MLP network with only one hidden layer has been proven to be a universal approximation of this function type. In more detail, the proposed ANN is made of a single hidden layer with eight neurons. The values of the learning period, size of data used in the training, neurons on the hidden layer and learning rate were set on a trial and error basis.

Then, to improve the performance of the ANN proposed here, another input representing the clear sky solar radiation has been added. In this case 15 neurons have been considered in the hidden layer.

In particular, two multi-layer-perceptron ANNs have been developed; the first one uses two input, the ambient temperature and the global solar radiation, for the training phase, while in the second ANN the clear sky solar radiation has been added as further input. Available data to train and test the ANNs were relative to seven years, from 2006 to 2012. The ANNs have been trained using only one year of the available data and then tested on data relative to the remaining years.

To evaluate the effectiveness of the approximation done using the ANNs in the evaluation of the AC power, three statistical coefficients have been evaluated: the relative mean square error (RMSE), the relative mean bias error (MBE) and the correlation coefficient (CC).

As shown in Table 4, using two inputs, the relative mean bias error varied between 1.51% and 4.94%, the relative root mean square error between 6.12% and 9.54%, while the correlation coefficient was between 0.9851 and 0.9936. On the other hand in the case of three inputs and 15 neurons in the hidden layer, as shown in Table 5, the relative mean bias error ranged between 1.36% and 5.20%, the relative root mean square error between 5.26% and 8.99%, while the correlation coefficient was between 0.9862 and 0.9954. The worst results were obtained for the year in which available data quantity was the lowest.

Table 4: Values of statistical coefficients relative to the ANN with two inputs (T_{amb} and GHI) values (source: ENEA)

Year for Test	MBE/A (%)	RMSE/A (%)	CC
2006	-4.5829	8.8055	0.9919
2008	1.7732	6.1224	0.9936
2009	2.2403	6.4724	0.9926
2010	1.5120	9.5372	0.9851
2011	3.8415	8.8910	0.9879
2012	4.9430	8.8840	0.9885

Table 5: Values of statistical coefficients relative to the ANN with three inputs (T_{amb}, GHI and CSM) values (source: ENEA)

Year for Test	MBE/A (%)	RMSE/A (%)	CC
2006	4.7613	8.4191	0.9938
2008	1.8174	5.2695	0.9954
2009	1.7453	5.7658	0.9934
2010	1.3649	8.9905	0.9862
2011	3.5926	7.7605	0.9901
2012	5.2045	8.1317	0.9918

3.2. Microgrid management (Greece)

During the evaluation of the MIRABEL concept [23] which was done at the Greek Center for Renewable Energy Studies' (CRES) experimental microgrid on Khytnos, the Engle, Granger, Ramanathan and Vahid-Arraghi (EGRV) model for the PV forecasting was used. In this implementation and despite the fact that no weather data were used to improve accuracy (i.e. only PV power was used), the short and long term accuracies were good enough to provide a fine-grained balancing in the microgrid's power. The diagrams in Figure 5 show that the used forecasting method achieved a high accuracy in the short-term and mid-term scales but also that it presented high discrepancies in the very-short-term. This is due to frequent changes in the solar irradiance profile, which led to fast drops of the PV production during the day. As a result the imbalances present some high (yet short in duration) peaks during the day when these phenomena are more intense.

The inclusion of weather parameter observation can therefore largely improve the very-short-term response and in turn the balancing of the system. Also, according to the analysis above, a larger scale spatiotemporal aggregation would substantially improve the model's performance because in the specific implementation the microgrid size was too small and as a result the aggregation's influence was negligible.

To sum up, the main conclusions from this analysis are the following:

- The specific methodological approach and the used forecast model are appropriate for market balancing mechanisms and can be used by actors responsible for balancing processes (i.e. Balance Responsible Parties) in a Balance Group thanks to its short and mid-term accuracy.
- The long-term accuracy of the model facilitates system and network operators in terms of long-term investment planning and technical upgrades of the grids.
- The very-short-term accuracy, once refined by using meteorological data and higher levels of aggregation, can assist grid operators in the management of reserves and frequency/voltage control.

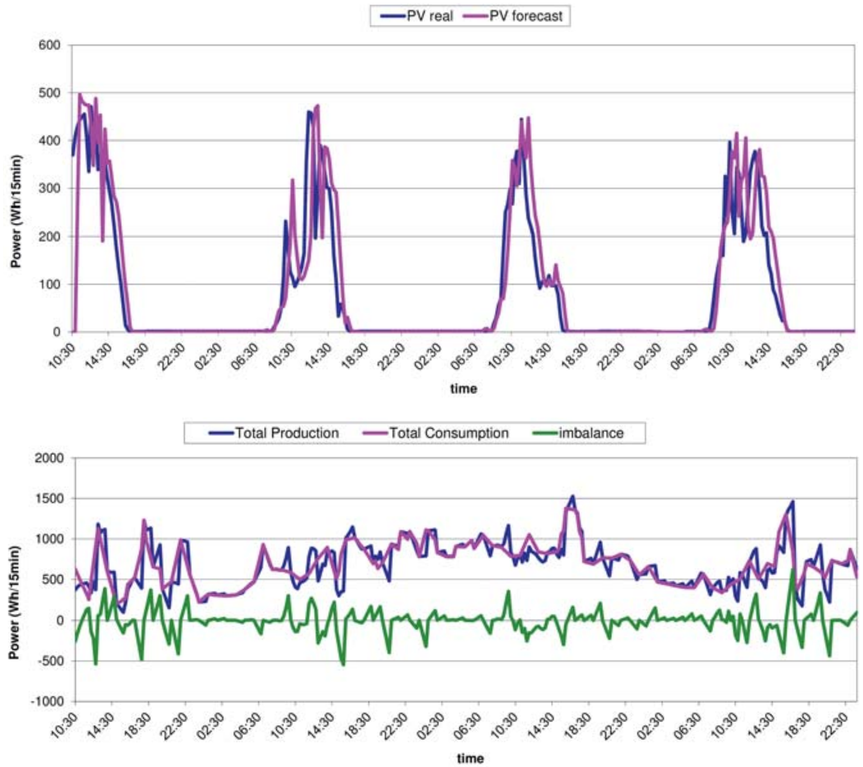

Figure 5: Experimental implementation of the EGRV model for PV forecasting in conjunction with overall balancing of the microgrid in which the method was implemented [24]

3. Case studies

3.3. Regional forecasting of PV production for distribution system operations (Cyprus)

The DSO of Cyprus is working closely with the FOSS Research Centre for Sustainable Energy to develop a complete forecasting tool covering all PV installations in Cyprus. The tool is to be utilised for the long term prediction needs of the DSO as well as day ahead and intraday needs for the development and operation of the grid in Cyprus.

RES power forecasting (mainly PV and wind in the case of Cyprus) plays an important role for the secure, economic and balanced operation of power systems with increased RES in the energy mix. At the moment, allocation by the local TSOs of the required reserves is made in a deterministic way which is based on historical data and the point forecast of a RES aggregate output. This methodology is not adequate as renewables are increasingly integrated into the grid and this may lead to a large provision of reserves, create operational security problems and operational shortcomings as a result of dynamic power variability (ramp-rates, reactive compensation etc). These limitations can be addressed by advanced forecasting tools that utilise spatiotemporal forecasting algorithms that enable relatively accurate prediction of the power produced by photovoltaic (PV) systems. An example of such a system is under development in Cyprus and results to date are promising.

The development of a spatiotemporal forecasting tool primarily requires that the existing and future PV systems are grouped spatially. The division of Cyprus into logical areas considered with the same irradiance (pixilation) is based on the locations of the existing meteorological stations. Connected smart-meters supply real time generation data from systems spread all over the island and provide a good basis to simplify the process without losing the required accuracy. In addition all installed PV systems are linked through their distribution substation to the nearest pixilation area. Figure 6 shows the locations of the 17 meteorological stations and the 100 smart-meters. The capacity and location of the installed PV systems is stored in a geodatabase, which includes all relevant additional parameters: technology, inclination angle, orientation and mounting system. The database is updated daily with information of the newly installed PV systems.

Figure 6: Location of meteorological stations (pins) and smart-meters (circles).
Source: FOSS, University of Cyprus

The forecasting tool receives day ahead Numerical Weather Prediction (NWP) datasets of the solar irradiance and ambient temperature from the Meteorological Service of Cyprus (MSC). The NWP and the data collected from the ground stations (meteorological stations, smart-meters and PV production at the distribution substation level) are fed into a physical and statistical model through which the day and hour-ahead dispatch forecasts of the PV power production of the island are calculated. A brief summary of the physical and statistical approach of the method used is presented in Figure 7.

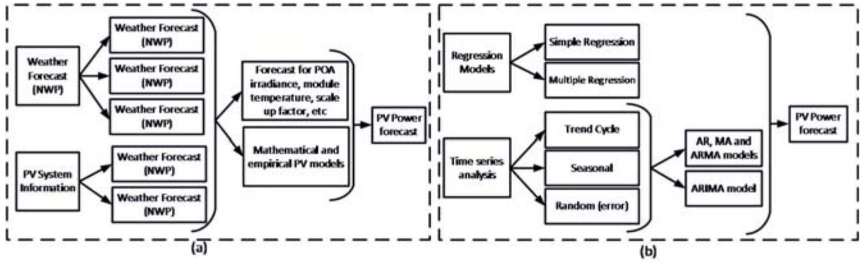

Figure 7: Physical (a) and statistical (b) model approaches to PV power production forecasting.
Source: FOSS, University of Cyprus

The overall approach governing the development of the forecasting tool is summarised in Figure 8. The datasets received from the Meteorological service, the meteorological stations, smart-meters, the local DSO and the distribution substations are initially stored in a database (after an initial quality check). The three following operations of the tool concern three different time horizons: the day ahead, the hour ahead and real time. The “long-term” prediction of the PV production is considered as day-ahead. For this forecast historical time series of all datasets collected are analysed using physical and statistical processes that deliver an accurate forecast. The second evaluation process addresses the “short-term” prediction of PV production, also known as “now casting”. This process uses historical and real time data with statistical valuations delivering PV production between 1 and 6 hours ahead. The last process concerns the real time PV production. Real time data from the meteorological stations, smart meters and the distribution substations are utilised with statistical valuations and scale-up mechanisms that estimate the real time production of PV systems in Cyprus at time intervals ranging from 15 to 30 minutes. The outputs from all operations of the forecasting tool are stored at the central database and are further used to train the developed processes and to continuously improve the effected calculations.

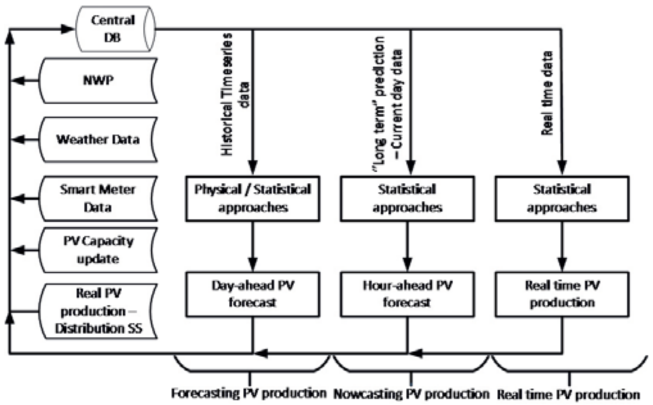

Figure 8: Overall approach in the development of the PV forecasting tool.
Source: FOSS, University of Cyprus

3.4. Regional forecasting of PV production for distribution system operations (Italy)

In a project carried out by EURAC in collaboration with the local DSO, Edyna (former AEW), ANNs were used to estimate and forecast the distributed generation of 1985 PV plants in a small part of the South Tyrol Region in Northern Italy with an installed capacity at the end of 2015 of 68.2 MWp and a PV penetration of 7%. This region of around 800 km2 has a complex orography and variable weather conditions (see Figure 9 (A)).

In this case the forecast method consists in applying spatial clustering of PV plants and then use satellite derived irradiance and numerical weather prediction (NWP) data (centered on each cluster centroids) as inputs for Ensemble of Artificial Neural Networks (ANNsE) that estimates or predict the regional PV power output. The clustering algorithm aggregates all the PV plants in six areas corresponding approximately to the municipality of Naturno, Tirolo-Merano, Lana, Nalles, Collalbo-Sopra Bolzano, and Bolzano (see Figure 9 (B)).

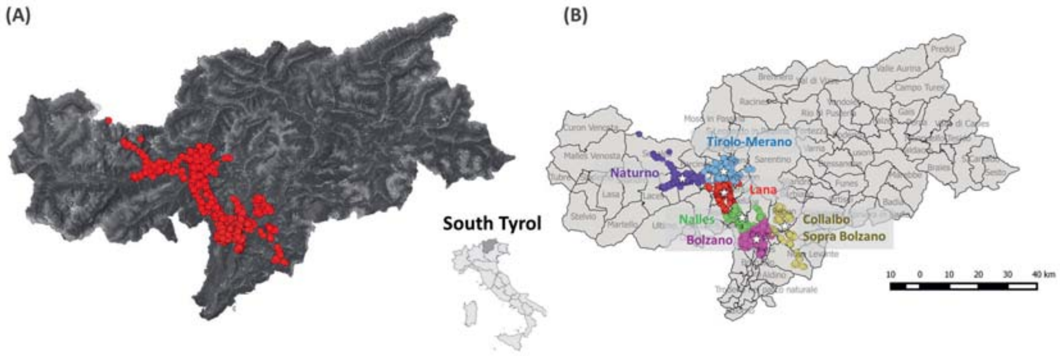

Figure 9: (A) PV plants in the region of interest, (B) PV plants spatial clustering. Source: EURAC Research

3.4.1. Data

The two years of PV power generation data used to train and test the models together with the load and transmission data used to analyse the benefit of PV forecast were provided with 15-minute resolution by the local DSOs (Edyna, AEW). The models were trained with data from 2014 and tested with data from 2015.

The satellite derived irradiance with a spatial resolution of 2 km and an hourly granularity comes from the geostationary radiative fluxes products, under Météo-France responsibility. It was obtained by OSI SAF SSI algorithm applied to the satellite images provided by METEOSAT-9 (MSG-3) at 0° longitude, covering 60S-60N and 60W-60E, at 0.05° latitude-longitude.

The numerical weather predictions were generated by the Weather Research and Forecasting (WRF–NWP 3.6.1) mesoscale model with 20 minute time resolution and 3 km spatial resolution centered on the region of interest.

3.4.2. Estimation and forecast accuracy

The main metric used to measure the power output forecast accuracy is the root mean square error evaluated over the number of sun hours N_{sun} (Table 3). Moreover to evaluate the performance of a forecast model the accuracy is compared to the accuracy obtained by a simple reference model. Usually the simplest reference model adopted as benchmark is the persistence (simple persistence or smart persistence) that considers the persistence of the weather conditions. For a fixed forecast horizon, the accuracy of this model can be considered a measure of the irradiance variability in a specific site or area. On one hand, the more stable the weather conditions or the smaller the forecast horizon, the higher the accuracy of the persistence model and the more difficult outperforming it with a forecasting model. On the other hand, the lower the accuracy of the persistence model, the higher the irradiance variability so that more sophisticated models should be used to provide an accurate prediction.

Figure 10 shows the accuracy of the regional estimation and mid-term forecast models and of the smart persistence model (clear sky persistence) for different forecast horizons.

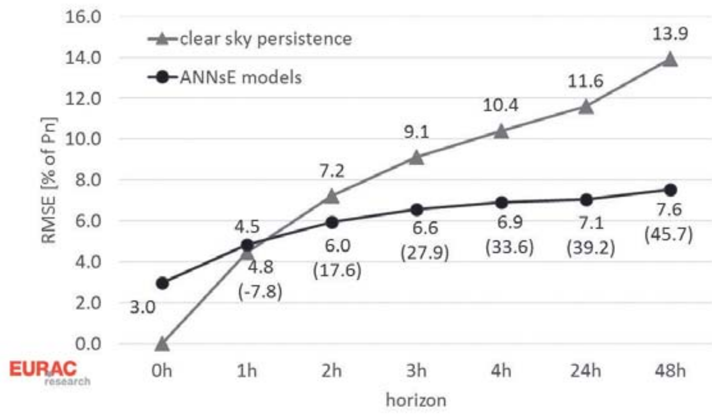

Figure 10: RMSE of regional PV power estimation and forecast vs forecast horizon evaluated for the year 2015 for the area located in the South Tyrol region, Italy. Source: EURAC Research

The trend of forecast accuracy (RMSE) as function of forecast horizon reflects the results reported in literature [25].

The accuracy of power estimation modelling achieved using satellite derived irradiance is around 3% of installed capacity thus the upscaling method could be adopted for real time power supervision.

The intra-day forecast performs with a RMSE between 4.8% and 6.9% of installed capacity while the persistence obtains an error between 4.5% and 10.4%. The day-ahead forecast performs with a RMSE of 7.1% and 7.6% with respect to 11.6% and 13.9% of the clear sky persistence model, or 12.3 % and 13.9% of the simple persistence. These results are compared with the state of the art of Table 6.

Table 6: Comparison of state-of-the-art forecast performance with the results achieved in this case study

Country	Time horizon	Spatial resolution	RMSE [% Pn]	Skill score [%]	Reference
Germany	intra-day	Two Regions (214120 km² - 103890 km²)	3.9- 4.3	40.0 - 42.3 (persistence)	Lorenz et al [26]
Germany	1 h to 4 h	Region (349223 km²)	1.8- 3.8	0 - 11.6 (smart persis-tence)	Wolff et al. [25]
Italy	intra-day	800 km²	5- 7	(-8)-34 (smart persis-tence)	This case study
Germany	24 h	Region (214120 km² - 103890 km²)	4.1- 4.3	48.0 - 52.8 (persistence)	Lorenz et al. [26]
Japan	24 h	Two regions (32424 km² – 72572 km²)	6- 7	50 - 60 (persistence)	Fonseca et al. [27] [28]
France	24 h	Two french counties (7000 km² each)	6- 5.8	-	Zamo et al. [29]
Italy	24 h	800 km²	7.1	42.8 (persistence)	This case study

Moreover, the accuracy of regional forecast leads to a reduction of RMSE between 30%-50% with respect to the performance obtained for the forecast of a single PV plant generation since the spatial averaging reduces the errors (ensemble smoothing effect). This effect depends on the size of the considered geographic area [17, 28]. The obtained RMSE of 7.1% for the regional day-ahead forecast can be compared with the RMSE of 11.8% achieved in the forecast of the power output of an optimal tilted PV plant located in Bolzano [16]. Thus the regional forecast provides a RMSE reduction of 40% with respect to the single site power output prediction, coherently with the literature results.

It should be remarked that the intra-day forecast model improves the day-ahead forecast by making use of past power estimation (based on satellite data). For time horizons longer than 4 hours the intra-day forecast is no longer able to improve the accuracy of the day-ahead forecast (based on NWP data). This means that after 4 hours satellite data becomes less accurate than the numerical weather prediction. Similar result can be found in literature [25]. Moreover in this case, the one hour forecast is slightly less accurate than the clear sky persistence.

3.4.3. Prediction interval estimation

Figure 11 shows the reliability plot i.e., the frequency of observation that lies inside of each prediction interval versus the respective confidence levels (expected probability). A prediction of the forecast errors is completely reliable if the observed frequency is equal to the corresponding confidence level (grey dash line e.g., 50% frequency at 50% confidence level).

In the present case, the model provides a correct estimation of the prediction interval since the observed frequency is almost equal to the expected one.

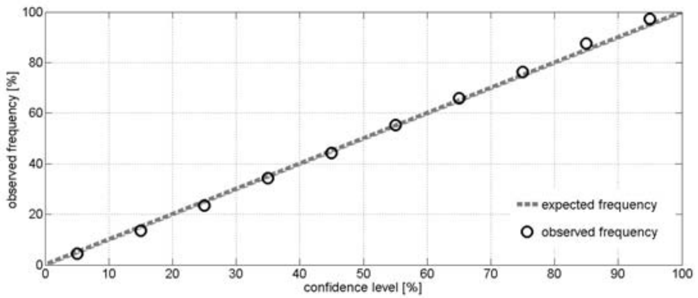

Figure 11: Reliability plot evaluated during the years 2015. Source: EURAC Research.

Figure 12 reports the trend of the prediction interval for five days of February 2015. It can be observed that the width of the interval is reduced when passing from overcast to clear sky days.

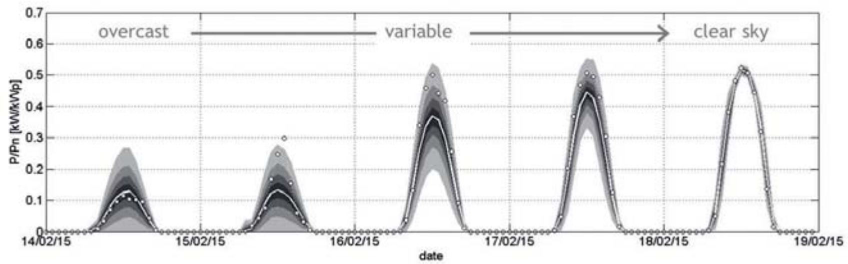

Figure 12: Example of prediction interval trend for five days of February 2015. The grey colours correspond to different confidence levels: 95%-75% -50%-25%; from the clearest grey showing the interval with 95% of confidence to the darker grey showing the interval with 25% of confidence. The dots represent the observed values while the white line is the forecast. Source: EURAC Research.

3.4.4. Impact of high PV penetration on the electric grid and benefits of accurate PV power forecast

The photovoltaic production in the considered region provided in 2015 6.9% of the electric consumption. Since the PV penetration is very similar to the one observed at the national level (7.9%), this is a good case study for analysing the impact of PV generation on the electric grid and the effects of PV power forecast on transmission scheduling and on secondary reserve estimation.

3.4.5. Transmission scheduling

Figure 13 shows how the distributed PV generation can affect the residual load and consequently the secondary reserve, which should be predicted for the day-ahead. The stochastic behaviour of the residual load induced by the irradiance variability introduces an additional error in the prediction of the power that should be supplied by TSOs to DSOs to fulfil the electricity demand of a region.

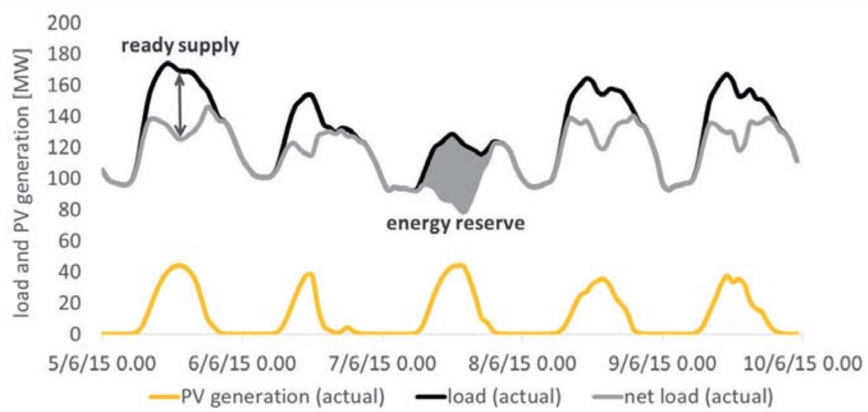

Figure 13: Impact of PV production on the residual load of a distribution network. Source: EURAC Research.

Figure 14 (A) shows the monthly daily average of the actual power transmission that should be provided by the Italian TSO (Terna) to the local area of interest and the expected transmission scheduling with and without considering the PV power forecast. Figure 14 (B) reports the monthly daily average of the absolute energy imbalance ($|transmission^{expected}$ - $transmission^{actual}|$) with and without the PV power forecast.

In 2015, without considering the PV power forecast the absolute energy imbalance was 11% of the power supplied by Terna (654 GWh). It should however be noted that, in the period from April to July around noon, the imbalance is between 60% and 75% of the TSO power supply. This means that in this period, for some hours of the day, the photovoltaic generation can provide from 60% to 75% of the energy that the local DSO should buy from Terna to cover the residual load.

3. Case studies

Thus with 7% of penetration, the distributed PV generation could have a great impact both on the transmission scheduling capability (high energy imbalance) and on the DSO energy needs.

Mid-term PV forecast could greatly reduce this energy imbalance. The day-ahead forecast reduces the absolute imbalance to 2.1% of the DSO energy needs, of which 1.1% is positive (transmission over-estimation) and 0.96% is negative (transmission under-estimation). The two hour ahead forecast brings the absolute imbalance to 1.8%, of which 1.1% is positive and 0.7% is negative. For both forecasts, the maximum imbalance is around 14% and is reached at noon in June.

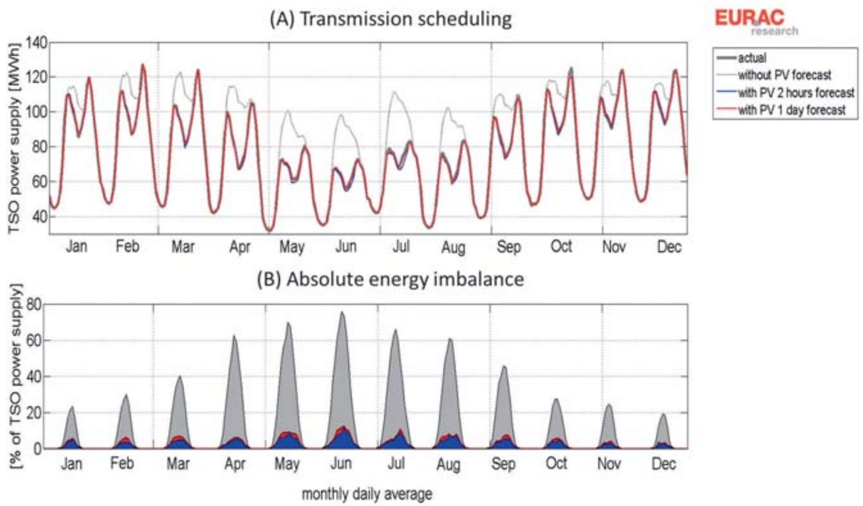

Figure 14: (A) Monthly daily average of the transmission scheduling during the year 2015 with and without considering the PV power forecast; (B) Monthly daily average of the absolute energy imbalance with (red and blue areas) and without (grey area) considering the PV power forecast. Source: EURAC Research

3.4.6. Energy reserve

Not only the PV power forecast but also the prediction of the forecast errors could be very important. Indeed prediction intervals could be used not only to estimate the probability of a specific PV generation bid on the energy market, but also to reduce the energy reserve predicted for the next day.

A conservative algorithm to estimate secondary reserve in a certain area could be built considering that the residual load of the next day will be surely between a minimum: *NetLoad(clear sky)=Load-*PV_{cs} and a maximum: *NetLoad(overcast)=Load*. PV_{cs} is the PV generation during clear sky conditions. It could be easily calculated by rescaling the clear sky global horizontal irradiance so that 95% of the PV production should be below the PV_{cs} curve (see figure 15 A).

A more effective algorithm can be built using the day-ahead prediction interval with a confidence level of 95% defined by the two forecast trajectories: $PV1dF_{low}(95\%)$ and $PV1dF_{up}(95\%)$. $PV1dF_{low}(95\%)$ is the lower forecast and $PV1dF_{up}(95\%)$ is the upper forecast so that the PV production of the day ahead will lie between these two limits with a probability of 95% (see figure 15 A). In this case, the residual load of the next day would be between $NetLoad_{low}(95\%)=Load\text{-}PV1dF_{up}(95\%)$ and $NetLoad_{up}(95\%)=Load\text{-}PV1dF_{low}(95\%)$ with a 95% probability.

Figure 15 shows both the PV production curves and the resulting residual load and energy reserve for five days of 2015. It can be observed that the curve $PV1dF_{low}$ reduces the reserve during clear sky days and variable conditions while $PV1dF_{up}$ reduces the reserve during overcast conditions.

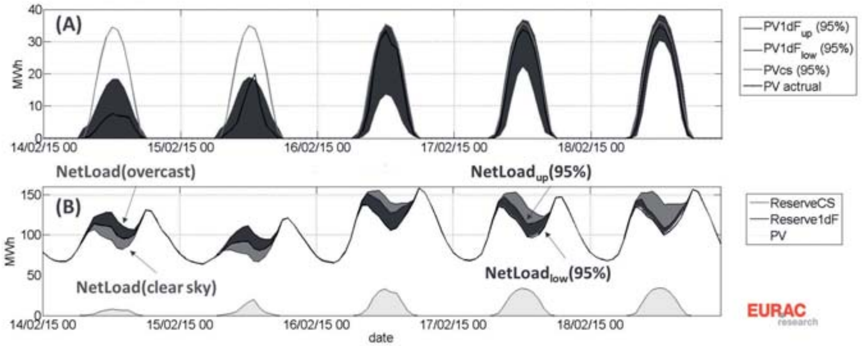

Figure 15: (A) PV production curves for five days of 2015; (B) residual load and energy reserve. The red area is representative of the 95% prediction interval and overlaps the grey area.

Figure 16 shows the monthly daily average of reserves (calculated with the two algorithms) and the PV production in 2015. The energy reserves estimated using the prediction intervals are 36.6% lower than the reserves calculated by the PV_{cs}. It can be noted that the greatest reduction could be achieved from May to September, mainly due to the use of the $PV1dF_{low}$ curve during clear sky days.

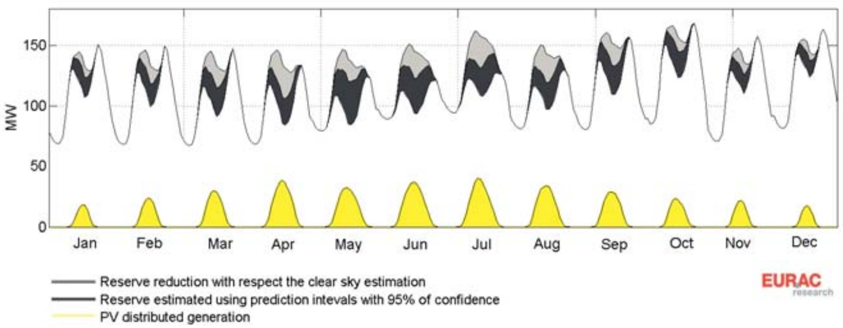

Figure 16: Monthly daily average of reserve and PV production in 2015

3.5. Real-time estimation of PV production for distribution system operations (Cyprus)

Accurate monitoring of PV systems is directly related to the observability of these systems. It is in practice not possible to gather information or monitor all PV systems in an electrical grid due to the continuously increasing number of systems, the huge amount of information and the implied high cost. Therefore, data gathered from a selected number of systems using smart-meters and weather data from meteorological ground stations are used to accurately estimate the output of PV systems. This implies that adequate information about the PV installations is available, such as the capacity, technology, orientation and year of implementation. In practice the findings of the simulations can be compared with real measurements from distribution or transmission substations, allowing to assess the accuracy of the simulation results. The detected errors can then be used to train the models in order to provide more accurate results. Figure 17 shows an outline of the described system.

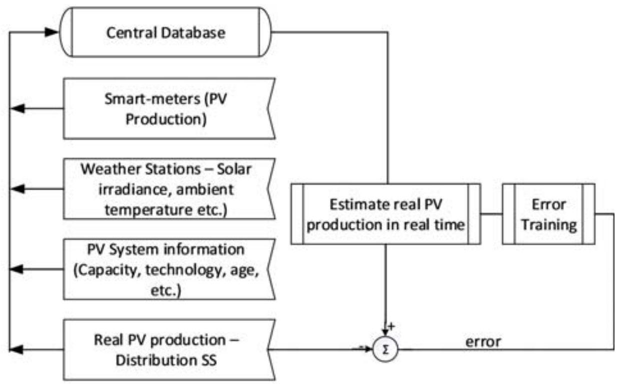

Figure 17: Simplified procedure for estimating in real time the distributed PV production. Source: FOSS, University of Cyprus.

However, the accuracy of the simulations depends heavily on the observability of PV systems. As mentioned above, PV power production depends on weather variations. Consequently, different PV production curve patterns exist due to the different weather conditions of each day. Therefore, in order to increase the observability of PV systems a clustering approach is required in order to distinguish the different patterns of the PV production curve for a region. The clustering of data aims to identify specific characteristics in a dataset and then the grouping of those characteristics into clusters. This process will group similar objects in different clusters. For the analysis presented here, the *K-POP* clustering method [30, 31] is used to classify and characterise the daily solar irradiance of a region into 9 different classes based on the quantity and quality (degree of cloudiness, etc) of solar irradiance.

The *K-POP* method uses two indices to quantify the quantity and quality of the solar irradiance at the earth's surface. The sky clearness index K_d and the probability of persistence POP_d are used respectively. The sky clearness index is the ratio of the global horizontal irradiance (GHI) to the extraterrestrial irradiance E_a as shown in Eq. 2.

$$K_d = \frac{GHI}{E_a} \tag{2}$$

This index captures the instantaneous fluctuations of the solar irradiance and can further be used to calculate the daily solar irradiance that the surface of the Earth receives during a day, K_{day}, depicted in Eq. 3.

$$K_{day} = \frac{\int_{day} GHI\,dt}{\int_{day} E_a dt} \tag{3}$$

However, the sky clearance index cannot capture the quality of solar irradiance, which during a day can be calculated using a probabilistic approach. Firstly, an array, ΔK_{day}, containing the differences between consecutive values of K_d within a day is calculated as depicted in Eq. 4. The POP_{day} index for the day is the probability the values of the ΔK_{day} array to be equal to zero (Eq 5.). Therefore, a high value of POP_{day} for a day demonstrates a low fluctuation probability (low solar irradiance ramp rate during that day).

$$\Delta K_{day} = \{|K_{d_1} - K_{d_2}|, \dots\dots\dots, |K_{d_{n-1}} - K_{d_n}|\} \tag{4}$$

$$POP_{day} = P(\Delta K_{day} = 0) \tag{5}$$

As a result, employing this method for each day yields the daily value of the clearness index and the probability of persistence. The daily solar irradiance can be represented on a two-dimensional plot, where the *x* and *y* axes are the daily values of K_{day} and POP_{day} respectively. The plot of K_{day} vs POP_{day} is divided into 9 classes as demonstrated in Figure 18 (a). The *x*-axis is divided into three sections based on the quantity of solar irradiance. The right column represents days with high solar quantity, the centre column days with medium solar quantity and the left column days with low solar quantity. Similarly, the *y*-axis representing the quality of solar irradiance is divided into 3 sections based on the quality of the daily sky conditions: clear or totally overcast sky (top row), relatively small and infrequent fluctuations (centre row) and large and frequent fluctuations (bottom row). Solar irradiance example profiles are shown in the cell of each class in (b).

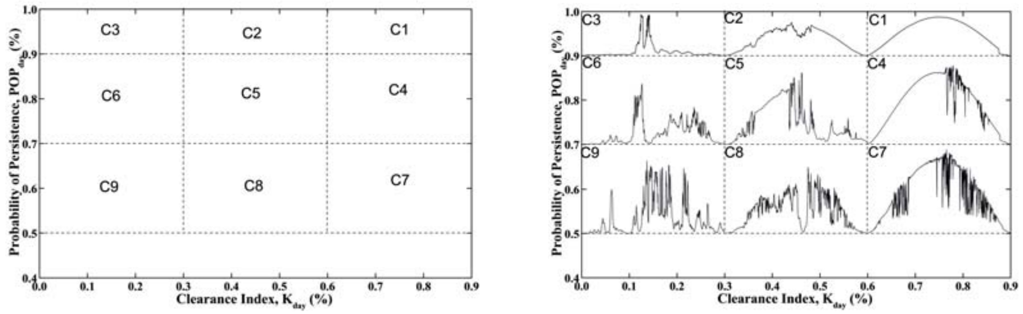

Figure 18: The daily solar irradiance classes (a) and examples of the daily solar irradiance plots (b) for each class [32]

For the above presented analysis four years of global horizontal irradiance (GHI) data with a resolution of 1 minute are used. The data was collected from 2011 to 2014 from a weather station located in the area of Akrotiri, Limassol, Cyprus, which is operated by the FOSS Research Centre of the University of Cyprus. The solar irradiance was measured using a "Kipp Zonen CMP 6" pyrometer. The extraterrestrial irradiance data was simulated using the online "Solar Position and Intensity Calculator" tool of NREL [33].

The solar irradiance patterns for the *K-POP* plots for the data collected in Cyprus are presented in Figure 19. The rectangular symbols represent the *K-POP* data points for each day of the year and the larger circular symbols indicate the centroid of the *K-POP* points of each year. Each colour represents a different year as indicated in the legend of the plot. A visual inspection of the results clearly shows that the distribution patterns of the data points throughout the years are very similar. This is an indication of the consistency of the yearly solar irradiance patterns.

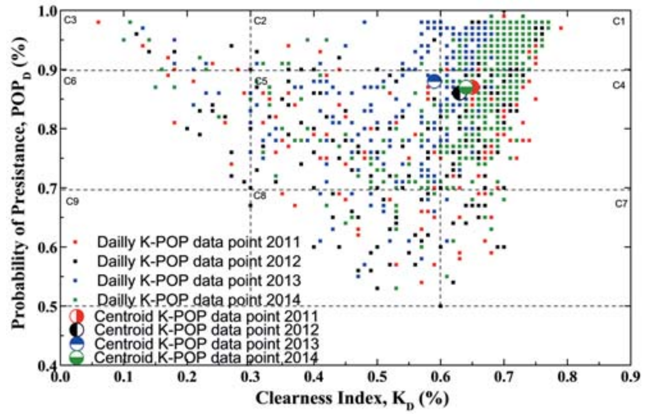

Figure 19: Daily solar irradiance distribution in Cyprus for all years of investigation. Source: FOSS, University of Cyprus

The consistency of the results through the years can be confirmed also from the results presented in Table 7. The average yearly centroid is around 0.87 for the and 0.63 for the indices, showing that the distribution patterns of the data points exhibit high correlation, a characteristic which enhances observability.

Table 7: Daily solar irradiance percentile distribution into the 9 classes and statistical average of the K-POP data points for each year. Source: FOSS, University of Cyprus

		Year				Average
		2011	2012	2013	2014	
Clases (%)	1	45.8	49.3	51.2	47.1	48.4
	2	1.1	2.2	6.1	1.6	2.8
	3	1.6	0.5	1.5	1.6	1.3
	4	30.4	22.7	12.5	30.7	24.1
	5	11.8	14.2	22.7	13.4	15.5
	6	1.6	2.2	0.9	0.8	1.4
	7	2.7	3.8	0.6	2.2	2.3
	8	4.9	4.9	4.7	2.5	4.3
	9	0	0	0	0	0
Centroid	K_{day}	0.65	0.63	0.59	0.64	0.63
Centroid	POP_{day}	0.87	0.86	0.88	0.87	0.87

The clustering of solar irradiance in Cyprus therefore revealed that grid connected solar systems in Cyprus are highly predictable and hence observable. As a result they can potentially deliver quality energy to the grid, without compromising its operational reliability. These findings lead to the conclusion that similar levels of reserves (primary, secondary and tertiary) that are currently enforced by national grid rules for systems with conventional generators can adequately offer similar levels of grid reliability and continuity when conventional generator capacity is replaced by equivalent solar generators [32].

4. CONCLUSION

Forecasting and observability of intermittent generation is becoming a critical requirement in the emerging energy mix. The nature of the emerging distributed technologies requires a fundamental change in the planning, development and operation of the integrated grid. This is the reality that affects the electricity grids through the high PV generation that introduces a stochastic variability dependent on meteorological conditions. Thus, a large share of PV power introduces new challenges for the stability of the electrical grid, both at the local and national level, requiring the need of revised reserve policy (more distributed than central) and utilization complementing technology solutions (flexible generation, flexible demand response, storage etc) to ensure electrical balancing and overcome the unpredictability and variability of demand and intermittent generation. PV power forecasts could mitigate the effects of high solar power injection into the electricity grid, both for grid management and on the energy market. This was well addressed in this white paper revealing the needs and practices for forecasts on all time horizons, and especially short-term forecasts (intra hours) and mid-term forecasts (intra-day and day-ahead).

The paper identifies that dealing with forecasting in the power system is of paramount importance since it affects the day ahead, hour to hour and minute to minute operation of electric grids with real substantial costs without adequate accuracy. Balance groups are playing a leading role in this direction. They can include generation and consumption units or be "virtual", when operated by financial actors who only trade. All balance groups report to a balancing authority, which in Europe at national level is generally the transmission system operator (TSO). This authority ensures that trades on the electricity market are balanced i.e., that contracted generation matches forecast consumption. Balance Group Managers (BGMs) are responsible to ensure that at each time step of market operations their contracted production and/or consumption matches the realised values. In case of mismatch between prediction and realisation, BGMs are penalised based on intraday market price; if the imbalance is in the same direction as the whole system (e.g., a producer under-delivering when there is a shortage in production), the penalty will be above the intraday market price and if the imbalance is in the opposite direction the penalty will be below. Hence, it is of critical importance that forecasting tools and methods are reliable and adaptive.

The paper goes on to address real use cases for all related stakeholders that include among others:

- Investors in solar PV power plants
- Operators of PV power plants
- Grid operators
- Electricity retailers/aggregators
- Balance group managers
- Balancing authorities

A forecasting tool requires capturing all parameters that have an influence in the delivered energy of systems. For this reason the paper gives proper attention to these parameters and identifies their relativeness to the accuracy of results. The annual production is usually calculated through formulas with different levels of complexity varying from fixed efficiency equations to equations which account for second order effects and derating (e.g. various losses during the energy conversion value chain). The current status is clarified by classifying the existing PV forecasting techniques into either physical or statistical methods. Physical methods use solar and PV models to generate PV forecasts, while statistical methods use past data combined with autoregressive or artificial intelligent models to forecast the PV output. A comparison of the two methods has shown that the statistical method slightly outperformed the physical. However, in practice these two methods are often combined.

Another approach for developing a solar PV forecasting model based on neural networks is presented in the paper, giving the design details which will allow the basic tool to be integrated in a utility software.

The paper presents in detail a forecasting tool developed for the needs of a DSO which shows promising results for all stakeholders. This was further elaborated through the presented results of the MIRABEL project, which addressed the optimality criteria to improve forecasting adequacy in different use cases. The analysis revealed the importance of input data characteristics and compared typical energy forecasting models. A brief overview of the applied method and results were presented.

The paper gives more information through another case study in a project carried out in collaboration with a local DSO, by utilizing ANNs to estimate and forecast the distributed generation of 1975 PV plants in a small part of the South Tyrol Region in Italy.

Similarly the paper addresses the current status of observability techniques based on the fact that accurate monitoring of PV systems is directly related to the observability of these systems. It is in practice not possible to gather information or monitor all PV systems in an electrical grid due to the continuously increasing number of systems, the huge amount of information and the implied high costs. Therefore, data gathered from a selected number of systems using smart-meters and weather data from meteorological ground stations are used to accurately estimate the output of PV systems. Moreover, the clustering of solar irradiance improves predictability and the methodology was adequately presented in the paper with examples from real systems. As is shown in the paper these findings lead to the conclusion that levels of reserves (primary, secondary and tertiary) similar to those currently enforced by national grid rules for systems with conventional generators can adequately offer similar levels of grid reliability and continuity when conventional generator capacity is replaced by equivalent solar generators.

Future steps

We believe that this paper is an important step towards more clarity on targets and objectives of improved observability and forecasting. It should be built upon to create a technology roadmap that will help guiding R&D efforts. To define such a roadmap, critical questions will need to be addressed:

- At present PV generators, balance group managers, distribution system operators and transmission system operators all have to provide some PV production forecasts or to include them in their net consumption forecasts, but in many cases the specifications of these forecasts are unclear. Whose responsibility will it be in the future to provide forecasts for (distributed) PV generation? What will be the incentives and penalties?
- Will the development of forecasting be driven by regulations (e.g., mandates from TSOs) or by market mechanisms?
- How will costs and deployment of energy storage technologies evolve, and how will that affect the need for power forecasting?
- Can the provision of confidence intervals with forecasts reduce the needs for reserve generation?

Based on our analysis, we can already point out that:

- Improvements in forecasts and in observability of both weather data and electrical quantities in the grid at a higher granularity are interdependent.
- Control logics (e.g., curtailment) that include forecast also need an improvement in observability.
- Common metrics need to be standardised, including the integration time and the normalization factors, so that the outcome of various methodologies (at single site and at regional level) can be quantitatively compared.
- A common benchmark (e.g. RMSE of persistence model) for both GHI and production forecasts (e.g. at optimal tilt angle and azimuth, similar to PVGIS) should be mapped across the EU as a reference.

5. LIST OF ACRONYMS

PV	Photovoltaics
IEA	International Energy Agency
WWS	Wind, Water (hydro) and Solar energy
DSO	Distribution system operator
TSO	Transmission system operator
BG	Balance group
BGM	Balance group manager
ETIP PV	European Technology & Innovation Platform for Photovoltaics
AC	Alternating current
O&M	Operation & maintenance
MV	Medium voltage
NWP	Numerical weather prediction
MOS	Model output statistics
MBE	Mean bias error
MAE	Mean absolute error
RMSE	Root-mean-square error
GHI	Global horizontal irradiance
RES	Renewable energy sources
DB	Database
EU	European Union
ANN	Artificial neural network
MLP	Multi-layer perceptron
GUI	Graphical user interface
POA	Plane of (solar PV) array

6. LIST OF SYMBOLS

G_{POA}	Global solar irradiance on PV plane	W/m²
GHI	Global horizontal irradiance	W/m²
E_a	Extraterrestrial irradiance	W/m²
V	Wind speed	m/s
S	Surface area of a PV power plant	m²
N	Number of aggregated PV power plants	-
N_{sun}	Number of sun hours	h
P_n	Nominal power output of a PV power plant	MW
P_{ac}	Instant AC power output of a PV plant	MW
$Y_{forecast}$	Forecast PV production	kWh
$Y_{realised}$	Actual PV production	kWh
T_{amb}	Ambient temperature	°C
K_d	Instantaneous sky clearness index	-
K_{day}	Daily sky clearness index	-
POP_d	Instantaneous probability of persistence	-
POP_{day}	Daily probability of persistence	-
PV_{cs}	PV generation during clear sky conditions	MW

7. References

7. REFERENCES

[1] IEA, “Technology Roadmap, Solar Photovoltaic Energy, 2014 Edition,” 2014.

[2] Mark Z. Jacobson et al., “100% Clean and Renewable Wind, Water, and Sunlight (WWS) All-Sector Energy Roadmaps for 139 Countries of the World,” Oct. 2016.

[3] N. Maluf, “Qnovo blog article: What is the interest in energy storage?,” 13 Jun. 2016. [Online]. Available: http://qnovo.com/94-interest-energy-storage/.

[4] P.-J. Alet et al, in 31st European PV Solar Energy Conference and Exhibition, 2015.

[5] S. Pelland et al, “Photovoltaic and Solar Forecasting: State of the Art,” IEA, 2013.

[6] E. Lorenz et al., “31st European Photovoltaic Solar Energy Conference and Exhibition,” 2015.

[7] Black & Veatch, “Cost and performance data for power generation technologies,” National Renewable Energy Laboratory, 2012.

[8] J. A. G. B. A. Grandjean, Renewable and Sustainable Energy Reviews, p. 16, 2012.

[9] J. Marcos et al., in 31st European Photovoltaic Solar Energy Conference and Exhibition.

[10] J. Munkhammar et al., in 31st European Photovoltaic Solar Energy Conference and Exhibition.

[11] TenneT TSO GmbH, „TenneT TSO,“ 2016.

[12] Autorità per l’energia elettrica il gas e il sistema idrico, “Disposizioni in materia di dispacciamento delle fonti rinnovabili non programmabili a seguito della sentenza del Consiglio di Stato,” 2014.

[13] Elexon, “BM Reports”.

[14] RTE, “Historique des tendances et des prix de règlements des écarts validés,” RTE Réseau de transport d’électricité, 2016.

[15] SDE TF, “Key support elements of RES in Europe: moving towards market integration,” Council of European Energy Regulators (CEER), 2016.

[16] M. Pierro et al., Solar Energy, p. 134, 2016.

[17] E. Lorenz et al., in 23rd European Photovoltaic Solar Energy Conference and Exhibition, 2008.

[18] R. Perez et al., Solar Energy, no. 305, p. 94, 2013.

[19] M. Pierro et al., Solar Energy, Nr. 99, p. 117, 2015.

[20] T. Cebecauer, M. Šúri, R. Perez, 2010.

[21] R. Perez et al., in ASES Annual Conference, 2010.

[22] J. Freeman et al., in IEEE 40th Photovoltaic Specialist Conference (PVSC), 2014.

[23] “MIRABEL project website,” [Online]. Available: http://mirabel-project.eu/ .

[24] Leading partner Technische Universität Dresden- TUD, “D4.1: State-of-the-Art Report on Forecasting,” MIRABEL, 2010.

[25] Björn Wolff et al., “Comparing support vector regression for PV power forecasting to a physical modeling approach using measurement, numerical weather prediction, and cloud motion data,” Solar Energy, no. 135, pp. 197-208, 2016.

[26] E. Lorenz et al., “Regional PV power prediction for improved grid integration” Progress in Photovoltaics: Research and Applications, no. 19, pp. 757-771, 2011.

[27] J.G.S Fonseca Jr et al., “Regional forecasts of photovoltaic power generation according to different data availability scenarios: a study of four methods” Progress in Photovoltaics: Research and Applications, no. 23, pp. 1203-1218, 2015.

[28] J.G.S Fonseca Jr et al., “Regional forecasts and smoothing effect of photovoltaic power generation in Japan: An approach with principal component analysis,” Renewable Energy, no. 68, pp. 403-413, 2014.

[29] M. Zamo et al., “benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production, part I: Deterministic forecast of hourly production” Solar Energy, no. 105, pp. 792-803, 2014.

[30] B. O. Kang and K.-S. Tam, “New and improved methods to estimate day-ahead quantity and quality of solar irradiance,” Applied Energy, no. 137, pp. 240-249, 2015.

[31] B. O. Kang and K.-S. Tam, “A new characterization and classification method for daily sky conditions based on ground-based solar irradiance measurement data,” Solar Energy, no. 94, pp. 102-118, 2013.

[32] I Koumparou et al., “Characterization and Classification of Daily Sky Conditions in Cyprus and France Based on Ground Measurements of Solar Irradiance,” in 31st European Photovoltaic Solar Energy Conference and Exhibition, 2015.

[33] National Renewable Energy Laboratory, “SOLAR and LUNAR POSITION CALCULATORS,” 2016. [Online]. Available: http://www.nrel.gov/midc/solpos/.


Saturday, 18 February 2017
07:25:02
50.05
SYSTEM DATA
50.05 Hz
492 MW
-44 MVAr
SYSTEM FREQUENCY
50.05 Hz

PhotoVoltaic
EUROPEAN TECHNOLOGY
& INNOVATION PLATFORM

www.etip-pv.eu

April 2017