Bio-Inspired System for Electricity Price Forecast in the Brazilian Market

1st Lídio Mauro Lima de Campos
Jherson Haryson Almeida Pereira,
Danilo Souza Duarte
Faculty of Computing - ICEN
Federal University of Pará
Rua Augusto Corrêa 01-Guamá,
Belém, Pará CEP 66075-100, Brasil
Email: lidio@ufpa.br
danilo.duarte8@hotmail.com

2rd Roberto Célio Limão de Oliveira
Faculty of Computer Engineering-ITEC
Federal University of Pará
Rua Augusto Corrêa 01-Guamá,
Belém, Pará CEP 66075-100, Brasil
Email: limao@ufpa.br

***Abstract*—The aim of this paper is to intoduce a biologically insipired methodology for energy price prediction multi-step ahead. The system combines a Genetic Algorithm (GA), Lindenmayer Systems, and Deep Neural Networks (DNNs). In the final section of the paper, we present some experiments to investigate the possibilities of the method, especially for scenarios in which DNNs should be evolved. The results for MLP networks show good ability to predict spikes and satisfactory accuracy according to error measures up to 9.8%. In 58.33% of the cases, ADEANN-Deep provides better results than other hybrid systems. Preliminary studies show that the LSTM network can predict the PLD value with a percentual error of up to 2.9%, which is lower than that obtained using multilayer perceptron networks and other methods.**

***Index Terms*—NeuroEvolution; Machine Learning,Time series forecasting.**

I. Introduction

Forecasting the price of electricity is an important matter for all market participants to decide on the most appropriate bidding strategies and establish bilateral contracts that maximize their profits and minimize their risks. Currently, some forecast methods have been proposed in the literature [1], [5]. We mention the following contribuitions and advantages of our method: ADEANN-Deep improved past applications for working with a population of deep neural networks instead of only one, all are trained, and those which have better predictive capacity are selected, that is, ADEANN-Deep makes the automatic design of deep neural networks, without the intervention of the human specialist. Moreover, the novelty from a neural network perspective is that ADEANN-Deep was designed as an accurate automatic method for designing direct and recurrent ANNs, thereby solving pattern recognition problems and dynamical systems simulation, such as the prediction of time series. In addition, our approach can be generalized for other electricity markets, including wind power [14] and Photovoltaics [13] and others.

Fig1 illustrates the general structure of ADEANN-Deep. The hybrid system is described in detail in section V. In relation to the previous version of ADEANN [4], this research presents the following improvements: system migration to Python language justified by its applicability and portability in the artificial intelligence area, which enabled integration with prominent frameworks used in the current market for data processing, as well as data science, such as Keras and Tensorflow.

This paper is organized as follows. Section II discusses related work. Section III presents the features of the Brazilian electricity market. In Section IV we describe a new approach to formalize the problem of ADANNs (artificial development and evolution of ANNs) as a local search based on rational agents. Section V introduces a biologically inspired method for automatic design of ANNs. Section VI presents Material and Methods. Lastly, simulation results and conclusions are presented in Sections VII and VIII, respectively.

II. Related Work

The paper [6] proposes a scheme for evolving multiple-input-multiple-output (MIMO) artificial neural networks (ANNs) using grammatical evolution (GE). GE is a well-known technique for program evolution. While it has also been used for the evolution of ANN structures in the past, little work is reported on the evolution of MIMO ANNs.

The paper [2] proposes a Deep evolutionary network structured representation (DENSER) which is a novel evolutionary approach for the automatic generation of deep neural networks (DNNs) which combines the principles of genetic algorithms (GAs) with those of dynamic structured grammatical evolution (DSGE). The GA-level encodes the macro structure of evolution, i.e., the layers, learning, and/or data augmentation

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

methods (among others); the DSGE-level specifies the parameters of each GA evolutionary unit and the valid range of the parameters. The use of a grammar makes DENSER a general purpose framework for generating DNNs: one just needs to adapt the grammar to be able to deal with different network and layer types, problems, or even to change the range of the parameters. DENSER is tested on the automatic generation of convolutional neural networks (CNNs) for the CIFAR-10 dataset, with the best performing networks reaching accuracies of up to 95.22%.

[15] used a Hybrid neural model (HIRA model) for Short-Term Electricity price forecasting, which was also applied and tested on two eletricity markets: German and Hungarian. In Hungary, thermal plants have 90% of installed capacity and represent a stable source of electricity, but generally involving a large installed capacity per unit. Hungary uses bid-based market, ther work calculates the electricity price forescast at three levels: Hourly Forecasting error, Peak load Forecasting Error and Base Load Forecasting Error. In section III, we briefly describe the particular characteristics of the Brazilian market.

This paper [3] proposes a new hybrid approach for short-term energy price prediction. The approach combines auto-regressive integrated moving average (ARIMA) and neural network (NN) models in a cascaded structure applying explanatory variables. A two-step procedure is applied. In the first step, the selected explanatory variables are predicted, while in the second one, the energy prices are forecasted by using the explanatory variables prediction.

Function SEARCH−ANN(SearchParam, TransitionModel, FitnessFunction) **return** an ANN topology

---

1:inputs: SearchParam, TransitionModel, FitnessFunction
2:vars: Pop, t, PopPerformances;
3:k← 0
4:$ANNsPop^k$←Generate-ANNs(SearchParam)
5:$ANNsPerformance$← Evaluate-ANNs($ANNsPop^k$, FitnessFunction)
6:loop do
7:if StopConditionTest(k, ANNsPerformance, SearchParam)
8:then return solution(Best-ANN($ANNsPop^k$, ANNsPerformance)
9:$ANNsPop^{(k+1)}$←($ANNsPop^{(k+1)}$, ANNsPerformance, TransitionModel, SearchParam)
10:ANNsPerformance←Evaluate-ANNs($ANNsPop^{(k+1)}$, FitnessFunction)
11:k ←k+1
12:end

III. Brazilian electricity market

Brazil uses a cost-based market instead of a bid-based market, and adopts a tight pool model with a centralized and least cost dispatch organized by National System Operator (ONS) [11]. This scheme is adopted due the country peculiarities, which has an installed capacity of 121 GW where 65.96% corresponds to hydro generation. The hydro system is composed of several reservoirs capable of multi-year regulation located at the same river with different owners [3].

The difference between the quantity of energy contracted and that effectively consumed or produced by the agents is accounted in the short-term market based on the spot price called PLD (settlement price for the differences) [3]. PLD is calculated weekly, based on the system marginal cost of operation obtained from an optimization process to dispatch generators, at three load levels (”Medium”,”Heavy”,”Light”). The PLD is established by the Brazilian Electricity Regulatory Agency (ANEEL) [10] and is evaluated to each submarket associated with the country regions: North, Northeast, Center-west/Southeast, and South.

IV. Outline of the Approach

Our approach involves the formulation of an artificial neural network design as an optimization problem (ANNDP), that is: given a set of L observations on the behavior of a particular process, Ψ={(xd^l, yd^l)}, 1 = 1...L , where xd^l represents a numeric vector defined in R^n and yd^l is a numeric vector defined in R^m, the goal is to find an ANN's topology, yc^l = ANN(w*, xd^l), which minimizes the mean square error between yd^l and yc^l, this is, between the desired values in the observations set and the computed values in the neurons' outputs situated in the ANN's output layer.

An ANNs topology can be described as a finite set of neurons, that is, nodes of an oriented graph Nodes ={n1, n2, ..nk}, and a finite set H $\subseteq$ N x N of connections between neurons, which means directed edges in graphs notation. An input layer is a set of input units, that is, a subset of n nodes whereas an output layer is a set of output units, namely a subset of m nodes. In feed-forward ANNs(FANNs), the k^{th} layer (k > 1) is the set of all nodes n_i ϵ Nodes . These types of nodes have an edge path of length k - 1 between some input unit and u. In fully connected recurrent ANNs (RANNs), all units have connections to all non-input units.

We approach the solution of ANNDP based on the methodology of Russell and Norvig [7] called problem-solving-agent, whose agent is named ADEANN (Artificial Development and Evolution of ANNs), which encapsulates a special scheme of solutions representation as well as a local search strategy based on genetic algorithms to solve the problem. Regarding the representation scheme, the approach adopts a generative representation, which means that, instead of an encoded ANN topology, each chromosome stores a set of production rules of a Lindenmayer system (Fig 1 a, k, h) which, in turn, generates ANN's, (Fig 1 f). The SEARCH-ANN function illustrates the structure of the program in the ADEANN agent.

Rule Identifier	Rule
1,2	S→. (axiom) (2) .→(f...f)n
3	(3.1) f→[f (3.2) f→fFf (3.3) f→fF (3.4) f→n
3,4,5	(3.5) f→f (3.6) f→fB (4) [→[Ff] (5) f→f*

TABLE I
THE PRODUCTION RULES OF THE PARAMETRIC L-SYSTEM WITH MEMORY

The SEARCH-ANN function starts the local search process aiming at achieving an artificial neural network topology yc^l=ANN$\{(w^*, xd^l)\}$, which minimizes the mean square error between yd^l and yc^l, for l = 1...L in the ANNDPs formulation (stop condition in Fig 1). This function employs information on the search parameters (SearchParam input term) as well as a transition model (TransitionModel input term) to describe how to modify current populations of ANNs and generate a new population, (Fig 1 i, j, k), in addition to an evaluation function (FitnessFunction input term), (Fig 1 g), to measure the value of each ANN in a current population.

Firstly, in the beginning of the process, Generate-ANNs function generates an initial population of ANNs, (Fig 1 a), in which each ANN is represented by a set of production rules codified in a chromosome (bit values 0 and 1). This function considers the information in the SearchParam input term on the desired number of ANNs in the populations as well as on the desired length for the chromosomes in the population. Evaluate-ANNs function stores in ANNsPerformace the computed performance value of each ANN topology in the current population based on the mean square error computed in the output layer of the ANN-SEARCH-ANN function, (Fig 1 g), which employs an iteration counter (k) and a condition named StopConditionTest boolean function to decide when to stop the local search process and return a solution to a problem (stop condition in Fig 1). The description of the stop condition is based on a proposition relating the information on the current iteration counter k and the information available in the SearchParam input term. This means that the max number of loops in its repetition scheme is central to the local search strategy in the approach, as well as an ideal performance value such that for an ANN to be considered a solution. Modify-ANN function is executed repeatedly seeking to transform a current population of ANNs in a new population of ANNs (Fig 1 i, j, k). In our approach, this function encapsulates the evolutionary principles of pairs selection and crossing over pairs and individual mutation (Fig 1 j) . Central to the approach, compact indirect encoding scheme (IES) conducts and controls the process of mapping a set of production rules of a Lindenmayer system (Fig 1 c) , codified in a chromosome to an associated ANN topology (Fig 1 e). In sections V.A and V.B the generative representation, which generates ANN's through production rules and the rule extraction process are explained in details.

V. Biologically Inspired NEA

The following subsections describe the three subsystems of ADEANN-Deep.

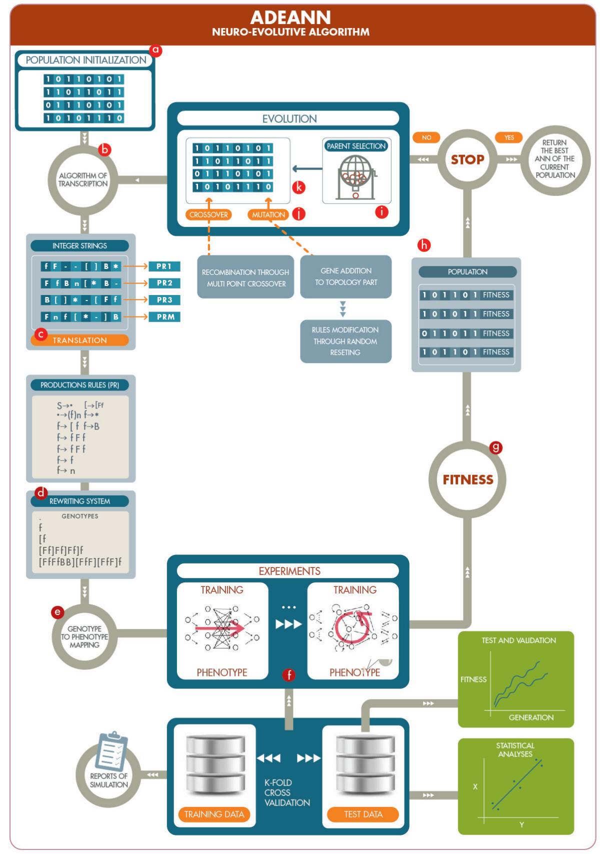

Fig. 1. The general structure of ADEANN-Deep.

A. L-system based artificial embryogenesis model

To mimic the mechanism of grown structures, including neurons, we adopt a parametric L-system with memory. It comprises a set of rules created from an alphabet. This system can be described as a grammar $G = \{\Sigma, \Pi, \alpha\}$, where the alphabet consists of the elements of the set $\Sigma = \{., f, F, n, [,], *, B\}$, and the production rules (Π) described in Table I.

The axiom α = . is the starting point of the developmental process, where f denotes a neuron and F is a connection between neurons, [and] indicate storage and recovery, respectively, of the current state of the development, * denotes that the string is recovered from storage, and B is the connection of a neuron with a block of neurons. The second rule .→(f...f)n, means replace the start point by the neurons of the input layer. Rule 3.1 (f→[f) means to store the position of the current neuron, so as to start a new ramification from it. Rule 3.2 (f→fFf) means establish a connection between two neurons. Rule 3.3 (f→fF) means establishing a connection from a specific neuron. Rule 3.4 (f→n) means replace a provisional neuron with a permanent neuron. Rule 3.5 (f→f) means to maintain a specific neuron during development. Rule 3.6 (f→fB) means connect a neuron to a block of neurons. Rule 4 ([→[Ff]) means start the development of a new ramification

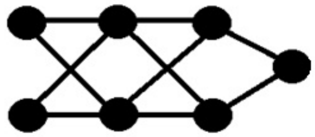

Fig. 2. A simple example of the construction process of a branch of an iterated ANN using the rules of the L-system is illustrated in Table **??**.

from a specific neuron and recover the previous state. Rule 5(f→f*) means recover a previous ramification stored for use.

As a simple example, suppose that starting with the axiom ($\alpha = .$) twice , and applying the second production rule . $\to f$ to the axiom twice, the resulting string is **ff**. Applying the third rule (3.1) $f \to [f$ to string **ff** yields a new string, **[f[f]**.

After one, two, three and applications of the fourth rule $[\to [Ff]$ on string **[f[f**, the string becomes **[Ff]Ff]f[Ff]Ff]f**. Applying the rule (3.3) f $\to$**fFf** eight time to the previous string, the resulting string is **[FfFfFf]FfFfFf]f [FfFfFf]FfFfFf]f** . After Applying the (3.6) $f \to$**fB** to the previous string the resulting string is **[FfBFfBFfB]FfBFfBFfB]f [FfBFfBFfB]FfBFfBFfB]f**. Finally applying the rule (3.4) f $\to$**n** to the previous string the resulting string is **[FfnFfnFfn]FfnFfnFfn]f [FfBFfBFfB]FfBFfBFfB]f**. This phenotype represents the RNA structure shown in Fig 2.

B. Rule extraction by genetic algorithms

The neurons generated in the previous subsection are developed after the following process. To formulate a biologically realistic GA, we let the genes of the chromosomes (sequences of hypothetical DNA) encode a recipe (the production rules of the L-system described in subsection V.A and illustrated in Table I). The recursive rules in Table I drive the developmental stages of the neurons , see Fig 2.

In biological genetic processing Fig 3 (b), DNA is transcribed into ribonucleic acid (RNA), and the RNA is translated into proteins. The proteins are derived from linear sequences of amino acids encoded by codons (groups of three nucleotides selected among U, G, A, and G of the genetic code (Table II). In Fig 3 (b), the protein is formed by a sequence of amino acids starting with methionine (Met) and ending with proline (Pro). Such protein synthesis triggers all stages of the neuronal development (phenotypic effects), as shown in Fig 3 (b). The elements of the alphabet $\Sigma = \{., f, F, n, [,], *, B, \}$ of the L-system, described in subsection V.A and displayed in bold font in Table II, are a metaphor of the genetic code. Each two-bit sequence represents one nucleotide; for example, the set (00, 01, 10, 11) symbolizes (U, C, A, G) in the original genetic code. Accordingly, six bits represent three nucleotides; that is, (000000, 011111) symbolizes (UUU, CGG).

In the figure Fig3 (a), We show an example of extracting rules, the transcription process (from binary sequence to string) yields the string B.f[Ff.nB]Bf. We seek the shortest string containing all valid rules; in this case, (**.f[Ff *nB]**). After finding the minimum string, we identify the positions at which the rules were found, see Fig3 (a)). For example, Rule 2 (.→f),

	00 (U)	01 (C)	10 (G)	11 (A)	
00 (U)	f (UUU)	F (UCU)	n (UAU)	. (UGU)	00 (U)
00 (U)	n (UUC)	. (UCC)	f (UAC)	F (UGC)	01 (C)
00 (U)	F (UUA)	f (UCA)	B (UAA)	f (UGA)	10 (A)
00 (U)	[(UUG)	n (UCG)	[(UAG)	* (UGG)	11 (G)
01 (C)	f (CUU)	] (CCU)	n (CAU)	* (CGU)	00 (U)
01 (C)	* (CUC)	F (CCC)	f (CAC)	F (CGC)	01 (C)
01 (C)	] (CUA)	f (CCA)	* (CAA)	[(CGA)	10 (A)
01 (C)	f (CUG)	* (CCG)	B (CAG)	] (CGG)	11 (G)
10 (A)	* (AUU)	] (ACU)	n (AAU)	f (AGU)	00 (U)
10 (A)	f (AUC))	B (ACC)	f (AAC)	B (AGC)	01 (C)
10 (A)	F (AUA)	[(ACA)	B (AAA)	n (AGA)	10 (A)
10 (A)	* (AUG)	f (ACG)	* (AAG)	] (AGG)	11 (G)
11 (G)	] (GUU)	[(GCU)	F (GAU)	n (GGU)	00 (U)
11 (G)	n (GUC)	B (GCC)	[(GAC)	. (GGC)	01 (C)
11 (G)	f (GUA)	] (CGA)	B (GAA)	F (GGA)	10 (A)
11 (G)	B (GUG)	f (GCG)	* (GAG)	[(GGG)	11 (G)

TABLE II

THE GENETIC CODE FROM THE PERSPECTIVE OF MRNA, TRANSLATED AS IN FIG 3 (B). IN THE SAME TABLE, THE DNA'S METAPHOR

symbolically represented by (.f), is found at positions 1 and 2 of the string **.f[Ff *nB]**. Rule 3.1 f→[f is found at positions 1 and 2, and 3 and 5.

C. Neural Network

Neural Networks have been successfully aplied to a variety of complex problems due to its ability to learn non-linear relationships between input and output patterns, which would be difficult to model conventional methods [3]. In this research, ADEANN-Deep enables automatic design of different recurrent and deep neural network architectures that yields the best generalization accuracy for each submarket.

Recurrent neural networks have gained widespread use in modeling sequential data. The Long Short-Term Memory network (LSTM network) is a type of RNN in deep learning because very large architectures can be successfully trained. The work [8] shows the robustness of LSTM networks in handling unbalanced data. The empirical studies conducted and reported in the paper [9] show that deep learning based algorithms such as LSTM outperform traditional-based algorithms such as ARIMA model

VI. MATERIAL AND METHODS

A. Explanatory Variable Selection in Prediction Models

In prediction models, the explanatory variable can explain or cause differences in a response variable. After the identification of the explanatory variables, an explanatory variable selection method is applied to find the optimal set of input variables required to describe the behavior of the energy price, which should contain a minimum degree of redundancy. The aim is to test how two or more variables act togheter to affect the output variable and determine whether they improve the prediction of the desired value. The PLDs prediction has the following referenced variables: stored energy in reservoirs (%MLT), inflow energy in reservoirs (%MLT), total hydro generation (MWmed), total thermal generation (MWmed), system power load (MWmed). Below, we detail the variables chosen in the work of [3] for each submarket, : North:Stored energy, Inflow

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

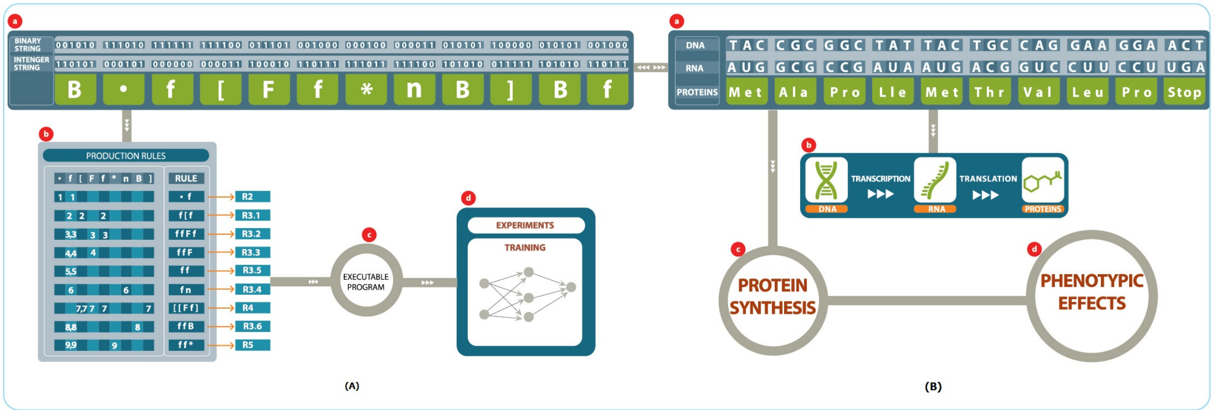

Fig. 3. (b) DNA transcription into RNA and translation of RNA into protein. (a) In the analogous artificial process, a binary string is transcribed into an integer string and the string is translated into the production rules of the L-system.

energy and Load, Northeast:Stored energy, Inflow energy , Thermal generation and Load. Center-Wets/Southeast:Stored energy, Hydro generation, Thermal generation and Load, South:Stored Enregy, Hydro generation, Thermal generation and load. We have used the same explanatory variables, selected in the work [3] for each submarket.

B. Dataset Description

The dataset used in this research contains the electricity prices data taken from Brazilian Electrical Energy Commercialization Chamber website [10] presented on a weekly basis, in addition to the explanatory variables data taken from Brazilain National System Operator website [11]. In the simulations we applied the dataset constructed by [3] (period from 2002 to 2009) to each submarket:North, Northeast, South and Center-West/Southeast.

C. Used Metrics

The hybrid system proposed in this system is applied to the Brazilian electricity market. Some metrics commonly used to evaluate proce forecasting accuracy are employed in this paper , Root mean squared error (RMSE) , Mean absolute error (MAE) and Mean absolute percentage error (MAPE). These quantities are calculate by:

$$MAE = \frac{1}{N}\Sigma_{i=1}^{N}|p_i true - p_i forecast| \tag{1}$$

$$RMSE = \sqrt{\frac{1}{N}\Sigma_{i=1}^{N}\Big(p_i true - p_i forecast\Big)^2} \tag{2}$$

$$MAPE = \frac{100\%}{N}\sum_{I=1}^{N}\left|\frac{p_i true - p_i forecast}{p_i true}\right| \tag{3}$$

Where N is the number of samples, $p_i true$ is the actual price and $p_i forecast$ is the forecasted price.

The fitness function, given by equation 6, selects economical deep neural networks.

$$FitnessI = \exp(-RMSE) \times \prod_{i=0}^{n}\exp(-NNHL) \tag{4}$$

$$FitnessII = \frac{1}{RMSE \times \prod_{i=0}^{n}\exp(-NNHL)} \tag{5}$$

$$Fitness = FitnessI + FitnessII \tag{6}$$

where NNHL is the number of neurons in the hidden layer.

D. Statistics

We repeated each experiment five times. The independence of the events was assured since the runs were independent and had randomly generated initial seeds. Furthermore, to evaluate the significance of the results obtained from ADEANN and the other NEAs, we carried out t-tests with a confidence level of 95% (i.e., a p-value under 0.05). To statistically compare the performances of two NEAs for the prediction problems, we considered the three following criteria: RMSE as the primary criterion, MAPE as the second one, and MAE as the third one. Similarly to [12], without the occurrence of any significant statistical differences between the RMSE, MAPE, and MAE values of two NEAs on a given dataset, it was considered that both algorithms performed equally well. In this case, both algorithms receive 1 point. In contrast, if two algorithms obtain significantly different RMSE or MAPE or MAE scores, the better performing algorithm receive two points and the other zero points. Consequently, H_0 and accept the alternative hypothesis H_1. In deciding whether two performances differ, we test the significance of the difference between u_1 and u_2 ($p < 0.05$). The overall performance of each NEA is then calculated by summing all points achieved in the pairwise comparisons.

E. Data Preparation

Many problems are involved in the analysis of rare patterns of ocurrences. As an example, Figure 4 shows the histogram of the PLD series for the North region. It shows that some patterns occur more often than others. In addition, most of the time the price remains at low values, under R$100.00 and the energy price rarely reaches values above R$300.00. However, neural networks are sensitive to imbalanced data sets since it

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

Variable	Minimun	Mean	Maximum	Std.Deviation
Stored energy	20.71	65.30	87.64	16.7
Inflow energy	47.86	103.87	182.00	24.38
Hydro generation	8854.14	17.635.02	23.378.14	3253.84
Thermal generation	192.86	1112.57	3258.86	615.93
Load	19.295.57	28.048.88	34.668.00	3148.61
PLD	4.0	74.65	684.00	117.53

TABLE III
SUMMARY STATISTICS OF THE VARIABLES FROM CENTER-WEST/SOUTHEAST REGION. FONT: [3]

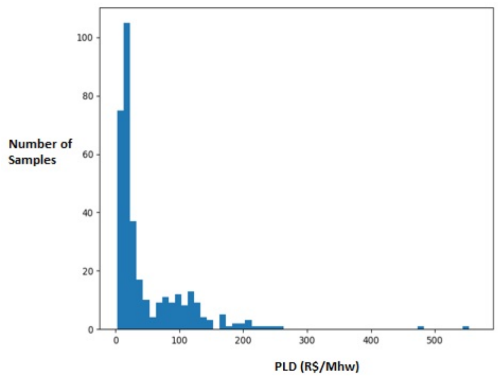

Fig. 4. Histrogram of the PLD series to North Region.

causes difficulties in the learning process and can deteriorate the model performance. Then the data balancing was applied in this paper during data preparation process. Figures 5 shows the histogram of the PLD series, before and after data balancing, for the North region.

VII. SIMULATION RESULTS

The methodology proposed in this paper is applied to the Brazilian electricity market and some criteria commonly used to evaluate price forecasting accuracy are employed, such as RMSE, MAE and MAPE.

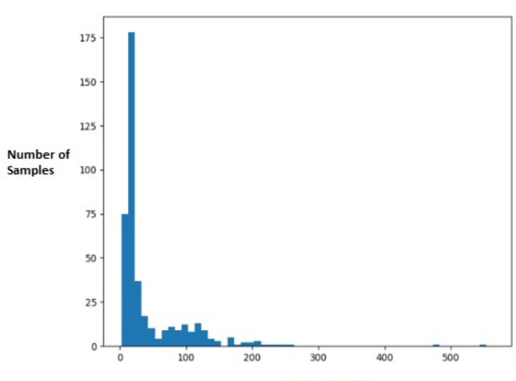

Fig. 5. Histrogram of the PLD series to North Region (after data balancing).

	South Region			North Region		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE
Medium PLD	11.40	8.05	0.21	9.35	6.61	0.22
Light PLD	11.07	7.82	0.20	9.27	6.55	0.20
Heavy PLD	11.59	8.20	0.21	10.48	7.41	0.22
Average	11.35	8.02	0.21	9.7	6.86	0.22
	Northeast Region			Southeast Region		
Medium PLD	9.03	6.38	0.19	15.47	10.94	0.20
Light PLD	10.21	7.22	0.21	15.47	10.94	0.20
Heavy PLD	9.01	6.37	0.18	17.15	12.13	0.22
Average	9.41	6.66	0.19	16.12	7.69	0.20

TABLE IV
ENERGY PRICE ERROR MEASURES OBTAINED WITH THE PROPOSED HYBRID SYSTEM 36-WEEKS AHEAD - (UNBALANCED DATA)

A. Multi layer perceptron netwokws assessment for energy price forecasting

*1) **Unbalanced Data**:* Tables IV illustrates the values of RMSE, MAE, and MAPE obtained from ADEANN-Deep for the South, North, Northeast and Southeast regions. Figure 6 shows the short-term price observed and predicted with the proposed hybrid system 36-weeks ahead to Northeast region. For the Southeast region, the RMSE and MAE obtained from ADEANN-Deep 16.12 R$/Mwh and 7.69 R$/ Mwh were higher than those obtained from the Hybrid System 9 R$/Mwh and 3 R$/Mwh [3]. However, the value of MAPE 0.20 R$/Mwh reached using ADEANN-Deep was below that obtained from the Hybrid System [3], which generated a MAPE value equal to 5 R$/ Mwh. For the Northeast region, the RMSE and MAE achieved using ADEANN-Deep 9.41 R$/Mwh and 6.66 R$/ Mwh were higher than those from the Hybrid System 9 R$/Mwh and 3 R$/Mwh [3]. However, the value of MAPE 0.19 R$/Mwh generated from ADEANN-Deep was below that from the Hybrid System [3], which reached a MAPE value equal to 4.5 R$/ Mwh.

Figure 7 shows the short-term price observed and predicted with the proposed hybrid system 36-weeks ahead to North region. For the South region, the RMSE and MAE generated from ADEANN-Deep 11.35 R$/Mwh and 8.02 R$/ Mwh were above those from the Hybrid System 9 R$/Mwh and 3 R$/Mwh [3]. However, the value of MAPE 0.21 R$/Mwh reached using ADEANN-Deep was lower than that achieved using the Hybrid System [3], which generated a MAPE value equal to 5 R$/ Mwh. For the North region, the RMSE and MAE obtained from ADEANN-Deep 9.70 R$/Mwh and 6.86 R$/ Mwh were higher than those from the Hybrid System 7.5 R$/Mwh and 2.5 R$/Mwh [3]. However, the value of MAPE 0.20 R$/Mwh generated using ADEANN-Deep was below that obtained from the Hybrid System [3], which reached a MAPE value equal to 4.8 R$/ Mwh.

*2) **Balanced Data**:* Table V illustrates the values of RMSE, MAE, and MAPE obtained from ADEANN-Deep for the South, North, Northeast and Southeast regions. The analysis of the results for the South region revealed a mean square error (RMSE) generated using ADEANN-Deep of 8.20 R$ Mwh, lower than 9 R$/Mwh, obtained from the Hybrid model [3]. The MAE and MAPE generated using ADEANN-Deep

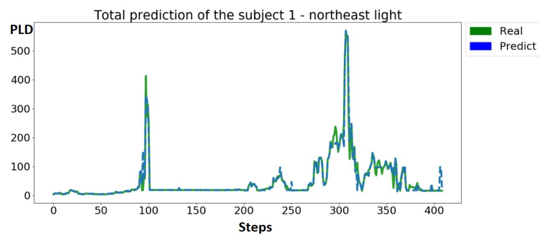

Fig. 6. Energy price observed and predicted with the hybrid model to Northeast Region (Light PLD) - UnBalanced Data

	South Region			North Region		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE
Medium PLD	7.52	5.31	0.18	4.22	2.98	0.11
Light PLD	8.52	6.02	0.20	5.2	3.68	0.13
Heavy PLD	8.57	6.06	0.19	5.64	3.99	0.14
Average	8.20	5.80	0.19	5.02	3.55	0.12
	Northeast Region			Southeast Region		
Medium PLD	5.86	4.14	0.17	9.91	7.01	0.15
Light PLD	5.70	4.03	0.16	9.95	7.03	0.15
Heavy PLD	5.84	4.13	0.17	9.55	6.75	0.14
Average	5.80	4.10	0.17	9.8	6.93	0.14

TABLE V
ENERGY PRICE ERROR MEASURES OBTAINED WITH THE PROPOSED HYBRID SYSTEM 36-WEEKS AHEAD - (BALANCED DATA)

were 5.80 R$/ Mwh and 0.19 R$/Mwh, respectively, while the values obtained from the hybrid method [3] were 3 R$/ Mwh and 5 R$/ Mwh, respectively. Therefore, our MAE value was higher than that obtained from the hybrid system, while our MAPE value was below it. For the North region, the mean square error (RMSE) generated from ADEANN-Deep was 5.02 R$/Mwh, lower than 7.5 R$/Mwh, reached using the hybrid system [3]. The MAE and MAPE obtained using ADEANN-Deep were 3.55 R$/ Mwh and 0.12 R$/ Mwh, respectively, while the values generated using the hybrid model [3] were 2.5 R$/Mwh and 4.8 R$/Mwh, respectively. Therefore, our MAE value was higher than that from the hybrid system and our MAPE value was below it. The analysis of the results for the Southeast region revealed a mean square error (RMSE) using the ADEANN-Deep of 9.8 R$ Mwh, higher than 9.4 R$/Mwh, obtained from the Hybrid model [3]. The MAE and MAPE reached using ADEANN-Deep were 6.93 R$/ Mwh and 0.14 R$/Mwh, respectively, while the values from the hybrid method [3] were 5 R$/ Mwh and 4 R$/ Mwh, respectively. Therefore, our MAE value was higher than that obtained from the hybrid system and our MAPE value was below it. For the Northeast region, the mean square error (RMSE) generated from ADEANN-Deep was 5.8 R$/Mwh, lower than 8 R$/Mwh, obtained from the hybrid model [3]. The MAE and MAPE generated using ADEANN-Deep were 4.1 R$/ Mwh and 0.17 R$/ Mwh, respectively, while the values obtained from the hybrid model [3] were 4 R$/Mwh and 5 R$/Mwh, respectively. Therefore, our MAE value was higher than that reached using the hybrid system and our MAPE value was below it. It is noteworthy that all other models are forecasting 12 weeks ahead, however, ADEANN-Deep is

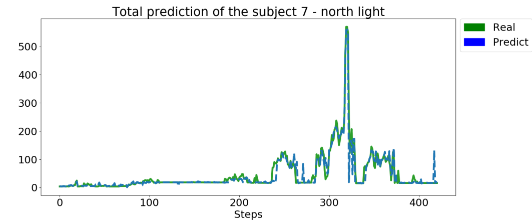

Fig. 7. Energy price observed and predicted with ADEANN-DEPP for the North region (Light PLD) - Unbalanced Data.

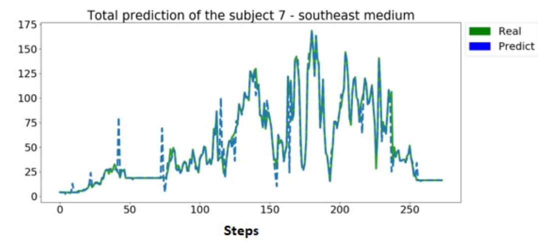

Fig. 8. Energy price observed and predicted applying ADEANN-DEPP for the Southeast region (Heavy PLD) - Balanced Data - MLP network

forecasting 36 weeks ahead.

B. LSTM networks assessment for energy price forecasting

After simulations with feedforward neural networks, presented in sections VII.A.1 and VII.A.2, using unbalanced and balanced data, respectively, we started simulations with LSTM networks, which are more robust for time series prediction. For purposes of verifying the robustness of LSTM networks, our simulations did not use data balancing, since it is more relevant for the MLP neural networks, which are sensitive to imbalanced data sets for causing difficulties in the learning process and possibly deteriorating the model performance. Table VI illustrates the values of %Error, MAE, in addition to the MAPE obtained from ADEANN-Deep for the Northeast and Southeast regions.

The analysis of the results for the Northeast region, considering only heavy PLD, revealed a MAE and MAPE using the ADEANN-Deep of 5.17 R$/ Mwh and 0.19 R$/ Mwh , respectively, lower than the values from simulations with multilayer perceptron networks (MLPs), see Table IV , 6.66 R$/ Mwh and 0.19 R$/ Mwh, respectively. The neural network architecture returned by ADEANN-Deep had the following parameters: Number of inputs = 5, number of neurons in the first hidden layer = 100, number of neurons in the second hidden layer = 100, number of outputs = 1, function of activation of the neurons of the first hidden layer = hyperbolic tangent, function of activation of the neurons of the second hidden layer = hyperbolic tangent, function of activation of the output layer = LINEAR, optimizer = RMSPROP, Batch=32 and number of epochs = 2000.

The analysis of the results for the Southeast region, considering only heavy PLD a MAE and MAPE using the ADEANN-Deep of 7.38 R$/Mwh and 0.21 R$/Mwh, respec-

	Southeast Region			Northeast Region		
	%Error	MAE	MAPE	%Error	MAE	MAPE
12 weeks ahead	0.16%	2.5	0.15	2.9%	0.12	0.007
24 weeks ahead	4.6%	8.73	0.52	10.4%	3.43	0.21
36 weeks ahead	6.9%	7.38	0.21	8.83%	5.17	0.19

TABLE VI
ENERGY PRICE ERROR MEASURES OBTAINED WITH THE PROPOSED SYSTEM FREE SIMULATIONS FOR 12, 24 AND 36 WEEKS AHEAD - HEAVY PLD - (UNBALANCED DATA)

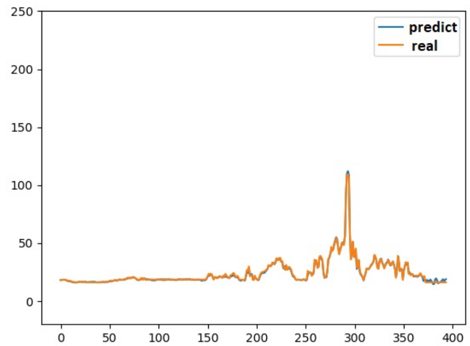

Fig. 9. Energy price observed and predicted applying ADEANN-DEPP for the Southeast region (Heavy PLD) - Unbalanced Data - LSTM network

tively, lower than the values from simulations with multilayer perceptron networks (MLPs), see Table IV , 7.69 R$/Mwh and 0.20 R$/Mwh, respectively. Figure 9 shows the short-term price observed and predicted with the proposed hybrid system 36-weeks ahead to Southeast region using LSTM network.

The results using HIRA model [15] in the Hungarian market, three days in advance for day d, at the 3rd interval for base load energy price (is the average hourly price for all 24 h) and peak (is the delivery of the same amount of energy in the period from 08:00 to 20:00) indicate a percentage error around 0.56% to 2.96%. The percentage hourly price forecast error is around 4.55% to 11.37%. Our results for the Northeast and Southeast regions, twelve weeks in advance for week w, considering only heavy PLD, revealed a percentage error around 0.16% to 2.9%, lower than the values from HIRA model [15]. In addition, our predictions encompass the period of 12 weeks ahead, while the predictions of the HIRA method correspond to three days ahead.

VIII. CONCLUSIONS

This paper proposes a hybrid approach for a short-term energy price prediction. The model considers multi-step ahead price prediction and is applied to the Brazilian electricity market. The results obtained using MPLs networks are compared with the study of [3] and others methods in section VII.A. The results obtained from ADEANN-Deep applied to the Brazilian market presented a sufficiently good accuracy level compared to other methods. Data balancing and the use of explanatory variables proved essential for having improved the results generated using ADEANN-Deep, according to the results presented in section VII.A.2. Statistical test with a confidence level of 95% shows that in 58.33% of the cases, ADEANN-Deep provides better results than the hybrid system [3].

Our simulations with LSTM networks are still at an early stage. It can be concluded from the results presented in section VII.B for the southeastern region (heavy PLD) that the LSTM Network can predict the value of the PLD with an error of up to 2.9%, which is lower than with the value generated from multilayer perceptron networks and HIRA model [15]. In addition, the new version of the hybrid system (ADEANN-Deep) using Keras enhances the possibilities of using multiple deep recurring Neural Network architectures. We proved the applicability of our method to forecast electricity prices and enable effective hedging of price risk in the production.

ACKNOWLEDGMENT

We would like to thank the Federal University of Pará for having supported this study.

REFERENCES

[1] Sharma M. ;Gupta S. ; Sharma P.;Gupta D,Bio-inspired algorithms for diagnosis of breast cancer. International Journal of Innovative Computing and Applications, 2019, Vol.10 No.3/4.

[2] Assuno, F., Loureno, N., Machado, P. et al. Genet Program Evolvable Mach (2019) 20: 5. https://doi.org/10.1007/s10710-018-9339-y

[3] Jose C. Reston Filho, Carolina de M. Affonso, Roberto C.L. de Oliveira, Energy price prediction multi-step ahead using hybrid model in the Brazilian market, Electric Power Systems Research, Volume 117, 2014, Pages 115-122.

[4] De Campos L.M.L. , De Oliveira, R.C.L. , Roisenberg M.Evolving artificial neural networks through l-system and evolutionary computation, International Joint Conference on Neural Networks (IJCNN), 2015, Pags. 1-9, IEEE.

[5] Oluwarotimi W.S., Grace M.A., Arun K.S., Guanglin L., Computational Intelligence Enabling the Development of Efficient Clinical Decision Support Systems: Case Study of Heart Failure.In Intelligent Data-Centric Systems, Computational Intelligence for Multimedia Big Data on the Cloud with Engineering Applications, Academic Press, 2018.

[6] Q Ahmad, A Rafiq, MA Raja, N Javed. Evolving MIMO multi-layered artificial neural networks using grammatical evolution. Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing, 1278-1285. 2019.

[7] RUSSEL, S.; NORVIG, P.; Artificial Inteligence: A Modern Approach. Prentice-Hall, Second Edition, 2013.

[8] S. Akkaradamrongrat, P. Kachamas and S. Sinthupinyo, "Text Generation for Imbalanced Text Classification," 2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE), Chonburi, Thailand, 2019, pp. 181-186.

[9] S. Siami-Namini, N. Tavakoli and A. Siami Namin, "A Comparison of ARIMA and LSTM in Forecasting Time Series," 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, 2018, pp. 1394-1401.

[10] Website of Brazilian Electrical Energy Commercialization Chamber http://www.aneel.gov.br/, March 2020.

[11] Website of Brazilain National System Operator http://www.ons.org.br/, March 2020.

[12] M. Castellani 2013, "Evolutionary generation of neural network classifiers—An empirical comparison," Neurocomput. 99, 214–229.

[13] Pengtao Li, Kaile Zhou, Xinhui Lu, Shanlin Yang,"A hybrid deep learning model for short-term PV power forecasting", Applied Energy, vol. 259,2020.

[14] Zhiyun Peng, Sui Peng, Lidan Fu, Binchun Lu, Junjie Tang, Ke Wang, Wenyuan Li, "A novel deep learning ensemble model with data denoising for short-term wind speed forecasting", Energy Conversion and Management, vol. 207, 2020.

[15] Cerjan Marin, Petricic Ana, Delimar Marko. HIRA Model for Short-Term Electricity Price Forecasting. Energies, MDPI, Volume: 12, Issue:3, Pages:586.