Applied Energy 228 (2018) 546–555

Contents lists available at ScienceDirect

Applied Energy

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

A novel data-driven scenario generation framework for transmission expansion planning with high renewable energy penetration

Mingyang Sun*, Jochen Cremer, Goran Strbac

Department of Electrical and Electronic Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

HIGHLIGHTS

- A novel data-driven load and wind power scenario generation framework is proposed.
- A sufficient number of unseen but important scenarios are generated for TEP.
- Interspatial dependencies between loads and wind power output are accurately captured.
- The method can identify near optimal investment decisions with higher net benefits.

ARTICLE INFO

Keywords:
Renewable energy resources
Multivariate dependence
Scenario generation
Wind power
Transmission expansion planning
Vine copulas

ABSTRACT

Transmission expansion planning (TEP) is facing unprecedented challenges with the rise of integrated renewable energy resources (RES), flexible load elements, and the potential electrification of transport and heat sectors. Under this reality, the inadequate information of the stochastic parameters' behavior may lead to inefficient expansion decisions, especially in the context of very high renewable penetration. This paper proposes a novel data-driven scenario generation framework for the TEP problem to generate unseen but important load and wind power scenarios while capturing inter-spatial dependencies between loads and wind generation units' output in various locations, using a vine-copula based high-dimensional stochastic variable modeling approach. The superior performance of the proposed model is demonstrated through a case study on a modified IEEE 118-bus system. The expected result of using the expected value problem solution (EEV) and the net benefits of transmission expansion (NBTE) are used as the evaluation metrics to quantitatively illustrate the advantages of the proposed approach. In addition, the case of very high wind penetration is carried out to further highlight the importance of the multivariate stochastic dependence of load and wind power generation. The results demonstrate that the proposed scenario generation method can result in near-optimal investment decisions for the TEP problem that make more net benefits than using limited number of historical data.

1. Introduction

The increasing penetration of Renewable Energy Resources (RES) and flexible demand introduces more operational variability and uncertainty in Transmission Expansion Planning (TEP) in power systems. Meanwhile, in order to reduce carbon emissions, it becomes imperative to further decrease the use of fossil fuel in energy supply through the electrification of transport and heating sectors, which may introduce radical changes in power demand when it is associated with electric vehicles (EVs) and heat pumps (HPs) [1–3]. This new reality renders it imperative to carry out transmission investment that considers multiple uncertainties (e.g., demand growth, location, timing and amount of connection of renewable generation) on a cost-benefit basis. The optimal decision should minimize the total cost, including the system operation and transmission investment costs, especially in the context of very high renewable penetration. In particular, uncertainties in wind generation arise from the large-scale deployment and the inherent variability in production. Consequently, beyond the classical TEP challenges of non-convexity, nonlinearity and the mixed integer problem (MIP) (e.g. linearized with direct current (DC) power flow [4]), the focus of recent researchers has shifted towards incorporating appropriate models that deal with uncertainty.

Conventionally, the TEP problem has been solved by deterministic approaches (e.g., linear programming [5]). However, it may not be appropriate to perform TEP only based on one operating profile, especially in the context of substantial power flow fluctuations

* Corresponding author.
E-mail address: mingyang.sun11@imperial.ac.uk (M. Sun).

https://doi.org/10.1016/j.apenergy.2018.06.095
Received 13 April 2018; Received in revised form 4 June 2018; Accepted 20 June 2018

Nomenclature

Sets and indices

$\gamma \in \Gamma$	indices of thermal generators, $\Gamma \subset G$
$d \in N_d$	indices of load
$g \in G$	indices of all generators
$i, j \in \Lambda$	existing transmission lines
$i, j \in \Lambda_{new}$	candidate transmission lines
$i, j \in \Omega_B$	indices of the network buses
$t \in T$	scenarios
$w \in W$	indices of wind generators, $W \subset G$

Input parameters

$\boldsymbol{D}$	matrix of sampled load [MW]
$\boldsymbol{P}_{\Gamma}^{max}$	matrix of maximum conventional generators capacities [MW]
$\boldsymbol{P}_{W}^{max}$	matrix of sampled wind power [MW]
$\boldsymbol{Z}$	branch-node incidence matrix with elements z_{ij}
π_g	generation cost of generator g [$/MWh]
Ψ	curtailed wind power penalty cost [$/MWh]
τ	duration of each scenario [hour]
Υ	value of lost load [$/MWh]
B_{ij}	susceptance of existing line from bus i to j [p.u.]
B_{ij}^{new}	susceptance of candidate line from bus i to j [p.u.]
f_{ij}^{max}	power flow limit of the transmission line from bus i to j [MW]
$I_{i,j}$	cost of building a line from bus i to j [$/year]
p^t	probability of scenario t

Decision variables

$\boldsymbol{F}$	matrix of power flow with elements f_{ij} [MW]
$\boldsymbol{P}_{\Gamma}$	matrix of conventional generator output [MW]
$\boldsymbol{P}_{\boldsymbol{W}}$	matrix of wind generator output [MW]
$\boldsymbol{U}$	matrix of load curtailment with elements U_d [MW]
$\boldsymbol{V}$	matrix of wind curtailment with elements V_W [MW]
θ_i	phase angle at bus i [rad]
f_{ij}	power flow of transmission line from bus i to j [MW]

introduced by the stochastic variables beyond the operator's control. To this end, probabilistic TEP models (e.g. two-point estimation method [6] and Monte Carlo simulation [7]) have been proposed to consider the uncertainties stemming from various sources. In [8], a novel method based on a combination of robust and stochastic optimization strategies has been proposed to solve the bundled generation and transmission planning problem. Contrasted with stochastic programming methods, dynamic robust optimization approach has been implemented in [9] to address the challenges of the resolution of the dynamic TEP problem when considering the year-by- year representation of uncertainties and investment decisions in an integrated manner. In [10], a novel TEP model considering the consumption-based carbon emission accounting is proposed. In addition, robust optimization based on reformulation (i.e., linear decision rules) and decomposition (i.e., column-and-constraints generation method) has been introduced in [11] as an efficient technique in developing models considering flexibility and resiliency of power systems. Note that the focus of this paper is not investigation of alternative advanced TEP models but on a novel scenario generation framework that can provide sufficient and effective input data for solving a Monte Carlo approximation of the stochastic TEP problem.

Recently, the uncertainties in both demand and wind have been simultaneously investigated in the literature. In general, three types of assumptions are considered: (i) no correlation (i.e. uncorrelated demand and total wind generation [12]); (ii) bivariate correlation (i.e. correlated between total load and total wind [13,14]); (iii) multivariate correlation (i.e. inter-spatial correlated between demand and wind generation at various locations [15]). As illustrated in [13], load and wind power generation may not have statistically independent magnitudes (e.g., relatively high wind generation outputs may correspond to low values of load during the night). Neglecting correlation between load and wind-generation may result in suboptimal and inefficient investment decisions. Nevertheless, it is inappropriate to assume that the individual variables (i.e. electricity consumption at each node and power output of each wind farm) are perfectly correlated. This substantial simplification is due to high computational burden related to modeling of the stochastic variables at the individual level. However, in the context of high penetration of renewable generation, it is crucial to capture the inter-spatial correlations of load and wind output at various locations. As stated in [16], a stochastic program with the real optimal solution can be approximately solved by sampling a certain number of the program's stochastic parameters. With the increasing number of samples, the solution of the approximated problem will be improved and converged to the real optimal solution if distributions of the stochastic variables are available and a sufficient number of samples can be generated. To this end, carrying out Monte Carlo simulation-based TEP entirely depending on the limited number of real historical observations [7] is insufficient in the high RES penetration case because not all potential future scenarios can be derived from historical observations. Therefore, it becomes crucial to identify the underlying distribution of the existing data and generate as many samples as required to enhance the investment decisions for TEP. In the literature, a series of wind power scenario generation approaches have been proposed such as [17,18]. Note that some of the proposed scenario generation approaches (e.g., [13,19]) are actually scenario reduction methods, which aim to reduce the computational time for solving the TEP problem by selecting representative scenarios from historical data that can describe the uncertainty of and the correlation between load and wind-power production with corresponding probabilities, which is not the objective of this paper. In this paper, we propose a framework for generating unseen but important scenarios (e.g., high demand and low wind-power production) that may drive transmission investment for future electricity systems.

To obtain an accurate representative distribution, multivariate statistical models have been broadly implemented to interpolate and extrapolate historical observations. The constructed model can generate samples that are similar but not identical to what has already been encountered [20]. For example, Gaussian Mixture Model (GMM) has been employed in [21] to represent distribution system loads. Recently, the authors in [22] have proposed the generalized dynamic factor model (GDFM) to represent the load and wind generation as the sum of a periodic component, common component, as well as idiosyncratic noise component, with the capability of retaining the correlation structure between load and wind. Beyond the correlation structure, another challenging task of a multivariate statistical framework is in its ability to model large-scale measurement data by accurately capturing the nonlinear dependence structure of data and the non-Gaussian marginal distributions. Given this point, copulas have been demonstrated as a powerful tool to accomplish this task [23,24]. For example, in [25], multivariate Gaussian copula is used to model historical measurement and to generate synthetic wind power output from 15 sites in the Netherlands. Authors in [26] employ the high-dimensional copula theory and discrete convolution method to conduct a high-dimensional dependent discrete convolution calculation for analyzing power system