

PAPER PRESENTED AT AWEA WINDPOWER 2010

'Good' or 'bad' wind power forecasts: a relative concept

R. J. Bessa¹, V. Miranda¹, A. Botterud² and J. Wang²

¹ INESC Porto—Instituto de Engenharia de Sistemas e Computadores do Porto, and FEUP—Faculty of Engineering, University of Porto, Porto, Portugal

² Argonne National Laboratory, Decision and Information Sciences Division, Argonne, Illinois, USA

ABSTRACT

This paper reports a study on the importance of the training criteria for wind power forecasting and calls into question the generally assumed neutrality of the 'goodness' of particular forecasts. The study, focused on the Spanish Electricity Market as a representative example, combines different training criteria and different users of the forecasts to compare them in terms of the benefits obtained. In addition to more classical criteria, an information theoretic learning training criterion, called parametric correntropy, is introduced as a means to correct problems detected in other criteria and achieve more satisfactory compromises among conflicting criteria, namely forecasting value and quality. We show that the interests of wind farm owners may lead to a preference for biased forecasts, which may be in conflict with the larger needs of secure operating policies. The ideas and conclusions are supported by results from three real wind farms. Copyright © 2010 John Wiley & Sons, Ltd.

KEYWORDS

wind power forecasting; neural networks; correntropy; electricity markets; good forecasts; bad forecasts

Correspondence

R. J. Bessa, INESC Porto, Campus da FEUP, Rua Dr Roberto Frias, 378, 4200-465 Porto, Portugal.

E-mails: rbessa@inescporto.pt

Received 26 April 2010; Revised 28 October 2010; Accepted 29 October 2010

1. INTRODUCTION

The demand for more accurate short-term wind power forecasting (WPF) models has led to solid and impressive development in recent years.^{1,2} Increasing the value of wind generation by improving WPF systems' performance—and furthering its integration into operational management—is one of the priorities in current research.³

On many past occasions, energy market participants have taken for granted that a 'good' forecast would be something that turns out to be as close to the actual values as possible. This view is called the forecaster paradigm, where the concern is to fit the prediction to the observed values in the signal processing sense. A good example of a WPF evaluation protocol in agreement with this paradigm is that of Madsen *et al.*:⁴ the main concern is minimizing the forecasting error, and no importance is given to the nature and consequences of the errors.

However, it may be argued that the definition of the *goodness* of a forecast should be strongly related to the use given to such a forecast. In the forecasting pilot project by Alberta Electric System Operator,⁵ an important idea is that a forecast should be oriented by the user's interests. Forecasting errors may be seen through many perspectives (e.g. mean, bias, maximum deviation, phase error, etc.), and the nature of the errors may be too broad—one particular forecast is not likely to be considered optimal simultaneously in distinct applications, such as market participation or power system operation.

This situation has been recognized in the domain of weather forecast. Several publications discuss what a *good* or *bad* forecast is. Murphy⁶ distinguished three different types of *goodness* in forecasts: (i) the correspondence between forecasts and forecasters' judgments (*consistency*); (ii) correspondence between forecasts and observations (*quality*); and (iii) the incremental benefits (economic/or other) when employed by users as an input into their decision-making processes (*value*).

The factor of *consistency* in WPF is not discussed in this paper: it is assumed to be included by the forecaster in the model during the development phase.

In order to assess the *quality* of WPF, the standard approach is based on statistical error measures (e.g. mean absolute error).⁴ However, this approach does not guarantee that the forecast with the best score is the one with best *quality*. An alternative approach, called *verification framework*, was presented by Murphy and Winkler.⁷ This approach is based on the joint distribution of forecasts and observations and factorization of this distribution into conditional and marginal distributions. Pinson⁸ used this framework, named the 'distribution-oriented approach', and applied it to WPF. His objective was to study the main characteristics of WPF uncertainty.

The third type of forecast goodness, forecast *value*, is often very difficult to quantify, as in some cases, it is related to non-economic factors. However, some decision makers may give a high relevance to these criteria.⁹ Some researchers have reported incremental economic benefit from the use of WPF in the participation of wind power in electricity markets. Usaola *et al.*¹⁰ quantified the economic advantages of using an advanced forecasting approach that takes as input numerical weather predictions (NWP) instead of persistence. Barthelmie *et al.*¹¹ estimated the wind farm size for which WPF outweighs the cost of installing or implementing short-term forecasts. The authors conclude that the economic benefit of using WPF depends on the accuracy and cost of purchasing the forecast. Angarita-Márquez *et al.*¹² estimated the income obtained by a wind farm in the Spanish and British markets by using an advanced forecasting system (Sipreólico) compared with persistence. In the British market, the Sipreólico system produced results similar to persistence because the period between gate closure and actual delivery was very short. For the Spanish market, because the period between gate closure and delivery was longer, Sipreólico improved the income amount(s) when compared to persistence.

Pinson *et al.*¹³ studied the use of WPF probabilistic forecasts to support the participation of wind generation in the electricity market. The authors showed that with the information about uncertainty, it is possible to derive advanced strategies for market participation leading to a significant increase in income when compared to persistence and point forecasts obtained with fuzzy-neural networks. Fabbri *et al.*¹⁴ proposed a probabilistic methodology for estimating the WPF error costs in the electricity market. Other related works can be found in literature.^{15–17}

One crucial aspect is that the relationship between *quality* and *value* is nonlinear and differs from problem to problem, as well as from user to user.¹⁸ Hence, it is important to recognize that there are different users for WPF, the two most important being wind generation companies (WGENCOs) that bid their wind generation into the electricity market and system operators (SOs) that must guarantee the security of the system.

The output of a forecasting process or model can therefore be evaluated according to two paradigms, which we will call the *forecaster paradigm* and the *forecast consumer paradigm*. The former is mainly concerned with the time series in itself, in the signal processing sense, and is therefore related with quality; the latter focuses on the use given to the forecasts and is therefore related to value.

The main idea driving this paper is to promote the recognition (many times forgotten) that distinct user groups may have conflicting interests in a restructured market environment, and therefore, the choice of a model is not neutral. For instance, SOs are interested in minimizing operational costs while maintaining a high level of reliability. In contrast, WGENCOs are mainly interested in maximizing their income levels in the electricity market. Hence, in some conditions, the definition of a 'good' forecast varies with different forecast users.

This paper contributes to the discussion of this topic by comparing the forecast errors and economic consequences of adopting several criteria for the statistical training of WPF models. To illustrate the relevance of the concepts discussed, the paper presents a case study of the Iberian day-ahead electricity market (MIBEL). The study could be reproduced for other markets; however, the objective here is not to study and compare markets but to discuss models, and the example presented is sufficiently representative. Our main objective is not to provide an alternative to the approaches based on probabilistic forecasts^{13–17} but rather to highlight a new view of the relative importance of the forecast error in a multi-criteria and multi-perspective paradigm, in which biased point forecasts could be produced to increase the market income, and to stress the conflicting objectives that may exist between different users (or stakeholders). To compensate for some of the problems identified and analysed, we introduce parametric correntropy, which is a new training criterion for WPF that aims to arrive at a satisfactory compromise between forecast *value* and *quality*.

We first discuss the WGENCOs' and SOs' perspectives on WPF. Next, on the basis of actual day-ahead market data from MIBEL, we illustrate the conflict between the viewpoints and also the applications of different training criteria. Finally, we analyse the consequences of the different training criteria for the two WPF users.

2. THE 'FORECAST CONSUMER PARADIGM'

2.1. Wind generation companies' viewpoint

In a restructured market environment, forecast errors may have an associated cost or penalty, as the wind power producer usually ends up supplying an amount of energy that differs from the market bid. The market remuneration scheme adopted in this paper is for operating in the day-ahead market and is analogous to the one formulated in literature.^{13–17}

In this mechanism, wind power producers offer energy quantity bids (E^b) at day D (typically till 12:00 noon) for every hour of day $D + 1$ in the day-ahead market, and they are paid at the market clearing price (p^s). Then, if the actual wind generation is above or below their bid, they will be subject to a penalty as follows:

- If the actual generation (E^p) is above the market bid, the excess is paid at a discounted price (p^{surplus}). This price can also be negative, which represents a cost.
- If the actual generation remains below the market bid, a penalty is applied according to the price (p^{shortage}) of the generation that the SO has to purchase to compensate for the insufficient generation.

Each market specifies its own rules for the calculation of deviation costs, and therefore, no general conclusion can be offered. However, it seems obvious—and this paper confirms this finding—that if the penalties are asymmetric, then an opportunity arises for a gambler to benefit from using a biased model instead of a neutral model.

A wind farm's income (I) from selling for a given look-ahead time $t + k$ can be formulated as:

$$I_{t+k} = \begin{cases} p_{t+k}^s \cdot E_{t+k}^b + p_{t+k}^{\text{surplus}} \cdot (E_{t+k}^p - E_{t+k}^b), & E_{t+k}^p > E_{t+k}^b \\ p_{t+k}^s \cdot E_{t+k}^p - p_{t+k}^{\text{shortage}} \cdot (E_{t+k}^b - E_{t+k}^p), & E_{t+k}^p < E_{t+k}^b \end{cases} \quad (1)$$

This formulation can be rearranged by considering the balancing prices, $c_{\text{down_reg}} = p_s - p_{\text{surplus}}$ and $c_{\text{up_reg}} = p_{\text{shortage}} - p_s$, as:

$$I_{t+k} = \begin{cases} p_{t+k}^s \cdot E_{t+k}^p - c_{t+k}^{\text{down_reg}} \cdot (E_{t+k}^p - E_{t+k}^b), & E_{t+k}^p > E_{t+k}^b \\ p_{t+k}^s \cdot E_{t+k}^p - c_{t+k}^{\text{up_reg}} \cdot (E_{t+k}^b - E_{t+k}^p), & E_{t+k}^p < E_{t+k}^b \end{cases} \quad (2)$$

It is important to emphasize that in this market formulation, the WGENCOs are appointed balance responsibility; hence, they have a financial responsibility for any imbalance.

In some markets, if the deviation helps the SO balancing measure, then the market agent is not penalized. Note that the intra-day markets are not taken into account but may reduce the forecast error—and consequently the deviations—considerably.

From the WGENCO viewpoint, the *value* of the WPF in the market is translated by the additional income provided by such forecasts over the income obtained with an alternative method. The first component of equation (2) is constant and corresponds to the income obtained with perfect forecasts. Hence, maximizing the market income is analogous to minimizing the expected costs of balancing (second component in equation (2)). Therefore, in order to meet its goal, the WGENCO will follow a particular bidding strategy depending on the market balancing costs.

Because of the uncertainty in prices and wind generation, resolving this matter is a decision problem under uncertainty. Several decision rules can be used for finding the preferred bid. The decision rule used most often is to find the bid E_{t+k}^b that minimizes the expected balancing costs (i.e. to maximize expected income). This rule was first introduced by Bremnes¹⁹ and has been used by Pinson *et al.*¹³ and Linnet.¹⁷ This decision rule describes a risk-neutral attitude (i.e. a linear utility function). Bourry *et al.*¹⁶ presented an alternative strategy based on portfolio theory that balances between the expected value and the risk of the balancing costs distribution. Botterud *et al.*²⁰ present a methodology to derive optimal day-ahead bids for a wind power producer under uncertainty and compare the results of different bidding strategies (e.g. maximizing the utility). The authors discussed how the optimal bids depend on the electricity market design.

Under the market assumptions outlined above, the 'optimal' bid of the expected value decision rule possesses an analytical solution; for a proof, see Bremnes.¹⁹ The bid that maximizes the expected income is the $C_{\text{down_reg}}(C_{\text{down_reg}} + C_{\text{up_reg}})^{-1}$ quantile of the forecasted probability distribution of wind generation. It is found that the 'optimal' bid depends only on the quantile proportion computed from the balancing prices. For instance, if $c_{\text{down_reg}} > c_{\text{up_reg}}$, the 'optimal' quantile is above 0.5 (median), and this result constitutes overestimation; conversely, if $c_{\text{down_reg}} < c_{\text{up_reg}}$, the 'optimal' quantile is below the median and the result is underestimation.

Based on this idea and as carried out by Pinson *et al.*,¹³ the forecast with higher economic *value* is not necessarily the one with lower forecast error (in accordance with the 'forecaster's paradigm'). The results presented by the authors support this idea: a point forecast with 40.55% of imbalances (with respect to produced energy) achieved an income of 1.15 M€ over 1 year, whereas an advanced strategy based on a probabilistic forecast reached a higher 55.46% of imbalances—but also an income of 1.21 M€.

As mentioned before, if the deviation is in favor of the system, then the WGENCO does not pay balancing costs under some markets' rules. In Linnet,¹⁷ the bidding strategy proposed by Bremnes¹⁹ was extended to include this possibility, assuming that the system balance is unrelated to the wind power imbalance or that all wind power producers follow the same bidding strategy.

Finding the decision rule and the best bidding strategy is clearly a complex problem that requires wind power and price forecasting information. One alternative is to develop economic-oriented training criterion for training wind-to-power (W2P) models. Ravn²¹ analysed a training criterion that penalizes according to the balancing costs and the sign of the forecast errors:

$$\min \sum_{j^{\text{down}}} w_{\text{down_reg}} \cdot |\hat{P}_{t+k} - P_{t+k}| + \sum_{j^{\text{up}}} w_{\text{up_reg}} \cdot |\hat{P}_{t+k} - P_{t+k}| \quad (3)$$

The expression is divided into terms of overestimation, $J^{\text{up}} (\hat{P}_{t+k} > P_{t+k})$, and underestimation, $J^{\text{down}} (\hat{P}_{t+k} < P_{t+k})$; $w_{\text{down_reg}}$ and $w_{\text{up_reg}}$ are penalization factors related to the market balancing costs.

By applying this training criterion, an economic bias is introduced, which can not only increase the forecast error, but also increases the income from the market. Therefore, the economic value of a wind power forecast can be increased by information provided by either probabilistic forecasts or economic-oriented deterministic (point) forecasts.

2.2. System operator's viewpoint

In our discussion, the viewpoint of the power SO is not linked to how the SO should make the forecasts. In Botterud *et al.*,²² the current use of WPF in U.S. independent system operator/regional transmission organization markets is discussed, and recommendations for how to make efficient use of the information in state-of-the-art forecasts are offered.

In this paper, the discussion is related instead to the impact on the operational policy from the economic-oriented market bids provided by the WGENCOs. As stated in the previous section, a WPF system that provides the maximization of income to a WGENCO should not be trusted or used by SOs.

The SO needs to manage the combination of generation and consumption variability considering unit outages. Therefore, in a restructured electricity market environment, the SO acquires all of the reserve needed for the control area, in order to maintain a minimum reliability level.

In the daily market, the SO at day D is in charge of defining and contracting the operating reserve (i.e. the reserve related to loss of load or generation surplus attributable to forecast errors—both in load and wind—and unit outages, e.g. spinning and non-spinning reserve) needed for the next day (day $D + 1$).²³ These operating reserve levels are generally settled by deterministic criteria, such as the capacity of the largest unit plus an empirical formula, such as the Union for the Coordination of Transmission of Electricity (UCTE, now ENTSO-E) rule.²⁴ However, in the literature, probabilistic approaches are proposed by several authors; see, e.g., Doherty and O'Malley²⁵ and Matos and Bessa.²⁶

The integration of large shares of wind generation requires an increase in the amount of operating reserve that is needed to balance generation and load in order to maintain an acceptable level of reliability. Results from Doherty and O'Malley²⁵ and Holttinen²⁷ support this idea. Despite reducing wind power variability as a result of the aggregation of several wind farms, the variability and uncertainty of wind generation tend to be a source of stress for the operations personnel.

Because the load and generation are dispatched by the market mechanism, and assuming that WGENCOs participate with bids in the market, the strategic bidding followed by WGENCOs may influence significantly the needs for operating reserves and also the overall cost of balancing the system. As mentioned in the previous section, a WGENCO's income does not necessarily increase by reducing forecast deviations. Therefore, the forecast error should be considered to be of secondary concern for the SO, because the relevant error in this case is introduced by the market bids.

For instance, if the market bids overestimate the wind generation, then the load will be met by 'untrue' generation in the day-ahead market. In this case, the SO with its own neutral WPF system can forecast the wind generation uncertainty and by using, for instance, an approach that is similar to that used by Matos and Bessa²⁶ to detect a high risk of not meeting the load. Hence, the SO will need to commit more generating units (e.g. using the reliability assessment commitment to commit more units²²) or to contract more upward reserve, which increases the operational costs.

The need for upward reserves is the most critical situation, even for planning the power system capacity, as it requires that traditional power station capacities must be available as operating reserves. This situation may limit the ability of wind power plants to replace the conventional power station capacities.

Of course, this system balancing cost depends on the generation mix (e.g. the reserve provided by gas turbines or hydro is less expensive) and on the operating strategy.²⁸ For instance, in cases where the SO defines the reserve requirements on the basis of deterministic criteria, the awareness of this overestimation may lead to a highly conservative attitude and an adoption of high safety margins to minimize the risk—and will therefore heavily increase the operational costs. Also, as mentioned in Ilic *et al.*,²⁹ imbalances may aggravate potentially large out-of-merit automatic generation control costs. All of these operational costs will be part of the tariff paid by all customers.

Nevertheless, the SO may have to shed load if it has offline units because of the expected wind power availability (as bid into the market). It may also happen that the deviation between load and generation cannot be met by all of the available operating reserve because of transmission constraints.

The same behavior is valid for the case where the wind power bids are underestimated. This problem is of less concern for the SO, because it does not need to shed load or have additional power stations available to cover the energy deficit. However, the SO also needs to contract downward reserve (e.g. reserve to deal with generation surplus due to forecast errors, e.g. pumped storage) and may have to curtail wind generation. This situation leads to what is called an 'over-commitment' of conventional generation, which will reduce the economic and environmental welfare of the power system.

At this point, the key question is, from an SO viewpoint, what is a 'good' forecast (or bid) for the market made by WGENCOs? The 'UCTE NetWork of Experts on Wind Power', composed of experts from several SOs such as Transpower (Germany) and REE (Spain), stated in UCTE NetWork of experts on Wind Power³⁰ that the crucial issue is the expected maximum forecast deviation, and not the mean forecast error. The group's finding is because the extreme deviations may lead to load shedding. Consequently, for an SO, a good criterion will be to minimize the expected maximum forecast error. However, this training criterion may be difficult to translate to a function, so the minimum mean square error (MSE) could be an acceptable criterion because it weights the large deviations more heavily.

However, this criterion must be viewed from the perspective of maintaining a desired reliability standard based on energy offered in the market by the market participants, including WGENCOS. In this case, a 'good' WGENCO forecast for an SO should be the one that leads to: (i) less use of reserves, in particular upward reserves; (ii) a lower expected maximum forecast error; and (iii) deviations in favor of the system.

The first two criteria may lead to a conflict with the WGENCO's viewpoint. With respect to the third item, it is important to point out that if all WGENCOs use a bidding strategy based on the system deviation, it is probable that their deviations will be against the system deviation. Even in a situation where forecast errors help the system deviation, the SO must (in the day before) contract reserve and schedule the generation according to a probabilistic algorithm or deterministic rule without knowing this information *a priori*.

2.3. The compromise viewpoint, or the search for Utopia

The forecasting and market bidding analysis must be performed carefully, because we are dealing with two generally conflicting viewpoints, as outlined above. In this case, there is no direct application of the trade-off concept, because there is no single agent that evaluates its gains in one criterion for giving up some value in another criterion—with disjoint or partially disjoint sets of criteria, a loss in one criterion is felt by an agent by just a loss. This analysis pertains to the field of multiple agent decision making, where it is recognized that the 'Utopia' solution (i.e. one that optimizes all criteria for all agents, even those with conflicting interests and who may not even share the same criteria) is generally an infeasible point. Utopia corresponds to the highest level of aspiration for all agents. In contrast, the 'ideal' solution refers to the maximum aspiration of a single agent and is therefore an individualistic concept.

Market participants therefore seek solutions that may be seen as an acceptable compromise by the several agents. These problems are solved either by negotiation or by arbitration (if one excludes the use of force); in the former case, the mode is unsupervised, and in the latter, it is a supervised mode. The relations between Utopia and the negotiation solutions have been extensively studied: the Nash bargaining solution (a Pareto-set solution that is the most equitable), the Kalai-Smorodinsky solution (providing more weighting to the 'needier' player), the Gupta-Liven solution (a Pareto-set solution in the line connecting the reference or focal point solution with Utopia), are just examples.^{31,32}

In our study, this function is not carried out through a negotiation process but by adopting the concept of a 'compromise viewpoint' through the external assumption of what a 'good' compromise should look like—this concept resembles the arbitration process, where an external independent agent with supervisory powers (e.g. a regulator) may decide on the characteristics of compromises among agents by considering criteria (or utility functions) pertaining to all agents in a joint framework.

It is natural to measure the forecast value relative to the income obtained by the market participant. However, the 'optimal' forecast for a WGENCO may not be desired by the SO, because this agent considers, in its own evaluation of solutions, distinct criteria from those that the WGENCOs considers, namely, costs that are not included in the deviation costs of equation (2): for instance, cost of energy not supplied, contracted reserve capacity and environmental costs. The aim in our approach is to offer a WPF training criterion function that suggests an acceptable compromise between the two viewpoints. In the following section, several training criteria described in the literature are presented.

It must be added, however, that an alternative exists: the decoupling of the problem—that is, if the distinct agents work with separate forecasts and do not take into account, for their own decisions, the forecasts used by the other parties. In that situation, each agent can adapt the models and training processes to accommodate its own interests without entering into direct conflict with the others.

3. TRAINING W2P MODELS

3.1. Aspects of training W2P models

A W2P model is a term used to designate a neural network, a fuzzy inference system or, in general, any system that emulates an input–output transfer function and whose performance depends on the tuning of internal weights or parameters, which are used to translate wind characteristics (e.g. wind speed, direction) into electric power.

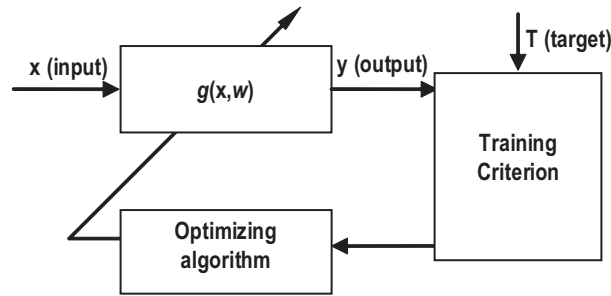


Figure 1. Basic arrangement of a W2P identifying its three main modules.

We can identify three basic modules in a W2P: its internal structure, the training criterion and the optimization algorithm. Three types of actions can be applied in order to constrain the training of a W2P model (see Figure 1) in:

- The internal structure, by modifying the number of weights;
- The training criterion, by selecting an adequate measure of performance related with the forecast error between the WP2 output (y) and target value (T) (e.g. mean square error);
- The optimization algorithm, by choosing a mechanism or procedure (e.g. gradient descent) to close a feedback loop that updates the weights as a function of the training criterion computed for a training dataset.

This paper addresses the aspect of training criterion.

3.2. Training criteria

The traditional MSE and the mean absolute error (MAE) are training criteria with generalized adoption under the ‘forecaster’s paradigm’. Information theoretic learning (ITL) concepts especially challenge the unquestioned adoption of MSE, because this criterion is equivalent to minimizing the variance of the error distribution, and MSE is only an optimal criterion if this distribution is Gaussian—which is not true in most cases, especially not in WPF models.

Bessa *et al.*³³ first introduced the idea of exploring ITL criteria³⁴ to use in training neural networks for WPF (W2P model), where two training criteria from the ITL paradigm were considered. The first training criterion is the minimum error entropy (MEE).³⁵ The basic idea is as follows: if the error distribution of the output would become a Dirac function (meaning that all errors would be equal), we would have reached a predictor whose output would reproduce exactly the actual data series—by just adding to the results a bias corresponding to the mean of the probability density function (*pdf*) of the errors (i.e. the deviation from zero). But it so happens that the Dirac function has minimum entropy. Renyi’s entropy is combined with a Parzen Windows estimation of the error *pdf* to form the basis of MEE under which neural networks are trained.

The second training criterion was derived from an ITL measure named correntropy.³⁶ Correntropy is a generalized similarity measure between two arbitrary scalar random variables, X and Y , defined by:

$$V_{\sigma}(X, Y) = E[k_{\sigma}(X - Y)] \quad (4)$$

where k_{σ} is the kernel function (usually Gaussian).

Correntropy is directly related to the probability of how similar two random variables are along the line $y = x$ in a neighborhood of the joint space defined by the kernel bandwidth σ , and it provides the probability density of the event $p(X = Y)$. The bandwidth controls the observation window in which the similarity is assessed but makes it impossible to assess similarity in the whole joint space. This limitation is actually good, because it leads to the rejection of outliers whose consideration contaminates models. The training criterion consists of maximizing the correntropy between the output y and the target T and was called maximum correntropy criterion (MCC).

These two ITL criteria were only studied in the context of the ‘forecaster’s paradigm’, and no attention was given to the value of the forecast.

Viewing the WPF problem from the ‘forecast user’s paradigm’, for a WGENCO, the training criterion should be translated into economic *value*, which is related with profit in the market and penalization of forecast errors. Hence, the cost function presented in equation (3) could in theory lead to a better ‘forecast’ under this paradigm. An alternative that we propose is to use the *parametric correntropy*,³⁷ which has been used so far in signal processing problems. The idea is

similar to that of correntropy; however, the comparison between X and Y is performed along the line, $aX + b = y$, where a and b are parameters.

The training criteria that we use to train the W2P model (e.g. neural networks) in the WPF problem are summarized as follows:

1. *MSE*. This value is the classical criterion that minimizes the variance of the error distribution and has the form:

$$MSE(e) \Leftrightarrow \min \frac{1}{N} \sum_{i=1}^N \frac{1}{2} e_i^2 \quad (5)$$

where $e = (T_i - y_i)$ is the error of sample i relative to the target value T_i , and N the number of training samples.

2. *MCC*. This criterion is based on a correntropy and may be expressed as:

$$MCC(e) \Leftrightarrow \max \frac{1}{N} \sum_{i=1}^N k_{\sigma}(e_i) \quad (6)$$

where k_{σ} is the Gaussian Kernel with bandwidth σ .

3. *MAE*. This criterion minimizes the absolute error without taking into account the sign of the error:

$$MAE(e) \Leftrightarrow \min \frac{1}{N} \sum_{i=1}^N |e_i| \quad (7)$$

4. *MPC—Minimum Penalty Costs*. This criterion minimizes the error penalties, as shown in equation (3).
5. *MPCC—Maximum Parametric Correntropy Criterion*. This criterion is based on parametric correntropy and may be expressed as:

$$MPCC(e, a) \Leftrightarrow \max \frac{1}{N} \sum_{i=1}^N k_{\sigma}(a \cdot y_i - T_i) \quad (8)$$

where a is a parameter defined by the forecaster or that is computed from the market balancing costs, and b is set to zero because the bias of the forecast can be introduced only with parameter a , without increasing the number of parameters of the problem.

4. CASE STUDIES

4.1. Data characteristics

Three real wind farms with different rated power (ranging from 20 to 50 MW) and situated in different types of terrain are used as case studies.

Data collected in the wind farms include supervisory control and data acquisition registers with an average 'package' containing 10 min of power delivered by the wind farms to the grid. We also have available forecasts produced for the same period by a mesoscale weather model for mean wind speed and wind direction for a reference point in the wind farm, with forecasting horizons ranging from 0 to 48 h in 1 h intervals. These data cannot be discussed in further detail for reasons of confidentiality.

To organize the tests, the available data were divided into three sets. The first set, with several months' worth of data, was used as a training set. The second set, with one month's worth of data, was used as a validation set. And the third set, with the remaining months, was used as a testing set for comparative purposes. Table I presents the months used for training, testing and validation of the neural network.

Table I. Training, validation, and testing sets.

Wind farm	Training	Validation	Testing
A	Feb–Dec (2007)	Jan (2007)	Jan–Dec (2008)
B	Mar–Dec (2007)	Feb (2007)	Apr–Dec (2008)
C	Feb–Aug (2007)	Jan (2007)	Sep–Dec (2007)

Table II. Annual mean market prices.

	2007	2008
Spot price	39.34	64.47
Up balancing	2.97	4.04
Down balancing	7.99	8.02

Data from the Iberian electricity market were used in this study. The data consist of hourly spot and balancing prices from 2 years, 2007 and 2008, and was collected from e-sios (the information system of the system operator of Red Eléctrica de España, www.esios.ree.es).

A very complete statistical analysis of the price data used in this study can be found in Bludszuweit³⁸ (see chapter 5). Moreover, the author analysed the correlation between the market prices (spot and balancing) and the wind generation for the same data used in this paper. The author concluded that wind power is still having only a limited impact on the spot prices in the Iberian market; however, it has already made an impact on the balancing prices. Table II shows the annual mean energy prices in 2007 and 2008 for the Spanish market (in €/MWh).

The balancing prices are very asymmetric, where the down balancing is more expensive. Also evident is a considerable change in prices from one year to the next, but this variation does not affect the results of our study—what is relevant is the asymmetry in balancing prices.

4.2. Training characteristics

We have trained a feedforward multilayer perceptron (MLP) neural network with only one hidden layer consisting of seven neurons by using a hyperbolic tangent activation function. The inputs of the MLP are NWP forecasted values: mean wind speed values and mean wind direction values. Because of the cyclic character of the wind direction, this variable comprised two components (i.e. the sine and cosine components). This means a total of three input variables, which were standardized by using the min–max method, were used.

Because the purpose is to compare results, we do not describe details of the training here; however, the training details can be found in Bessa *et al.*³³ Moreover, determining the best topology of the neural network is beyond the scope of this work.

The penalization factors of the training criterion MPC (equation (3)) consist in average values of balancing prices for each look-ahead time step (in this case, there are 24 time steps) computed from year 2007 data. Hence, the diurnal cycle of the prices is included in the model. Regarding the parameter a of the MPCC criterion, the adopted methodology is:

- 1 In the beginning of each training epoch, the parameter a is estimated from the training dataset by using the average value of balancing prices of 1 year for each look-ahead time step as penalization factors of underestimation (w_{down_reg}) and overestimation (w_{up_reg}). The training criterion is:

$$\max_a \begin{cases} k_{\sigma'}((a \cdot y_i - T_i) \cdot w_{down_reg}), y_i < T_i \\ k_{\sigma'}((a \cdot y_i - T_i) \cdot w_{up_reg}), y_i > T_i \end{cases} \quad (9)$$

where σ' is the kernel bandwidth and depends on the size of the penalization factors.

- 2 After estimating the parameter a , the following training criterion is maximized:

$$\max_w k_{\sigma}(a' \cdot y_i - T_i) \quad (10)$$

where σ is a kernel bandwidth much smaller than σ' , and a' is equal to $1/a$.

The kernel size σ used by MCC is the same size used in Bessa *et al.*³³ The kernel size σ in MPCC was found in the validation set to be equal to 0.1, and σ' was equal to 3. As expected, the kernel size does not vary widely from one wind farm to another. Yet, σ' presents a major importance in this problem, because small values indicate small sensitiveness to asymmetric costs. What is desired is a value that gives importance to the asymmetric costs but that also maintains a balance between over- and underestimation.

The same values are used for the three wind farms since what is desired in WPF is to train a neural network for any wind farm without concern about tuning parameters.

4.3. Testing methodology

The neural network was used to predict the power $p_{t+k|t}$ produced by the wind farm at time stamp 7:00 GMT (when new NWP predictions become available) for each look-ahead $t + k$ of the next day. The wind power prediction was performed for each day of the test dataset.

We have taken each wind farm and simulated its participation in the electricity market in the year 2008, offering the power prediction generated by prediction systems trained with different training criteria and using the hourly spot and balancing prices from 2008.

4.4. Analysis of the results

In this section, we present results for training of neural networks in three real wind farms and one electricity market, when training under the different criteria.

Table III–V summarize the results obtained from the participation of wind farms A, B and C in the Spanish market with bids equal to the forecasts provided by each training criterion. $Q_{95\%}$ is the quantile 95% of the overestimation errors distribution, which means that the probability of having an upward deviation larger or equal to this value is 5%; the conditional tail expectation (CTE) of this quantile can be interpreted as the expected maximum forecast deviation.

4.4.1. Wind generation companies' viewpoint

The first three training criteria are not economically oriented. One may see that wind farm A realized a lower level of income when using them than it did with MPC and MPCC. The comparison of the results from MPC with the first three

Table III. Simulation of wind farm A's participation in the Iberian market.

	MSE	MCC	MAE	MPC	MPCC
Surplus (% of produced energy)	33.79	35.55	39.08	19.71	30.89
Shortage (% of produced energy)	33.74	29.64	24.37	58.07	35.56
Total deviations (% of produced energy)	67.54	65.21	63.47	77.79	66.46
Deviations against the system (% of produced energy)	30.09	28.68	28.21	34.97	29.02
$Q_{95\%}$ (% of rated power)	41.74	41.70	43.58	61.50	48.30
CTE (% of rated power)	53.67	58.05	57.89	73.28	63.26
Total income (deviation from the best) (k€)	−10.16	−7.61	−12.47	−4.1	0.0

Table IV. Simulation of wind farm B's participation in the Iberian market.

	MSE	MCC	MAE	MPC	MPCC
Surplus (% of produced energy)	34.09	34.13	39.29	18.21	25.85
Shortage (% of produced energy)	25.16	24.14	19.67	48.31	34.37
Total deviations (% of produced energy)	59.25	58.28	58.96	66.52	60.23
Deviations against the system (% of produced energy)	26.33	26.10	26.41	28.83	26.49
$Q_{95\%}$ (% of rated power)	28.19	30.22	26.98	40.67	32.92
CTE (% of rated power)	34.21	40.27	34.78	49.78	42.13
Total income (deviation from the best) (k€)	−20.09	−20.43	−30.92	0.0	−7.45

Table V. Simulation of wind farm C's participation in the Iberian market.

	MSE	MCC	MAE	MPC	MPCC
Surplus (% of produced energy)	34.46	36.91	37.63	19.58	24.62
Shortage (% of produced energy)	19.71	16.09	15.02	33.12	27.42
Total deviations (% of produced energy)	54.17	53.00	52.64	52.70	52.04
Deviations against the system (% of produced energy)	23.07	22.88	23.14	20.54	21.13
$Q_{95\%}$ (% of rated power)	31.08	34.00	38.34	46.07	39.44
CTE (% of rated power)	38.33	44.42	46.88	52.42	48.21
Total income (deviation from the best) (k€)	−14.2	−15.35	−16.59	0.0	−4.78

training criteria (which are typical of the forecaster paradigm) illustrates that a forecast of worse quality (i.e. larger deviations) results nevertheless in higher income for a WGENCO. By comparing MCC with MPCC, it is evident that accepting a small increment in the error leads to an increase of around 7600 € in revenue. In addition, for wind farm A and over the 1 year test period, the forecasts generated by the MPCC led to the highest profits (around 4100 € higher) than did those forecasts obtained with MPC.

For wind farm B, the first three training criteria also led to lower income levels. In this case, MPC presents a higher profit potential than does MPCC (by around 7500 €).

The same conclusions for wind farm B are also valid for wind farm C. The economic-oriented training criteria led to higher levels of income, and once more, the MPC realized a higher level of income than did the MPCC.

From the results of the three wind farms, it becomes clear that using MPCC instead of any statistical criterion enables increases to the income without significantly increasing the deviations resulting from forecast errors.

4.4.2. System operator viewpoint

For wind farm A, the forecasts generated with MCC and MAE produce fewer total deviations and deviations against the system. MCC and MAE underestimate the wind generation, but despite this underestimation, a higher income was achieved with MCC when compared with MSE. The MSE criterion places a heavy penalty on large errors, and therefore, the CTE for MSE presents the lowest value of all of the training criteria—this result means that wind farm A is usually perceived as a good property by the SOs.

It is therefore satisfactory that the new criterion, MPCC, could lead as well to a high level of income without significantly increasing the total deviations and the deviations against the system. On the other hand, the MPC has a higher bias on the side with lower balancing costs (e.g. a shortage of generation), which is in opposition to the interests of the SO in terms of security of the system.

The MPCC for wind farm A is near the 'compromise viewpoint', as it allows some increase in the income without a major increase in the total deviation. The comparison of MPC and MPCC in the quantile 95% and CTE values also show that MPCC is better for the SO when compared with MPC.

For wind farm B, the first three training criteria led to lower total deviations and deviations against the system. The tendency for this wind farm is to underestimate the wind generation. For this wind farm, the training criterion with fewer total deviations was MCC, which is surprising because MAE was expected to achieve fewer total deviations. Once more, MSE presents the lowest value of CTE. MPC presents a higher realized income than does MPCC; however, MPCC allowed (or enabled) increases in realized income, while the total number of deviations and deviations against the system are maintained at an acceptable level.

The same behavior is verified for wind farm C, where, while the MPCC realized a lower income than did the MPC, the total deviation values and CTE are lower. In this case, MPC presented a lower value for deviations against the system.

5. CONCLUSIONS

This paper addresses the WPF problem by recognizing that the value and quality of forecasts must be appreciated in the framework that decision making is carried out by multiple agents and that agreement may not be possible among the agents (WGENCOs or SOs) regarding what constitutes a 'good' forecast. The issues discussed—the relation between forecast error and market profit—also occur in situations involving more complex forecasting systems and electricity market rules. The results in this paper provide evidence that forecasts with higher accuracy (in the forecaster's paradigm sense) do not necessarily lead to higher incomes for a WGENCO. As such, a WGENCO may be willing to reduce the forecasts' accuracy in exchange for an increased income.

The main contributions of the paper are summarized in the following:

- (i) A challenge to the assumption of the mathematical neutrality of forecasting models.
- (ii) A thorough discussion of conflicting interests associated with the 'goodness' of forecasts for two types of agents in an electricity market.
- (iii) A comparison of the relative advantages and drawbacks of several criteria referred to in the literature and used in training WPF models.
- (iv) A new approach proposed to achieve acceptable 'compromise' solutions to market participants in the choice of a forecasting model by introducing the ITL concept of parametric correntropy.
- (v) The verification of the theoretical concepts in a representative example, namely, the Iberian market, MIBEL.

What is the significance of the results shown in this paper for future forecasting practices and market policies?

First, these results reveal the importance of the choice of a training criterion for a forecasting tool. In this sense, it is not necessarily true that it is impossible to achieve a high remuneration in the electricity market and maintain the forecast error within an acceptable level. Methods that produce acceptable compromises are available.

Second, in systems with high wind penetration levels, it is important to look at the impact of the wind generation market bids on the balancing needs. In fact, the market rules and the way that wind power is remunerated in the market should be revisited and possibly redesigned to better facilitate the decisions made by the SO regarding the security of the system. This paper shows that an asymmetry in balancing prices makes possible the opportunity for profit by those forecaster's paradigm WGENCOs that make careful selection of the characteristics of the output of the forecasting systems. Whenever there is an asymmetry in the price structure, whether in balancing prices or spot prices (which is more likely to occur in cases of large penetration of wind power, with larger impacts on price volatility), a neutral model (in the forecaster's paradigm sense) may not be the model that leads to maximizing the profits of WGENCOs that bid in the market.

This paper also points toward areas for wider discussion, namely, toward an analysis on possible reactions of other agents in the market to wind power bids, especially if indications appear that bids do not correspond to a 'neutral' forecast (in the *forecaster paradigm* sense), which could lead to speculation against admittedly biased wind power forecasts.

One possible evolution that the results in this paper suggest would be to decouple the concepts of forecasting and bidding. In some markets these days, SOs demand that WGENCOs bid the forecast they have produced—and then trust it as a valid forecast. However, WGENCOs may wish to develop a bidding strategy, and this impetus leads to the possibility of adopting models that, while seeming to be mathematically 'neutral', will have characteristics that serve their particular market strategy—and will possibly be contrary to SOs' interests. So it follows naturally that WGENCOs may be allowed to bid in an unrestricted manner on whatever power value they prefer, but that SOs will also have their own independent forecasts and will act according to those and not solely to the forecasts announced by WGENCOs in order to ensure reliability in the power system.

ACKNOWLEDGEMENTS

The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory ('Argonne'). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up non-exclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.

The author Ricardo J. Bessa acknowledges Fundação para a Ciência e a Tecnologia (FCT) for PhD Scholarship SFRH/BD/33738/2009.

REFERENCES

1. Monteiro C, Bessa R, Miranda V, Botterud A, Wang J, Conzelmann G. Wind power forecasting: state-of-the-art 2009. Technical report ANL/DIS-10-1. Argonne National Laboratory, 2009. [Online]. Available: <http://www.dis.anl.gov/projects/windpowerforecasting.html>. (Accessed 10 October 2010)
2. Costa A, Crespo A, Navarro J, Lizcano G, Madsen H, Feitosa E. A review on the young history of the wind power short-term prediction. *Renewable and Sustainable Energy Reviews* 2008; **12**: 1725–1744.
3. Ernst B, Oakleaf B, Ahlstrom ML, Lange M, Moehren C, Lange B, Focken U, Rohrig K. Predicting the wind. *IEEE Power & Energy Magazine* 2007; **5**: 78–88.
4. Madsen H, Pinson P, Kariniotakis H, Nielsen HA, Nielsen TS. Standardizing the performance evaluation of short-term wind prediction models. *Wind Engineering* 2005; **29**: 475–489.
5. Industry Work Group. Wind power forecasting pilot project—industry work group report. Technical Report, 2008. [Online]. Available: <http://www.aeso.ca/gridoperations/13825.html>. (Accessed 10 October 2010)
6. Murphy AH. What is a good forecast? An essay on the nature of goodness in weather forecasting. *Weather and Forecasting* 1993; **8**: 281–293.
7. Murphy AH, Winkler RL. A general framework for forecast verification. *Monthly Weather Review* 1987; **115**: 1330–1338.
8. Pinson P. Estimation of the uncertainty in wind power forecasting. PhD Thesis, Ecole des Mines de Paris, Paris, France, 2006.
9. Stewart TR, Katz RW, Murphy AH. Value of weather information: a descriptive study of the fruit-frost problem. *Bulletin of the American Meteorological Society* 1984; **65**: 126–137.
10. Usaola J, Ravelo O, González G, Soto F, Dávila MC, Díaz-Guerra B. Benefits for wind energy in electricity markets from using short term wind power prediction tools; a simulation study. *Wind Engineering* 2004; **28**: 119–127.
11. Barthelmie RJ, Murray F, Pryor SC. The economic benefit of short-term forecasting for wind energy in the UK electricity market. *Energy Policy* 2008; **36**: 1687–1696.

12. Angarita-Márqueza JL, Hernandez-Aramburo CA, Usaola J. Analysis of a wind farm's revenue in the British and Spanish markets. *Energy Policy* 2007; **35**: 5051–5059.
13. Pinson P, Chevallier C, Kariniotakis G. Trading wind generation with short-term probabilistic forecasts of wind power. *IEEE Transactions on Power Systems* 2007; **22**: 1148–1156.
14. Fabbri A, Gomez T, Abbad JR, Quezada VH. Assessment of the cost associated with wind generation prediction errors in a liberalized electricity market. *IEEE Transactions on Power Systems* 2005; **20**: 1440–1446.
15. Matevosyan J, Söder L. Minimization of imbalance cost trading wind power on the short-term power market. *IEEE Transactions on Power Systems* 2006; **21**: 1396–1404.
16. Bourry F, Juban J, Costa LM, Kariniotakis G. Advanced strategies for wind power trading in short-term electricity markets. In *Proceeding of the European Wind Energy Conference & Exhibition EWECE 08*, Brussels, Belgium, March 31–April 3, 2008.
17. Linnet U. Tools supporting wind energy trade in deregulated markets. MSc Thesis, Technical University of Denmark, Lyngby, Denmark, July 2005.
18. Katz RW, Murphy AH. Quality/value relationships for imperfect weather forecasts in a prototype multistage decision-making model. *Journal of Forecasting* 2006; **9**: 75–86.
19. Bremnes JB. Probabilistic wind power forecasts using local quantile regression. *Wind Energy* 2004; **7**: 47–54.
20. Botterud A, Wang J, Bessa R, Keko H, Miranda V. Risk management and optimal bidding for a wind power producer. In *Proceedings of the IEEE PES General Meeting 2010*, Minneapolis, Minnesota, USA, 2010.
21. Ravn HF. Short term wind power prognosis with different success criteria. In *Proceedings of International Conference on Probabilistic Methods Applied to Power Systems—PMAPS 2006*, Stockholm, Sweden, June 11–15.
22. Botterud A, Wang J, Miranda V, Bessa R. Wind power forecasting in U.S. electricity markets. *The Electricity Journal* 2010; **23**: 71–82.
23. Verhaegen K, Meeus L, Belmans R. Development of balancing in the internal electricity market in Europe. In *Proceedings of the European Wind Energy Conference & Exhibition EWECE 06*, Athens, Greece, 2006.
24. Rebours Y, Kirschen DS. A survey of definitions and specifications of reserve services. Technical Report, University of Manchester, Oct 2005.
25. Doherty R, O'Malley M. New approach to quantify reserve demand in systems with significant installed wind capacity. *IEEE Transactions on Power Systems* 2005; **20**: 587–595.
26. Matos MA, Bessa R. Operating reserve adequacy evaluation using uncertainties of wind power forecast. In *Proceedings of the IEEE PowerTech Conference*, Bucharest, Romania, 2009.
27. Holttinen H. Impact of hourly wind power variations on the system operation in the Nordic countries. *Wind Energy* 2004; **8**: 197–218.
28. Doherty R, Denny E, O'Malley M. System operation with a significant wind power penetration. In *Proceedings of IEEE Power Engineering Society General Meeting*, 2004.
29. Ilic MD, Makarov Y, Hawkins D. Operations of electric power systems with high penetration of wind power: risks and possible solutions. In *Proceedings of IEEE Power Engineering Society General Meeting*, Tampa, Florida, USA, Jun 2007.
30. UCTE NetWork of experts on Wind Power. Wind power in the UCTE interconnected system. UCTE Position Paper, 2004.
31. Van Damme E. *Stability and Perfection of Nash Equilibria*. Springer-Verlag: Berlin, 1991.
32. Faratin P, Sierra C, Jennings NR. Using similarity criteria to make issue trade-offs in automated negotiations. *Artificial Intelligence* 2002; **142**: 205–237.
33. Bessa R, Miranda V, Gama J. Entropy and correntropy against minimum square error in off-line and on-line 3-day ahead wind power forecasting. *IEEE Transactions on Power Systems* 2009; **24**: 1657–1666.
34. Principe JC, Xu D, Fisher J. Information theoretic learning. In *Unsupervised Adaptive Filtering*. Simon H (ed.). Wiley: New York, 2000. 265–319.
35. Erdogmus D, Principe JC. Generalized information potential criterion for adaptive system training. *IEEE Transactions on Neural Networks* 2002; **13**: 1035–1044.
36. Liu W, Pokharel P, Principe JC. Correntropy: properties and applications in non-Gaussian signal processing. *IEEE Transactions on Signal Processing* 2007; **55**: 5286–5298.
37. Xu J. Nonlinear signal processing based on reproducing Kernel Hilbert space. PhD Thesis, University of Florida, Gainesville, Florida, USA, 2007.
38. Bludszuweit H. Reduction of the uncertainty of wind power predictions using energy storage. PhD Thesis, Universidad de Zaragoza, Spain, 2009.