

# Wind Power Scenario Tree Tool: Development and Methodology

Colm Lowery and Mark O'Malley

## 1 Introduction

The nature of power system operation is changing worldwide. Plans are in place to increase the proportion of demand met through wind power throughout the European continent [1], Ireland [2], the Great Britain [3] and the United States [4]. In Europe, for example, this change is driven by EU policy which aims to reduce CO<sub>2</sub> emissions and dependency on imported fuel. Primarily, this policy encourages the growth of renewable generation including wind power [5]. It is expected that this increase in renewable generation will displace conventional generation—causing a decrease in system operation costs as less conventional fuel is consumed in meeting system demand.

However, displacing conventional generation with non-synchronous renewable generation is creating new system operation challenges. Wind is by its nature variable. It is of limited predictability, and control of the resource is limited to curtailment. As wind cannot be forecast with perfect accuracy, additional reserve must be carried and conventional units must be operated in a more flexible, adaptable manner—leading to reduced efficiency through partial loading and an increase in the number of start-ups required for conventional power plants [6]. While the effects of this uncertainty can be disregarded for low wind penetrations, with high wind penetrations, the uncertainty associated with wind forecasting error will impact upon the reliability, efficiency and economic performance of unit commitment [7, 8].

Ideally, the system will be scheduled in such a way that the expected value of the operating costs is minimized given the uncertainty of wind generation and the

---

C. Lowery (✉) · M. O'Malley

Electricity Research Centre, University College Dublin, Dublin 4, Ireland  
e-mail: Colm.Lowery@ucdconnect.ie

M. O'Malley

e-mail: mark.omalley@ucd.ie

constraints on the system. This can be approximated by treating unit commitment as a stochastic problem where the distribution of possible wind generation is represented by a number of probability-weighted time series scenarios. In many forms of stochastic unit commitment, the operational cost of a unit commitment is evaluated for each of these scenario tree branches, similar to deterministic treatments of wind generation [9]. However, it is the expected cost of the entire scenario tree which is the objective to be minimized for the unit commitment.

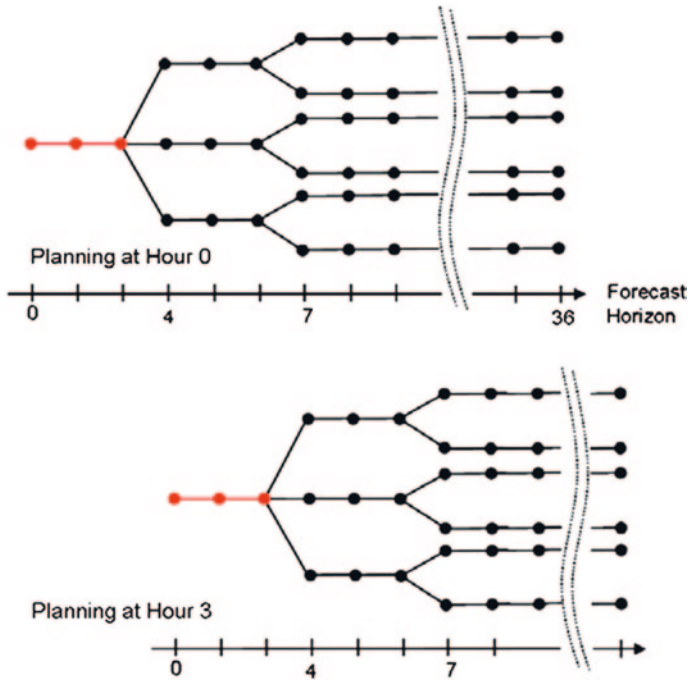
While stochastic treatments of wind uncertainty have previously been investigated, the contributions to solution quality of specific statistical properties of forecast error have not been considered. These properties are available from analysis of the error between historical forecasts and realized data. In ideal circumstances, a system operator will have the necessary information prioritized in the form best suited for decision making. In practice, and in many prior studies, quantification of forecast error has often been limited to simple statistical properties such as variance and mean error and an assumed Gaussian distribution [10]. While traditional measures of forecast performance such as mean average error or root-mean-square error do not distinguish between distributions with the same mean and variance, more sophisticated methods of analysis of information content such as Renyi entropies indicate that significant information is missed by these simplifications [11, 12].

The WILMAR project developed a stochastic unit commitment scheduling model to analyse the integration of wind power in a large liberalized electricity system [13]. The WILMAR scheduling model uses a rolling sequence of scenario tree forecasts to model the impact of error information (see Fig. 1) This model has previously been used to study the benefits of stochastic treatment of wind uncertainty over deterministic treatments as well as the impact of accounting for more of the wind uncertainty by increasing the frequency of rolling planning [8].

This chapter presents the methodology adopted for use in a Scenario Tree Tool (STT) constructed to allow for closer examination of the effect of forecast error assumptions and properties when combined with a suitable unit commitment model such as WILMAR [14]. Due to the requirement for direct control of the statistical properties in several areas of interest, ARMA series/scenario reduction methodologies were not considered suitable for this tool. A new STT was designed using a methodology based on a moment-matching technique where each time period, within a tree, has a defined variance, skewness and kurtosis. These statistics together with the autocorrelations of the scenario determine the values of the scenarios at that time period. While alternative heuristic methodologies have been proposed for deriving a scenario tree that matches specific moments [15], the methodology used in this tool is based on a nonlinear optimization moment-matching method [16, 17].

Section 2 briefly describes the sections of the STT and details the structure of the programme.

Section 3 describes the methodologies related to the STT. It details the methodologies (moment matching and scenario grouping) used to generate wind and load scenarios within the current STT as well as the method used to account for diurnal/seasonal variations. In addition to describing the new methodology, it also briefly describes the preceding scenario reduction/ARMA series method which was used by older versions of the STT.



**Fig. 1** Illustration of rolling planning scenario trees

[Section 4](#) covers the methodologies used for calculating the replacement reserve, the amount of reserve required to be available after a short delay, from these trees as well as detailing, for completeness, the assumptions about spinning reserve, the amount of reserve required to be available from online generators at a given time and the semi-Markov chain methodology used to simulate forced outages as required by the replacement reserve methodology.

[Section 5](#) displays example graphical outputs from the preceding sections.

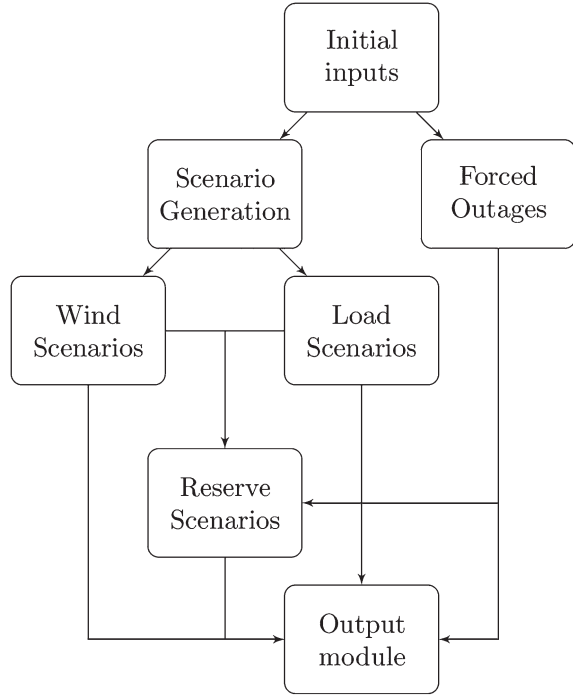
A conclusion is presented in [Sect. 6](#).

## 2 Structure of the Scenario Tree Tool

The overall structure of the STT can be seen in [Fig. 2](#). Each functional box represents a separate module or submodule in the STT, which will be covered in the subsequent sections. Entries on the same level represent parallel operations which can be run simultaneously. The detail of these modules is as follows:

- **Initial Inputs:** this module loads and prepares data from input libraries.
- **Scenario Generation:** this module utilizes forecast error information to generate wind and load power scenarios and branching structures. This module contains multiple submodules detailed in [Sect. 3](#) and [Fig. 3](#).

**Fig. 2** Structure of the scenario tree tool



- **Forced Outages:** this module utilizes semi-Markov Chains to calculate a random sequence of plant availability (Sect. 4.2).
- **Reserve Scenarios:** this module uses the outputs of the scenario generation and forced outage modules, along with plant capacity information to calculate replacement and spinning reserves (Sect. 4.1).
- **Output Module:** this module collates and converts the information from the other module into input files for the WILMAR model.

### 3 Development and Methodology of the Scenario Tree Tool

In power plant scheduling, decisions must be made on both information known with certainty and uncertainty which must be forecasted. In stochastic unit commitments, wind forecasts, demand forecasts and their forecast error can be accounted for by representative branching trees consisting of probability-weighted scenarios for available wind generation (Fig. 1). The number of scenarios has been restricted to minimize dimensionality in the unit commitment solution while retaining accuracy in the specified statistical information. For computational reasons, the first stage of these trees is assumed to be known with perfect certainty.

Each scenario in the tree consists of an assigned probability and two time series values—one for load and one for wind. Each of these time series can be

represented as a forecast time series, common to all scenarios within the tree, and an error time series which, together with the error time series of the other scenarios, represents the error distribution of the forecast.

In the All-Island Grid Study (AIGS), these scenario trees were generated by an STT using a methodology based on scenario reduction in ARMA series [18]. Specifically, the wind forecast error was based on the assumption of a Gaussian distribution of wind speed error with standard deviation dependent upon the forecast horizon [19, 20].

In brief, scenario reduction aims to take an initially large number of scenarios and then remove scenarios until only the desired number remains while maintaining as close a representation of the original distribution as possible according to the methodology given here:

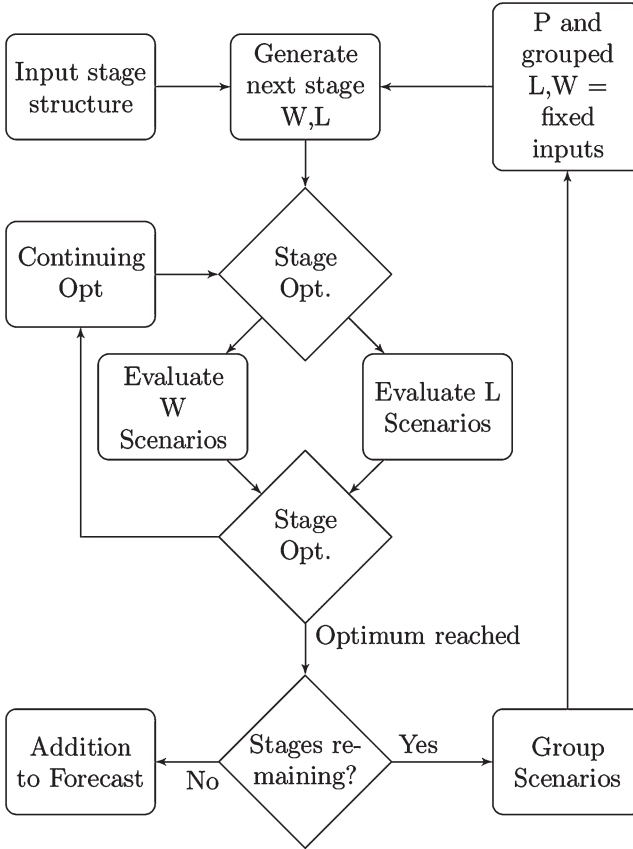
1. Generate 1,000 wind speed scenarios, with a length of 36 periods, using ARMA (1,1) series of equal probability [21].
2. Add these scenarios to the wind speed forecast.
3. Convert these wind speed scenarios to wind power forecasts using a normalized aggregated multi-turbine curve.
4. Scale these scenarios to the required level taking account of the spacial correlation of the wind power forecast [22].
5. Calculate the Kantorovich distance between each pair of scenarios [23, 24].
6. Merge the probability of the scenarios with the lowest Kantorovich distance and delete the scenario of smaller probability.
7. Repeat steps 2 and 3 until the desired number of scenarios is achieved.
8. To create the branching scenario tree, this process is repeated omitting the hours in the third stage to reduce the scenarios further.

Due to the requirement for direct control of the statistical properties in several areas of interest, ARMA series/scenario reduction methodologies were not considered suitable for this tool. A new STT was designed using a methodology based on a moment-matching technique where each time period, within a tree, has a defined variance, skewness and kurtosis.

These statistics together with the autocorrelations of the scenario determine the values of the scenarios at that time period. While alternative heuristic methodologies have been proposed for deriving a scenario tree that matches specific moments [15], the methodology used in this tool is based on a nonlinear optimization moment-matching method [16, 17].

The overall structure of scenario generation can be seen in Fig. 3 where W is wind scenarios, L is load scenarios and P is the probability of each scenario. The steps in brief are as follows:

1. Set initial parameters and begin from the final stage of the tree.
2. Generate the scenarios for the current stage of the tree using moment matching (II.A).
3. Calculate the degree of similarity between each subset (size determined by the ratio of branching) in terms of Kantorovich distance and autocorrelations (II.B).



**Fig. 3** Structure of the scenario generation

4. Generate an optimal branching structure between the current stage and its predecessor using the subsets which minimize the disparity between merging scenario branches.
5. Repeat the preceding steps for the earlier stages using the branching structure and scenario values as further inputs into the objective function (1).

### 3.1 Moment Matching

For each stage in the tree, a nonlinear optimization was used to produce a matrix of scenarios consisting of wind, demand and probability values matching the specified statistics as closely as possible. In addition, stages which have already been

determined are used to provide additional autocorrelation information for subsequent stages.

In each stage, the optimization acts to minimize the following objective function (1) while meeting the constraints (4, 5, 6) given the stages and branch connections that have already been defined:

$$\sum_{n=1}^S W_n (f_n(x_t, P_t) - S_n)^2 \quad (1)$$

where  $S$  is the set of statistical properties under consideration.  $W_n$  is the optimization weight assigned to the statistical property.  $S_n$  are the components of the set  $S$  (mean, variance, kurtosis, skewness, the first four autocorrelations).  $x_{t,j}$  is the set of values for scenario  $j$  during the given time step  $t$ .  $P_t$  are the probabilities assigned to the scenarios in time step  $t$ .

The individual components ( $f_n(x)$ ) of the objective function (1) are calculated using the following formulas for the moment (2) and autocorrelation (3):

$$M_l = \sum_{j=1}^s P_j (x_{j,t} - \mu_t)^l \quad (2)$$

$$AC_{\tau,j} = \sum_{j=1}^{N-\tau} \frac{(Y_i - \mu_j)(Y_{i+\tau} - \mu_j)}{\sum_{j=1}^N (Y_i - \mu_j)^2} \quad (3)$$

where  $M_l$  is the  $l$ th moment at time  $t$ .  $AC_{\tau,j}$  is the autocorrelation at time lag  $\tau$  for scenario  $j$ .  $\mu_t$  is the mean of the scenario set at time  $t$ .  $\mu_k$  is the mean of the scenario  $k$ , and  $\tau$  the time lag of the autocorrelation.  $k$  is the index of a given scenario.  $jk$  are the indices of the scenarios which branch from scenario  $k$ .

In order to ensure probability remains consistent across scenarios and time periods, the constraints upon the objective function are as follows:

$$\sum P_{jk,t} = P_{k,t-1} \quad (4)$$

$$\sum P_t = 1 \quad (5)$$

$$P_{jk,t} > 0 \quad (6)$$

### 3.2 Scenario Grouping

The scenario stages produced by moment matching are divided into groupings according to the branching of the scenario tree and the similarity of their autocorrelations. This is done to ensure that when joined to the relevant scenario in the preceding stage, the autocorrelation of each branch of the tree is consistent across the stage boundaries. This prevents the creation of suboptimal scenario branching as a result of joining two disparate scenarios. Scenario ordering is determined by finding the ordering which minimizes the difference in autocorrelation and the Euclidean distance between the short term (stage II) values of each scenario.

$$D(j1, j2) = E_D(j1, j2) + AC_D(j1, j2) \quad (7)$$

$$E_D(j1, j2) = \sum_{t=1}^T \frac{(x_{j1,t} - x_{j2,t})^2}{(x_{j2,t})^2} \quad (8)$$

$$AC_D(j1, j2) = \sum_{z=1}^{\tau} \frac{(AC_{\tau,j1} - AC_{\tau,j2})^2}{(AC_{\tau,j1})^2} \quad (9)$$

where  $D(j1, j2)$  is the measure of similarity,  $E_D(j1, j2)$  is the Euclidean distance and  $AC_D(j1, j2)$  is the difference in autocorrelation.

These groupings and scenarios are then used in the optimization of any subsequent moment-matching steps to provide consistent autocorrelations for each scenario throughout the tree stages.

While the WILMAR model uses tree structures which branch evenly from each node at a time point and never merge, the scenario tree methodology allows for uneven branching and hence can generate topologies which have different scenario resolutions at different time horizons and in different regions of the tree.

### 3.3 Diurnal and Seasonal Variation

The word “data” is plural, not singular. Wind and load forecasts exhibit different behaviour at different times of the year and day. In order to account for seasonal and diurnal variation, the error trees can be altered before being added to the individual forecasts by the addition of a mean error adjustment time series,  $\mu(j)$  to each scenario and the use of a scaling constant  $\varepsilon$ .

The values of  $\mu(j)$  and  $\varepsilon$  are chosen through fitting the general forecast error to the forecast error of each seasonal period and time period within the day. This is achieved by subdividing the known forecast information by seasonal period, calculating the same metrics as the main data and deriving the necessary adjustments from the proportional difference in standard deviation and mean between the general case and the specific case. This process is repeated for the starting hour of the forecast to derive the adjustment for time of day.



## 4 Replacement Reserve and Forced Outages

### 4.1 Replacement Reserve

Replacement reserve is estimated for each individual scenario branch of wind and demand within the tree and is calculated using a similar methodology to that presented in the AIGS [18]. However, the AIGS used methods assuming prior scenario reduction, which calculate reserve from the 90th percentile of the unreduced set of scenarios used to generate the scenario tree. This method can therefore not be used with the new STT as the new methodology uses moment matching and does not generate the large initial number of scenarios required by the AIGS methodology. Instead, the reserve is calculated from the 90th percentile of a distribution based on the variation in wind and demand error represented by those scenarios (10) according to the following method:

The  $n$  scenario combinations are mapped onto the wind scenario  $s$  from which they came, and the reserve is taken to be the 90th percentile of the difference between the reference power balance  $P_{Ref}$  (11) and the forecasted power balance,  $P$ , for the scenario combinations  $n$  at each wind scenario  $s$  (12):

$$\Delta P(t_0, f, s) = P_{Ref}(t) - P(t_0, f, s) \quad (10)$$

The power balance  $P_{Ref}$  at time  $t$  is calculated for the realized values of wind  $W_R$ , load  $L_R$  and available conventional capacity  $C$ :

$$P_{Ref}(t) = \sum_{g \in G} C(g) + W_R(t) - L_R(t) \quad (11)$$

The power balance  $P$  of each combination of wind and load scenarios  $n$  is calculated for the scenario wind  $W_E$  and load  $L_E$  as well as for the forced outages. Time  $t$  is calculated for the realized values of wind  $W_R$ , load and available conventional capacity  $C$ :

$$P(r, t_0, f, n) = \sum_{g \in G} C(g)Y(g, t) + W_E(t_0, f, n) - L_E(t_0, f, n) \quad (12)$$

In addition to this, required spinning reserve is estimated from the largest infeed to the system and forecasted wind using the methodology presented in the AIGS [10].

### 4.2 Forced Outages

While no new changes have been made to this methodology, it is included here for completeness as it is necessary for the reserve calculations as well as being a required input for the unit commitments it is currently in use with. Within the model, distinction is drawn between two types of outage (forced outages and

scheduled outages). As scheduled outages are defined by the user and by the test system, the STT simulates forced outages from the information provided about the test system including the information provided on scheduled outages.

The time series of forced outages for each conventional unit are simulated using semi-Markov chains [25], where the failure and repair rates are expressed by the mean time to failure and the mean time to repair. The methodology used in the STT is derived from that presented in the AIGS [18] and is detailed in brief below:

$$P_{\text{Available}} = \frac{\text{MTTF}}{\text{MTTR} + \text{MTTF}} \quad (13)$$

$P_{\text{Available}}$  is the probability that a plant will be available derived from the mean time to failure (MTTF) and mean time to repair (MTTR) for each plant.

$$P_{\text{Unavailable}} = \frac{\text{MTTR}}{\text{MTTR} + \text{MTTF}} = \text{FOR} \quad (14)$$

$P_{\text{Unavailable}}$  is the probability that a plant will be unavailable and is equivalent to the forced outage rate (FOR.)

$$\text{FOR} = \frac{t_{fo}}{t} \quad (15)$$

$t_{fo}$  is the total number of hours the plant is unavailable, and  $t$  is the total time period.

Figure 4 shows the structure of the forced outage algorithm. The initial state of the unit is determined by drawing a random number from a uniform distribution and comparing it with the full outage percentage (FOP) calculated according to Eq. (16).

$$\text{FOP} = \frac{\text{FOR}}{\text{MTTR}} \quad (16)$$

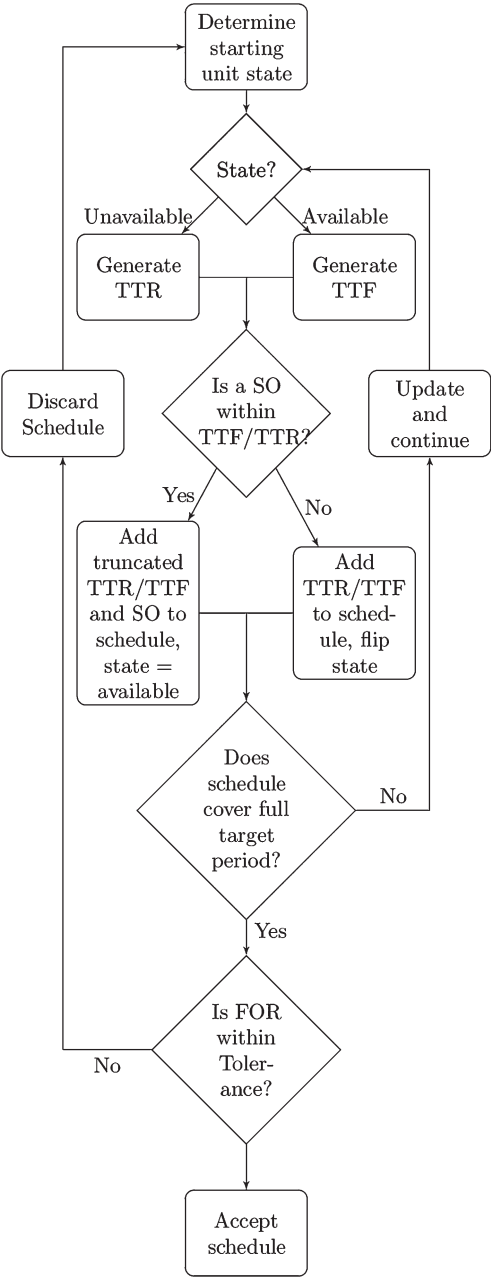
The result of this determines whether or not the model will draw a time to repair (TTR) or a time to failure ((TTF)) from the Weibull distribution.

$$f(t, \mu, k) = \frac{k}{\mu} \left( \frac{t}{\mu} \right)^{k-1} \cdot e^{-\left( \frac{t}{\mu} \right)^k} \quad (17)$$

where the shape factor  $k = 1$  is used for the TTF (as time to failure is captured by an exponential distribution),  $k = 5$  is used for the TTR (time to failure represented by a bell-shaped distribution) and  $\mu$  is calculated from MTTF and MTTR, respectively.

In the event that a forced outage extends into the period of a scheduled outage, the scheduled outage is run for its full time period as the scheduled repairs cannot be assumed to be equivalent in nature to those forced by the unscheduled outage.

**Fig. 4** Algorithm for forced outage calculation

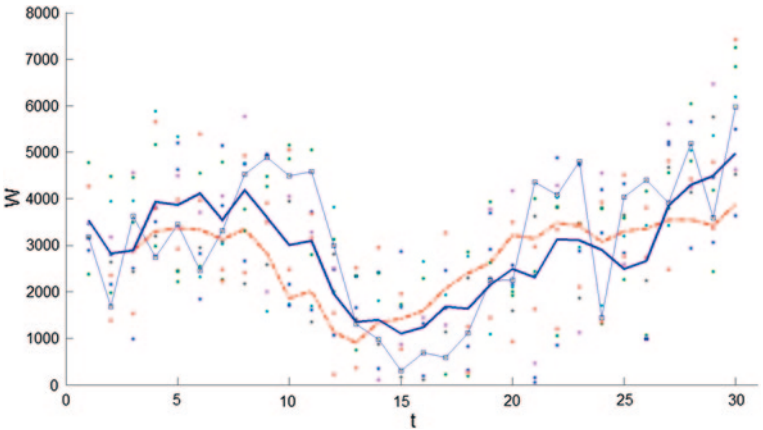


5 Results

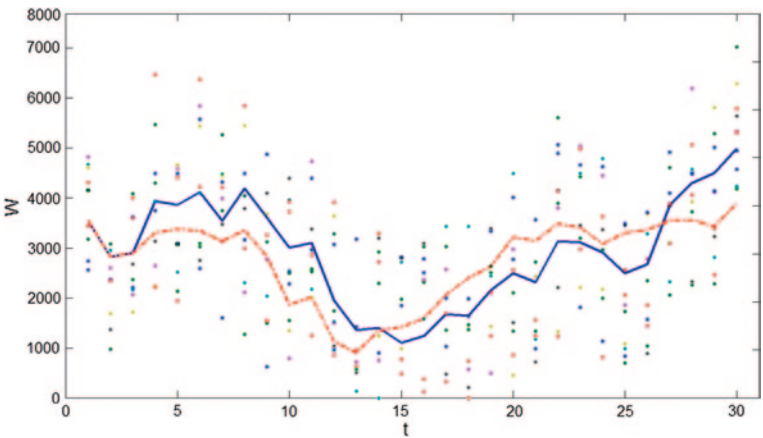
This section presents example scenario tree outputs from the STT which demonstrate the impact of changes in forecast error statistics such as kurtosis, skewness and autocorrelation between time periods.

Figure 5 gives an example of the points contained in a statistically accurate scenario tree overlayed upon the realized and forecasted values (thick dashed line and line, respectively). In addition, a single scenario is shown (thin line with squared points) to demonstrate the impact of including autocorrelation in the optimization.

Figure 6 presents the same tree with inverted skewness. While it is similar in shape to Fig. 5, its shape appears transposed around the forecast as its skewness



**Fig. 5** Example autocorrelated scenario, *thin line*, imposed onto scenario tree points. Deterministic forecast is shown by the *thick line*, and realized value is shown by the *thick dashed line*

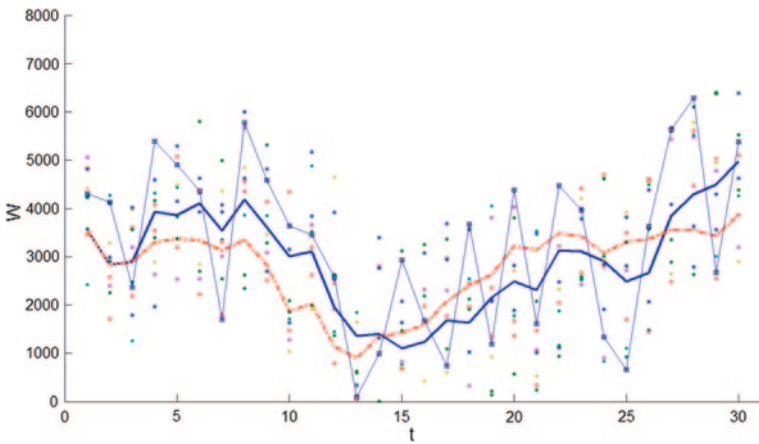


**Fig. 6** Example of a figure caption. Scenario tree with inverted skewness. Note the tree is skewed to the opposite side of the deterministic forecast

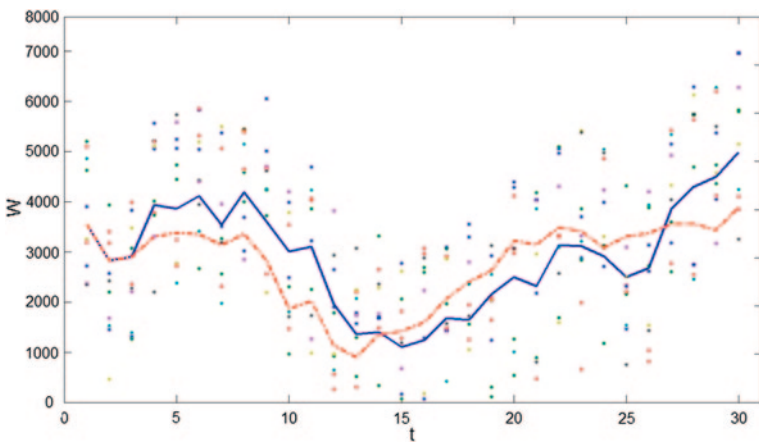
is of the opposite sign. The scenario tree constructed in this fashion assumes the opposite trend in terms of overestimating or underestimating the forecast.

Figure 7 presents a tree constructed from identical information without autocorrelation considered. Due to this, the included scenario in this version contains considerably more sharp changes than the reality entails. As a result, despite the statistics of each individual time period being correct, the grouped scenarios do not present useful information due to prediction of frequent unrealistic ramp events.

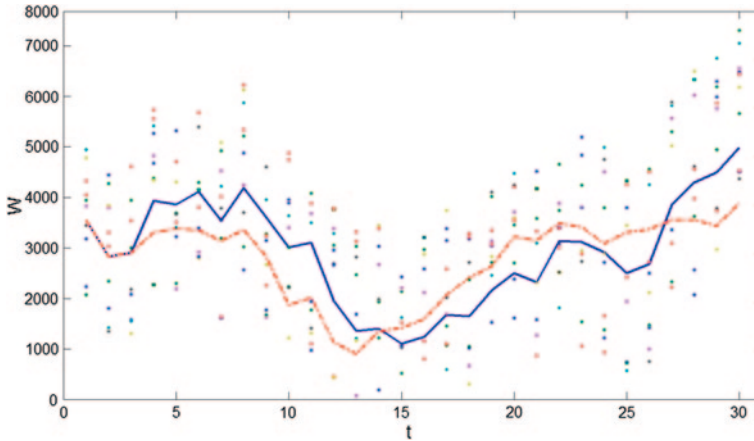
Figure 8 demonstrates the effect of increasing kurtosis—this tree clusters close to the forecast and contains an increased number of extreme values.



**Fig. 7** Example unautocorrelated scenario, *thin line*, imposed onto scenario tree points. Deterministic forecast is shown by the *thick line*, and realized value is shown by the *thick dashed line*



**Fig. 8** Scenario tree with *increased kurtosis*. Note the increase in points clustered at the mean and at the extremes



**Fig. 9** Scenario tree with *decreased kurtosis*. Note the reduction in clustering and the more even spread of points

In contrast, Fig. 9 demonstrates the effect of reducing kurtosis—a tree which is comparatively evenly spread throughout the distribution without many extreme values or much clustering around the forecast.

## 6 Conclusion

This chapter presented the methodology of the WILMAR STT based on moment matching. Example outputs demonstrating some of the converted to a graphical format were also included. This tool is under development to examine the impact of wind forecast error statistics on unit commitment for high wind penetration test systems.

**Acknowledgments** This work was conducted in the Electricity Research Centre, University College Dublin, Ireland, which is supported by the Commission for Energy Regulation, Bord Gáis Energy, Bord na Móna Energy, Cylon Controls, EirGrid, Electric Ireland, EPRI, ESB International, ESB Networks, Gaelectric, Intel, SSE Renewables, UTRC and Viridian Power and Energy.

This publication has emanated from research conducted with the financial support of Science Foundation Ireland under grant number 06/CP/E005.

## References

1. Pure power: wind energy targets for 2020 and 2030, EWEA, (2010) Available: <http://www.ewe.org/>
2. Delivering a sustainable energy future for Ireland, Government White Paper (2007) Available: <http://www.dcmnr.gov.ie>

3. Meeting the energy challenge, U.K. Government White Paper (2007) Available: <http://www.dti.gov.co.uk>
4. Second Quarter 2011 Market Report, AWEA. Available: <http://www.awea.org/>
5. Directive 2009/72/EC of the European Parliament. Available: <http://eur-lex.europa.eu>
6. Troy N, Denny E, O'Malley M (2010) Base-load cycling on a system with significant wind penetration. *IEEE Trans Power Syst* 25(2):1088–1097
7. Marti I et al. (2006) Evaluation of advanced wind power forecasting model results of the anemos project. Available: <http://anemos.cma.fr>. pp 1088–1097 (2010)
8. Tuohy A, Meibom P, Denny E, O'Malley M (2009) Unit commitment for systems with significant wind penetration. *IEEE Trans Power Syst* 24(2):592–601
9. Takriti S, Birge J, Long E (1996) A stochastic model for the unit commitment problem. *IEEE Trans Power Syst* 11(3):1497–1508
10. Doherty R, O'Malley M (2005) A new approach to quantify reserve demand in systems with significant installed wind capacity. *IEEE Trans Power Syst* 20:587–595
11. Hodge B, Orwig K, Milligan M (2012) Examining information entropy approaches as wind power forecasting performance metrics. In: *Proceedings of PMAPS Istanbul*
12. Hodge B, Milligan M *Wind Power Forecasting Error Distributions over Multiple Timescales* (2011) NREL Report No. CP-5500-50614. Available: <http://www.nrel.gov/>
13. Wind Power Integration in Liberalised Electricity Markets (WILMAR) Project. Available: <http://www.wilmar.risoe.dk>
14. Lowery C, O'Malley M (2012) Impact of wind forecast error statistics upon unit commitment. *IEEE Trans Sust Energy* 3(4):760–768
15. Hoyland K, Kaut M, Wallace SW (2003) A heuristic for moment-matching scenario generation. *Comput Appl Optim* 24:169–185
16. Hoyland K, Wallace SW (2001) Generating scenario trees for multistage decision problems. *Manag Sci* 47(2):295–307
17. Hasche B (2008) Scenario tree generation for an electricity market model. Available: <http://www.netmod.org/>
18. Meibom P, Barth R, Brand H, Swider D, Ravn H, Weber C (2007) All Island Renewable Grid Study Work stream 2b—Wind Variability Management Studies. Available: <http://www.dcmnr.gov.ie>
19. Lange M, Focken U (2006) *Physical approach to short-term wind power prediction*. Springer, Berlin
20. Lange M (2005) On the uncertainty of wind power predictions: analysis of the forecast accuracy and statistical distribution of errors. *J Sol Energy Eng* 127:177–184
21. Barth R, Soder L, Weber C, Brand H, Swider D (2006) Methodology of the Scenario Tree Tool. Available: <http://www.wilmar.risoe.dk>
22. Söder L (2004) Simulation of wind speed forecast errors for operation planning of multi-area power systems. In: *Proceedings of 8th international conference on probabilistic methods applied to power systems*, Ames
23. Dupacova J, Growe-Kuska N, Romisch W (2003) Scenario reduction in stochastic programming: an approach using probability metrics. *Math Program Ser. A* 95, vol 3, pp 493–511
24. Heitsch H, Roemisch W (2003) Scenario reduction algorithms in stochastic programming. *Comput Appl Optim* 24(2–3):187–206
25. Valenzuela J, Mazumdar M (2001) Monte-Carlo computation of power generation production costs under operating constraints. *IEEE Trans Power Syst* 16:671–677

Reliability and Risk Evaluation of Wind Integrated Power  
Systems

Billinton, R.; Karki, R.; Verma, A.K. (Eds.)

2013, XIV, 128 p., Hardcover

ISBN: 978-81-322-0986-7