Applying Copulas Functions for Wind and Hydro Complementarity Evaluation: a Brazilian Case

Submitted to the12th International Conference on European Energy Market (EEM), 2015

Peter Molnár
Department of Industrial Economics and Technology Management
Norwegian University of Science and Technology
Trondheim, Norway
Peto.Molnar@gmail.com

Luiz A. Steinle Camargo
Dorel Soares Ramos
Department of Power Engineering and Electrical Automation
University of São Paulo (USP).
São Paulo, Brazil
luiz.steinle@usp.br; dorelram@usp.br

***Abstract*— Wind energy has been expanding in many countries typically hydro electricity producers in the last decades and some investments take advantage from the complementarity among hydro and wind productions as a strategy to increase profits and minimize risks. In this paper, our focus is applying copulas functions to evaluate a complementarity between wind and hydro production in different regions in Brazil. The goal is evaluate which regions are most suitable for development of wind energy considering the existing hydro electricity production. Copulas functions can capture the dependence structure among random variables offering a great flexibility in building multivariate stochastic models while statistic correlation does not capture nonlinear effects sufficiently. For this reason, we apply nine different copulas functions in our study finding those that best capture the complementarity.**

***Index Terms*- Complementarity, Copulas Functions, Evaluation, Hydro, Wind.**

I. INTRODUCTION

Hydropower is the oldest and most utilized green energy in the world. In the recent years, wind power has emerged in many countries as alternative source of green energy. Some countries, e.g. Brazil, Norway and Canada are big producers of hydroelectricity and have large potential for wind power [1]. The Brazilian wind power has been rapidly expanded mostly due to local government incentives, and its installed capacity is estimated to grow from 2.2 GW in 2013 up to 22.4 GW in 2023, over half in the Northeastern region [2].

Wind and hydro powers are well known as renewable sources with seasonal generation patterns, dependent of weather conditions and hence with uncertain production. For generation companies, production uncertainty is one of the main sources of risk. Companies are therefore creating strategies to hedge their positions by investing in other renewable sources [3], [4]. From the perspective of the power system operator, the diversification of power technologies is beneficial, because system dependent on more energy sources with different production profiles is safer.

However, benefits of diversification depend on complementarity among the different power sources generation. [5] explains how complementarity of hydro and wind can improve the risk profile of energy inflows for the province of Québec (Canada), since this area is producing electricity mostly from hydropower and adding wind results in a reduction of market risk. In its turn, [6] describes a mechanism to attract wind generators to the wholesale market of U.K. by mitigating individual exposure to intermittent outputs based on an adaptation of the Brazilian experience with hydro generation. Moreover, [7] discuss the increase of different sources in the Brazilian energy market, highlighting the role, that wind energy can play based on its complementarity with hydro source.

Brazil is a huge country, which regions have wide variety of weather patterns and abundance of water resources. Over 90% of electricity comes from hydro sources. However, most of the hydropower potential is already developed or in area of environmental protection. The Brazilian Power System has circa 124 GW of total installed capacity (68.9% of hydro powers, 17.1% nonrenewable - thermal and nuclear - and 14 % alternatives - biomass, solar and wind) and the federal government aims to increase it in 60% during this decade. In this plan, wind power capacity should increases 10 times [2].

Hydro and wind power plants have different generation profiles. In Brazil, wind plants located in the northeast region (best wind potential conditions) monthly average production tends to increase from June to November whereas hydro plants production increases during the wet season (December to April). During the high period of wind plants productions, spot price is in general also high and the opposite occurs

We recognize the Norwegian research centre CenSES, Centre for Sustainable Energy Studies (RCN grant 209697), and acknowledge financial support from the Research Council of Norway through project 228811.
We also acknowledge the financial support from *Duke Energy International Geração Paranapanema S.A.* through the R&D project PD-0387-0512/2012 (ANEEL Code).

during the wet period (as the Brazilian generation mix has huge hydro participation, hydro production is negatively correlated to spot price) [4].

Since both hydro and wind production are seasonal and partly unpredictable, it is important to understand whether and to which extend are these electricity sources complementary.

The aim of this paper is to study the complementarity of hydro wind power among Brazilian submarkets. We use Brazilian historical data for wind and hydro sources. In order to measure wind and hydro in same units, we work with energy production (affluent natural energy).

One possible way how to evaluate complementarity among wind and hydro in different regions in Brazil would be to focus on statistic correlation among these resources. However, correlation does not capture nonlinear effects sufficiently. Based on [8], Copulas are mathematical objects that capture the dependence structure among random variables and hence, offer a great flexibility in building multivariate stochastic models. We therefore use copulas functions to allow for a more flexible relationship between wind and hydro.

This paper evaluates which regions in Brazil are most suitable for development of wind energy in terms of being complementary to hydro power production which already exist both in those regions and in Brazil as a whole.

II. Data

The Brazilian Power system (BPS) is a hydrothermal system with centralized scheduling operated by The National Operator System (NOS). In such system, also named 'tight pool' model, operational planning has per objective to minimize the expected total cost, as a variable decisions the hydro and thermoelectricity generation dispatches. The operation is computed by a multi-stage stochastic optimization model (named Newave) that takes into account a detailed representation of hydro plant operation and inflow uncertainties in order to determine the thermal and hydro optimal dispatch [9], [10]. We used information from Newave in our study, as it is the same model used by the operator. In order to adequate the data, we worked with energy production (affluent natural energy) instead of wind speed and water inflow, by converting them into the same base and assuming that this approximation is necessary to fit in our propose: to study the correlation among wind and hydro production. In this section, we explained how it was done for both cases.

BPS is divided into two blocks: The National Interconnect System (NIS), which covers almost all the Brazilian territory, and the Isolates Systems, mainly in the North region. In turn, the NIS is divided in four main geoelectric regions (submarkets): North (N), Northeast (NE), Midwest-Southeast (MW-SE) and South (S), as shown in Fig. 1.

A. *Affluent Natural Energy of Hydro inflows*

In the BPS, Newave simulates the system conditions by considering a detailed representation of the whole hydro system. One output provided by this model is the Affluent Natural Energy (ANE), which is the electricity energy that can be generated from natural inflow of hydroelectric exploitation under specific conditional of 65% reservoir level. This can be found by integrating a watershed or a subsystem.

In our study, we used the ANE historical series (from 1948 to 2010) of all subsystems: North, South, Northeast and Southeast/Midwest, given by the NOS [11]. The ANE approach is relevant since in the BPS hydro power plants are dispatch by the NOS. It means that the production of each plant depends on the whole system generation and the energy allocated to each one of them does not represent their real physical generation, but their correspondent production generation capacity.

Fig. 1 Energy Submarkets in Brazil

B. *Affluent Natural Energy of Wind inflows*

Since the Newave proposes to simulate hydrothermal systems, it is not applicable for the case of wind energy production. In this sense, we transformed wind inflows into wind production (hereafter called Affluent Natural Wind Energy: ANE-W) by collecting some points of Brazilian regions, where we have a long-term historical wind data from 1948 to 2010, and crossing it with one commercial wind turbine power curve. The wind data used to develop this study was obtained by Vestas Meso-scale Model [12] and NOAA (National Oceanic and Atmospheric Administration) [13]. The energy production representation was performed considering the Vestas V112 wind turbine (3.0 MW of Potential, 110m hub height, class II-IEC).

With the purpose of better represent the Brazilian wind energy potential, 13 points in different States were chosen: two in Ceará; two in Rio Grande do Norte; five in Bahia; one in Sergipe and three in the Rio Grande do Sul. These States concentrate investments in wind energy in the country. Mainly investments in wind energy are in Northeast submarket following by South submarket. In both cases, there are wind power plants along the coast and in the countryside. Others states in Brazil have no significant investments in wind energy for a while, so that we confine ourselves in the points shown in Table I (States locations per submarket).

III. Method and Results

In order to study complementarity between hydro and wind energy production, it is important to distinguish between expected/seasonal/predictable part of production and unexpected/unpredictable part of production. This distinction is important not only from statistical perspective, but from practical perspective also. System operator can more easily deal with predictable part of production, which is given mostly by yearly seasonality.

Table I

Index	Coordinates (°)		State	Submarket
	Lat	Long		
NE 1	-3.00	-40.00	Ceará	Northeast
NE 2	-5.00	-40.00	Ceará	Northeast
NE 3	-5.00	-37.50	Rio G. do Norte	Northeast
NE 4	-5.00	-35.00	Rio G. do Norte	Northeast
NE 5	-10.00	-42.50	Bahia	Northeast
NE 6	-10.00	-40.00	Bahia	Northeast
NE 7	-10.00	-37.50	Sergipe	Northeast
NE 8	-12.50	-42.50	Bahia	Northeast
NE 9	-12.50	-40.00	Bahia	Northeast
NE 10	-15.00	-42.50	Bahia	Northeast
S 1	-30.00	-55.00	Rio G. do Sul	South
S 2	-30.00	-50.00	Rio G. do Sul	South
S 3	-32.50	-52.50	Rio G. do Sul	South

Fig. 2 shows yearly submarket production pattern for the hydro and wind considered in our study. Value of each month was calculated as average over all the months of our sample. As we can see, there is strong seasonality in the production data. However, the seasonality depends on the region, e.g. one place might have rather stable wind power production (South), whereas in other might be highly seasonal (Northeast).

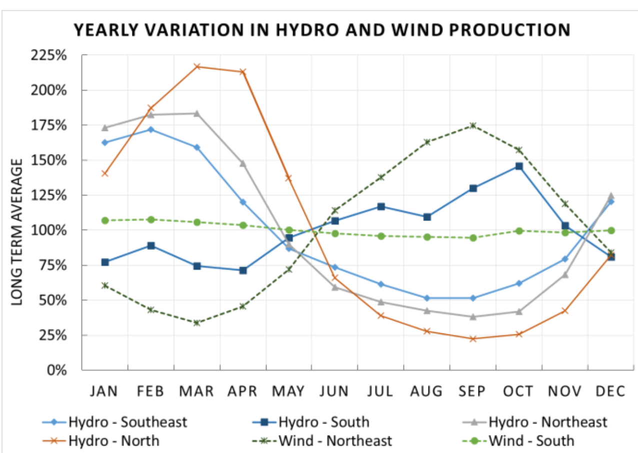

Fig. 2 Yearly variation in hydro and wind production

A. Complementarity in the seasonal part of production

Complementarity between seasonal pattern in the hydro and wind production in various regions is rather easy to see and analyze. If we want to see which region's wind production complements Brazilian hydro production, we can e.g. simply calculate correlation coefficient among the total Brazilian hydro production and regions' possible wind production.

As we can see from Table II, wind production in the entire Northeast region is highly complementary to the hydro production in the Northeast, North and Southeast. Wind production in the South region is not complementary to the Brazilian hydro production as a whole, but it is still complementary to the hydro production in South.

Table II

Wind	Hydro				
	Northeast	North	Southeast	South	All Hydro
NE 1	-0.57	-0.78	-0.55	0.33	-0.58
NE 2	-0.61	-0.74	-0.64	0.28	-0.65
NE 3	-0.55	-0.76	-0.56	0.28	-0.59
NE 4	-0.51	-0.74	-0.52	0.27	-0.55
NE 5	-0.63	-0.61	-0.69	0.23	-0.67
NE 6	-0.66	-0.69	-0.69	0.26	-0.69
NE 7	-0.52	-0.67	-0.49	0.19	-0.54
NE 8	-0.61	-0.60	-0.65	0.26	-0.63
NE 9	-0.53	-0.60	-0.48	0.28	-0.49
NE 10	-0.50	-0.51	-0.52	0.29	-0.50
S 1	-0.03	0.00	-0.01	-0.01	-0.02
S 2	0.38	0.36	0.34	-0.23	0.33
S 3	0.23	0.24	0.16	-0.19	0.16

Locations NE 5, NE 6, NE 8 and NE 2 presented the strongest correlation with all hydro submarket. These regions are in Bahia state (NE 5, 6 and 8) and Ceará state (NE2), where several wind farms are being built. NE 1 and NE2, both in Ceará state, and NE 3 and NE4, both in R. G. do Norte state have highest correlation with North hydro submarket.

Results shown the correlation among different wind locations and hydro submarkets. These can aid investors in defining the location their investments should be done. For instance, the Bahia state (NE 6) presented the strongest correlation with the overall hydro submarket. That is an unknown place in Brazil where several wind farms have been built. In turns, wind in the South has weak correlation with the hydro system as it has a smooth seasonal profile. However, this profile can be commercially strategy for a hydro-wind complementarity approach when considering the generation portfolio profile against spot prices movements [4].

B. Complementarity in the unpredictable part of production

However, main focus of this article is complementarity in the unpredictable part of the hydro and wind production. In order to study this, we need to first remove the predictable part of production. We model the predictable part of production for all the data series in the following way:

$$P_t = \alpha_0 + \Sigma\alpha_i M_i + \alpha_{12}P_{t-1} + \alpha_{13}P_{t-2} + \alpha_{14}P_{yearly,t-1} \quad (1)$$

Where regression coefficients are denoted as α's, P_t is production in month t, M_i are monthly dummies, P_{t-1} is production in previous months and $P_{yearly,t-1}$ is average production in the previous year. Monthly dummies are here to capture yearly seasonality. Production during previous month and year is here to capture autoregressive features of the time series. Residuals from these regressions are the unpredictable part of production.

Next, we study the complementarity between the unpredictable parts of the hydro-wind productions. One possibility is to again check the correlations. Table III reveals that there is generally very little correlation in the unpredictable part of the hydro and wind production. However, there might be some nonlinear dependence and therefore we use copula approach. Copula is a function that "joins" marginal distribution functions into the joint distribution function. Formally, this is stated by Sklar's theorem:

"*Let H be a joint distribution function with margins F and G. Then there exists a copula C such that for all x, y in R:*

$$H(x,y) = C(F(x), G(y)) \tag{2}$$

Conversely, if C is copula and F and G are distribution functions, then the function H is a joint distribution function with margins F and G."

Table III

Wind	Hydro				
	Northeast	**North**	**Southeast**	**South**	**All Hydro**
NE 1	-0.02	0.12	0.02	-0.09	0.04
NE 2	-0.03	0.03	0.02	-0.02	0.00
NE 3	-0.06	0.03	0.02	-0.06	-0.03
NE 4	-0.09	0.07	-0.01	-0.08	-0.04
NE 5	-0.13	0.00	-0.02	0.01	-0.09
NE 6	-0.17	0.03	-0.06	-0.04	-0.12
NE 7	-0.13	-0.03	-0.03	-0.05	-0.11
NE 8	-0.12	0.01	-0.03	0.02	-0.09
NE 9	-0.15	0.05	-0.09	-0.06	-0.11
NE 10	-0.08	0.05	-0.01	0.02	-0.03
S 1	0.01	-0.05	-0.02	-0.01	-0.02
S 2	-0.03	-0.11	0.03	0.06	-0.05
S 3	-0.05	-0.09	-0.02	0.00	-0.08

This means that there is one-to-one relationship between joint distribution function and copula. This approach has two big advantages. First, this approach allows us to cover very wide class of joint distribution functions. Second, this approach allows us decreases the dimension of the problem by making it possible to study separately marginal distributions and the copula. Loosely speaking, marginal distributions capture the probability of large or small production and copula captures the comovement between productions from two different resources. Copula captures much more than correlation. Two time series (e.g. hydro and wind power production) might have the same correlation, but the comovement might be very different (e.g., these might be highly correlated for unusually small production, but weakly correlated for unusually high production).

It is not obvious which copula will fit the data best. Therefore, we use many different copulas. These copulas are taken from [14] and their functional forms can be found there. We denote these copulas as indicated in Table IV. Some copulas are parametrizes by 1 parameter (normal copula), whereas others are parametrized by 2 parameters (Student's t copula). For illustration, Normal and Student's t copulas are plotted in Fig. 3 and Fig. 4.

Even though these two figures look similar, there is a big difference between them. Normal copula has zero, whereas Student's t copula has nonzero tail dependence. This means that joint extreme events (e.g. very low wind and hydro production) are very unlikely to happen under the Normal copula, but rather likely to happen under Student's t copula.

Table IV

Index	Copula Function Name	Parameters
C1	Normal	1
C2	Clayton	1
C3	Rotated Clayton	1
C4	Plackett	1
C5	Frank	1
C6	Gumbel	1
C7	Rotated Gumbel	1
C8	Student's t	2
C9	Symmetrized Joe-Clayton	2

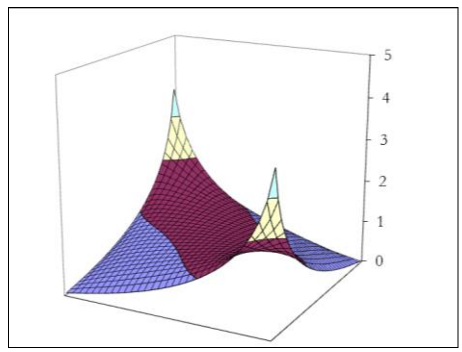

Fig. 3 Normal copulas example

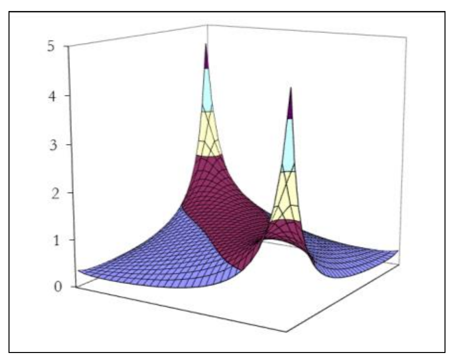

Fig. 4 Student's t Copulas example

We want to determine the parameters of the copula that describes the real data the best. I.e. we want to find the parameters of the joined density function, which can be rewritten as the product of marginal and copula densities:

$$h(x,y) = c(f(x), g(y),\alpha)\, f(x)g(y) \tag{3}$$

Where α is the parameter (or set of parameters) of the considered copula.

Now we should jointly estimate parameters of copula and marginal distributions by Maximum Likelihood. However, our focus is the copula part and we do not want to impose any restriction on the marginal densities. Therefore, we use Canonical Maximum Likelihood (CML). CML is a semi-parametric approach with two steps. First, we estimate marginal distributions using empirical cumulative distribution (a non-parametric step). Second, we use Canonical Maximum Likelihood to estimate joint density by copula (a parametric step). We put estimated distribution functions from the first stage to maximum likelihood function, which we want to maximize over the parameter of the considered copula.

After estimating joint distribution with different copulas we want to know which one can estimate the joint distribution of data best. There are two commonly used criteria based on Maximum Likelihood function: Akaike information criterion (AIC) and Bayes Information Criterion (BIC). The best copula is the one with the minimal value of AIC or BIC. However, for our data set AIC and BIC always select the same copula, so we do not need to distinguish between them.

Next, we discuss the results. We cannot present the results in the same way as we did in Table II and Table III, because instead of each correlation coefficient we now estimate nine different copulas, i.e. we estimate 9 copulas for each of 65 wind-hydro pairs (5 hydro x 13 wind). We therefore first present results for combination of different wind location with the total hydro production in Brazil, see Table VI.

Table V

Wind	**Ranking**								
	1.	**2.**	**3.**	**4.**	**5.**	**6.**	**7.**	**8.**	**9.**
NE 1	C8	C2	C9	C1	C4	C5	C3	C7	C6
NE 2	C2	C4	C1	C5	C3	C8	C9	C7	C6
NE 3	C8	C4	C1	C2	C5	C3	C9	C7	C6
NE 4	C8	C4	C1	C5	C3	C2	C9	C6	C7
NE 5	C1	C8	C4	C5	C2	C3	C9	C7	C6
NE 6	C8	C4	C1	C5	C2	C3	C9	C6	C7
NE 7	C8	C1	C4	C5	C2	C3	C9	C6	C7
NE 8	C1	C4	C8	C5	C3	C2	C9	C6	C7
NE 9	C4	C8	C1	C5	C3	C2	C9	C6	C7
NE 10	C1	C4	C8	C5	C3	C2	C9	C6	C7
S 1	C8	C4	C1	C2	C5	C3	C9	C7	C6
S 2	C4	C8	C1	C5	C2	C3	C9	C7	C6
S 3	C1	C8	C4	C5	C2	C3	C9	C7	C6

For each pair we ranked copulas according to their fit (minimum value of AIC or BIC), e.g. for a NE1-total hydro pair is the best fitting copula C8, next best fitting copula C2, and so on. These results mean that for example for the joint modeling of NE1 wind location with the total hydro production is the most suitable copula C8, next C2 etc. We can see that some copulas are ranked at the bottom most of the time (C6, C7 and C9), whereas other copulas are consistently on the first places (C1, C8).

Next, we present summary of the results for all the 65 pairs in Table VI. Number 31% at the intersection of row C1 and column “1.” mean that copula C1 was ranked as the best fitting copula in 31% of the cases (20 out of 65).

Copulas that are most suitable in fitting unpredictable part of wind and hydro production are Normal (C1), Plackett (C4) and Student’s t copula (C8). Results suggest that normal copula is generally fairly good fit for the co-movement between unpredictable part of the wind and hydro production.

Table VI

Copula Index	**Ranking**								
	1.	**2.**	**3.**	**4.**	**5.**	**6.**	**7.**	**8.**	**9.**
C1	31%	26%	25%	6%	3%	5%	5%	0%	0%
C2	5%	8%	2%	12%	31%	35%	5%	3%	0%
C3	2%	3%	6%	9%	37%	29%	6%	8%	0%
C4	20%	32%	22%	9%	11%	5%	2%	0%	0%
C5	3%	2%	11%	57%	15%	11%	2%	0%	0%
C6	0%	0%	0%	2%	0%	0%	2%	52%	45%
C7	2%	3%	0%	0%	0%	0%	3%	37%	55%
C8	35%	23%	23%	2%	2%	15%	0%	0%	0%
C9	3%	3%	12%	3%	2%	0%	77%	0%	0%

It is almost always ranked as either first best, second best or third best cupula. Therefore, in many practical applications is copula approach not necessary and it is sufficient to assume normality. However, the best fitting copula (copula that fits best in most cases) is Student’s t copula. Moreover, normal copula is a special case of Student’s t copula. Therefore, we recommend using copula approach in cases when it is important to model mutual dependence carefully.

There is another reason why to prefer Student’s t copula. The main difference between Normal and Student’s t copula is probably in tail dependence. Tail dependence is a concept that describes of often events in “tails” happen together. In our case, it can be interpreted as how often we observe unusually low (or high) production for both wind and hydro. Normal copula has zero tail dependence by definition, whereas Student’s t copula has nonzero. Therefore, if the true copula is Student’s t, but we use in our model normal copula, then we will underestimate joint occurrence of very high wind and very high hydro (or very low wind and very low hydro). On the contrary, if we use Student’s t copula, but the true copula is Normal, then it is not a problem. After estimation of Student’s t copula we will find low tail dependence.

IV. CONCLUSION

In this paper, we study complementarity between wind and hydro power in Brazil. Wind resources in the Northeast region are complementary to the Brazilian hydropower production. Since most of this complementarity comes from yearly seasonality of these resources, we decompose the wind and hydro power production into the predictable and unpredictable components. Next, we study the complementarity between the unpredictable component of the wind and hydro production. We find that these two are generally not correlated. Moreover, we conclude that copula approach can often capture mutual relationship better than implicit assumption of normality and that Student’s t copula fits data the best.

As avenue for future research, we suggest further investigation of the significance of our findings. A financial analysis is also recommended to evaluate the spot price movements against hydro-wind seasonal profile, as in Brazil spot price it strongly negative correlated with hydro generation because of the system operation characteristic.

ACKNOWLEDGMENT

The authors gratefully acknowledge the contributions of Lucas Witzler (University of São Paulo; MRTS Consultoria).

REFERENCES

[1] REN21. 2012. “Renewables 2012 global status report”. Renewable Policy Network for the 21st century. Paris. Available: http://www.ren21.net

[2] EPE. 2014 “Plano decenal de expansão energética - 2023”. Empresa de Pesquisa Energética. Retrieved January, 2015. Available in: http://www.epe.gov.br

[3] F. Ralston, S. Granville, M. Pereira, L. A. Barroso, A. Veiga. 2010. “Risk constrained contracting strategies of renewable portfolios”. In: 7th International Conference on the European Energy Market, Madrid.

[4] D. S. Ramos, L. A. S. Camargo, E. Guarnier, L. T. Witzler. Minimizing Market Risk by Trading Hydro-Wind Portfolio: A Complementarity Approach. In: 10th International Conference on the European Energy Market, 2013, Stockholm, Sweden.

[5] M. Denault, D. Dupuis, S. Couture-Cardinal, “Complementarity of hydro and wind power: Improving the risk profile of energy inflows,” Energy Policy, vol. 37, pp. 5376–5384, 2009.

[6] G. L. Susteras, D. S. Ramos, J. R. Chaves, A. C. V. J. Susteras. 2011. “Attracting wind generators to the wholesale market by mitigating individual exposure to intermittent outputs: an adaptation of the Brazilian experience with hydro generation”. In: 8th International Conference on the European Energy Market, Zagreb.

[7] J. C. Ricosti and I. L. Sauer, “An assessment of wind power prospects in the Brazilian hydrothermal system”, Renewable and Sustainable Energy Reviews, vol 19, pp. 742–753, 2013.

[8] P. Jaworski, F. Durante, W. Härdle, T. Rychlik, “Copula theory and its application”. Proceeding of the workshop held in Warsaw, 25–26 September 2009, vol. 198, Ed. New York: Springer.

[9] M. V. Pereira, N. Campodónico, R. Kelman.1998. “Long-term hydro scheduling based on stochastic models”. Proceedings of EPSOM Conference, Zurich, 1998 – available at http://www.psr-inc.com

[10] 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.

[11] ONS—Operador Nacional do Sistema. Available from: www.ons.org.br/historico/energia_natural_afluente_out.aspxs.

[12] Data provided by Vestas (www.vestas.com) for academic purposes.

[13] NOAA – National Oceanic and Atmospheric Administration. Available from: www.noaa.gov.

[14] A. J. Patton, “On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation”. Journal of Financial Econometrics, vol 2, pp. 130–168, 2004.