#### Energy 120 (2017) 507-517

Contents lists available at ScienceDirect

# Energy

journal homepage: www.elsevier.com/locate/energy

# Synthetic wind speed scenarios generation for probabilistic analysis of hybrid energy systems



Idaho National Laboratory, Idaho Falls, ID 83415, USA

#### ARTICLE INFO

Article history: Received 25 August 2016 Received in revised form 26 September 2016 Accepted 17 November 2016 Available online 25 November 2016

Keywords: Hybrid energy systems Renewable energy integration Synthetic data generation Autoregressive moving average

# ABSTRACT

Hybrid energy systems consisting of multiple energy inputs and multiple energy outputs have been proposed to be an effective element to enable ever increasing penetration of clean energy. In order to better understand the dynamic and probabilistic behavior of hybrid energy systems, this paper proposes a model combining Fourier series and autoregressive moving average (ARMA) to characterize historical weather measurements and to generate synthetic weather (e.g., wind speed) data. In particular, Fourier series is used to characterize the seasonal trend in historical data, while ARMA is applied to capture the autocorrelation in residue time series (e.g., measurements with seasonal trends subtracted). The generated synthetic wind speed data is then utilized to perform probabilistic analysis of a particular hybrid energy system configuration, which consists of nuclear power plant, wind farm, battery storage, natural gas boiler, and chemical plant. Requirements on component ramping rate, economic and environmental impacts of hybrid energy systems, and the effects of deploying different sizes of batteries in smoothing renewable variability, are all investigated.

© 2016 Elsevier Ltd. All rights reserved.

## 1. Introduction

Hybrid energy systems (HES) consisting of multiple energy generations/utilizations have been proposed to be an effective element to increase renewable energy penetration, [1–10]. Prior works have been focused on dynamic modeling, simulation, control, and optimization for HES [6–10]. These results suggest that, from both technical and economic point of view, HES can be operated under flexible operations schedules to accommodate the variability introduced from renewable generation, modern loads (such as electric vehicles), and markets. However, these prior works model the renewable generations using historical weather condition, for which only a limited number of measurements database are available.

To address this limitation, the objective of this paper is to develop a mechanism for generating synthetic weather scenarios, e.g., wind speed data, which are statistically conformed to the actual measurements but posses different temporal profile. In particular, a combined model with Fourier series and autoregressive moving average (ARMA) is utilized to characterize the

\* Corresponding author. *E-mail addresses*: jun.chen@inl.gov, jchenee2015@gmail.com (J. Chen). seasonal trends in the yearly measurements and the autocorrelation in the residues (i.e., measurements with seasonal trends substracted). The model with identified parameters is then able to generate synthetic wind speed data, where the synthesis step consists of generating independent white noise for each time step, utilizing ARMA model and the synthesized white noise to compute residues for each time step, and finally adding the Fourier series representing seasonal trends. An illustration of the process can be found in Fig. 1. The model is implemented as part of the Risk Analysis Virtue ENvironment (RAVEN) [11] developed at Idaho National Laboratory. To validate the trained model and the corresponding synthetic wind speed data, key statistics computed from actual wind speed measurements (database) as well as those from synthetic data are compared, including mean, variance, quantiles, and empirical cumulative distribution function.

The synthetic wind speed data will in turn be utilized to analyze a particular hybrid energy system configuration to understand its technical, economical, and environmental values. The specific HES configuration studied in this paper includes a nuclear power plant, wind farm, battery storage, natural gas boiler, and chemical plant. The HES model will be simulated with 3000 synthetic wind speed scenarios, while various probabilistic analysis are performed to understand: (1) requirement on component power output and ramping rate in order to accommodate variability introduced by





ENERGY Transfer



Fig. 1. Implemented flow diagram in RAVEN (Risk Analysis Virtual ENvironment).

renewable generation; (2) effects of employing different sizes of batteries for variability smoothing; and (3) variance on environmental and economic impacts of HES. The contribution of this paper is highlighted as follows: (1) a computational model, combining Fourier series and ARMA, is proposed to synthesize artificial wind speed data with consistent characteristics with data base; (2) Monte Carlo simulation and probabilistic analysis are performed for hybrid energy systems; (3) the effect of battery in smoothing variability of wind power generation is investigated; and (4) financial viability of hybrid energy systems is proven to be robust with respect to uncertainties in wind power generation.

The rest of this paper is organized as follows. Section 2 reviews related work in literature. Section 3 presents the theoretical foundation of the proposed model, while Section 4 formally presents the algorithm for synthetic scenarios generation and its implementation. The HES configuration under study is presented in Section 5, and results on model validation and probabilistic analysis of HES are given in Section 6. Finally, Section 7 concludes the paper.

#### 2. Literature review

This section reviews related works on synthetic wind speed data generation as well as hybrid energy systems.

#### 2.1. Point forecast v.s. scenario generation

Two types of research have been conducted in literature concerning weather data analysis, i.e., point forecast and scenario generation. The former focuses on predicting exact future values for a given time span [12], i.e., predicting exact waveform, while the objective of the latter is to generate waveform that has same statistical characteristics (mean, volatility, autocorrelation, etc) with those from database [13]. The time span for point forecast can be very short term (from seconds to 30 min) [14], short term (30 min to 6 h) [15,16], medium term (6 h to 1 day) [17] or long term (1 day to 1 week or more) [18], while the time span for scenario generation is usually long term (1 day to 1 year) [13,19].

#### 2.2. Wind speed prediction

Even though this paper doesn't focus on point forecast, existing work in this area are summarized below for the sake of completeness. Point forecast techniques are categorized into physical approach and statistical approaches, [12,20–22]. Physical approaches use meteorological data, which is good for long term prediction but lack of short term accuracy [20]. Statistical approaches provide accurate short-term results, but their reliability in long-term prediction is questionable [21]. Statistical approaches are further categorized into artificial neural network (ANN) [12] and time-series models [22], both of which can be trained over historical data and metrological data.

Some latest research combines a variety of techniques for more accurate prediction. Reference [23] combines wavelet transform with autoregressive (AR) model. Firstly wavelet method is applied to decompose original time series into a number of different subseries. Then AR model is built for each sub-series for prediction in its domain. Finally aggregating prediction in each sub-series provides the final forecasting. Reference [24,25] use Kalman filter to improve available NWP forecast on wind speed. Let  $m_t$  be the NWP forecast,  $y_t$  be its error (the difference between  $m_t$  and the real). The relation between  $y_t$  and  $m_t$  is assumed to be:  $y_i = x_{0,t} + x_{1,t}m_t + x_{2,t}m_t^2 + \cdots + x_{n,t}m_t^n + v_t$ . Then the goal is to estimate parameter  $X_t$  to fine tune future NWP forecast  $y_t$ . The state and output equations become:

$$X_t = X_{t-1} + w_t$$

$$y_t = H_t x_t + v_t,$$

where  $H_t$  is  $[1, m_t, m_t^2, ..., m_t^n]$ . Kalman filter is readily applied here.

Other approach in point forecast of wind speed includes Mycielski approach [26] and grey model [27], where [26] predicts the next value in currently ongoing random process by the longest repeating data chain that has shown up in the past data sequence. Reference [27] uses Grey model for wind speed prediction. Let *X* be the time series, then the grey model is expressed as  $\frac{dX}{dt} + aX = b$ . The only two parameters *a* and *b* can be estimated from available historical measurement data, and the prediction of next value is

given by 
$$\widehat{X}(i+1) = \left(X(1) - \frac{b}{a}\right)e^{-ai} + \frac{b}{a}$$
.

#### 2.3. Wind speed scenario generation

The synthetic data generation has been studied in the literature [28], where different synthesis algorithms have been proposed. The simplest yet intuitive algorithm is the empirical approach, which

multiplies the empirical time series by a constant. Another approach combined empirical approach with ARMA model, where the ARMA model is not used to generate the data of interests directly but the noise term that will later be added to the empirical time series. For example [29], uses ARMA to generate prediction errors, and adds the sampled prediction errors to the historical data. The ARMA model itself has also been proposed to generate wind speed scenarios. For example [30–32], use ARMA or AR model to fit available wind speed data, and then use the fitted ARMA model, together with sampled white noise, to generate scenarios. The original time series may need to be transformed into Gaussian distribution by its marginal distribution (either empirical or fit to a parametric distribution) to satisfy the assumption of ARMA. Reference [31] further normalizes the transformed time series with respect to each hour for each month. Reference [33] uses AR-GARCH (autoregressive generalized autoregressive conditional heteroskedasticity) model for wind speed prediction, which allows the regression of both mean and variance. Seasonal terms were added in the mean regression to account for seasonal effect. The prediction output, in terms of the mean and variance of the wind speed/power, can then be sampled to generate both sample paths and point forecast.

Gaussian process and neural network have also been found in literature for scenario generation. Reference [34] uses Gaussian process for point prediction, which takes metrological data as input and wind prediction as output. Reference [35] applied similar methodology over historical wind speed data. After choosing proper covariance function, the hyper-parameters can be estimated by maximum likelihood function. Note that these work focus on point prediction instead of scenario generation, but since the estimated model is a distribution, sample path can also be synthesized.

Authors in Ref. [36] use factor analysis for scenario synthesis, which is a statistical method to describe variability among observed correlated variables in terms of a potentially lower number of unobserved variables (called *factors*). Let X be the observed data, its linear representation over factors can be written as X = LF + E, where L is called loading matrix and F is called factors. The dimension of F represents the number of factors. E is random noises with covariance  $\Phi$ . Values for *L*,  $\Phi$  and *F* can be found by fitting the observation covariance S = cov(X) into  $LL^T + \Phi$ . In wind application, data needs to be normalized before applying factor analysis. Sample path generations follow as simulating the random noise E to generate randomized X and de-normalizing. Reference [37] computes power spectrum density (PSD) through measurement data or through AR(MA) model of the data, predicts PSD based on future capacity, and generates sample waveform by inverse FFT of predicted PSD. When doing so, phase angles are determined by genetic algorithm to fit certain statistical characteristics including: uniform distributed phase angle, minute-tominute difference distributed as Laplace, ramping distribution, etc.

#### 2.4. Hybrid energy systems

Energy systems utilizing more than one type of resources or generating more than one type of energy output have been widely studied in literature [1-10,38-63]. For example, combined heat and power (CHP) systems [38-41] produce both thermal and electric outputs, while hybrid renewable energy systems (HRES) [42-49] integrate different energy resources (e.g., wind, solar, or baseload generation) to generate electricity. Technical and/or economic analysis of hybrid systems can be found in Refs. [1,2,10,50-52]. For example, it is found in Ref. [50] that grid connected hybrid systems have a higher level of adaptation than standalone ones. Accordingly, optimization problems for integrated

systems are also investigated in the literature. Design and analysis of HRES are studied in Refs. [53–55] by using different optimization strategies (i.e., simulated annealing, response surface methodology, etc). Reference [56–59] study the design optimization problem for hybrid systems, in which the sizes of different components are optimized, given an economic/technical objective function. On the other hand, operations optimization of distributed energy systems were studied in Refs. [10,60–62], where [10] proposes a computational framework for operations optimization of hybrid energy system considering both day ahead and real time electricity markets as well as ancillary service market, while [60] includes also the exergy efficiency in the optimization process. Reference [63] reviews different optimizations techniques that have been applied to sustainable and renewable energy systems, including wind, solar, hydropower, bioenergy, geothermal, and hybrid systems.

### 3. Theoretical foundation

#### 3.1. ARMA model and identification

Autoregressive moving average (ARMA) model provides a mathematical framework to characterize stationary processes by imposing a linear dependence among the variables and a series of white noise [30]. An ARMA model with orders p and q, often referred to as ARMA(p,q), is given as [64]:

$$x_{t} = \sum_{i=1}^{p} \phi_{i} x_{t-i} + \alpha_{t} + \sum_{j=1}^{q} \theta_{j} \alpha_{t-j},$$
(1)

where the process variable *x* is a vector of dimension *n*, and parameters  $\phi_i$  for i = 1, ..., p and  $\theta_j$  for j = 1, ..., q are both *n* by *n* matrices. The noise term  $\alpha$  is usually assumed to be white noise. When q = 0, the model reduces to autoregressive (AR) model; when p = 0, it is called moving average (MA) model.

Given an ARMA(p,q) model, identification of its parameters  $\phi_i$ 's and  $\theta_j$ 's can be done by computing either maximum likelihood estimator (MLE) or nonlinear least square estimator. Below is an overview of the parameter estimation procedure. A more detailed treatment of time series and ARMA model can be found in Ref. [64].

Suppose there are *T* number of measurements of the process variable *x*, denoted as  $x_1, x_2, ..., x_T$ . Denote the covariance of  $\alpha$  to be  $\Sigma$ , and

$$\Sigma = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & & \sigma_n^2 \end{bmatrix}.$$

Define the parameters to be estimated as  $\eta := (\phi_1, ..., \phi_p \cdot \theta_1, ..., \theta_q, \sigma_1^2, ..., \sigma_n^2)$ . Given the *T* measurements, the likelihood function over  $\eta$  can be written as:

$$\begin{split} L(\eta) &= f(x_1, \dots, x_T | \eta) \\ &= f(x_1 | \eta) f(x_2 | \eta, x_1) \cdots f(x_T | \eta, x_1, \dots, x_{T-1}) \\ &= \prod_{t=1}^T \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\bigg( -\frac{1}{2} \widehat{\alpha}_t' \Sigma^{-1} \widehat{\alpha}_t \bigg), \end{split}$$

where the estimation of the error term at time *t*, i.e.,  $\hat{\alpha}_t$  is recursively computed as:

$$\widehat{\alpha}_1 := x_1 \tag{2}$$

$$\widehat{\alpha}_2 := x_2 - (\phi_1 x_1 + \theta_1 \widehat{\alpha}_1) \tag{3}$$

$$\widehat{\alpha}_3 := x_3 - (\phi_1 x_2 + \phi_2 x_1 + \theta_1 \widehat{\alpha}_2 + \theta_2 \widehat{\alpha}_1) \tag{4}$$

÷

$$\widehat{\alpha}_t := \mathbf{x}_t - \left(\sum_{i=1}^{\min(p,t-1)} \phi_i \mathbf{x}_{t-i} + \sum_{j=1}^{\min(q,t-1)} \theta_j \widehat{\alpha}_{t-j}\right).$$
(5)

Furthermore, the log-likelihood function, i.e.,  $logL(\eta)$ , can be expressed as:

$$\log L(\eta) = -\frac{nT}{2}\log(2\pi) - \frac{T}{2}\log|\Sigma| - \frac{1}{2}\sum_{t=1}^{T} \widehat{\alpha}_t' \Sigma^{-1} \widehat{\alpha}_t.$$
 (6)

The MLE of  $\eta$  is then given by  $\hat{\eta}$  that maximizes the above log-likelihood function, i.e.,

$$\widehat{\eta} := \arg \max \log L(\eta).$$
 (7)

On the other hand, given the estimators (2)–(6) of the noise term, the nonlinear least square estimate of  $\eta$  is given by finding the optimal  $\hat{\eta}$  that minimizes the sum of squares of the errors. In other words:

$$\widehat{\eta} := \arg \max r(\eta),$$

where  $r(\eta)$  is defined as:

$$r(\eta) := \sum_{t=1}^t \widehat{\alpha}_t' \widehat{\alpha}_t.$$

In this paper, the MLE as given in (8) is adopted for model identification. Given measurement data, the parameters p and q are chosen as following. Though larger values of p and q will generally result in a better fitting of the data, such "better fit" can be a result of overfitting. To balance between good data fitting and prevention of overfitting, two selection criteria are widely used, namely Akaike information criterion (AIC) and Bayesian information criterion (BIC), defined as:

$$AIC(p,q) = \log(\widehat{\sigma}^2) + \frac{2(p+q)}{T}$$
  
 $BIC(p,q) = \log(\widehat{\sigma}^2) + \frac{\log(T)(p+q)}{T}$ 

where  $\hat{\sigma}^2$  is the determinant of  $\Sigma^T \Sigma$ . It is not hard to see that the first term on the right hand size captures how the model fits to the data, while the second term penalizes larger model to prevent overfitting. Note that usually BIC penalizes large models more than AIC. In this paper, BIC is chosen for model selection.

#### 3.2. Seasonal trend and normality transform

Two assumptions of ARMA limit its usage to model data. One is that the process (1) is stationary, and the other being that the process variable x shall be normally distributed. Regarding the stationarity, seasonal ARMA model is used to mitigate the difficulty. In other words, the data is first "detrended" by a combination of Fourier series, and the residues between original database and the Fourier series are then used to train the ARMA model. The Fourier series used in this paper is given as:

$$F_t := \sum_k \{ a_k \sin(2\pi f_k t) + b_k \cos(2\pi f_k t) \}.$$
 (8)

The set of frequency  $\{f_k\}$  is user-defined parameters, and the coefficient  $\{a_k\}$  and  $\{b_k\}$  can be estimated by linear regression.

In general, renewable source and load profile do not satisfy the normality assumption, even after seasonal trends being extracted. However, this problem can be moderately mitigated by properly transform the data so that the transformed data has Gaussian property [30,31,64]. Define a new stochastic process *y*, which has a standard normal distribution, as follows:

$$y_t := \Phi^{-1}[f(x_t - F_t)], \tag{9}$$

where f is the empirical cumulative distribution function (CDF) of the residues and  $\Phi$  is the CDF function of the standard normal distribution.

Note that the above procedure removes the seasonal trends from the original data *x* and transform the residues into stochastic process *y* that is normally distributed. The transformed data is then used to find the ARMA model parameters  $\phi_i$ 's and  $\theta_j$ 's. The trained ARMA model can then be used to simulate process *y*, which is in turn used to generate the scenarios by an inverse transformation, as following:

$$x_t := f^{-1}[\Phi(y_t)] + F_t.$$
(10)

#### 3.3. Data pre-processing

The database used in this study spans over multiple years. To train an ARMA model, one year data is enough to achieve parameter convergence. However, to ensure that the trained model possesses the typical statistical characteristics from the database, a preprocessing step is applied to translate multi-year data into "typical"-year data that spans exactly one year.

The pre-processing is a simplified procedure of [65], which selects individual months from different years in the database. For example, all the Januarys are examined and the one determined as most typical would be selected (see below). This procedure continues until all the months of the year are examined and then the twelve selected typical months are concatenated to form a complete typical year. To select a typical month, the empirical CDF of the interested variable (i.e., wind speed in this case) of the entire multi-year database is computed, termed as long term CDF. Then the CDF over each candidate month is also computed, and the typical month is selected based on the Finkelstein-Schafer (FS) statistics [66] between the candidate month's CDF and the long term CDF. The FS statistics is defined as following:

$$FS = \sum_{x_i} FS_{x_i}$$
$$FS_{x_i} = \frac{1}{N} \sum_{n=1}^N \delta_n$$

where *N* is the number of value reading in the empirical CDF and  $\delta_n$  is the absolute difference between the long term CDF and the candidate month's CDF at value  $x_i^n$ . The candidate month with minimal FS statistics will be selected. Such process repeats for every month of the year, and then all the typical months will be

concatenated to yield a complete year.

#### 4. Algorithm and implementation

In this section, the procedure used to identify seasonal trends and ARMA model, and to generate scenarios is described. As can be seen from Fig. 1, the process consists of three major steps, namely, data pre-processing and detrending, model characterization, and scenarios generation.

#### 4.1. Algorithm 1

Step 1 Data pre-processing and detrending. This step includes two micro-steps.

- (a) For the given database (e.g., wind speed measurements), generate the typical year time series from the multi-year data according to the procedure in Section 3.3.
- (b) Compute coefficient for Fourier series of equation (8) by linear regression.

Step 2 Model characterization. This step includes two microsteps.

- (a) Compute the empirical CDF from residues  $\{x_t F_t\}$ , and transform the residues into normal dataset  $\{y_t\}$  according to equation (9);
- (b) Fit the ARMA model into normal dataset  $\{y_t\}$  according to equation (7) and Bayesian information criterion, yielding optimal model ARMA(p,q) with parameters p, q,  $\phi_i$  for i = 1, ..., p and  $\theta_i$  for j = 1, ..., q.

Step 3 Scenarios generation. This step includes the following micro-steps.

- (a) Simulate N number of independent standard normal errors {α<sub>t</sub>}, where N is the number of measurements in a one-year time series;
- (b) Simulate N transformed normal variable  $\{y_t\}$  according to the identified ARMA(p,q) specified in equation (1);
- (c) Transform the normal dataset  $\{y_t\}$  into residues according to (11);
- (d) Add the seasonal trends  $F_t$  identified in Step 1(b) into the residues to yield the complete scenario;
- (e) Repeat Step 3(a)-(d) until the desired number of scenarios is achieved.

The above algorithm is implemented as part of RAVEN (Risk Analysis Virtual ENvironment) [11] that is currently developed in Idaho National Laboratory (INL) using Python programming language.

#### 5. HES configuration and economic model

A hybrid energy system (HES) that integrates wind power generation and nuclear power plant was studied in Refs. [6,7,9,10], which is analyzed in this paper. The HES under study is shown in Fig. 2, which includes the following components:

- a heat generation plant with 180 MW capacity,<sup>1</sup> consisting of a small modular reactor (SMR) and a steam generator, denoted as primary heat generation (PHG),
- a series of steam turbines, feedwater systems, and heaters, paired with an electric generator that converts steam into electricity, denoted as thermal-to-electrical conversion (TEC),

- a series of wind turbines as renewable power generation source, denoted as REN (renewable), with total capability of 45 MW,
- electrical storage (i.e., a system scale battery set) used for power smoothing of the electricity generated by REN, denoted as energy storage element (ESE),
- a gasoline production plant (GPP) able to utilize process steam up to 45 MW and convert natural gas (NG) and water into gasoline (and liquefied petroleum gas [LPG]),
- an auxiliary heat generation (AHG), which is a NG fired steam generator boiler plant of up to 45 MW capacity that generates additional on demand steam for GPP,
- electric grid connected to HES at a point of common coupling and consuming electricity up to 180 MW, and
- operations optimizer that computes operations schedule and energy distribution among HES components, according to various market dynamics and renewable generations, for maximal economic performance.

In the following, we briefly describe the modeling of wind power generation unit and the economic model. For the detailed modeling of the rest components, please refer to [6,7,9,10]. The wind farms consist of 30 wind turbines, each rated at 1.5 MW and located on a 2 square kilometer site for a maximum of about 45 MW generation at full production. Each wind turbine is then modeled as a mapping function from wind speed to power output, as shown in Fig. 3, which includes four operating regimes separated by critical wind speed values. When wind speed is below a minimum cut-in velocity (e.g., 3 m/s), the kinetic energy is insufficient to cause any blade rotation, thus there is no electricity being produced. When wind speed is above a cut-out velocity (e.g., 25 m/s), for safety reasons a braking system is activated to prevent blade rotation, thus again there is no electricity being produced. When the wind speed is between the rated (e.g., 14 m/s) and cut-out velocity values, the turbine provides a steady maximum electrical power output, referred to as rated power. Finally, for wind speed between the cut-in and rated speeds, the power is calculated according to:

$$E_{REN} := 0.5 \eta \rho U^3 \frac{\pi d^2}{4},$$
 (11)

where  $\eta$  is the conversion efficiency of the wind turbine,  $\rho$  is the density of the air at the site, *U* is the wind speed, and *d* is the diameter of the turbine blades. In this study the values used for each parameter in equation (11) are:  $\eta = 35\%$ ,  $\rho = 1.17682 \text{ g/m}^3$ , d = 58.13 m.

An economic model to evaluate the financial viability of HES is studied in Refs. [9,10], in which both payback time and internal rate of return are analyzed. In particular, net present value (NPV) is defined as follows [67]:

$$NPV = \sum_{k=0}^{N} \frac{FCFF_{R,k}}{(1+r_R)^k},$$
(12)

where *N* is the years of operations of HES,  $r_R$  denotes the discount rate used in computing weighted average cost of capital (WACC), and *FCFF*<sub>*R,k*</sub> is the real discounted Free Cash Flow to Firm for year *k*. For a fixed discount rate, payback time, the period of time required to recoup the expense of an investment [68], is defined as the years of operations such that the net present value (NPV) equals 0, i.e.,

$$T_{pb} = \arg_N[NPV = 0]. \tag{13}$$

On the other hand, internal rate of return (IRR) is used to measure and compare the profitability of investments. For a fixed N

<sup>&</sup>lt;sup>1</sup> For simplicity, all power calculations will be expressed using the electrical equivalence (in MW) of the particular power stream, assuming fixed thermal-to-electrical conversion efficiency.



Fig. 2. Topology of the hybrid energy system configuration considered in this paper.



Fig. 3. Turbine power output vs. wind speed.

years of operations, it is defined as the value of discount rate such that NPV equals 0 [67], i.e.,

 $IRR = \arg_{r_{R}}[NPV = 0]. \tag{14}$ 

Please refer to [6,7] for more details of this HES configuration and to [9,10] for its economic evaluation.

# 6. Results and discussion

# 6.1. Model characterization and validation

This section presents the results on model characterization. The



Fig. 4. Wind speed data used to train the model: (a) Typical year data; (b) Seasonal trend extracted from typical year data.

wind speed database<sup>2</sup> used to train the model consists of three years data, namely the years of 2004, 2005, and 2006. The typical year data is formed by the procedure discussed in Section 3.3, and is shown in Fig. 4(a). Furthermore, Table 1 shows the constitution of the typical year data. Seasonal trends, extracted from typical year

<sup>&</sup>lt;sup>2</sup> Downloaded from the Eastern Wind dataset maintained by NREL (National Renewable Energy Laboratory) at http://www.nrel.gov/electricity/transmission/ eastern\_wind\_dataset.html on November 21, 2014.

**Table 1**Constitution of typical year data.

Month	From the year of	
January	2005	
February	2006	
March	2004	
April	2004	
May	2006	
June	2006	
July	2006	
August	2006	
September	2004	
October	2004	
November	2004	
December	2005	

Table 2

Optimal parameter for ARMA of equation (1).

Parameter	Identified value	
р	1	
q	2	
$\phi_1$	0.96492115	
$\theta_1$	0.35438419	
$\theta_2$	0.03969797	
$\sigma_1^2$	0.03603672	

data and modeled as Fourier series of equation (8), is given in Fig. 4(b). Finally, Table 2 provides parameters for ARMA model of equation (1).

Fig. 5 plots the synthetic wind speed scenario (generated by the identified model) and the actual database for selected 7 days. As can be seen, the synthetic scenario and actual database exhibit similar dynamics and volatility. Furthermore, Table 3 compares several key statistics (mean, standard deviation, etc) between the synthetic and actual database, showing identical statistics between these two. The Quantile-Quantile plot (qq-plot) between the synthetic scenario and actual database is given in Fig. 6, while Fig. 7 compares the CDF of synthetic scenario and actual database, both suggesting good match between synthetic wind speed scenarios and actual database. These results suggest that the presented model can produce synthetic wind speed scenario with same statistical characteristics as the actual database.

Furthermore, to demonstrate the benefit of the synthetic scenarios, Fig. 8 plots the actual wind speed together with 100 synthetic scenarios, where each synthetic scenario possesses very different time profile from the actual database. Considering that the synthetic data possesses same statistical characteristics with database while having different temporal profile, they can be used for Monte Carlo simulation and risk analysis of energy integration systems, while avoiding bias introduced by using the same database. In the following sections we utilize the synthetic scenarios for probabilistic analysis of the HES configuration introduced in Section 5.

#### 6.2. Probabilistic analysis of HES

In this section 3000 synthetic wind speed scenarios are generated, which are then used to simulate the HES configuration introduced in Section 5.

Fig. 9 plots the histogram of the wind turbine power output. As can be seen, frequency decreases exponentially with the increase of power, while there is a sudden increase at 45 MW (i.e., full production rate of the wind farm) due to the saturation of wind turbine power output (see Fig. 3).

In all of the simulations, the HES is operated in such a way that it



Fig. 5. Synthetic wind speed scenario and the actual database for selected 7 days.

Table 3Comparison between synthetic scenario and database.

Variable	Database	Synthetic
Mean (wind speed)	8.078	8.088
Standard deviation (wind speed)	3.392	3.372
Mean (step to step difference)	0	0
Standard deviation (step to step difference)	0.659	0.642



Fig. 6. Quantile-Quantile plot (qq-plot) between the synthetic scenario and actual database.

provides constant electricity output of 180 MW to the power grid, while the output of AHG is fluctuated to compensate the energy needs of GPP. For more information on the dynamic operations of this HES configuration, please refer to [7,9,10]. Fig. 10 shows the wind turbines output and AHG output for selected 7 days, based on actual wind speed database. It can be seen that it requires the AHG to dynamically ramp up or down its power output in order to maintain the GPP at its constant production rate. Fig. 11 shows the histogram of average ramping rate of each yearly synthetic wind



Fig. 7. CDF of synthetic scenario and actual database.



Fig. 8. 100 synthetic scenarios and actual database for 3 days.



Fig. 9. Histogram of turbine power output for 3000 synthetic scenarios.



Fig. 10. Turbine power output and corresponding AHG power output for selected 7 days.

speed scenario for 3000 simulations. In particular, it requires AHG to ramp averagely 0.167 MW/min in order to absorb the volatility in the wind farm production. As discussed in Refs. [7,9,10], a battery storage can be used to smooth the variability of the wind farm production before sending the renewable power to HES. To analyze the effects of such battery storage, multiple Monte Carlo simulations are performed, each with different battery size, namely, no battery, 20 MWh, 40 MWh, and 80 MWh. Fig. 12 plots the box plots of maximum yearly rates on AHG ramping up and down, suggesting less AHG ramping is needed if larger battery is employed to smooth renewable generation. Fig. 12 also suggests that, under current HES configuration, the battery has better effects on relaxing the AHG ramping down requirement than that of ramping up.

Fig. 13 plots the box plots of annual natural gas usage by AHG for different battery sizes. Since the annual natural gas usage is dependent on the total wind power production instead of its volatility, it is not affected by the battery size, as can be seen from Fig. 13.

An economic model to evaluate the financial viability of HES is



Fig. 11. Histogram of yearly average ramping rate of AHG for 3000 synthetic scenarios.



**Fig. 12.** Box plots of yearly maximum AHG ramping up and down rates for 3000 synthetic scenarios with different battery size.



Fig. 13. Box plots of yearly natural gas usage for 3000 synthetic scenarios with different battery size.

studied in Refs. [9,10] and briefly discussed in Section 5. To measure the economic risk of investing on this particular HES configuration, two economic figures of merits are analyzed, namely, payback time and internal rate of return. Payback time is defined as the period of time required to recoup the expense of an investment [68], while internal rate of return (IRR) is used to measure and compare the profitability of investments. For more details in computing payback time and IRR for HES, please refer to Section 5 and [9,10]. In particular, 3000 synthetic wind speed scenarios are generated to simulate the HES for economic value computation. Consequently, Fig. 14 plots the histogram of payback time, while Fig. 15 presents the box plot of the IRR. Both results suggest that only a small deviation is present regarding the economic value of HES in these 3000 Monte Carlo simulations, further concluding its economic viability and robustness against uncertainty in renewable generation.

#### 7. Conclusion

This paper proposed a computational model to generate



Fig. 14. Histogram of payback time for 3000 synthetic scenarios.



Fig. 15. Box plot of IRR for 3000 synthetic scenarios.

synthetic wind speed scenarios, which are shown to possess the same statistical characteristics with historical measurements. The proposed model is based on Fourier series and autoregressive moving average model, the former of which is used to capture the seasonal trends while the latter to characterize the autocorrelation among residues. The synthetic wind speed data was then used to analyze a particular HES configuration to understand its technical and economic values. Probabilistic analysis shows that the utilization of a battery storage can effectively reduce the ramping rate requirement on auxiliary heat generator, and the larger battery the better relaxation on ramping requirement. Two key economic values, i.e., payback time and internal rate of return, are also computed for 3000 synthetic scenarios, where the results show that the economic value of this HES configuration is robust against the uncertainty in the renewable generation, proving its economic viability.

#### Acknowledge

This research is supported by the Energy Security Initiative and the N-R HES (Nuclear-Renewable Hybrid Energy Systems) Program at INL (Idaho National Laboratory) under the U.S. Department of Energy contract DE-AC07-05ID14517.

penetrations in Ireland. IEEE Trans Power Syst 2011;26(3):1367-79.

- [30] Morales JM, Minguez R, Conejo AJ. A methodology to generate statistically dependent wind speed scenarios. Appl Energy 2010;87(3):843-55.
- [31] Papavasiliou A, Oren SS, O'Neill RP. Reserve requirements for wind power integration: a scenario-based stochastic programming framework. IEEE Trans Power Syst 2011;26(4):2197-206.
- [32] Ma X-Y, Sun Y-Z, Fang H-L. Scenario generation of wind power based on statistical uncertainty and variability. IEEE Trans Sustain Energy 2013;4(4): 894-904
- [33] Taylor JW, McSharry PE, Buizza R. Wind power density forecasting using ensemble predictions and time series models. IEEE Trans Energy Convers 2009;24(3):775-82.
- [34] Mori H, Kurata E. Application of gaussian process to wind speed forecasting for wind power generation. In: Proc. 2008 IEEE int. conf. sustainable energy techn. IEEE; 2008. p. 956-9.
- [35] Lee D. Baldick R. Short-term wind power ensemble prediction based on Gaussian processes and neural networks. IEEE Trans Smart Grid 2014;5(1): 501 - 10.
- [36] Lee D, Baldick R. Synthesis of sample paths of wind power through factor analysis & cluster analysis. In: Proc. 2013 North American power symposium (NAPS), Manhattan, KS; 2013. p. 1–6.
- [37] Lee D, Baldick R. Future wind power scenario synthesis through power spectral density analysis. IEEE Trans Smart Grid 2014;5(1):490–500.
- Fumo N, Mago PJ, Chamra LM. Emission operational strategy for combined [38] cooling, heating, and power systems. Appl Energy 2009;86(11):2344–50.
- [39] Nuytten T, Claessens B, Paredis K, Van Bael J, Six D. Flexibility of a combined heat and power system with thermal energy storage for district heating. Appl Energy 2013:104:583-91.
- [40] Rolfsman B. Combined heat-and-power plants and district heating in a deregulated electricity market. Appl Energy 2004;78(1):37-52.
- [41] Zhao H, Holst J, Arvastson L. Optimal operation of coproduction with storage. Energy 1998;23(10):859-66.
- [42] Bernal-Agustín JL, Dufo-López R. Simulation and optimization of stand-alone hybrid renewable energy systems. Renew Sustain Energy Rev 2009;13(8): 2111 - 8
- [43] Ibrahim H, Younès R, Basbous T, Ilinca A, Dimitrova M. Optimization of diesel engine performances for a hybrid wind-diesel system with compressed air energy storage. Energy 2011;36(5):3079-91.
- [44] Deshmukh M, Deshmukh S. Modeling of hybrid renewable energy systems. Renew Sustain Energy Rev 2008;12(1):235-49.
- [45] Wang X, Teichgraeber H, Palazoglu A, El-Farra NH. An economic receding horizon optimization approach for energy management in the chlor-alkali process with hybrid renewable energy generation. J Process Control 2014;24(8):1318-27.
- [46] Zhu B, Tazvinga H, Xia X. Switched model predictive control for energy dispatching of a photovoltaic-diesel-battery hybrid power system. IEEE Trans Control Syst Tech May 2015;23(3):1229-36.
- Chong W, Naghavi M, Poh S, Mahlia T, Pan K. Techno-economic analysis of a [47] wind-solar hybrid renewable energy system with rainwater collection feature for urban high-rise application. Appl Energy 2011;88(11):4067-77.
- Graditi G, Ippolito MG, Telaretti E, Zizzo G. An innovative conversion device to [48] the grid interface of combined RES-based generators and electric storage systems. IEEE Trans Ind Electron 2015;62(4):2540-50.
- [49] Bekele G, Palm B. Feasibility study for a standalone solar-wind-based hybrid energy system for application in Ethiopia. Appl Energy 2010;87(2):487-95.
- Türkay BE, Telli AY. Economic analysis of standalone and grid connected [50] hybrid energy systems. Renew Energy 2011;36(7):1931-43.
- [51] Fabbri A, Román TGS, Abbad JR, Quezada VHM. Assessment of the cost associated with wind generation prediction errors in a liberalized electricity market. IEEE Trans Power Syst 2005;20(3):1440-6.
- [52] Graditi G, Ippolito M, Telaretti E, Zizzo G. Technical and economical assessment of distributed electrochemical storages for load shifting applications: an Italian case study. Renew Sustain Energy Rev 2016;57:515-23.
- [53] Ekren O, Ekren BY. Size optimization of a PV/wind hybrid energy conversion system with battery storage using simulated annealing. Appl Energy 2010;87(2):592-8.
- [54] Ekren BY, Ekren O. Simulation based size optimization of a PV/wind hybrid energy conversion system with battery storage under various load and auxiliary energy conditions. Appl Energy 2009;86(9):1387-94.
- [55] Ekren O, Ekren BY. Size optimization of a PV/wind hybrid energy conversion system with battery storage using response surface methodology. Appl Energy 2008;85(11):1086-101.
- [56] Yang H, Wei Z, Chengzhi L. Optimal design and techno-economic analysis of a hybrid solar-wind power generation system. Appl Energy 2009;86(2):163-9.
- [57] Kaabeche A, Belhamel M, Ibtiouen R. Sizing optimization of grid-independent hybrid photovoltaic/wind power generation system. Energy 2011;36(2): 1214-22.
- [58] Hagos DA, Gebremedhin A, Zethraeus B. Towards a flexible energy system-A case study for Inland Norway. Appl Energy 2014;130:41-50.
- [59] Di Silvestre ML, Graditi G, Sanseverino ER. A generalized framework for optimal sizing of distributed energy resources in micro-grids using an indicator-based swarm approach. IEEE Trans Ind Inf 2014;10(1):152-62.
- [60] Di Somma M, Yan B, Bianco N, Graditi G, Luh PB, Mongibello L, et al. Operation optimization of a distributed energy system considering energy costs and energy efficiency. Energy Convers Manag 2015;103:739-51.

#### References

- [1] Garcia HE, Mohanty A, Lin W, Cherry R. Dynamic analysis of hybrid energy systems under flexible operation and variable renewable generation - part i: dynamic performance analysis. Energy 2013;52:1-16.
- [2] Garcia HE, Mohanty A, Lin W, Cherry R. Dynamic analysis of hybrid energy systems under flexible operation and variable renewable generation - part ii: dynamic cost analysis. Energy 2013;52:17-26.
- [3] Chen J, Garcia HE. Operations optimization of hybrid energy systems under variable markets. In: Proc. 2016 American control conference. Boston, MA: IEEE; July, 2016. p. 3212-8.
- [4] Binder WR, Paredis CJ, Garcia HE. Hybrid energy system modeling in modelica. In: Proc. 10th int. Modelica conf., Lund, Sweden; Mar. 10-12, 2014. p. 779-88.
- [5] Du W, Garcia HE, Paredis CJJ. An optimization framework for dynamic hybrid energy systems. In: Proc. 10th int. Modelica conf., Lund, Sweden; Mar. 10-12, 2014. p. 767-76.
- [6] Garcia HE, Chen J, Kim JS, McKellar MG, Deason WR, Vilim RB, et al. Nuclear hybrid energy systems - regional studies: west Texas & northeastern Arizona. Tech. Rep. Idaho Falls, ID: Idaho National Laboratory, INL/EXT-15-34503; Apr. 2015.
- [7] Garcia HE, Chen J, Kim JS, Vilim RB, Binder WR, Sitton SMB, et al. Dynamic performance analysis of two regional nuclear hybrid energy systems. Energy 2016:107:234-58
- [8] Kim JS, Chen J, Garcia HE. Modeling, control, and dynamic performance analysis of a reverse osmosis desalination plant integrated within hybrid energy systems. Energy 2016;112:52-66.
- [9] Chen J, Garcia HE, Kim JS, Bragg-Sitton SM. Operations optimization of nuclear hybrid energy systems. Nucl Technol 2016;195(2):143-56.
- [10] Chen J, Garcia HE. Economic optimization of operations for hybrid energy systems under variable markets. Appl Energy 2016;177:11-24.
- [11] Rabiti C, Alfonsi A, Mandelli D, Cogliati J, Martinueau R, Smith C. Deployment and overview of RAVEN capabilities for a probabilistic risk assessment demo for a PWR station blackout. Tech. Rep. Idaho Falls, ID: Idaho National Laboratory; 2013.
- [12] Lei M, Shiyan L, Chuanwen J, Hongling L, Yan Z. A review on the forecasting of wind speed and generated power. Renew Sustain Energy Rev 2009;13(4): 915-20.
- [13] Pinson P, Girard R. Evaluating the quality of scenarios of short-term wind power generation. Appl Energy 2012;96:12-20.
- [14] Mendes J, Sumaili J, Bessa R, Keko H, Miranda V, Botterud A, et al. Very shortterm wind power forecasting: state-of-the-art. Tech. Rep. Argonne National Laboratory (ANL), ANL/DIS-14/6; 2014.
- [15] Kusiak A, Zheng H, Song Z. Short-term prediction of wind farm power: a data mining approach. IEEE Trans Energy Convers 2009;24(1):125-36.
- [16] Pinson P, Madsen H, Nielsen HA, Papaefthymiou G, Klöckl B. From probabilistic forecasts to statistical scenarios of short-term wind power production. Wind energy 2009;12(1):51-62.
- [17] Soman SS, Zareipour H, Malik O, Mandal P. A review of wind power and wind speed forecasting methods with different time horizons. In: North American power symposium (NAPS), 2010. IEEE; 2010. p. 1–8.
- [18] Barbounis TG, Theocharis JB, Alexiadis MC, Dokopoulos PS. Long-term wind speed and power forecasting using local recurrent neural network models. IEEE Trans Energy Convers 2006;21(1):273-84.
- [19] Lee D, Baldick R. Analyzing the variability of wind power output through the power spectral density. In: 2012 IEEE power and energy society general meeting, IEEE: 2012, p. 1-8.
- [20] Tascikaraoglu A, Uzunoglu M. A review of combined approaches for prediction of short-term wind speed and power. Renew Sustain Energy Rev 2014;34: 243-54
- [21] Bludszuweit H, Domínguez-Navarro JA, Llombart A. Statistical analysis of wind power forecast error. IEEE Trans Power Syst 2008;23(3):983-91.
- [22] Zhang Y, Wang J, Wang X. Review on probabilistic forecasting of wind power generation. Renew Sustain Energy Rev 2014;32:255-70.
- Liu H, Tian H-Q, Chen C, Li Y-f. A hybrid statistical method to predict wind [23] speed and wind power. Renew Energy 2010;35(8):1857-61.
- [24] Cassola F, Burlando M. Wind speed and wind energy forecast through Kalman filtering of numerical weather prediction model output. Appl Energy 2012;99: 154 - 66
- [25] Louka P, Galanis G, Siebert N, Kariniotakis G, Katsafados P, Pytharoulis I, et al. Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. J Wind Eng Ind Aerodyn 2008;96(12):2348–62. Hocaoğlu FO, Fidan M, Gerek ÖN. Mycielski approach for wind speed pre-
- [26] diction. Energy Convers Manag 2009;50(6):1436-43.
- [27] El-Fouly T, El-Saadany E, Salama M. Grey predictor for wind energy conversion systems output power prediction. IEEE Trans Power Syst 2006;3(21): 1450-2
- [28] Suganthi L, Samuel AA. Energy models for demand forecasting-A review. Renew Sustain Energy Rev 2012;16(2):1223-40.
- [29] Meibom P, Barth R, Hasche B, Brand H, Weber C, O'Malley M. Stochastic optimization model to study the operational impacts of high wind

- [61] Ippolito M, Di Silvestre M, Sanseverino ER, Zizzo G, Graditi G. Multi-objective optimized management of electrical energy storage systems in an islanded network with renewable energy sources under different design scenarios. Energy 2014;64:648-62.
- [62] Graditi G, Di Silvestre ML, Gallea R, Riva Sanseverino E. Heuristic-based shiftable loads optimal management in smart micro-grids. IEEE Trans Ind Inf 2015;11(1):271-80.
- [63] Banos R, Manzano-Agugliaro F, Montoya F, Gil C, Alcayde A, Gómez J. Optimization methods applied to renewable and sustainable energy: a review. Renew Sustain Energy Rev 2011;15(4):1753–66.
- Box GE, Jenkins GM, Reinsel GC, Ljung GM. Time series analysis: forecasting and control. John Wiley & Sons; 2015. [64]
- [65] Wilcox S, Marion W. Users manual for TMY3 data sets. Golden, CO: National Renewable Energy Laboratory; 2008. Finkelstein JM, Schafer RE, Improved goodness-of-fit tests. Biometrika
- [66] 1971;58(3):641-5.
- [67] Rabiti C, Cherry RS, Deason WR, Sabharwall P, Bragg-Sitton SM, Boardman RD. Framework for the economic analysis of hybrid systems based on exergy consumption. Tech. Rep. Idaho Falls, ID: Idaho National Laboratory, INL/EXT-14-32934; Aug. 2014.
- [68] Farris PW, Bendle NT, Pfeifer PE, Reibstein DJ. Marketing metrics: the defin-itive guide to measuring marketing performance. Upper Saddle River, New Jersey: Pearson Education, Inc.; 2010.