G. Díaz and E. Planas, "A Note on the Normalization of Spanish Electricity Spot Prices," in I**EEE Transactions on Power Systems**, vol. 31, no. 3, pp. 2499-2500, May 2016, https://doi.org/10.1109/TPWRS.2015.2449757

A Note on the Normalization of Spanish Electricity Spot Prices

Guzmán Díaz, *Senior Member, IEEE,* and Estefanía Planas

***Abstract*—Merit order of renewable generation has produced an increasing, relevant number of zero values in Spanish electricity spot prices since 2010, in contrast with the rest of markets which rarely present zero values. This letter shows that existing methodologies for normalizing the price distribution fail to yield a proper normalization. This letter proposes a new algorithm based on Nataf transformation which addresses this issue and ensures a correct normalization. Numerical results validate the proposed algorithm.**

***Index Terms*—Normal distributions, Probability, Autoregressive processes, Time series**

I. Introduction

THE electricity spot price has been modeled following different statistical approaches, mainly based on mean reverting diffusion and auto-regressive models [1]. The specifications of these models require, or at least benefit from, the assumption of normality of the underlying stochastic process. This was emphasized by Conejo *et al.* in [2] in the context of uncertain market analysis. Conejo and colleagues argued that the normality assumption is of major importance in building Auto-Regressive Integrated Moving Average (ARIMA) models for scenario generation in stochastic programming. (Indeed, ARIMA time series analysis in the formulation of Box–Jenkins does not require strict but weak stationarity, and, thus, the normality assumption may be dropped for basic analysis maybe without relevant loss of accuracy. But other analyses often require the normality assumption, such as the use of AIC for order selection or the asymptotic prediction of confidence bounds [3, §5.4].) Typically in the case of energy spot markets a variance stabilizing transformation such as a logarithmic transformation—or more broadly a Box–Cox transformation—offers some normalization to the data; at the cost of dropping the non-positive prices [4]. These non-positive prices can be typically ignored since they are usually very limited in number. However, we have detected that Spanish spot market often presents these non-positive prices (particularly zero prices) specially since 2010, mainly due to the high penetration of wind energy with high merit order [5]. Furthermore, the frequency of these zero prices has been continually rising since. Thus, these zero prices can not be disregarded. This letter proposes a new modified transformation which addresses this drawback and obtains a suitable normalization of spot prices with an important quantity of non-positive prices. Based on Nataf transformation, it does not increase the complexity of the normalization process.

G. Díaz is with the Univ. of Oviedo, Campus de Viesques, 33204, Spain (phone: 34 985182592; fax: 34 985182068; e-mail: guzman@uniovi.es).

E. Planas is with the UPV/EHU, Alaameda de Urquijo, 48013, Spain (e-mail: estefania.planas@ehu.eus).

Manuscript received May 26, 2015. Work inside UFI11/16 (UPV/EHU); supported by the BUS IT394-10 and by the Gov. of the Basque Country (ETORTEK) as FUTUREGRIDS-2020 (IE14-389).

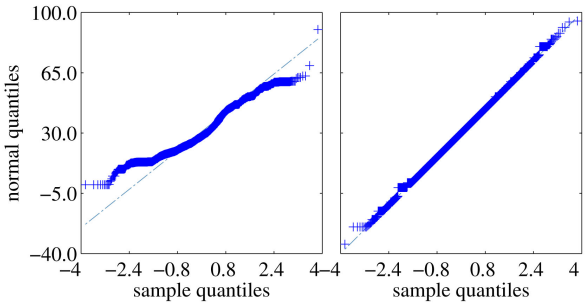

Fig. 1: Year 2000 Normal probability plots (QQ-plots). (a) Untransformed prices. (b) Basic Nataf transformation.

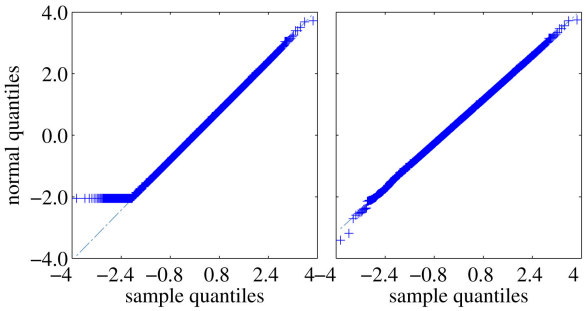

Fig. 2: Year 2014 Normal probability plots. (a) Basic Nataf transformation. (b) Modified Nataf transformation.

II. Basic Nataf price transformation

In the context of scenario generation for market-clearing procedures in systems with a high penetration of wind power, reference [6] provided a detailed discussion on how non-Gaussian processes could be translated into Gaussian before specifying an ARIMA model without resorting to dropping nonpositive data. The proposal was to use a double transformation based on the non-Gaussian cumulative distribution function (CDF) and the standard normal CDF. Reference [6] provides a graphical explanation of the procedure; the first step of Nataf transformation (NT) before eliminating the correlation, and that was popularized through [7]. It consists in defining a new stochastic process, $\mathbf{Z}$, with a standard Normal marginal distribution given by the transformation of the spot price $\mathbf{S}$ as $\mathbf{Z} = \Phi^{-1}[F_{\mathbf{S}}(\mathbf{S})]$, where $F_{\mathbf{S}}$ is the spot price CDF and Φ is the standard Normal CDF. (In what follows we employed a kernel estimation of the empirical CDF; see Fig. 3b.)

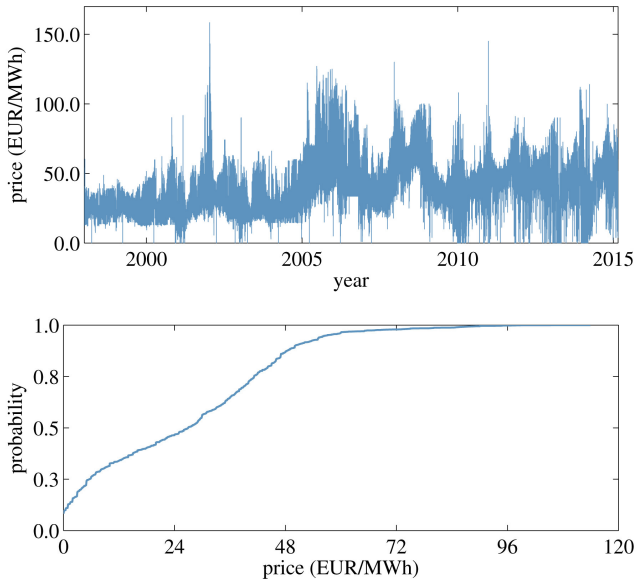

Fig. 3: (a) Spanish spot price from January 1, 1998 through February 25, 2015. (b) $F_{\mathbf{S}}(\mathbf{S})$ of 2014-Q1.

In [8] Contreras *et al.* focused their analysis on the Spanish and Californian markets, proposing high degree ARMA models rather than Hull-White processes. They published their paper in 2003 using data of 2000. In order to normalize the data and take advantage of the above commented properties of normalization before specifying an ARIMA model, we employed also the data of year 2000. The unprocessed data is remarkably non-Gaussian (Fig. 1), but the transformed prices using the basic NT in [6] reasonable agree to a Gaussian distribution except for some outliers in the lowest quantiles (Fig. 1). However, when we applied the same reasoning to the prices of 2014, the results were remarkably different, with the lowest quantiles clearly departing from the normality (Fig. 1). That is, the basic NT in [6] fails to produce a Gaussian process (because of the high quantity of zero values).

III. Modified Nataf transformation

Fig. 3a shows the evolution of the prices over the last 17 years. When Contreras *et al.* published their paper, the market followed an evolution similar to those that follow other European markets nowadays. However, it can be appreciated a change in the market evolution five years ago. Driven by its higher merit order, wind generation in Spain was favored at the time of dispatching [5]. Because of the high wind power penetration in Spain and the inelastic behavior of demand, this has entailed periods of zero prices in the intra-day market. And what is more remarkably is that this trend has been exacerbated over the last five years. We argue that those increased number of zeroes account for the lack of capability of NT to produce a Gaussian distribution of prices. We alternatively propose a modification of the basic transformation, based on [9], by redefining the spot price CDF as follows:

$$F^*_{\mathbf{S}}(\mathbf{S}) = \begin{cases} \mathrm{U}(0, \mathrm{pr}\{\mathbf{S}=0\}), & \text{if } S_t = 0 \\ F_{\mathbf{S}}(\mathbf{S}), & \text{otherwise} \end{cases} \quad (1)$$

TABLE I: Shapiro–Wilk hypothesis test, year 2014.

Price series	W	p-value	h
Untransformed series	0.9569	0	1
Basic NT	0.9953	2.2×10^{-16}	1
Modified NT	0.9998	0.4665	0

The first line of (1) establishes that the transformation must not employ $F_{\mathbf{S}}$ when the prices are zero, but alternatively it must be the result of a uniform CDF sample draw. Note that the uniform CDF must be upper bounded by the frequency with which the zeroes are present in the sample. That is, for years previous to 2010, the frequency is as low as to neglect the contribution of the uniform CDF. But in 2013 the frequency of occurrence was as high as a 5.5% and therefore the CDF of $\mathrm{U}(0, 0.055)$ must be used if that year is analyzed.

The result of applying (1) to the year 2014 is plotted in Fig. 2. The normalization of the lower quantiles is evident.

Quantitative assessment is given in Table I, where the results of the Shapiro–Wilk test of year 2014 prices are listed. The test checks the null hypothesis that the distribution of the stochastic process is Normal with unspecified mean and variance. As shown in the table, the basic NT fails to pass the test. Only the transformed series by means of the modification proposed in this paper is considered to be Gaussian to a significance level $\alpha = 5\%$.

IV. Conclusions

This paper deals with the normalization of Spanish electricity market spot prices. This market has shown the particularity of a relevant quantity of zero values since 2010, which makes the traditional normalization process failing. A new transformation is proposed which normalizes the data suitably and without increasing the complexity of the normalization process. The validity of the proposed transformation is confirmed by means of numerical results.

References

[1] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *International Journal of Forecasting*, vol. 30, no. 4, pp. 1030–1081, Oct. 2014.

[2] A. J. Conejo, M. Carrion, and J. M. Morales, *Decision making under uncertainty in electricity markets*. New York: Springer, 2010.

[3] P. J. Brockwell and R. A. Davis, *Introduction to time series and forecasting*, 2nd ed. New York: Springer, 2002.

[4] C. R. Knittel and M. R. Roberts, "An empirical examination of restructured electricity prices," *Energy Economics*, vol. 27, no. 5, pp. 791–817, Sep. 2005.

[5] J. M. Morales, M. Zugno, S. Pineda, and P. Pinson, "Redefining the Merit Order of Stochastic Generation in Forward Markets," *IEEE Transactions on Power Systems*, vol. 29, no. 2, pp. 992–993, Mar. 2014.

[6] J. Morales, R. Mínguez, and A. Conejo, "A methodology to generate statistically dependent wind speed scenarios," *Applied Energy*, vol. 87, no. 3, pp. 843–855, Mar. 2010.

[7] P.-L. Liu and A. Der Kiureghian, "Multivariate distribution models with prescribed marginals and covariances," *Probabilistic Engineering Mechanics*, vol. 1, no. 2, pp. 105–112, Jun. 1986.

[8] J. Contreras, R. Espinola, F. Nogales, and A. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE Transactions on Power Systems*, vol. 18, no. 3, pp. 1014–1020, Aug. 2003.

[9] L. Rüschendorf, "On the distributional transform, Sklar's theorem, and the empirical copula process," *Journal of Statistical Planning and Inference*, 2009.