Reserve Setting and Steady-State Security Assessment Using Wind Power Uncertainty Forecast: a Case Study

R.J. Bessa, M.A. Matos, *Member, IEEE*, I.C. Costa, L. Bremermann, I.G. Franchin, R. Pestana, N. Machado, H-P. Waldl, C. Wichmann

***Abstract*--This paper reports results and an evaluation methodology from two new decision-aid tools that were demonstrated at a Transmission System Operator (REN, Portugal) during several months in the framework of the EU project Anemos.plus. The first tool is a probabilistic method intended to support the definition of the operating reserve requirements. The second is a fuzzy power flow tool that identifies possible congestion situations and voltage violations in the transmission network. Both tools use as input probabilistic wind power predictions.**

***Index Terms*—Wind power, reserve, probabilistic forecasts, risk, fuzzy power flow, system operator, congestion.**

I. Introduction

THE integration of electrical energy from renewable sources, in particular wind energy, into the power system has motivated research of new algorithms for power system planning and operation. A comprehensive review of computational tools for analyzing the integration of renewable energy can be found in [1].

Furthermore, in the last twenty years, research has been conducted for developing wind power forecasting (WPF) algorithms. Recent and comprehensive literature reviews on WPF can be found in [2] and [3]. Several authors explored the idea of including information from Numerical Weather Prediction (NWP) systems in WPF algorithms, aiming to produce forecasts with acceptable accuracy several days-ahead [4][5].

The subsequent step was to develop algorithms for wind power uncertainty forecast. Two relevant publications on statistical algorithms are the time-adaptive quantile regression [6] and the adapted resampling [7]. In parallel, NWP systems were also explored to produce uncertainty forecasts represented by ensemble of meteorological predictions [8].

Presently, one of the new research priorities is a ful integration of wind power forecasts and information on their uncertainties in the algorithms of the power system operational management tools. Some examples of related work consist of setting the power system operating reserves [9][10], probabilistic power flow [11] and stochastic unit commitment [12][13].

Aiming to establish a synergy between wind power forecast and power system management, this paper presents two decision-aid tools that were installed and demonstrated for the Portuguese Transmission System Operator (Redes Energéticas Nacionais, REN) in the framework of a demonstration task of the ANEMOS.plus project. Both tools were compared with the current procedures followed by the TSO. Results from other demonstration cases of the project can be found in [14].

The first tool is a probabilistic reserve setting tool [10] intended to support the transmission system operator (TSO) in defining the operating reserve requirements for the next day or hours.

For generation adequacy assessment and operating reserve planning, probabilistic methods are already commercially available and commonly used by TSO. One example is the chronological Monte Carlo simulation described in [15]. Regarding the operational domain (topic addressed in this paper), in a recent past, a large number of TSO defined the operating reserve requirements based on deterministic rules such as a percentage of the load and/plus the largest unit in the system [16]. Presently, with the increasing penetration of wind power, TSO are starting to account WPF errors when setting the operating reserve. For instance, REE (Spanish TSO) started to study the inclusion of the load and wind forecast error distributions in setting the operating reserve requirements [17]; ERCOT (Texas Independent System Operator) is also changing to a probabilistic rule [18].

The algorithm used in this paper was already presented in a previous paper [10], so the contributions from this paper are: details and results from an operational demonstration for the Portuguese TSO; methodology for evaluating the results quality, as well as the *value* (benefits) of the wind power

R.J. Bessa, M.A. Matos, I.C. Costa, L. Bremermann and I.G. Franchin are with INESC TEC - INESC Technology and Science (formerly INESC Porto) and FEUP - Faculty of Engineering, University of Porto, Portugal (emails: rbessa@inescporto.pt, mam@fe.up.pt, ivo.c.costa@inescporto.pt, leb@inescporto.pt, ivan.franchin@inescporto.pt). R. Pestana and N. Machado are with REN, Redes Energéticas Nacionais, Portugal (e-mails: rui.pestana@ren.pt, nelio.machado@ren.pt). H-P. Waldl and C. Wichmann are with Overspeed GmbH & Co. KG, Germany (e-mails: h.p.waldl@overspeed.de, c.wichmann@overspeed.de).

This work was performed in the framework of the project ANEMOS.plus (contract no 038692) funded in part by the European Commission under the 6th RTD Framework Programme. The work of R.J. Bessa and L. Bremermann was also supported by Fundação para a Ciência e Tecnologia (FCT) Ph.D. Scholarships SFRH/BD/33738/2009 and SFRH/BD/65355/2009.

uncertainty forecasts.

The second tool is a fuzzy power flow (PF) [19] intended to support day-ahead network constraints management by identifying possible branch congestions and voltage limit violations.

Presently in order to ensure a feasible generation program (i.e. security constrained dispatch), the TSO uses deterministic power flow with wind power point forecasts for detecting possible congestion and voltage. So far, in the literature, the fuzzy PF has been used for planning studies [20], thus this paper reports the first combination of fuzzy PF with wind power uncertainty forecasts in an operational demonstration for a TSO. A further contribution is an evaluation framework for the results quality. Note that, in what concerns probabilistic PF, some authors already extended the algorithm to include wind power uncertainty forecasts [11].

The paper is organized as follows: sections 2 and 3 describe the operational implementation of the probabilistic reserve setting and fuzzy power flow tools, respectively; section 4 describes an evaluation methodology for both tools; section 5 presents the results from the demonstration; section 6 summarizes the conclusions.

II. ANEMOS Wind Power Forecasts

The probabilistic wind power predictions used as input to both tools were generated with the Anemos wind power forecasting system [21]. Anemos is a high-availability platform for prediction services, hosting a variety of specialized prediction modules for particular purposes [22].

For the REN application, NWP from GFS (NOAA, USA) and the Skiron model (CENER, Spain) with a higher horizontal resolution were used. On base of these NWP data streams, a statistical (WPPT [23], ENFOR, Denmark) and a physical/hybrid forecast model (ROGI, Overspeed, Germany) generated deterministic wind power predictions (WPP) twice a day with Skiron and four times per day with GFS. Both models were tuned with the help of historic and on-line wind power SCADA data available for the different nodes of the REN grid.

Using a historical analysis of WPP and realized wind power production, a statistical post-processor model, the so called uncertainty model [7] (ARMINES, France), has been tuned. As final step of the modeling chain, this module generates probabilistic wind power predictions from the deterministic forecasts and the historic tuning. The forecast is delivered to the two tools as quantiles of wind power from 0 to 100 % with 5 % steps with a prediction horizon from 6 to 48 hours for Skiron and from 6 hours to one week with GFS.

III. Robust Reserve Setting (RRS) Tool

A. Methodology

The Robust Reserve Setting (RRS) tool is a probabilistic approach to build the system generation margin distribution (amount of available generating capacity that exceeds the system load) that results from convolving the probability mass functions associated to generation (conventional and wind power) and load.

Risk indices describing the possible consequences of each downward and upward operating reserve level are computed from the system generation margin distribution. Risk in this problem is represented by the probability of failing to satisfy the load in the short-term, as the result of loss of generation due to forced outages, load and wind power forecast errors.

Two risk indices, related to the surplus and shortage of generation in the system, and meaningful for the TSO, are computed from the system generation margin distribution. The loss of load probability (LOLP) for the upward reserve given by

$$\sum_{(m+R_{UP})\leq 0} f_M\left(m+R_{UP}\right) \quad (1)$$

where f_M is the probability mass function of the margin M (generation minus load) and R_{UP} is the upward reserve level; and, the probability of wasting renewable energy (PWRE) for the downward reserve given by

$$\sum_{(m-R_{DOWN})>0} f_M\left(m-R_{DOWN}\right) \quad (2)$$

where R_{DOWN} is the downward reserve level.

The next step is the determination of the risk/reserve curve, as a basis to the application of decision-making methods incorporating the preferences of the decision maker (in this case the TSO).

The decision-making method adopted for this paper consists in setting a reference threshold for the maximum acceptable risk. Note that more elaborated decision-making methods to balance risk against reserve cost and incorporating more complex structures of the decision maker preferences could be used (see [10] for details).

The output is the downward and upward operating reserve hourly levels determined at a specific time instant for each hour of the programming period.

B. Operational Implementation

In Portugal and many other countries, the TSO, after the day-ahead market settlement, determines the secondary reserve levels for the next day and receives secondary reserve bids from the market agents. Note that this is a sequential market where the reserve levels are defined after the electrical energy market.

The secondary reserve margin remains fixed during the whole operational period, and it is mobilized when it is necessary to bring back the frequency and the interchange programs to their steady-state values [24].

After each intraday session (and six times per day) the TSO defines the tertiary reserve (also named regulation reserve) levels for the remaining hours of the operational period, and receives tertiary reserve bids. The tertiary reserve covers the loss of a generation unit or wind power and load forecast errors.

The tool was installed operationally at the TSO. It provides suggestions of the total reserve (secondary plus tertiary) for the day-ahead market session and of tertiary reserve for the intraday market sessions. The output was produced after each

market session.

Fig. 1 depicts the tool's inputs and outputs. The tool has the following inputs and corresponding update frequency: forecasts for the total load updated seven times per day; upscaled (up to 3.8 GW) wind power probabilistic forecasts from the ANEMOS forecasting system; conventional generation dispatched by the market pool in the day-ahead and six intraday sessions; forecasts for the non-wind generation remunerated with feed-in tariffs (e.g. mini-hydro, PV, CHP) updated seven times per day; interconnection levels between Portugal and Spain. In this demonstration, both wind power forecasts from GFS and Skiron were considered and compared.

Wind power uncertainty is modeled through probabilistic forecasts represented by a set of quantiles.

Load uncertainty is characterized through a Gaussian distribution with a given standard deviation and zero mean. The conventional generation uncertainty is characterized by the capacity outage probability table, including information from the unit's outage replacement rate [10].

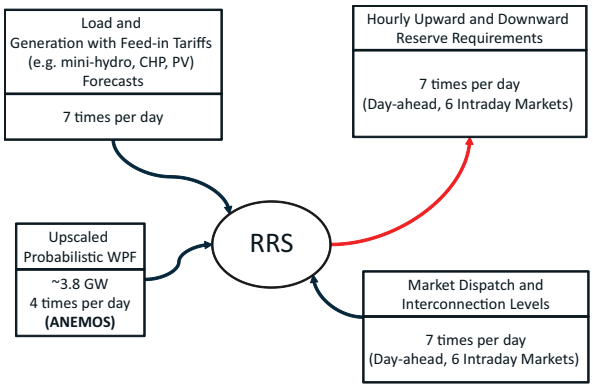

Fig. 1. Operational implementation of the RRS tool.

The interconnection levels and the non-wind generation remunerated with feed-in tariffs were assumed to be deterministic variables. However, no significant changes are necessary in the tool for incorporating the uncertainty of these two variables.

The outputs are the upward and downward reserve levels for the day-ahead and intraday sessions obtained by setting a reference threshold for the risk. For demonstration purposes, three different values for the LOLP and PWRE were considered: 0.1%, 0.5% and 1%.

IV. Fuzzy Power Flow (PF) Tool

A. Methodology

Fuzzy PF allows the inclusion of generation and load uncertainty, incorporated by fuzzy numbers (e.g. triangular).

First, the deterministic values (the central value of the fuzzy number) are calculated by solving an AC power flow with the Newthon-Raphson method, considering a deterministic set of data. The algorithm takes as input the central values of the input data and delivers as output the central values of the power flow results.

Computation of the deviations (ΔV and $\Delta\theta$) regarding the central values (V^{ctr} and θ^{ctr}) results from the propagation of the uncertainty described by the fuzzy numbers for active (P) and reactive (Q) injections. A similar process leads to the characterization of the uncertainty in the branch power flows. The following paragraphs describe the fuzzy PF algorithm steps.

After computing the AC power flow for the central values, $\Delta\tilde{\theta}$ and $\Delta\tilde{V}$ deviations are calculated using Eq. 3 with the inverted Jacobian matrix $[J]^{-1}$, using fuzzy arithmetic. Adding these deviations to the crisp power flow values, the fuzzy results $\tilde{\theta}$ and $\tilde{V}$ are computed with Eq. 4 and 5.

$$\begin{bmatrix}\Delta\tilde{\theta}\\ \Delta\tilde{V}\end{bmatrix} = [J]^{-1}\cdot\begin{bmatrix}\Delta\tilde{P}\\ \Delta\tilde{Q}\end{bmatrix} \quad (3)$$

$$\left[\tilde{\theta}\right] = \left[\theta^{ctr}\right] + \left[\Delta\tilde{\theta}\right] \quad (4)$$

$$\left[\tilde{V}\right] = \left[V^{ctr}\right] + \left[\Delta\tilde{V}\right] \quad (5)$$

Active and reactive branch flows, losses and branch currents are non-linear functions of voltage magnitudes and phases at the extreme nodes of each branch. However, their deviations can be approximately linearized using Taylor Series expansion.

For example, the active power flow P_{ik} between nodes i and k can be expressed by Eq. 6, and its deviation is given by Eq. 7, where g_{ik} and b_{ik} are the conductance and susceptance of branch i-k and θ_{ik} is the phase difference between i and k.

$$P_{ik} = -g_{ik}.V_i^2 + V_iV_k\left(g_{ik}\cos\theta_{ik} + b_{ik}\sin\theta_{ik}\right) \quad (6)$$

$$\Delta\tilde{P}_{ik} \cong \frac{\partial P_{ik}}{\partial V_i}\cdot\Delta\tilde{V}_i + \frac{\partial P_{ik}}{\partial V_k}\cdot\Delta\tilde{V}_k + \frac{\partial P_{ik}}{\partial\theta_i}\cdot\Delta\tilde{\theta}_i + \frac{\partial P_{ik}}{\partial\theta_k}\cdot\Delta\tilde{\theta}_k \quad (7)$$

Substituting the $\Delta\tilde{V}_i$, $\Delta\tilde{V}_k$, $\Delta\tilde{\theta}_i$ and $\Delta\tilde{\theta}_k$ in Eq. 7 by their linear expressions obtained directly from Eq. 3, the active power flow deviation is obtained from Eq. 8, where $SP_{ik,j}$ and $SQ_{ik,j}$ represent the sensitivity coefficients relating the power flow branch i-k with active and reactive injections in node j; the fuzzy number for the active power flow is given by Eq. 9.

$$\Delta\tilde{P}_{ik} = \sum\nolimits_j SP_{ik,j}\cdot\Delta\tilde{P}_j + \sum\nolimits_j SQ_{ik,j}\cdot\Delta\tilde{Q}_j \quad (8)$$

$$\tilde{P}_{ik} = P_{ik}^{ctr} + \Delta\tilde{P}_{ik} \quad (9)$$

A similar non-iterative process is followed to calculate the fuzzy values of the active generation in the slack bus, as well as the reactive generation in PV nodes. A detailed explanation of the mathematical model can be found in [19].

The outputs are the traditional ones (e.g. active and reactive power flows), but they are fuzzy numbers that incorporate the uncertainty of the wind power probabilistic forecasts.

B. Operational Implementation

In Portugal, the TSO after the day-ahead market pool performs a validation of the market settlement to detect congestion and voltage violations. In case of an unfeasible market dispatch, the TSO performs adjustments in the generation dispatch, in order to obtain a technically feasible

dispatch.

In order to include the influence of wind power uncertainty in this analysis, triangular fuzzy numbers built after the probabilistic forecasts were used. The lower limit is the 10% quantile, the central point is the point forecast value, and the upper limit is the 90% quantile. Approaches that are more complex could be adopted [25], but for this end-user, the three-point representation of uncertainty was found to be more suggestive. The tool provides the standard outputs of the power flow, but in the form of triangular fuzzy numbers. This information is then used to detect *possible risk situations*, not detectable by the deterministic PF. The tool provides an index $\alpha_{congest}$ measuring the possibility of congested branch and another index $\alpha_{voltage}$ for the possibility of voltage violation.

The two indices provide alarms on possible constraints violations, and correspond to the maximum value of possibility for which there is no congestion or voltage violation. A zero degree means no congestion/voltage deviation. A degree equal to one means that even the deterministic value (fuzzy number midpoint) is greater than the limit.

The tool was installed operationally at REN and provides fuzzy power flow results for the day-ahead market session. It runs around 2 PM for the day-ahead and a fuzzy PF is performed for each hourly interval of a 24 hours period.

Fig. 2 depicts the tool's inputs and outputs. The input data consists of a set of standard network parameters of a classical power flow (e.g. nodes type, specific voltages); forecasted load in each network node; conventional generation dispatched by the market pool; point and probabilistic wind power forecasts from GFS launched at 6 AM for 18 individual and 6 network nodes (giving a total of 39 wind farms).

The load and other generation sources are represented by their point forecast, but a fuzzy number representation could also be adopted.

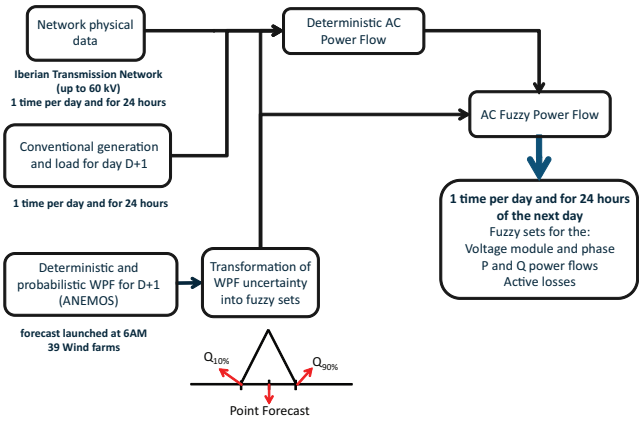

Fig. 2. Operational implementation of the Fuzzy PF tool.

V. Evaluation Methodology

A. Reserves Setting Problem

1) Rules for Comparison

The reserve that deals with unplanned unit's outages, wind power and load forecast error is the tertiary reserve (or regulation reserve). Thus, in this paper only the tertiary reserve suggestions are evaluated.

The tool will be compared with the following deterministic rules for tertiary reserve:

- rule 1 (presently used in Portugal): upward reserve equal to 2% of the forecasted load plus the largest unit in the system (435 MW); downward reserve equal to 2% of the forecasted load;
- rule 2: rule 1 plus 20% of the wind power point forecast for both upward and downward reserve.

2) Criteria for Evaluation

The tertiary reserve level suggested by the tool is compared with the mobilized tertiary reserve in each hour. Since for each intraday session a different tertiary reserve suggestion is provided and there is an overlapping of hours, only the more recent values are used for comparison. For example, intraday session number 2 covers all the 24 hours of the next day, and intraday session 3 covers only the period between 5 AM and 12 PM, so the suggested reserve level evaluated for hour 5 AM corresponds to the one estimated for intraday session 3.

The first requirement for a probabilistic reserve setting tool is to have a complete match between the predefined risk threshold value β and the observed number of insufficient reserve situations (i.e. mobilized reserve greater than suggested reserve) for an evaluation period.

First, it is necessary to define an index on whether or not the suggested reserve R_t^{β} for hour t and risk threshold β is lower than the mobilized reserve R_t^m,for both upward and downward reserves:

$$\theta_t^{\beta} = \begin{cases} 1 & if \quad R_t^m > R_t^{\beta} \\ 0 & otherwise \end{cases} \tag{10}$$

Then, the number of insufficient reserve situations, which can be computed for both upward and downward directions, is given by:

$$\hat{\beta} = \sum_{t=1}^{N} \theta_t^{\beta} \Big/ N \tag{11}$$

where N is the number of hours in the evaluation dataset.

For the deterministic rules, this gives an idea of the embedded risk.

In addition to a single number for $\hat{\beta}$, 95% confidence intervals computed with non-parametric bootstrapping are also presented. The bootstrapping method consists in drawing with replacement from the original set of pair samples with mobilized and suggested reserve. The empirical distribution function is used in the sampling process, and places a probability of *1/n* in each sample of the original dataset. The statistic of Eq. 11 is computed for each resample dataset (with the same size of the original dataset). This process is repeated B times (300 in this paper) for getting an estimate of the distribution of β. Then, the confidence intervals are constructed using the "basic" bootstrap confidence limits method [26].

If $\hat{\beta}$ is higher than β, it means that the reserve values are not sufficient for ensuring the adopted risk level. On other hand,

if $\hat{\beta}$ is lower than β, this would mean higher reserve costs than necessary.

Thus, for a complete evaluation of the suggested reserve quality it is necessary to measure the amount of suggested reserve and the ability to provide different reserve levels for different conditions of the power system operation.

Following this objective, the quantiles 0.1% and 99.9% are computed from the distribution of the difference between suggested and mobilized reserve.

In the literature for evaluating quantile forecasts, several authors use scoring metrics encompassing three important characteristics: calibration, sharpness and resolution [27]. In the reserve setting problem, calibration is the match between β and $\hat{\beta}$. Sharpness and resolution are related with the amount of suggested reserve and the ability to provide different reserve levels for different conditions of the power system operation. Reserve suggestions almost flat along the whole day are expected to have a bad performance in terms of resolution.

The scoring metric used in [27] for quantiles evaluation was borrowed for comparing reserve's suggestions obtained from different forecasts (GFS and Skiron):

$$S_t^{\beta} = \left(\theta_k^{\beta_i} - \beta\right) \cdot \left(R_t^{\beta} - R_t^{m}\right) \quad (12)$$

The maximum value is zero, and less negative values mean better performance.

The average of the scoring rule is taken for the whole evaluation period. Note that the scoring rule can only be applied to probabilistic reserve setting approaches because the deterministic rules do not have an associated risk value β.

For comparing the resolution of the RRS and deterministic rules, the standard deviation of the reserve requirements are computed from the data of the evaluation period.

B. Steady-state Security Assessment Problem

1) Model for Comparison

The fuzzy PF will be compared with the AC deterministic power flow, which uses only wind power point forecasts as input.

2) Criteria for Evaluation

The detection problem is a binary classification problem with two non-overlapping classes: congestion/voltage violation (1); no congestion/voltage violation (0).

The evaluation results can be presented by a contingency table like the one depicted in Table 1. True Positive (TP) is the number of positive correctly classified cases (e.g. detection of a congestion that occurred). True Negative (TN) is the number of negative correctly classified cases (e.g. no detection of congestion in a branch which did not occurred). There are also cases that either were incorrectly assigned to positive (False Positives, FP) or that were not recognized as positive situations (False Negatives, FN).

TABLE I
CONTINGENCY TABLE

	Classified as Positive	Classified as Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

To build this table it is necessary to know whether or not a congestion/voltage violation occurred. The ideal would be to run a power flow with the realized values of both generation and loads. However, the generation values already have corrective and preventive actions for avoiding such problems. Therefore, the approach adopted for this paper was to replace the triangular fuzzy numbers by the realized wind power generation values in each wind farm.

Afterwards, the power flows are computed with the realized values and the detected congestions are compared with the ones obtained from the fuzzy PF.

From this contingency table it is possible to compute several quality metrics [28]. Since the number of lines and buses without congestion or voltage violation is very high, two metrics that do not use the true negative (TN) term were adopted: *% of false alarms* and *% of overlooked violations*.

The percentage of false alarms is given by:

$$\%\ \text{of false alarms} = \frac{FP}{TP + FP} \quad (13)$$

The percentage of overlooked violations is given by:

$$\%\ \text{of overlooked violations} = \frac{FN}{TP + FN} \quad (14)$$

Note that Eq. 13 do not measures the overlooked violations, while Eq. 14 does not take into consideration the false alarms. Therefore, the evaluation consists in building a *% of false alarms* vs *% of overlooked violations* curve for each cut-off level *c*. A cut-off level *c* means considering that: $\alpha \geq c$ means alarm for congestion or voltage violation; $\alpha < c$ means no alarm (α is the possibility of violation mentioned in section IV). Note that the fuzzy PF results with cut-off equal to 1 are analogous to the deterministic PF results.

Afterwards, the decision maker, based on his/her preferences, selects the preferred method (i.e. deterministic or FPF with a certain cut-off) and the preferred threshold for the degree of congestion and voltage violation. The following linear trade-off function can be used for evaluation:

$$\mu \times (\%\ \text{of false alarms}(c)) + \%\ \text{of overlooked violat.}(c) \quad (15)$$

where *c* is the cut-off value, μ is the trade-off value that indicates how much the decision maker is willing to accept increasing the *% of false alarms* in exchange for decreasing the *% of overlooked violations*. For instance, μ equal to 3 would mean that an increase of 3% in false alarms is accepted to decrease the overlooked violations in 1%.

VI. RESULTS FROM THE DEMONSTRATION CASE

A. Brief Description of the Portuguese Power System

The capacity mix in the Portuguese power system in the end of 2010 was: 4.578 GW of hydro generation, 7.407 GW of thermal generation, and 6.077 GW of generation remunerated with feed-in tariffs. The generation under a tariff remuneration scheme is: mini-hydro (0.414 GW), CHP (1.687 GW), wind generation (3.854 GW) and photovoltaic (0.123 GW).

The total consumption in Portugal in 2010 was 52.2 TWh divided in: 20% from gas power plants, 5% imported from Spain, 17% from generation under tariff remuneration scheme

(without wind generation), 17% of wind generation, 28% of hydro generation and 13% from coal power plants. In 2010 the maximum peak power was 9.403 GW.

Wind power is remunerated with a fixed feed-in tariff and does not participate in the electricity market.

B. Results of the RRS Tool

The results presented in this section correspond to a demonstration period between October 2010 and May 2011.

The first set of analyses regards the upward tertiary reserve. Fig. 3 depicts reserve suggested by the RRS and deterministic rules for a selected day with a high amount of mobilized of upward reserve. This plot shows, for the last hours of the day, that the sudden increase of the mobilized reserve was covered by RRS suggestion. On the contrary, the deterministic rules show lack of flexibility, resulting in insufficient reserve.

Fig. 4 presents the estimated percentage of situations with insufficient suggested reserve (computed with Eq. 11) for the RRS with GFS and Skiron forecasts, and for the two deterministic rules. The RRS (Skiron) and rule 2 lead to the lowest percentage of insufficient reserve cases. The reference thresholds of the LOLP are inside the confidence interval for both 0.1% and 0.5%, and outside the interval for 1.0%. The comparison between the results obtained with Skiron and GFS, shows that GFS leads to a higher number of insufficient reserve situations and it is very far from the reference LOLP. The confidence intervals for the RRS (Skiron) are also narrower.

Both deterministic rules 1 and 2 lead to a high percentage of insufficient reserve cases. The main limitation of the deterministic rules is the fact that they do not give any indication about the level of risk that the SO is taking. Hence, these two deterministic rules lead to a number above this reference value, and with the limitation of not knowing the risk associated to the rule.

Table II additional evaluation metrics: quantiles 0.1% and 99.9% from the suggested minus mobilized reserve distribution, scoring rule and standard deviation of reserve requirements. The results are only presented for a risk threshold equal to 0.1%, but the conclusions are analogous for the other values.

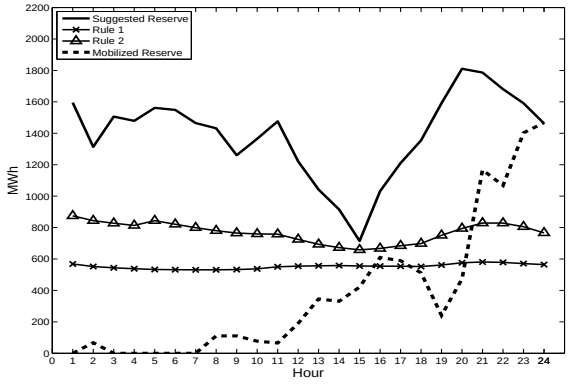

Fig. 3. Suggested upward reserve for each market session and mobilized reserve (27 Nov 2010) for LOLP equal to 0.5%.

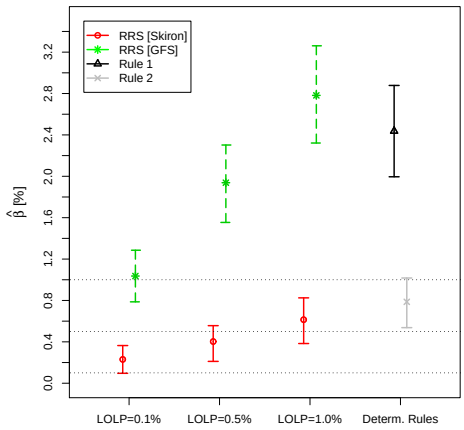

Fig. 4. Estimated percentage of situations with insufficient upward tertiary reserve and 95% confidence intervals.

The RRS (Skiron) leads to less negative 0.1% quantile compared to RRS (GFS) but it gives a higher value in the quantile 99.9%. Rule 1 and 2 lead to a high negative 0.1% quantile and to low 99.9%. The 99.9% quantile value suggests that RRS (Skiron) leads to higher suggested reserve requirements than RRS (GFS).

It is not an objective of this paper to present detailed analyses of the forecast quality (details on how to do this evaluation can be found in [29]). Nevertheless, this demonstration shows that different wind power forecasts lead to different performances. Fig. 5 and 6 present the calibration plots (observed against nominal quantiles/probabilities) for the wind power probabilistic forecasts obtained by using GFS and Skiron NWP.

The two calibration plots show that GFS based forecasts underestimate the quantiles and the Skiron based forecasts overestimate the quantiles. For example, in the GFS case, the forecasting system predicts that there is a probability of 20% that the wind power generation would be below the quantile value; however, the observed quantile computed from the evaluation data is higher. This translates directly to an underestimation of the risk. For the Skiron case, the situation corresponds to an overestimation of the risk. This explains a higher $\hat{\beta}$ than the risk threshold for the GFS based decisions, and a lower $\hat{\beta}$ than the risk threshold for the Skiron based decisions. The scoring metric of Eq. 12 quantifies the quality of the Skiron and GFS based reserve levels. This means that the reserve levels based on Skiron have better quality.

The two deterministic rules have a low standard deviation. This indicates that these rules do not give situation dependent reserve suggestions (i.e. meaning low resolution).

Fig. 7 presents the estimated percentage of situations with insufficient suggested downward reserve case. In this case, rules 1 and 2 have a very high embedded risk: rule 1 leads to 55.6% of situations with insufficient reserve and rule 2 leads to 34.36%. These results stress the idea that deterministic rules do not inform about the risk level that the decision maker is taking and may lead to situations of excessive risk.

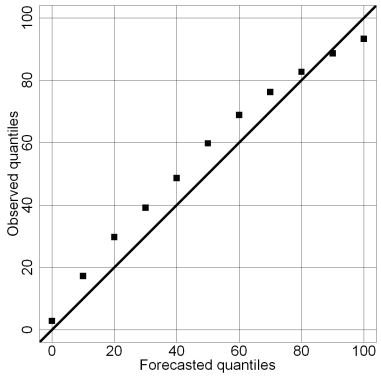

Fig. 5. Calibration plot for the GFS based forecasts.

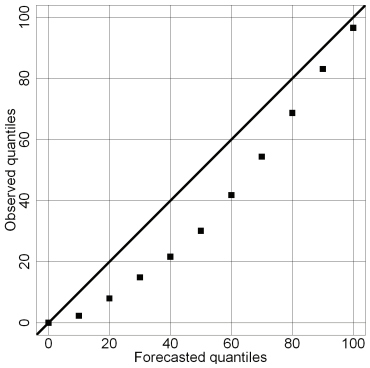

Fig. 6. Calibration plot for the Skiron based forecasts.

In contrast to the upward reserve results, the RRS (GFS) leads to a lower percentage of situations with insufficient downward reserve. The results for a reference PWRE of 0.1% are very similar for both RRS (GFS) and RRS (Skiron).

For the downward reserve case, the information provided by the calibration plots of Fig. 5 and 6 is also valuable. However, in this case the underestimation of quantiles leads to an overestimation of risk, and the overestimation of quantiles leads to an underestimation of risk. This explains the fact that the GFS based decisions presented a $\hat{\beta}$ lower than the Skiron based decisions.

TABLE II
EVALUATION METRICS FOR THE UPWARD RESERVE SUGGESTED BY RRS (LOLP EQUAL TO 0.1%) AND DETERMINISTIC RULES

	Q0.1% [MWh]	Q99.9% [MWh]	Scoring Metric	Std. Dev. [MWh]
RRS (Skiron)	-41.9	3268.6	-3.3	701.0
RRS (GFS)	-223.5	2780.5	-4.4	615.5
Rule 1	-517.5	614.5	-	24.0
Rule 2	-337.4	1207.2	-	176.0

Table III presents the additional evaluation metrics. An interesting observation from combining Fig. 7 and Table III is that the RRS (GFS) and RRS (Skrion) have almost the same value of $\hat{\beta}$ for β=0.1% but RRS (Skrion) presents a less negative 0.1% quantile and a lower 99.9% quantile, ascompared to RRS (GFS). The better performance of RRS (Skiron) is also confirmed by the scoring metric. This result underlines the importance of making the evaluation with different metrics.

The deterministic rules, as expected, have a very negative 0.1% quantile, a very low 99.9% quantile and a low resolution.

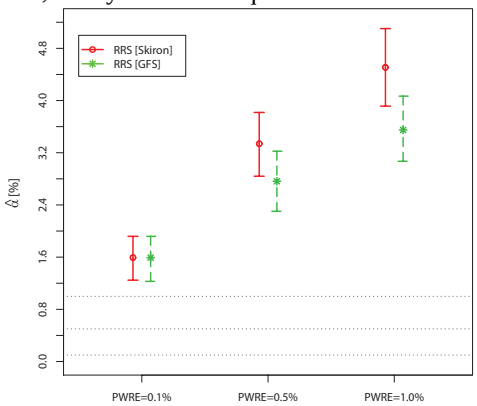

Fig. 7. Estimated percentage of situations with insufficient downward tertiary reserve and 95% confidence intervals.

C. Results of the Fuzzy PF Tool

The results presented in this section correspond to a demonstration period between December 2010 and May 2011.

TABLE III
EVALUATION METRICS FOR THE DOWNWARD RESERVE SUGGESTED BY RRS (PWRE EQUAL TO 0.5%) AND DETERMINISTIC RULES

	Q0.1%	Q99.9 %	Scoring Metric	Std. Dev.
RRS (Skiron)	-365.7	3164.9	-1.5	516.0
RRS (GFS)	-496.7	3394.4	-1.9	565.4
Rule 1	-1582.7	177.8	-	23.12
Rule 2	-1220.6	770.5	-	175.9

Fig. 8 depicts an example of a congestion detected on 14 May 2011: the realized wind power was greater than the forecasted, leading to a branch flow of 124.5 MVA (>121 MVA). In this case, the deterministic power flow was unable to detect the congestion (110.2 < 121 MVA) in the branch while the fuzzy PF, depending on the cut-off selected, is capable of detecting the possible line capacity limit violation. Note that if the decision maker defines a cut-off greater than 0.42, then the congestion would not be detected.

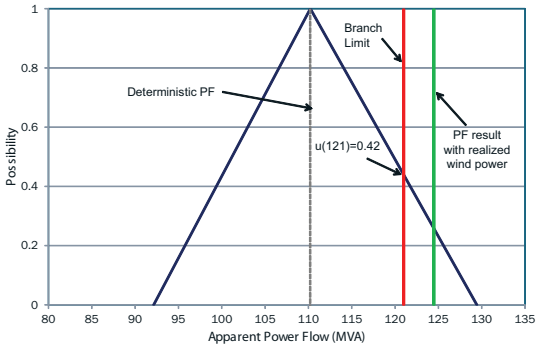

Fig. 8. Fuzzy power flow number for 14 May 2011 (15:00).

Fig. 9 depicts the *% of false alarms* vs *% of overlooked violations* curve, made with ROCR package [30] in R, for the congestion detection problem. The results for a cut-off equal to

zero are irrelevant, since it means violations in all the buses and lines. Note that the trade-off between the two metrics is not very pronounced. This is because the transmission network is very robust to congestion, and branches are normally operating far from the limit.

Three different linear trade-offs (Eq. 15) are depicted: μ=3, 12.5, 18. For a trade-off of 3 (and below) the preferred (intersected) point is cut-off equal to 1, i.e. deterministic PF. The deterministic PF (has the lowest value of *% of false alarms*, explained by a low number of false positives (FP). On the other hand, the deterministic PF has the highest *% of overlooked violations* value due to a high number of cases incorrectly classified as negative (FN).

For a trade-off of 12.5 the solution is more balanced between *% of false alarms* and *% of overlooked violations* with a cut-off of 0.88. If the decision maker gives more value to a low *% of overlooked violations*, a trade-off equal to 18 for example, it would lead to a low cut-off of 0.13.

The contingency tables for the deterministic PF and fuzzy PF with a cut-off of 0.13 are presented in tables IV and V. The tables confirm that the network is robust. The number of incorrectly classified cases in the deterministic PF is just 127 out of 4009. Therefore, the FPF can only improve over this number at the cost of increasing the *% of false alarms*. For example, a cut-off means a decrease of 40 in *% of overlooked violations*, but it increases the false alarms in 678.

If the network was being operated near the limits there would be more situations where the deterministic PF is near the line's limit, and the FPF will be very useful for detecting possible congestion when the forecasted wind power generation is underestimated. Nevertheless, results like the one depicted in Fig. 8 illustrate the fuzzy PF potential benefit.

Fig. 10 depicts the *% of false alarms* vs *% of overlooked violations* curve for the upper voltage limit, and for three different trade-off values: 0.8, 3, 10. The deterministic PF is preferred by a decision maker with a trade-off value equal or below to 0.8. Compared to the curve of Fig. 9, this curve shows a pronounced trade-off between the two metrics. And the deterministic PF is only prefered to a trade-off below 0.8, in contrast to a trade-off value below 3 for the congestion problem.

A trade-off value equal to 3 would lead to a cut-off equal to 0.68, and a trade-off value equal to 10 leads to a cut-off equal to 0.33. These two decision makers (μ equal to 3 and 10) are not very concerned with the *% of false alarms*, conversely they give more value to the decrease in false negative (FN) value. The FN with the deterministic approach were 365, and with a cut-off equal to 0.33 it was reduced to 116, as shown in tables VI and VII.

Fig. 11 depicts the *% of false alarms* vs *% of overlooked violations* curve for the lower voltage limit, and also for three different trade-off values: 1, 10, 15. In this case, the difference between a cut-off of 0.1 and 0.9 can be translated in a 0.022 difference in *% of false alarms* and 0.003 in *% of overlooked violations*. For a trade-off equal to 1, the preferred cut-off is 1, meaning that the decision-maker is giving much more value to a lower *% of false alarms* than to a lower *% of overlooked violations*. For a trade-off value of 10 the solution is more balanced between these to metrics, giving a cut-off equal to 0.6. If the decision maker gives more value to *% of false alarms* by using a trade-off value of 15, than the cut-off is 0.5.

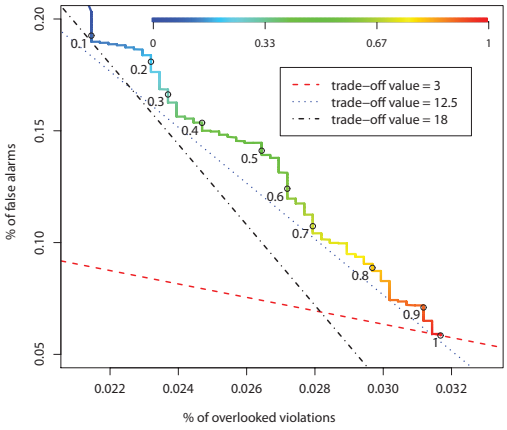

Fig. 9. % of false alarms vs % of overlooked violations curve with trade-off values (3, 12.5, 18) for the congestion problem.

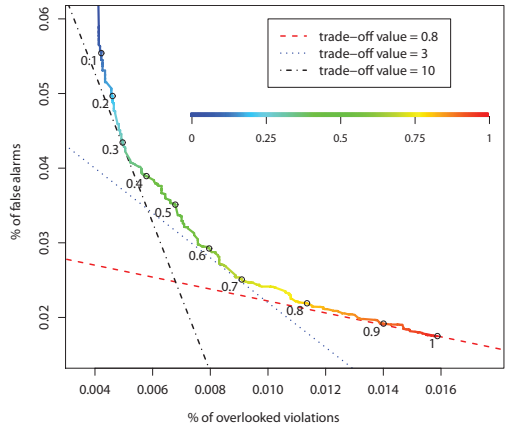

Fig. 10. % of false alarms vs % of overlooked violations curve with trade-off values (0.8, 3, 10) for the upper voltage limit problem.

Table VIII and IX present the results for the deterministic PF and for μ equal to 15 (cut-off of 0.5) respectively. Comparing a cut-off of 0.5 with the deterministic PF, the number of correct classified positive cases increase, and the number of FN was reduced. On the other hand, the number of false alarms increases.

TABLE IV
CONTINGENCY TABLE FOR CONGESTION DETECTION WITH DETERMINISTIC PF

	Classified as Positive	Classified as Negative
Positive	3882	127
Negative	241	1032415

TABLE V
CONTINGENCY TABLE FOR CONGESTION DETECTION WITH FUZZY PF (C=0.13)

	Classified as Positive	Classified as Negative
Positive	3922	87
Negative	919	1031737

TABLE VI
CONTINGENCY TABLE FOR UPPER VOLTAGE LIMIT WITH DETERMINISTIC PF

	Classified as Positive	Classified as Negative
Positive	22624	365
Negative	403	1141678

TABLE VII
CONTINGENCY TABLE FOR UPPER VOLTAGE LIMIT WITH FUZZY PF (C=0.33)

	Classified as Positive	Classified as Negative
Positive	22873	116
Negative	1009	1141072

TABLE VIII
CONTINGENCY TABLE FOR LOWER VOLTAGE LIMIT WITH DETERMINISTIC PF

	Classified as Positive	Classified as Negative
Positive	17225	223
Negative	105	1147517

TABLE IX
CONTINGENCY TABLE FOR LOWER VOLTAGE LIMIT WITH FUZZY PF (C=0.5)

	Classified as Positive	Classified as Negative
Positive	17274	174
Negative	326	1147296

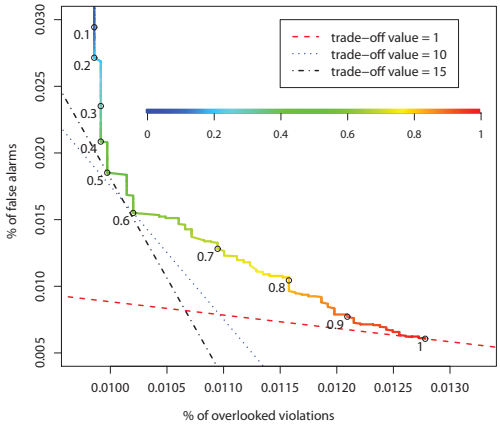

Fig. 11. % of false alarms vs % of overlooked violations curve with trade-off values (0.8, 3, 10) for the lower voltage limit problem.

VII. CONCLUSIONS

This paper presents implementation details and results from an operational demonstration of two tools for a Transmission System Operator. Moreover, the paper contributes with an evaluation methodology for both problems: reserve setting and power system steady-state security assessment.

The Robust Reserve Setting (RRS) tool outperformed the deterministic rules in place. The RRS informs the decision maker about the level of risk that he/she is taking, while deterministic rules do not inform on the level of risk and may lead to situations with high risk. This was particularly clear for downward reserve, which presented a high embedded risk.

The results showed that different forecasts lead to different performances of the management tools. Thus, it is imperative to evaluate and monitor wind power probabilistic forecast quality. "Good" wind power forecasts (i.e. point and probabilistic) may lead to a decrease in reserve requirements and also to better decisions.

The RRS tool was considered to be very useful by the end user, in particular for situations with high forecast error. For instance, a high wind power forecast error may lead to an assignment of tertiary reserve in a higher value than it should be. This could lead to start a thermal generation unit which has fixed price. Afterwards, the wind energy could have a different behavior than predicted and the thermal unit is disconnected. The tertiary reserve has a cost that will be reflected on the tariff and therefore should be an issue to develop more appropriate tools to help the TSO decisions in order to minimize the system operations costs.

The fuzzy PF tool identifies potential cases of congestion or voltage violation. In some situations the constraint violation would occur if preventive and corrective actions were not taken. It is clear that the use of wind power uncertainty forecasts in power flow calculations represents an additional benefit for the transmission network operation, in particular when the network is operating near the limits.

The results obtained with the fuzzy PF concluded that the transmission network is robust enough to accommodate the presently installed wind power capacity. These good characteristics limited the demonstration extension, although the potential of the tools has been fully demonstrated.

The end user recommended, as a future feature, the adaption of the fuzzy PF for performing "N-1" security evaluation as the ENTSO-e Policy 3 suggests.

As future recommendation, a similar tool could be also valuable at the distribution grid level, namely considering electric vehicles and demand response integration.

The demonstrations showed that "dialog" between tools and operators is crucial for the success of these tools. The operators presently receive too much information and raw wind power forecast uncertainty may be unproductive. Hence, future research should cover enhanced approaches for communicating uncertainty and risk to operators, in order to provide clear advices.

VIII. REFERENCES

[1] D. Connolly, H. Lund, B.V. Mathiesen, and M. Leahy, "A review of computer tools for analysing the integration of renewable energy into various energy systems," *Appl. Energ.*, vol. 87(4), pp. 1059-1082, Apr. 2010.

[2] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl, "The state-of-the-art in short-term prediction of wind power, a literature overview, 2nd Edition," Technical report, EU project ANEMOS.plus, Jan. 2011.

[3] C. Monteiro, R.J. Bessa, V. Miranda, A. Botterud, J. Wang, and G. Conzelmann, "Wind power forecasting: state-of-the-art 2009," Technical Report ANL/DIS-10-1, Argonne National Laboratory, Nov. 2009.

[4] G. Kariniotakis, E. Nogaret, and G. Stavrakakis, "Advanced short-term forecasting of wind power production," in *Proc. of the European Wind Energy Conference EWEC'97*, Dublin, Ireland, Oct. 1997.

[5] R.J. Bessa, V. Miranda, A. Botterud, and J. Wang, "'Good' or 'bad' wind power forecasts: a relative concept," *Wind Energy*, vol. 14 (5), pp. 625-636, Jul. 2011.

[6] J.K. Møller, H.A. Nielsen, and H. Madsen, "Time-adaptive quantile regression," *Comp. Stat. & Data Analysis*, vol. 52(3), pp. 1292-1303, Jan. 2008.

[7] P. Pinson and G. Kariniotakis, "Conditional prediction intervals of wind power generation," *IEEE Trans. on Power Sys.*, vol. 25(4), pp. 1845-1856, Nov. 2010.

[8] M. Leutbecher and T.N. Palmer, "Ensemble forecasting," *J Comput Phys*, vol. 227(7), pp. 3515-3539, Mar. 2008.

[9] R. Doherty and Mark O'Malley, "New approach to quantify reserve demand in systems with significant installed wind capacity," *IEEE Trans. on Power Sys.*, vol. 20(2), pp. 587-595, May 2005.

[10] M.A. Matos and R.J. Bessa, "Setting the operating reserve using probabilistic wind power forecasts," *IEEE Trans. on Power Sys.*, vol. 26(2), pp. 594-603, May 2011.
[11] J. Usaola, "Probabilistic load flow with correlated wind power injections," *Elect. Power Sys. Res.*, vol. 80(5), pp. 528-536, May 2010.
[12] A. Tuohy, P. Meibom, E. Denny, and Mark O'Malley, "Unit commitment for systems with significant wind penetration," *IEEE Trans. on Power Sys.*, vol. 24(2), pp. 592-601, May 2009.
[13] J. Wang, A. Botterud, R.J. Bessa, H. Keko, L. Carvalho, D. Issicaba, J. Sumaili, and V. Miranda, "Representing wind power forecasting uncertainty in unit commitment," *Appl. Energ.*, vol. 88, pp. 4014-4023, Nov. 2011.
[14] G. Kariniotakis, J. Lugaro, et al. "Proceedings of the workshop on towards smart wind integration," Deliverable 6.4 of the EU project ANEMOS.plus, Paris, France, 29 Jun. 2011.
[15] M.A. Matos, J.A. Peças Lopes, M. Rosa, R. Ferreira, A. Leite da Silva A, W. Sales, *et al.*, "Probabilistic evaluation of reserve requirements of generating systems with renewable power sources: The Portuguese and Spanish cases," *Intern. Journal of Elect. Power & Energy Sys.*, vol. 31(9), pp. 562-569, Oct. 2009.
[16] Y. Rebours and D.S. Kirschen, "A survey of definitions and specifications of reserve services," Technical Report, University of Manchester, Oct. 2005.
[17] M. De la Torre, "Managing uncertainty at REE," presentation at Utility Wind Integration Group (UWIG), Albuquerque, New Mexico, 10-11 Feb., 2010. Available online: http://www.uwig.org (accessed on Jul. 2011)
[18] D. Maggio, "Integrating wind forecasting into market operation-ERCOT," presentation at Utility Wind Integration Group (UWIG) Workshop, Phoenix, USA, 18-19 Feb. 2009. Available online: http://www.uwig.org (accessed on Jul. 2011)
[19] V. Miranda, M.A. Matos, and J.T. Saraiva, "Fuzzy load flow- new algorithms incorporating uncertain generation and load representation," in *Proc. 10th Power Sys. Comp. Conf.*, Graz, Austria, 1990, pp. 621–627.
[20] E.M. Gouveia and M.A. Matos, "Symmetric AC fuzzy power flow model," *Eur. Jour. of Oper. Res*., vol. 197(3), pp. 1012-1018. Sept. 2009.
[21] H-P. Waldl, F. Dierich, G. Kariniotakis, A. Bocquet, and S. Virlot "The ANEMOS wind power forecasting platform," in Proc. of the European Wind Energy Conference EWEC'06, Athens, Greece, 27/2-2/3 2006.
[22] www.windpowerpredictions.com (accessed on Jul. 2011)
[23] H. Madsen, H.A. Nielsen, and T.S. Nielsen, "A tool for predicting the wind power production of off-shore wind plants," in Proc. of the Copenhagen Offshore Wind Conference & Exhibition, Copenhagen, Denmark, 2005.
[24] UCTE Operating handbook - Policies P1: load-frequency control and performance, March 2009. Available online: https://www.entsoe.eu (accessed on Jul. 2011)
[25] M. Oussalah, "On the probability/possibility transformations: a comparative analysis," *Int. Jour. of Gen. Sys.*, vol. 29(5), pp. 671-718, 2000.
[26] A.C Davison and D.V. Hinkley, *Bootstrap Methods and Their Application*, Cambridge: Cambridge University Press, 1997.
[27] T. Gneiting and A.E. Raftery, "Strictly proper scoring rules, prediction, and estimation," *J. of the Amer. Stat. Assoc.*, vol. 102, pp. 359-378, 2007.
[28] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Infor. Proc. and Manag.*, vol. 45(4), pp. 427-437, Jul. 2009.
[29] P. Pinson, H.A. Nielsen, J.K. Moller, H. Madsen, and G. Kariniotakis, "Nonparametric probabilistic forecasts of wind power: required properties and evaluation," *Wind Energy*, vol. 10(6), pp. 497-516, 2007.
[30] T. Sing, O. Sander, N. Beerenwinkel, and T. Lengauer, "ROCR: visualizing classifier performance in R," *Bioinformatics*, vol. 21(20), pp. 3940-3941, Oct. 2005.

IX. BIOGRAPHIES

Ricardo Bessa received his Licenciado (five years) degree from the Faculty of Engineering of the University of Porto, Portugal (FEUP) in 2006 in Electrical and Computer Engineering. In 2008 he received his Master of Science degree in Data Analysis and Decision Support Systems on the Faculty of Economics of the University of Porto (FEP). Currently, he is a researcher at INESC TEC in its Power Systems Unit and a PhD student of the Doctoral program in Sustainable Energy Systems (MIT Portugal) at FEUP. His research interests include wind power forecasting, electric vehicles, data mining and decision-aid methods.

Manuel A. Matos (El. Eng., Ph.D., Aggregation, M'94) was born in 1955 in Porto, Portugal. He is with the Faculty of Engineering of the University of Porto (FEUP), Portugal, since 1978 (Full Professor since 2000). He is also coordinator of the Power Systems Unit of INESC TEC. His research interests include classical and fuzzy modeling of power systems, reliability, optimization and decision-aid methods.

Ivo Costa have a Master degree in Electrical and Computer Engineering from the Faculty of Engineering of the University of Porto (FEUP), Porto, Portugal, in 2010. He is he is a researcher at INESC TEC in its Power Systems Unit.

Leonardo Bremermann received his Graduation (five years) degree in Electric Engineering and Scientific Master degree in Energy Systems from the Pontifical Catholic University of Rio Grande do Sul (PUCRS), Porto Alegre, Brazil, in 2005 and 2008, respectively. Currently, he is a researcher at INESC TEC in its Power Systems Unit and a PhD student of the Doctoral program in Sustainable Energy Systems (MIT Portugal) at Faculty of Engineering of the University of Porto - FEUP, Portugal. His research interests are in line with reliability assessment of electrical power systems, including renewable energy sources and electric vehicles.

Ivan Gustavo Franchin received his Licenciado (four years) degree from the Federal University of São Carlos, Brazil (UFSCar) in 2003 in Computer Science. In 2007 he received his Master degree in Software Engineering on the São Paulo University (USP). Currently, he is a researcher and system analyst at INESC TEC in its Power Systems Unit.

Rui Pestana received his "Licenciatura" and "Mestrado" degrees in Electrical Engineering and computers from Instituto Superior Técnico, Lisbon, Portugal, in 1986 and 1990, respectively. He is with REN-Rede Eléctrica Nacional, S.A. since 1989, and he is assistant director of the System Operator division since 2008 and the head of the systems and development department. He is member of the Portuguese professional association of Engineers. He is a regular member of CIGRÉ C2 Study Committee. He is corresponding member of several WG on ENTSO-E. He is member of the ANEMOS.plus, PEGASE and WG2 of the SmartGrid EU projects. He is also invited Professor at the Polytechnic Institute of Lisbon since 1990.

Nélio Machado received his "Bacharel" degree in "Energy and Power Systems" on Instituto Militar dos Pupilos do Exército, Lisbon, Portugal in 1989. In 1996 completed his "Licenciatura" and in 2002 the "Pos-Graduação" degree, both on the Electrical Engineering and Computers Course on Instituto Superior Técnico (IST), Lisbon, Portugal. Currently he is completing his thesis of Master of Science on the Faculdade de Engenharia da Universidade do Porto on Electrical Engineering and Computers - Major in Energy Systems. He was on the Project and Construction Department in EDP - Energias de Portugal,SA since 1989. Now he is with REN - Rede Eléctrica Nacional,SA since 1995 on the Systems and Development department of the System Operator Division. He is the representative member in ENTSO-e Continental Europe Sub-Group "Network and Forecast Models" since 2005.

Hans-Peter Waldl is managing director of and partner in Overspeed GmbH & Co. KG. He was working until mid 2001 as a senior scientist at the University of Oldenburg, managing the wind energy activities of the energy and semiconductor group from 1992 to 2001. He received a Diploma degree in Physics at the University of Oldenburg and holds a PhD degree in Physics from the University of Oldenburg ("Modelling the power output from wind farms").

Christian Wichmann born 1974 in Wilhelmshaven, received his diploma in Computer Science at the University of Oldenburg in 2006. Since then, he has been working as System and Software Developer at Overspeed GmbH & Co. KG, being involved in the EU research projects Anemos and Anemos.plus which are dealing with wind power forecasting.