Electric Power Systems Research 134 (2016) 9–18

Contents lists available at ScienceDirect

Electric Power Systems Research

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

ANN-based scenario generation methodology for stochastic variables of electric power systems

Stylianos I. Vagropoulos [a], Evaggelos G. Kardakos [a], Christos K. Simoglou [a], Anastasios G. Bakirtzis [a,*], João P.S. Catalão [b,c,d]

[a] Department of Electrical & Computer Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece
[b] Faculty of Engineering of the University of Porto, R. Dr. Roberto Frias, 4200-465 Porto, Portugal
[c] University of Beira Interior, R. Fonte do Lameiro, 6201-001 Covilhã, Portugal
[d] INESC-ID, Inst. Super. Tecn., University of Lisbon, Av. Rovisco Pais, 1, 1049-001 Lisbon, Portugal

ARTICLE INFO

Article history:
Received 1 December 2014
Received in revised form 7 October 2015
Accepted 21 December 2015
Available online 21 January 2016

Keywords:
Artificial neural networks
Load forecasting
Photovoltaic generation
Scenario generation
Wind production

ABSTRACT

In this paper a novel scenario generation methodology based on artificial neural networks (ANNs) is proposed. The methodology is flexible and able to generate scenarios for various stochastic variables that are used as input parameters in the stochastic short-term scheduling models. Appropriate techniques for modeling the cross-correlation of the involved stochastic processes and scenario reduction techniques are also incorporated into the proposed approach. The applicability of the methodology is investigated through the creation of electric load, photovoltaic (PV) and wind production scenarios and the performance of the proposed ANN-based methodology is compared to time series-based scenario generation models. Test results on the real-world insular power system of Crete and mainland Greece present the effectiveness of the proposed ANN-based scenario generation methodology.

1. Introduction

The development of sustainable energy systems based on reduced fossil fuel emissions, improved energy efficiency and increased renewable energy sources (RES) penetration is a leading priority of the energy roadmaps in many countries worldwide [1,2]. However, the promotion of sustainability in power systems should take into account the inherent characteristic of uncertainty, which poses difficulties in predicting the exact values of many random variables that influence the power system operation in different time scales (long-term, short-term, real-time). For instance, the electric load, the generation unit availability and the RES production (characterized by high variability and uncertainty) are stochastic variables that have a strong impact on the secure, reliable and efficient power system operation and management. The great value for predicting these variables led to the development of appropriate forecasting tools that in some cases can be very accurate (e.g. hourly load forecasting for large regions).

Popular forecasting methods include ARMA models [3] that are used for stationary time series, ARIMA models for non-stationary processes, SARIMA models that capture seasonal patterns of the time series and ARMAX models that include input terms related to exogenous parameters [4]. Another forecasting approach includes probabilistic methods based on probability density functions (PDFs) [5]. Finally, another popular approach that is able to capture both linear and non-linear dependencies and has been widely used in power systems engineering and other sciences comprises the Artificial Neural Networks (ANN). Relevant bibliography dealing with the design and use of ANNs for load, photovoltaic (PV) and wind generation forecasting can be found in Refs. [6–8].

In many cases, and especially in power systems with a large share of variable RES production, typical deterministic scheduling procedures based on point forecasts are not adequate. A point forecast represents an estimation, a single summary statistics, for the examined random variable [9]. However, the probability that this event occurs in clearly close to zero, since a point forecast is always subject to an error. Therefore, stochastic approaches have been adopted lately using multiple scenarios as inputs that account for possible realizations of the random variable and not just the most likely outcome. Through the stochastic approaches more robust solutions are expected compared to the deterministic approaches based solely on the point forecast, since more information regarding the uncertainty of the random variable in incorporated in the optimization problem. However, the challenge

* Corresponding author. Tel.: +30 2310 996383; fax: +30 2310 996302.
E-mail address: bakiana@eng.auth.gr (A.G. Bakirtzis).

http://dx.doi.org/10.1016/j.epsr.2015.12.020
0378-7796/

of creating reliable scenario sets it is not an easy task. The effects of stochastic wind and load on the unit commitment and economic dispatch problems with high levels of wind power are investigated in [10], where it is shown that stochastic optimization results in better performing schedules than the deterministic optimization. A two-stage stochastic programming model for committing reserves in systems with large amounts of wind power is presented in [11], where the proposed model outperforms common deterministic rules for reserve quantification. Finally, the solution of the unit commitment problem under a two-stage stochastic programming formulation is investigated in [12] considering the effects of generation availability and load uncertainty.

For these approaches that require the presence of a scenario set, various methodologies have been proposed in the literature for the generation of a representative set of scenarios for random variables. A popular scenario generation technique is the moment matching method that was used in [13] to generate a limited number of scenarios that satisfy specified statistical properties. The basic idea is to minimize some measure of distance between the statistical properties of the generated outcomes and the specified properties. The same authors in [14] presented an algorithm that produces a discrete joint distribution consistent with specified values of the first four marginal moments and correlations. The joint distribution is constructed by decomposing the multivariate problem into univariate ones and using an iterative procedure to achieve the correct correlations without changing the marginal moments. An approach that relies on the moment-matching technique was proposed in [15]. The approach was based on the idea of integrating simulation and optimization techniques. In particular, simulation is used to generate outcomes associated with the nodes of the scenario tree, which, in turn, provide the input variables for an optimization model that aims at determining the scenarios' probabilities matching some prescribed targets. An algorithm based on heteroskedastic models and a moment matching approach to construct a scenario tree that is a calibrated representation of the randomness in risky asset returns was presented in [16]. Another widely used scenario generation technique is the path-based method [17]. This method evolves the stochastic process to generate complete paths in a 'fan' structure, which is transformed into a scenario tree using "clustering", also called "bucketing".

An optimization-based method to generate moment matching scenarios for numerical integration and how it can be used in stochastic programming has been proposed in [18]. The main advantage of the method is its flexibility: it can generate scenarios matching any prescribed set of moments of the underlying distribution rather than matching all moments up to a certain order, and the distribution can be defined over an arbitrary set. In the same framework, three approaches for generating scenario trees for financial portfolio problems have been presented in [19]. These are based on simulation, optimization and hybrid simulation/optimization. Finally, an optimal discretization method that seeks to find an approximation of the initial scenario set that minimizes an error based on the objective function was described in [20].

Other scenario generation methods can be found in papers that deal with wind power uncertainty. A method that allows for the generation of statistical scenarios from non-parametric probabilistic forecasts is described in [9], while a first-order autoregressive time-series model with an increasing noise to approximate the behavior of wind speed forecast errors is presented in [21]. This model allows for the creation of a large number of wind speed scenarios using Monte Carlo simulations, which are then transformed into wind power scenarios with the use of an aggregated power curve model. In addition, simple scenarios around point forecasts are generated in [22] for the optimal scheduling of the generators in a wind integrated power system considering the demand and wind power production uncertainty. Finally, a new scenario generation methodology that adopts the empirical distributions of a number of forecast bins to model the forecast error of wind power, which are used as inputs to scenario generation, is proposed in [23].

This paper proposes a novel scenario generation methodology suitable to account for various stochastic variables commonly used in power system studies (e.g. electric load, PV and wind production). The proposed technique combines ANNs that are able to capture the linear or non-linear dependencies of the time series under consideration with its historical values, as well as with exogenous variables (e.g. ambient temperature, wind speed, solar radiation, etc.), with an iterative process based on the assimilation of randomly generated uncorrelated error terms with specific statistical properties to the ANN outputs. Actually, the methodology presented in this paper is an extension of the methodology proposed in [25] that includes time series analysis to create scenarios. The extension of the methodology takes advantage of the easy modeling of time stamping, which generally presents correlations with the underlying variables, and the easy incorporation of exogenous inputs in the ANN-based scenario generation application. These features are very useful in creating more representative and well-defined sets of scenarios, further analyzed in the following through the comparison of the proposed approach with two relevant scenario generation approaches.

The remainder of the paper is organized as follows: Section 2 describes in detail the proposed scenario generation methodology, which is extended to account for cross-correlated stochastic processes and scenario reduction techniques. Section 3 presents results from the application of the proposed methodology for the creation of scenarios of electric load, PV and wind production. In Section 4 the performance of the proposed methodology is compared with two relevant scenario generation approaches, while in Section 5 the value of the proposed method for the optimal participation of a PV agent in day-ahead electricity market with respect to the other two scenario generation approaches is investigated. Finally, valuable conclusions are drawn in Section 6.

2. Methodology

2.1. Forecasting with ANNs

Neural Networks (NNs) may be seen as multivariate, nonlinear and nonparametric methods, and they could be expected to model complex nonlinear relationships much better than the traditional linear models [24]. Given a proper set of explanatory variables for a single stochastic variable, the NN is trained in order to capture the nonlinear dependency between the inputs and the respective output. Therefore, NNs can be seen as a "black-box", however a good engineering judgment is necessary for a successful NN setup (layers and neurons selection, explanatory variable selection, etc.). In contrast to traditional linear models, NNs are very flexible in integrating time stamping inputs, such as the hour-of-day, day-of-year, etc. Many stochastic processes are highly related to time stamping information and this is a valuable advantage of NNs.

A detailed analysis on ANN training properties and parameterization practices can be found in [24]. The ANN structure used in this paper includes a multi-input/single output design (one step-ahead forecasting) with three layers (one input, one output and one hidden layer) and feed-forward design (i.e. the outputs of one layer are used as inputs to the following layer). In general, the estimation of parameters (weights) is performed by minimizing a loss function (usually a quadratic function of the output error is used). For the parameters estimation, a back-propagation algorithm which uses a steepest-descent technique based on the computation of the gradient of the loss function with respect to the network parameters