

# Generating joint scenarios of renewable energy sources: the case for non-Gaussian models with time varying parameters

Henrique Hoeltgebaum, Cristiano Fernandes, and Alexandre Street, *Senior Member IEEE*

**Abstract**—The development of medium- and long-term studies for power-system operation and planning under uncertainty of renewable generation is one of the key challenges faced by planners and market players in most of the power system world-wide. There is a vast literature on stochastic optimization models devoted to address the relevant issues on power system operation and planning applications. Notwithstanding, few papers focus on addressing the gaps on the subject of joint scenario generation despite the high sensibility of the stochastic optimization models with regard to their input scenarios. The characterization wind power generation (WPG) stochastic processes to devise time- and spatial-dependent scenarios, based on simulation procedures, for one to few years horizon is a difficult task. Multiple regimes and non Gaussian distributions are one of the main issues faced in both centralized and private planning/investment studies. In this paper a new methodology to simulate long-term joint scenarios for multivariate WPG time series is presented. The proposed framework is derived based on a new class of time-series model with time-varying parameters and arbitrary non Gaussian distribution, known as Generalized Auto Regressive Score models (GAS). GAS models provide a relevant flexibility and control of the simulated scenarios in long-run chronological Monte Carlo simulation procedures. In this work, we study the case of a beta PDF and spacial dependence is captured by an elliptical copula with time-varying correlation matrix. Case studies based on the Brazilian power system show that the proposed methodology is capable to address relevant issues that arise in long-term simulation studies.

**Index Terms**—Score driven models, wind power, probabilistic forecasting, time varying parameters, dynamic copula.

## I. INTRODUCTION

**R**ENEWABLE energy expansion has been growing world-wide, mainly in response to governmental incentive for reducing greenhouse gas emissions. In particular, wind-power generation (WPG) is one of the largest sources of renewable energy, and according to the International Energy Agency, will respond to 18% of global power by 2050 [1]. However the uncertainty associated with its non dispatchable nature may jeopardize reliability of electricity supply. In attempting to minimize this kind of risk it is highly desirable to produce reliable forecasts and scenarios for WPG time series. The importance of such scenarios emerges in many instance, as for example: (i) energy trading, (ii) unit commitment, (iii)

grid expansion planning, and (iv) investment decisions (see ([2]–[5]) and references therein). There is a vast literature on stochastic optimization models devoted to address the relevant issues on the aforementioned power system applications. Notwithstanding, few papers focus on addressing the gaps on the subject of joint scenario generation despite the high sensibility of the stochastic optimization models with regard to their input scenarios.

In contrast to the main objective of WPG simulation methods, which aim to produce scenarios that characterize all the conditional density and its fit to the observed quantile data, much research has been devoted to devise short-term forecast methods for wind speed and WPG time series (for instance, see [6]). In some applications on medium and long-term forecast, the time series model is conducted using conventional ARMA models with seasonal lags, or SARIMA models, under a Gaussian distribution (see [7] for a high-dimension estimation process based on LASSO). In its original formulation, SARIMA models present a fixed conditional variance. To add extra flexibility to these models [8] proposes an ARIMA model with GARCH effect, allowing the conditional variance of the WPG distribution to vary over time. Also, to provide a description of the spatial correlation of wind-speed time series, [9] proposed an Autoregressive Fractionally Integrated-GARCH model (ARFIMA-GARCH). In these and other models, conditional on the past, the distribution of the response variable is assumed Gaussian.

The non-Gaussian nature of WPG time series is well reported in the literature (see [10]–[13] and references therein). However, only very limited classes of non-Gaussian time-series models are available for modeling WPG. The state of the art literature on short-term probabilistic modeling of WPG time series mainly rely on non-parametric models. For instance, in [10] and [13], non-parametric approaches were devised to forecast conditional distributions. While in [10] kernel density estimators were employed to build a predictive conditional density function, in [13] a quantile-regression-based model was used. Despite the virtues found in nonparametric methods, those models require a lot of data to be fitted and are mainly devised for the univariate case. Thus, the development of new parametric models capable to properly characterize the full multivariate non-Gaussian distribution for WPG time series is a relevant research theme, which did not received too much attention so far. Relevant applications arise from risk assessment in both medium and long-term planning/investment studies, which mainly rely on few data points to characterize

This work was partially supported by the R&D project ANEEL PD-7625-0001/2013.

Henrique Hoeltgebaum, Cristiano Fernandes and Alexandre Street are with the Electrical Engineering Department, Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Rio de Janeiro, RJ, Brazil (e-mail: hhhelfer@ele.puc-rio.br; cris@ele.puc-rio.br; street@ele.puc-rio.br).

conditional estimates of extreme quantiles [5], [14]–[17].

In [18] and [19], a general framework for time series models with time-varying coefficients was proposed considering any non-Gaussian conditional distribution, either discrete or continuous, univariate or multivariate. Such model has been named in the recent literature as Generalized Autoregressive Score (GAS) model [18] and Dynamic Conditional Score (DCS) [19]. In [18], it has been shown that several well-known time-series models from the econometric literature are a particular case of GAS models<sup>1</sup>. More specifically, a GAS model is built based on a user-defined conditional probability function whose parameters follow a data-driven dynamic equation, generally on the form of an autoregressive mechanism. Then, by construction, time-varying parameters can be accommodated according to an updating mechanism that uses the score as its driving force. The use of the score function for updating time-varying parameters is an intuitive choice. It is defined as the steepest ascent direction (gradient) for improving the local fit of the model in terms of likelihood. In such an updating mechanism, information from the whole density is used to track the evolution of time varying parameters.

Most of the applications using GAS models have been published in the recent literature of finance and economics. We refer the interested reader to the main on-line repository on GAS papers [23]. The virtue of GAS models to address non-linear and non-Gaussian time series notwithstanding, to the best of the authors' knowledge there is no publication in the power systems literature using such framework. The consideration of a seasonal structure in the updating mechanism for the time-varying parameters was also not addressed in the GAS literature so far, despite being a relevant structure shared by a wide range of time series such as WPG (see [7], [14], [24], [25] and references therein).

To generate joint scenarios for WPG time series based on GAS models for power plants belonging to different geographical areas, it is important to include spatial dependence information among these units [7]. In [26], the authors provided an empirical application for a panel of daily-equity returns updating both variance and correlation matrix by the proposed GAS mechanism for a multivariate Student t density function. They indicate the link between their framework to a time-varying copula specification, but no results and tests were provided. Furthermore, it is possible to combine linear models with copulas that also have time-varying parameters as shown in [27]–[29]. However, none of such works proposed a GAS mechanism to update time-varying parameters from both marginal densities and copula structures.

The objective of this paper is to propose a new methodology to address non-Gaussian WPG multivariate time-series through GAS models. We also propose a time-varying copula-based approach to capture the dynamics of spatial dependencies. Therefore, the contribution of this paper are twofold:

- A new simulation framework based on GAS models

<sup>1</sup>For instance, GARCH models that deal with heavy tailed distributions in finance [20], Autoregressive Conditional Duration models [21] to tackle asymmetric distributions in finance applications such as time duration, and the Poisson count models of Davis [22] are particular cases of GAS models.

capable to generate non-Gaussian joint scenarios for multivariate WPG time series. The proposed framework allows the simulation of non-Gaussian scenarios widely demanded in many power system applications such as planning, operation, and risk assessment.

- A new spatial-dependency method for GAS models through a time-varying elliptical copula model applied to the residuals of the univariate GAS models. In our proposed framework, the copula parameter, namely, a correlation matrix, is guided by a novel second GAS mechanism allowing the consideration of a spatial-dependence structure that changes along seasons. This is a significant step forward on the subject of modeling GAS models that allows a good characterization of multivariate WPG time series where all these effects are found. Thus, in our proposed framework, both time and spatial dependences are guided through GAS-based updating mechanisms.

The rest of this paper is organized as follows. In section II, the proposed univariate GAS model framework with seasonal dynamics in the updating mechanism is presented. Estimation, diagnoses, and forecast procedures are also discussed in this section. In section III, an elliptical-copula model is devised based on [26] to model the time-varying correlation matrix for the residuals. In section IV, a case study using realistic data from the Brazilian power system is presented to illustrate the application of the proposed framework. Finally, in section V conclusions are addressed.

## II. GAS MODELS

We consider a generic normalized WPG time series with a support in the range  $[0,100]$ , in percentage of the maximum power injected. The densities selection rely on the support range and shape of conditional distributions. In this work we select and study the case of a Beta distribution with the first-shape parameter varying over time. More precisely, in our application, the  $i$ -th, from a set of  $K$ , WPG time series is described by a beta probability density function (PDF) as follows:

$$p(y_{it}|f_{it}, \mathcal{F}_{i,t-1}; \theta_i) \sim \text{beta}(\beta_{it}, \alpha_i) \forall i \in K \text{ and } t \in T, \quad (1)$$

where  $y_{it}$  is the WPG injected by unit  $i \in K$ , at time  $t$ ,  $f_{it}$  is generally defined as a logarithmic transformation of the time-varying first shape parameter of the beta distribution,  $\beta_{it}$ , to avoid nonnegativity constraints in the estimation process, and  $\alpha_i$  is the second shape parameter, in this paper assumed to be fixed.  $\theta$  is the vector of hyperparameters, assumed to be static in time, encompassing  $\alpha_i$ , dummy variables to capture the effect of outliers when needed, and the rest of parameters to be estimated.

Thus, the time-varying parameter GAS( $p, q$ ) updating mechanism is devised for  $f_{it}$  as follows:

$$f_{i,t+1} = \omega_i + \sum_{l=1}^p A_{i,l} s_{i,t-l+1} + \sum_{l=1}^q B_{i,l} f_{i,t-l+1}. \quad (2)$$

In expression (2),  $s_{i,t-l+1}$  is the score of the beta PDF for unit  $i$  at time  $t-l+1$ , and  $\omega_i$ ,  $A_{i,l}$ , and  $B_{i,l}$  are fixed parameters, which are considered as part of vector  $\theta_i$ .

To complete the description of the the updating mechanism of GAS models it is necessary to define  $s_{i,t}$  in Equation (2). The authors in [18] proposed the following scheme:

$$s_{i,t} = \mathcal{I}_{i,t|t-1}^{-d} \cdot \nabla_{i,t} \quad (3)$$

$$\nabla_{i,t} = \frac{\partial \ln p(y_{i,t}|f_{i,t}, \mathcal{F}_{i,t-1}; \theta_i)}{\partial f_{i,t}} \quad (4)$$

$$\mathcal{I}_{i,t|t-1} = E_{t|t-1}[\nabla_{i,t}^T \nabla_{i,t}], \quad (5)$$

where  $p(y_{i,t}|f_{i,t}, \mathcal{F}_{i,t-1}; \theta_i)$  is the observation PDF,  $\mathcal{F}_{i,t-1}$  collects all relevant information of time series  $i$  up to time  $t-1$ , and  $\mathcal{I}_{i,t|t-1}$  is a scaling factor, which accounts for the variance of  $s_{i,t}$ . More details can be found in [18]. In practice, the choice of  $d$  has shown to be an empirical question: for a given model one chooses the value of  $d$  which produces best diagnostics and forecasting.

As mentioned before, because the beta parameter can only assume positive values, it is conveniently reparameterized as the exponential of an unconstrained parameter, namely,  $f_{i,t}$ . In addition, to remove the effect of possible outliers, dummy variables can be considered into parametrization. In this sense, if we define the time-varying beta parameter as

$$\beta_{i,t} = e^{\phi_i^T D_{i,t} + f_{i,t}}, \quad (6)$$

with

$$D_{i,t} = \begin{cases} 1, & \text{if } y_{i,t} \text{ is an outlier;} \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

#### A. Estimation

Regarding the estimation of the fixed parameter vector  $\theta_i$  for each time series  $i$ , the estimation procedure is based on the following log-likelihood maximization problem:

$$\hat{\theta}_i = \underset{\theta_i}{\operatorname{argmax}} \sum_{t=1}^n \ln[p(y_{i,t}|f_{i,t}, \mathcal{F}_{i,t-1}; \theta_i)]. \quad (8)$$

Because  $f_{i,t}$  is a function of  $\theta_i$  (we recall that the coefficients in equation (2) are all part of the vector of hyperparameters,  $\theta_i$ ), to solve the estimation problem (8), a numerical non-linear optimization procedure has to be used. In this work, we apply the Nelder Mead algorithm to find good initial points to the BFGS algorithm. Notwithstanding, the evaluation of the log-likelihood function requires initial values for  $f_{i,0}, f_{i,-1}, \dots, f_{i,1-q}$  and  $s_{i,0}, s_{i,-1}, \dots, s_{i,1-p}$ . These are obtained by fitting fixed beta distributions, one for each month, assuming both parameters fixed in time. More precisely, the WPG time series is split into 12 monthly time series, i.e.,  $\{Y_{Jan}^{(i)}, \dots, Y_{Dec}^{(i)}\}$ . Then, for each of these time series  $\{Y_v^{(i)}\}_{v=1}^{12}$ , a fixed beta density is estimated via maximum likelihood, resulting in the estimates  $\{\alpha_v^{(i)}\}_{v=1}^{12}$  and  $\{\beta_v^{(i)}\}_{v=1}^{12}$ . These are then used to initialize the recursion in (2) for each evaluation of the log-likelihood function. Thus, the function that is passed to the BFGS algorithm is a numerical function that receives as input a trial value for vector  $\theta_i$ , calculates the value of  $f_{i,t}$  according to expression (2), and returns as output the value of expression (8). Based on an initial value for  $\theta_i$ , the BFGS algorithm iterates until finding a local maximum point.

To enhance the solution quality, a Nelder Mead algorithm is used on the top of this process and its solution is used as initial points for the BFGS step and it iterates until the GAP between these solutions are smaller than 0.1.

#### B. Diagnostics

Diagnostics in GAS models can be obtained using an appropriate type of residuals for non linear and non Gaussian time series models known as quantile residuals. As described in [30], the observed quantile residual is given by

$$r_{t,\hat{\theta}_i} = \Phi^{-1}[F(y_{i,t}|f_{i,t}, \mathcal{F}_{i,t-1}; \hat{\theta}_i)]. \quad (9)$$

Where  $F(\cdot)$  is the cumulative distribution function (CDF) associated with the proposed beta density  $p(y_{i,t}|f_{i,t}, \mathcal{F}_{i,t-1}; \theta_i)$  and  $\Phi^{-1}[\cdot]$  is the quantile function of a standard Gaussian distribution. Under correct model specification, these residuals should be normally distributed and show no temporal dependence. These can be checked, for example, using the Jarque Bera test for normality, the Ljung Box test for absence of serial correlation, and Ljung Box (on squared residuals) to test for absence of non linear dependence, such as ARCH effect.

### III. SPATIAL DEPENDENCE MODEL

Once the proposed beta GAS models were individually applied to each of the  $K$  WPG time series, the observed dependence among different units is captured by the probability integral transformed (PIT) variables,  $\mathbf{u}_t \in \mathbb{R}^K$ , which are calculated as

$$u_{i,t} = F(y_{i,t}|f_{i,t}, \mathcal{F}_{i,t-1}, \hat{\theta}_i) \quad \forall i \in K \text{ and } t = 13, \dots, T. \quad (10)$$

The PIT multivariate time series denoted by  $\mathbf{u}_t$  belong to the  $[0, 1]$  support.

Assuming  $G_\nu$  as the CDF of a univariate Student t distribution with  $\nu$  degrees of freedom, a new vector of t-Student-distributed time series can be build as follows:  $\tilde{\mathbf{y}}_t = [G_\nu^{-1}(u_{1,t}), \dots, G_\nu^{-1}(u_{K,t})]^T$ . The elliptical copula used in this work is developed under the assumption that the transformed time series,  $\{\tilde{\mathbf{y}}_t\}_{t \in T}$ , can be modeled as a multivariate t-Student PDF given by:

$$g(\tilde{\mathbf{y}}_t|\Sigma_t; \nu) = \frac{\Gamma(\frac{\nu+K}{2})}{\Gamma(\frac{\nu}{2}) [(\nu-2)\pi]^{K/2} |\Sigma_t|^{1/2}} \left[ 1 + \frac{\tilde{\mathbf{y}}_t' \Sigma_t^{-1} \tilde{\mathbf{y}}_t}{(\nu-2)} \right], \quad (11)$$

where  $\Sigma_t$  is the correlation matrix which is assumed to follow a second GAS updating mechanism similar to that proposed in [31]. The key of such model is the use of a factorization scheme for  $\Sigma_t$  that allows the updating mechanism to always generate positive definite correlation matrices. Such factorization scheme, used in [31] and simplified here, is described in the sequel:

$$\Sigma_t = \operatorname{diag}(Q_t)^{-1/2} Q_t \operatorname{diag}(Q_t)^{-1/2}. \quad (12)$$

where  $Q_t$  is a symmetric positive definite matrix which guarantees a symmetric positive definite matrix  $\Sigma_t$  with elements outside the main diagonal lying inside the  $[-1, 1]$  range.

The matrix  $Q_t$  carries all the information regarding the spatial dependence structure among different WPG time series. The idea of adopting a time varying mechanism to update the elements of  $Q_t$  is to include the temporal variation of this dependence when generating joint scenarios of WPG. In order to update the quantities of the time-varying vector  $q_t = \text{vech}(Q_t)$ , which stands for the half-vectorization<sup>2</sup> of  $Q_t$ , we adopt the same GAS dynamic proposed in [26], i.e.,

$$\text{vech}(Q_{t+1}) = \Omega + \Pi s_t + \Upsilon \text{vech}(Q_t). \quad (13)$$

In (13),  $\Pi$  and  $\Upsilon$  are diagonal matrices,  $\Omega$  is a  $K \times 1$  vector with elements equals to one when updating the elements of the main diagonal of  $Q_t$  in order to identify either model. Hence the elements of  $\Omega$  corresponding to the diagonal elements of  $Q_t$  can be multiplied by any arbitrary positive number without changing the decomposition. Regarding the elements of the GAS mechanism  $s_t$ ,  $\nabla_t$  and  $\mathcal{I}_{t-1}^{-1}$  for the multivariate Student t distribution, they were calculated in [26] and can be found in Appendix A.

#### A. Estimation of the copula parameters

We now deal with the estimation method used in the second-step GAS procedure, i.e., the estimation of the static parameters of the elliptical-copula model,  $\Theta = (\nu, \Omega, \Pi, \Upsilon)$ . For this we applied the Inference for Margins (IFM) optimization procedure (see [32], [33]). In IFM estimation, the estimated vector of hyperparameters,  $\hat{\theta}_i \in K$ , of each WPG time series are first estimated individually using the same procedure described in Section II-A. Then, based on the individual estimates  $\hat{\theta}_i \in K$ , copula parameters are estimated via maximum likelihood. Considering the  $K$ -dimensional vector  $\tilde{y}_t = [G_\nu^{-1}(u_{1,t}), \dots, G_\nu^{-1}(u_{K,t})]^T$ , the log-likelihood of the multivariate Student t model is

$$\begin{aligned} \mathcal{L} = \sum_{t=1}^n \left\{ \log \left[ \Gamma \left( \frac{\nu + K}{2} \right) \right] - \log \left[ \Gamma \left( \frac{\nu}{2} \right) \right] \right. \\ \left. - \frac{1}{2} \log |\Sigma_t| - \frac{K}{2} \log [(\nu - 2)\pi] \right. \\ \left. - \frac{\nu + K}{2} \log \left[ 1 + \frac{\tilde{y}_t' \Sigma_t^{-1} \tilde{y}_t}{\nu - 2} \right] \right\} \end{aligned} \quad (14)$$

where the temporal variation of  $\Sigma_t$  is described by the GAS mechanism provided by Equation (13). Maximization of the log-likelihood function  $\mathcal{L}$  is analogous to the univariate ones given on Section II-A. The estimation problem (14) is solved through a non-linear optimization procedure which also uses Nelder Mead algorithm to find good initial points to the BFGS algorithm. The degrees of freedom  $\nu$  were estimated using profile likelihood. For more details see, [34].

<sup>2</sup>The half-vectorization is a linear transformation used to convert the lower triangular portion of  $K \times K$  symmetric matrices in to  $K(K+1)/2$  vectors.

#### B. Generating joint scenarios

The scenario generation procedure is built based on two steps that are repeated for every period within the simulation horizon. In the first step,  $M$  individually simulated scenarios, one for each WPG, are sampled from the conditional univariate PDF's and the values of the time-varying parameters are updated according to (2). In the second step, the spacial dependence is adjusted by means of the estimated copula and its time-varying parameters are also updated based on (13). After performing the two-step procedure for the whole simulation horizon, one multivariate path (scenario) is generated. Then, this process is repeated (or parallelized) as many times as the number of samples that one needs to generate for a given application. Formally, the simulation procedure is as follows:

##### Part I - Univariate

- 1) Given  $\hat{\theta}_i \in K$ , obtained by the maximization of (II-A) and the filtered vector  $\hat{f}_{i,T+1}$  obtained by applying the recursion (2), draw  $M$  vectors  $y_{i,T+1}^{(1)}, \dots, y_{i,T+1}^{(M)}$  from the estimated conditional density (1) for time  $T+1$ ;
- 2) Use  $y_{i,T+1}^{(1)}, \dots, y_{i,T+1}^{(M)}$  and the updating equation (2) to obtain  $\hat{f}_{i,T+2}^{(1)}, \dots, \hat{f}_{i,T+2}^{(M)}$  conditional on  $\hat{\theta}_i$  and  $\hat{f}_{i,T+1}$ ;
- 3) For each  $\hat{f}_{i,T+2}^{(m)}$ ,  $m = 1, \dots, M$ , redo steps 1 and 2, for periods  $T+2, \dots, T+h$ , producing vectors  $y_{i,T+h} = [y_{i,T+h}^{(1)}, \dots, y_{i,T+h}^{(m)}]^T \forall i \in K$ ;

##### Part II - Multivariate

- 1) Given  $\hat{\theta}_i \in K$ , apply equation (10) into the vectors  $y_{i,T+h}$ , which produces  $u_{i,T+h} = [u_{i,T+h}^{(1)}, \dots, u_{i,T+h}^{(m)}]^T$ , for  $h \geq 1$  and  $m = 1, 2, \dots, M$ ;
- 2) Given  $\hat{\Theta}$ , and the filtered vector  $\hat{q}_{T+1} = \text{vech}(\hat{Q}_{T+1})$  obtained by applying the recursion (13) into  $\tilde{y}_t = [G_\nu^{-1}(u_{1,t}), \dots, G_\nu^{-1}(u_{K,t})]^T$ , use  $\text{vech}(\hat{Q}_{T+1})$  as initial value for the  $M$  future recursions, for  $h \geq 1$ , of  $\tilde{y}_{T+h} = [G_\nu^{-1}(u_{1,T+h}^{(m)}), \dots, G_\nu^{-1}(u_{K,T+h}^{(m)})]^T$ ;
- 3) With  $\Sigma_{T+h}^{(m)}(\text{vech}(\hat{Q}_{T+h}^{(m)}))$ ,  $m = 1, 2, \dots, M$ ,  $h \geq 2$  and  $\hat{\Theta}$ , apply conditional sampling technique (see [35]) to produce conditional vectors of  $u_{i,T+h}^{(m)*}, \dots, u_{i,T+h}^{(m)*}$ ;
- 4) Since it is a fully parametric copula model, the values of the joint scenarios are given by  $y_{i,T+h}^{*(m)} = F^{-1}(u_{i,T+h}^{(m)*} | f_{i,t}, \mathcal{F}_{i,t}, \hat{\theta}_i) \forall i \in K$ , where  $\beta_{i,T+h} = e^{f_{i,T+h}}$ .

By following this procedure, one obtains  $M$  future joint scenarios of WPG,  $y_{t+h}^* \in \mathbb{R}^K$ , for each one of the  $K$  wind plants. Here and after this model will be denoted as  $t$ -GAS.

#### IV. CASE STUDY - SIMULATING MEDIUM AND LONG-TERM JOINT SCENARIOS

Our time series data comprises monthly WPG, from January 1981 to December 2011, measured at three wind plants located in the Brazilian Northeast, namely: Rio do Fogo (RF), Icaraizinho (IC) and Enacel (EN). The last year of these time series were kept for out of sample scenario evaluation. The intrinsic non-Gaussian nature of WPG time series can be easily shown by observing the shape of a qq-plot or by running a simple normality test, such as the Jarque-Bera test. Considering individually each of the three time series of WPG,

the null hypothesis of normality for each of these time series has been rejected ( $p$ -values  $< 0.001$ ). Likewise, considering QQ-plot depicted in Figure 1 of IC unit, it strongly suggest non normality through the observed positive skewness, which suggest strong adequacy of our proposed beta GAS model.

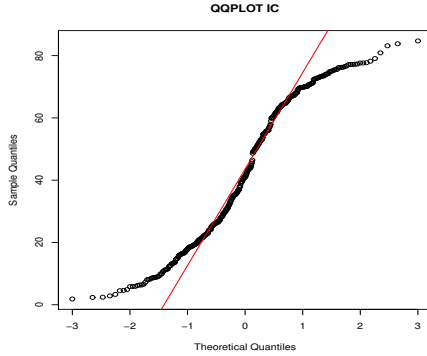


Fig. 1. QQ-plot of IC monthly capacity factor time series ranging from January 1981 to December 2011.

In this study, the same lag structure was used for the three WPG time series, as presented in Equation (2). After removing the first and last years of data for initial values estimation and scenario evaluation respectively, maximization of the log-likelihood function resulted in the estimated parameters presented in Table I. Dummy variables  $D_t$  were introduced until quantile residuals show no more auto correlation structure. The way to choose in which time index  $t$  a dummy variable  $D_t$  should be introduced is by checking, after model fit, where  $|r_{t,\hat{\theta}_t}| > 3$ , since under the Gaussianity of the residuals, 3 times the standard deviation reflects an abnormal value.

TABLE I  
MAXIMUM LIKELIHOOD ESTIMATION OF THE BETA GAS MODEL APPLIED TO THREE BRAZILIAN WIND PLANTS.

Parameter	Rio do Fogo (RF)		Icaraizinho (IC)		Enacel (EN)	
	Estim.	S.E.	Estim.	S.E.	Estim.	S.E.
$A_1$	0.551	0.006	0.539	0.071	0.540	0.060
$A_2$	0.231	0.004	-0.226	0.066	0.498	0.073
$A_3$	-0.147	0.007	0.352	0.070	0.041	0.053
$A_{11}$	0.008	0.007	0.219	0.038	0.169	0.034
$A_{12}$	0.110	0.004	-0.176	0.043	0.228	0.031
$B_1$	-0.080	0.002	1.058	0.055	-0.164	0.039
$B_2$	0.441	0.006	-0.818	0.043	-0.186	0.036
$B_3$	-0.369	0.011	0.156	0.031	-0.008	0.016
$B_{11}$	0.207	0.013	0.227	0.016	-0.004	0.005
$B_{12}$	0.325	0.011	0.121	0.037	0.721	0.035
$\omega$	3.139	0.019	2.571	0.260	5.485	0.780
$\alpha$	9.408	0.042	10.244	0.791	9.293	0.515
$D_{t=21}^{(Set,1983)}$	-	-	2.320	0.287	1.447	0.247
$D_{t=22}^{(Oct,1983)}$	-	-	2.021	0.309	3.610	0.481
$D_{t=29}^{(May,1984)}$	-	-	-0.031	0.641	-	-
$D_{t=129}^{(Sep,1992)}$	-	-	3.349	0.317	1.675	0.386
$D_{t=130}^{(Oct,1992)}$	-	-	-	-	1.189	0.345
$D_{t=165}^{(Sep,1995)}$	-	-	1.460	0.264	-	-
$D_{t=241}^{(Jan,2002)}$	-	-	-	-	-1.527	0.582
$D_{t=308}^{(Aug,2007)}$	1.771	0.346	-	-	-	-
$D_{t=309}^{(Sep,2007)}$	2.384	0.369	-	-	-	-
$D_{t=345}^{(Sep,2010)}$	3.004	0.277	-	-	-	-
$D_{t=346}^{(Oct,2010)}$	-	-	-1.462	0.350	-	-

Finally, to investigate correct model specification, a fully detailed analysis in the quantile residuals was undertaken and reported in Table II. Jarque-Bera test for normality (under the correct specification of the beta density, the residuals should be normally distributed), Ljung-Box test for absence of autocorrelation and Ljung-Box on the squared residuals to check for ARCH effects were considered (both were conducted using until lag 30). As reported in Table II, there were no rejections of the null hypothesis in none of the tests, hence it can be concluded that the proposed beta GAS models are adequate to describe the three time series of WPG.

TABLE II  
P-VALUES OF STANDARD DIAGNOSTIC TESTS.

Test	Rio do Fogo	Icaraizinho	Enacel
Normality	0.242	0.582	0.386
Autocorrelation	0.174	0.854	0.127
Arch effect	0.703	0.479	0.175

Although the forecasts of our beta GAS model are superior to the ones obtained using a simple SARIMA model, we advocate that the main contribution of the proposed framework here is not to improve the conditional mean forecast. Alternatively, the main goal of the proposed methodology is 1) to enhance the whole PDF description for individual time series, e.g., extreme quantiles are better approximated under the proposed framework, and 2) to enhance the multivariate description by capturing the spatial dependence among time series (generally located at different geographical areas) through a time-varying copula.

TABLE III  
FORECASTING EVALUATION BETWEEN BETA GAS(12,12) AND SARIMA MODEL.

Modelo	Medidas	RF	IC	EN
beta GAS(12,12)	RMSE	5.639	6.181	8.736
	MAE	5.209	5.021	7.346
	Pseudo R <sup>2</sup>	0.834	0.941	0.875
SARIMA	RMSE	7.566	10.668	9.737
	MAE	6.263	7.675	8.222
	Pseudo R <sup>2</sup>	0.809	0.831	0.859

In the following we will conduct the multivariate modeling of the three WPG time series using the second step GAS procedure previously described in 1-4 of Part II in III-B. The copula parameters were estimated and its values are displayed in Table IV. Regarding its degrees of freedom, it was estimated using profile likelihood process which resulted in  $\hat{\nu} = 340$ . Such value of degrees of freedom points out that the proposed elliptical copula reduces itself to a time-varying Gaussian copula, which makes perfect sense since we are dealing with monthly data.

From Table IV, it follows that all parameters are statistically significant at 5%, or less. The value of the parameter  $\Pi$  plays a key role here, since it reinforces the fact that dependence structure among these three WPG time series might not be static over the analyzed period. One should note that, if such values are zero or not significant, in the limit, the time varying correlation matrix will be constant, which means that our framework reduces itself to a constant copula model.

TABLE IV  
MAXIMUM LIKELIHOOD ESTIMATION OF THE SECOND STEP GAS MODEL  
APPLIED TO THE PIT VARIABLES OF THREE BRAZILIAN WIND PLANTS.

Parameter	t-GAS	
	Estim.	S.E.
$\Omega$	1.589	0.063
$\Pi$	0.354	0.017
$\Upsilon$	0.425	0.083

To introduce spatial dependence information into the joint scenarios of WPG, one should follow the steps provided in Section III-B to simulate scenarios of the  $t$ -GAS model. To evaluate the quality of the simulated values, we compared the empirical quantiles obtained from the  $t$ -GAS and SARIMA model 12 steps ahead against the historical monthly quantiles of the original WPG series (both in the original scale)<sup>3</sup>. The quantiles  $Q^{(\alpha\%)} \forall \alpha \in \{0.05, 0.10, 0.5, 0.9, 0.95\}$  of each set were calculated and plotted in Figures 2 and 3, where the lines are the monthly quantiles of the simulated set from one of the models, while the geometric figures represent the historic monthly quantiles of the real data set.

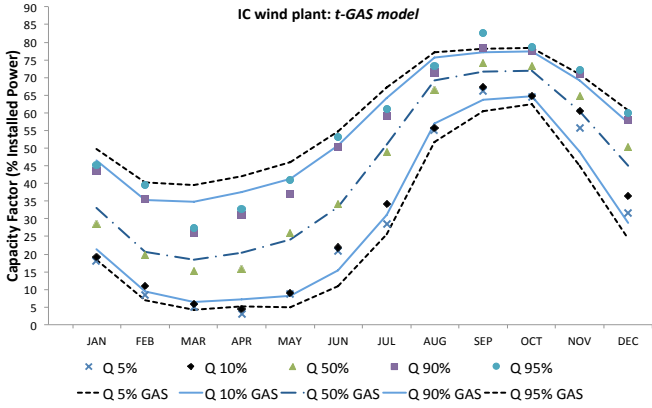


Fig. 2. Scenario evaluation of the year of 2011 through beta  $t$ -GAS model against the real data set.

Note that the quantiles from the simulated sets using our  $t$ -GAS model seems to be more reasonable and likely to the ones estimated from the real data set, while SARIMA ones are far beyond the historical quantiles. This provides evidences on the distortion brought by the anti log transformation, since we fitted a Gaussian SARIMA model to the log of the three WPG time series wind plants (EN, IC, RF).

To reinforce our findings, we calculate the reliability of our quantiles as described in [13]. Low reliability on the extreme quantiles could prejudice the decision-making process. Based on the definition of quantile, the empirical proportion of quantile  $q_t^\alpha$  is defined as,

$$\hat{P}_Q^\alpha = \frac{1}{T} \sum_{t=1}^T \eta_t \quad (15)$$

<sup>3</sup>The historical monthly quantiles of the original WPG was calculated by first segmenting the WPG time series of wind plant  $i \in K$  into 12 monthly time series, i.e.,  $\{Y_{Jan}^{(i)}, \dots, Y_{Dec}^{(i)}\}$ . The same was applied to the simulated sets from  $t$ -GAS and SARIMA models.

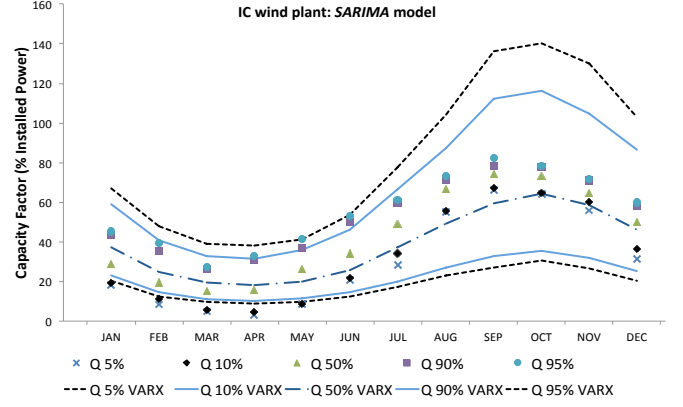


Fig. 3. Scenario evaluation of the year of 2011 through SARIMA model against the real data set.

where  $T$  is the total number of observations, and  $\eta_t$  denotes the indicator of quantiles  $\hat{P}_Q^\alpha$  defined by,

$$\eta_t = \begin{cases} 1, & y_t \leq q_t^\alpha \\ 0, & y_t > q_t^\alpha \end{cases} \quad (16)$$

The reability will provide the deviation between the empirical proportion  $\hat{P}_Q^\alpha$  and nominal proportion  $\alpha$  of the predicted quantiles. The average proportion deviation (APD) is then calculated by  $\hat{P}D_Q^\alpha = \hat{P}_Q^\alpha - \alpha$  and provided in Table V. The closest to zero, the higher reability of the estimated quantiles  $q_t^\alpha$ .

TABLE V  
AVERAGE PROPORTION DEVIATION (APD)

$\alpha$	GAS			SARIMA		
	RF	EN	IC	RF	EN	IC
5%	3.33%	3.33%	3.33%	11.67%	11.67%	11.67%
10%	-1.67%	6.67%	6.67%	15.00%	15.00%	6.67%
50%	3.33%	8.33%	8.33%	16.67%	16.67%	25.00%
90%	1.67%	1.67%	-6.67%	10.00%	10.00%	10.00%
95%	-3.33%	-3.33%	-3.33%	5.00%	5.00%	5.00%

By the results of Table V it is possible to argue that the scenarios generated by our proposed framework, specially on extreme quantiles, are very well estimated which is an expected outcome due to the fact that the copula model acts on the tail behaviour of the predictive density. In addition, we also implemented the conditional coverage Value at Risk exceedances test of Christoffersen [36] to check the coverage level of each predictive density against real data in Table VI. Non significant p-values reinforce the idea that the proposed model is adequate to fit the WPG time series.

TABLE VI  
P-VALUES CONDITIONAL COVERAGE VALUE AT RISK EXCEEDANCES TEST OF CRISTOFFERSEN CONSIDERING A COVERAGE LEVEL OF 95%.

Coverage level	RF	IC	EN
$t$ -GAS model	0.305	0.305	0.305
SARIMA	0.035	0.035	0.035

As expected, p-values of Cristoffersen test applied to SARIMA predictive densities are all significant, considering

a coverage level of 95%. In contrast, predictive densities produced by  $t$ -GAS model are all non-significant, using the same level of coverage.

## V. CONCLUSION

In this paper we proposed a framework to simulate joint scenarios of WPG time series taking into account the spatial dependence among different units. Such scenarios are deeply demanded for the development of medium- and long-term studies for power-system planning under uncertainty of renewable energy generation. Our methodology seems fruitful and have several benefits when comparing against a benchmark SARIMA model, regarding its parametric features and the straightforward way to estimate and generate extreme quantiles. The proposed framework was derived making use of a recently proposed class of time-series models with time-varying parameters and arbitrary non Gaussian distribution, known as Generalized Auto Regressive Score models (GAS). When adopting such models, one does not need to transform WPG and work with the  $\log$  of the observed data point. Our framework is tailor made for the range of values that WPG time series can assume. In addition, copula models were used to input spatial dependence into the generated joint scenarios. Its use is unarguably better when comparing the extreme scenarios to the ones generated by the SARIMA model.

## ACKNOWLEDGMENT

Henrique Hoeltgebaum, Cristiano Fernandes and Alexandre Street acknowledge support from CNPq/Brazil.

## REFERENCES

- [1] International energy agency. [Online]. Available: <https://www.iea.org/newsroom/news/2013/october/wind-power-seen-generating-up-to-18-of-global-power-by-2050.html>
- [2] A. Moreira, D. Pozo, A. Street, and E. Sauma, "Reliable renewable generation and transmission expansion planning: Co-optimizing system's resources for meeting renewable targets," *IEEE Transactions on Power Systems*, vol. PP, no. 99, pp. 1–1, 2016.
- [3] R. Jabr, "Robust transmission network expansion planning with uncertain renewable generation and loads," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4558–4567, 2013.
- [4] C. Zhao and Y. Guan, "Data-driven stochastic unit commitment for integrating wind generation," *IEEE Transactions on Power Systems*, vol. 31, no. 4, pp. 2587–2596, July 2016.
- [5] A. C. Passos, A. Street, and L. A. Barroso, "A dynamic real option-based investment model for renewable energy portfolios," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 883–895, March 2017.
- [6] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renewable and Sustainable Energy Reviews*, vol. 32, pp. 255–270, 2014.
- [7] M. Souto, A. Moreira, A. Veiga, A. Street, J. D. Garcia, and C. Epprecht, "A high-dimensional varx model to simulate monthly renewable energy supply," in *Power Systems Computation Conference (PSCC), 2014*. IEEE, 2014, pp. 1–7.
- [8] A. Lau and P. McSharry, "Approaches for multi-step density forecasts with application to aggregated wind power," *The Annals of Applied Statistics*, pp. 1311–1341, 2010.
- [9] J. W. Taylor, P. E. McSharry, and R. Buizza, "Wind power density forecasting using ensemble predictions and time series models," *Energy Conversion, IEEE Transactions on*, vol. 24, no. 3, pp. 775–782, 2009.
- [10] R. J. Bessa, V. Miranda, A. Botterud, J. Wang, and M. Constantinescu, "Time adaptive conditional kernel density estimation for wind power forecasting," *Sustainable Energy, IEEE Transactions on*, vol. 3, no. 4, pp. 660–669, 2012.
- [11] J. Jeon and J. W. Taylor, "Using conditional kernel density estimation for wind power density forecasting," *Journal of the American Statistical Association*, vol. 107, no. 497, pp. 66–79, 2012.
- [12] J. W. Taylor and J. Jeon, "Forecasting wind power quantiles using conditional kernel estimation," *Renewable Energy*, vol. 80, pp. 370–379, 2015.
- [13] C. Wan, J. Lin, J. Wang, Y. Song, and Z. Y. Dong, "Direct quantile regression for nonparametric probabilistic forecasting of wind power generation," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2767–2778, July 2017.
- [14] A. Street, D. A. Lima, A. Veiga, B. Fanzeres, L. Freire, and B. Amaral, "Fostering Wind Power Penetration into the Brazilian Forward-Contract Market," *IEEE PES General Meeting 2011*, pp. 1–8, Jul. 2011.
- [15] A. Shapiro, W. Tekaya, J. P. da Costa, and M. P. Soares, "Risk neutral and risk averse stochastic dual dynamic programming method," *European Journal of Operational Research*, vol. 224, no. 2, pp. 375 – 391, 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0377221712006455>
- [16] B. Fanzeres, A. Street, and L. A. Barroso, "Contracting strategies for renewable generators: A hybrid stochastic and robust optimization approach," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1825–1837, July 2015.
- [17] G. Kariniotakis, *Renewable Energy Forecasting: From Models to Applications*, ser. Woodhead Publishing Series in Energy. Elsevier Science, 2017.
- [18] D. Creal, S. J. Koopman, and A. Lucas, "Generalized autoregressive score models with applications," *Journal of Applied Econometrics*, vol. 28, no. 5, pp. 777–795, 2013.
- [19] A. C. Harvey, *Dynamic models for volatility and heavy tails: with applications to financial and economic time series*. Cambridge University Press, 2013, vol. 52.
- [20] R. F. Engle and T. Bollerslev, "Modelling the persistence of conditional variances," *Econometric reviews*, vol. 5, no. 1, pp. 1–50, 1986.
- [21] R. F. Engle and J. R. Russell, "Autoregressive conditional duration: a new model for irregularly spaced transaction data," *Econometrica*, pp. 1127–1162, 1998.
- [22] R. A. Davis, W. T. Dunsmuir, and S. B. Streett, "Observation-driven models for poisson counts," *Biometrika*, pp. 777–790, 2003.
- [23] Gas on-line repository. [Online]. Available: <http://www.gasmodel.com/>
- [24] F.-M. Tseng and G.-H. Tzeng, "A fuzzy seasonal arima model for forecasting," *Fuzzy Sets and Systems*, vol. 126, no. 3, pp. 367–376, 2002.
- [25] M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, and Z. Yan, "A review on the forecasting of wind speed and generated power," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 4, pp. 915–920, 2009.
- [26] D. Creal, S. J. Koopman, and A. Lucas, "A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations," *Journal of Business & Economic Statistics*, vol. 29, no. 4, pp. 552–563, 2011.
- [27] A. J. Patton, "Modelling asymmetric exchange rate dependence\*," *International economic review*, vol. 47, no. 2, pp. 527–556, 2006.
- [28] —, "Copula-based models for financial time series," in *Handbook of financial time series*. Springer, 2009, pp. 767–785.
- [29] —, "Copula methods for forecasting multivariate time series," *Handbook of economic forecasting*, vol. 2, pp. 899–960, 2012.
- [30] P. K. Dunn and G. K. Smyth, "Randomized quantile residuals," *Journal of Computational and Graphical Statistics*, vol. 5, no. 3, pp. 236–244, 1996.
- [31] R. Engle, "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models," *Journal of Business & Economic Statistics*, vol. 20, no. 3, pp. 339–350, 2002.
- [32] J. J. Xu, "Statistical modelling and inference for multivariate and longitudinal discrete response data," Ph.D. dissertation, University of British Columbia, 1996.
- [33] H. Joe, *Multivariate models and multivariate dependence concepts*. CRC Press, 1997.
- [34] S. A. Murphy and A. W. Van der Vaart, "On profile likelihood," *Journal of the American Statistical Association*, vol. 95, no. 450, pp. 449–465, 2000.
- [35] U. Cherubini, E. Luciano, and W. Vecchiato, *Copula methods in finance*. John Wiley & Sons, 2004.
- [36] P. F. Christoffersen, "Evaluating interval forecasts," *International economic review*, pp. 841–862, 1998.



APPENDIX A  
GAS MECHANISM FOR THE MULTIVARIATE STUDENT T  
DISTRIBUTION

In this appendix we present the elements of the GAS mechanism  $s_t$ ,  $\nabla_t$  and  $\mathcal{I}_{t|t-1}^{-1}$  for the multivariate Student t distribution, but first let us define a few matrix notations. The matrix  $\mathcal{D}_K$  is a duplication matrix given by  $\mathcal{D}_K \text{vech}(C) = \text{vec}(C)$  and the Kronecker product of two matrixes  $C$  and  $D$  is denoted by  $C \otimes D$ . In case  $C = D$ ,  $C_{\otimes}$ . By convenience the operator  $\text{vec}(C)$  creates a column vector from a matrix  $C$  by stacking the column vectors of  $C$  while  $\text{vech}(C)$  also creates a column vector stacking only column vectors of the inferior triangular part of  $C$ . For a full detailed review on linear algebra operators, please see [?].

Now, the proposed GAS mechanism in[26] to update the time varying correlation matrix  $\Sigma_t$  using a multivariate Student t is given by

$$\nabla_t = \frac{\partial \ln p(\tilde{y}_t | \Sigma_t; \nu)}{\partial f_t} = \frac{1}{2} \Psi_t' \mathcal{D}_K' \Sigma_{t\otimes}^{-1} [w_t \tilde{y}_{t\otimes} - \text{vec}(\Sigma_t)] \quad (17)$$

$$\mathcal{I}_{t|t-1} = E[\nabla_t \nabla_t'] = \frac{1}{4} \Psi_t' \mathcal{D}_K' J_{t\otimes}' [gZ - \text{vec}(I) \text{vec}(I)'] \quad (18)$$

$$J_{t\otimes} \mathcal{D}_K \Psi_t$$

where  $\Psi_t = \partial \text{vech}(\Sigma_t) / \partial f_t'$  which will be presented next in Equation (20), the scalar  $w_t = (\nu + K) / (\nu - 2 + \tilde{y}_t' \Sigma_t^{-1} \tilde{y}_t)$  reduces the effects of outliers, once the former quantity is divided by  $\tilde{y}_t' \Sigma_t^{-1} \tilde{y}_t$ . Matrix  $J_t$  is defined as  $\Sigma_t^{-1} = J_t' J_t$ , which can be obtained by any matrix decomposition procedure. In our application in Section IV, we adopted Cholesky decomposition. The scalar  $g = (\nu + K) / (\nu + 2 + K)$  and the elements of the square matrix  $Z$  of dimension  $K \times K$ , are given by

$$Z[(i-1) \cdot k + \ell, (j-1) \cdot k + m] = \delta_{ij} \delta_{\ell m} + \delta_{i\ell} \delta_{jm} + \delta_{im} \delta_{j\ell} \quad (19)$$

with  $i, j, \ell, m = 1, \dots, K$ , assuming  $\delta_{ij}$  equals to 1 if  $i = j$  and zero otherwise.

As previously discussed, the key to define a copula framework in [26] is to use the following specification of  $\Psi_t$  in Equations (17)-(18),

$$\Psi_t = \mathcal{B}_K \mathcal{D}_{t\otimes} \Delta_{t\otimes}^{-1} [\mathcal{D}_K - (\Delta_t \oplus Q_t) \Delta_{t\otimes}^{-1} W_{\Delta_t} \mathcal{S}_{\Delta}] \mathcal{S}_Q \quad (20)$$

where  $\mathcal{S}_Q = I$  since  $f_t = \text{vech}(Q_t)$ . The matrix  $W_{\Delta_t}$  is initially calculated using a diagonal matrix of dimension  $K^2 \times K^2$  with elements of the main diagonal equals to  $0.5 \text{vec}(\Delta_t^{-1})$ , and later the columns of zeros are eliminated. The matrix  $\mathcal{S}_{\Delta}$  is a selection matrix, with ones and zeros, given by  $\text{diag}(\Delta_t^2) = \mathcal{S}_{\Delta} \text{vech}(Q_t)$ . The operator  $\oplus$  is given by  $C \oplus D = (C \otimes D) + (D \otimes C)$ . In addition,  $\mathcal{B}_K = (\mathcal{D}_K' \mathcal{D}_K)^{-1} \mathcal{D}_K'$  is a duplication matrix composed by only zeros and ones, with more lines than columns.  $\Delta_t$  is a diagonal matrix which elements are the square root of the elements contained in the main diagonal of  $Q_t$