

Article

An Artificial Neural Network Approach to Forecast the Environmental Impact of Data Centers

Joao Ferreira ^{1,*},[†] , Gustavo Callou ²,[†], Albert Josua ³,[†], Dietmar Tutsch ⁴,[†] and Paulo Maciel ¹,[†]

¹ Informatics Center, Cidade Universitária, Federal University of Pernambuco, Recife 50740-560, Brazil; prmm@cin.ufpe.br

² Department of Computing, Federal Rural University of Pernambuco, Dom Manuel de Medeiros, Dois Irmãos, Recife 52171-900, Brazil; gustavo.callou@ufrpe.br

³ Department of Informatics, Federal University of Amazonas, Manaus 69020-120, Brazil; albert.costa@ifam.edu.br

⁴ Automation Technology, Bergische Universität Wuppertal, D-42119 Wuppertal, Germany; tutsch@uni-wuppertal.de

* Correspondence: jfsj3@cin.ufpe.br; Tel.: +55-81-2126-8430

† These authors contributed equally to this work.

Received: 24 December 2018; Accepted: 5 March 2019; Published: 14 March 2019



Abstract: Due to the high demands of new technologies such as social networks, e-commerce and cloud computing, more energy is being consumed in order to store all the data produced and provide the high availability required. Over the years, this increase in energy consumption has brought about a rise in both the environmental impacts and operational costs. Some companies have adopted the concept of a green data center, which is related to electricity consumption and CO₂ emissions, according to the utility power source adopted. In Brazil, almost 70% of electrical power is derived from clean electricity generation, whereas in China 65% of generated electricity comes from coal. In addition, the value per kWh in the US is much lower than in other countries surveyed. In the present work, we conducted an integrated evaluation of costs and CO₂ emissions of the electrical infrastructure in data centers, considering the different energy sources adopted by each country. We used a multi-layered artificial neural network, which could forecast consumption over the following months, based on the energy consumption history of the data center. All these features were supported by a tool, the applicability of which was demonstrated through a case study that computed the CO₂ emissions and operational costs of a data center using the energy mix adopted in Brazil, China, Germany and the US. China presented the highest CO₂ emissions, with 41,445 tons per year in 2014, followed by the US and Germany, with 37,177 and 35,883, respectively. Brazil, with 8459 tons, proved to be the cleanest. Additionally, this study also estimated the operational costs assuming that the same data center consumes energy as if it were in China, Germany and Brazil. China presented the highest kWh/year. Therefore, the best choice according to operational costs, considering the price of energy per kWh, is the US and the worst is China. Considering both operational costs and CO₂ emissions, Brazil would be the best option.

Keywords: energy flow model; sustainability; data center power architectures; optimization; multilayer perceptron

1. Introduction

Currently, more people than ever have access to the Internet. Social changes have transformed both the way people live and how the world works. There has been a significant increase in the growth of the mobile market and advances in cloud computing technology are generating a huge amount

of data, thereby implying unprecedented demands on energy consumption. This digital universe corresponds to 500 billion gigabytes of data [1], and since 25% of this is from on-line data, this value may therefore increase greatly over time.

Data center power consumption has increased significantly over recent years influenced by the increasing demand for storage capacity and data processing [2]. In 2013, data centers in the US consumed 91 billion kilowatt-hours of electricity [3], which is expected to continue rising. Moreover, critical elements in the performance of daily tasks, such as social networks, e-commerce and data storage, also contribute to the rise in energy consumption across these systems.

Due to the growing awareness of issues such as climate change, pollution, and environmental degradation, the scientific and industrial communities have paid increasing attention to the impact of human activity. The environmental impact caused by the production of some energy sources is massive. For example, the carbon dioxide (CO₂) released by the use of coal, petroleum, natural gas and other similar energy sources contributes significantly to global warming. According to estimates, CO₂ emissions may increase between 9% and 27% by 2030, depending on which policies are enacted [4].

The main focus of this paper is to propose an integrated strategy to evaluate the operational costs and estimate the environmental impacts (CO₂ emissions) of data centers.

In this strategy, an artificial neural network (ANN) is applied to the energy flow model (EFM) metrics, which forecasts the values in the future, based on the consumption of data center electrical architectures and by considering different energy sources. To demonstrate the applicability of the proposed strategy, a case study compared costs and environmental impacts, according to the Brazilian, Chinese German and US energy mixes adopted. Autoregressive integrated moving average (ARIMA) was used to validate perceptron predictions. The results demonstrate that RNA predictions were within the 90% confidence interval of ARIMA.

The cost used in this work was computed according to the time of operation of the data center, how much energy was consumed in this period and its financial cost. The calculation of the environmental impact considered the amount and source of energy consumed by the environment. Section 6 details these relationship.

A substantial amount of CO₂ emissions originates from the production of energy. In China, burning coal generates 65% of the electricity consumed [5], which is a very high level of non-green power generation when compared with Brazil, where the level is 3%. In Germany, 14.7% of the energy produced comes from nuclear fission [6], whereas in China this figure is only 1% [7]. Many countries now require that polluting energy sources should be replaced by cleaner alternatives, such as solar, wind or hydro plants.

Data center infrastructures entail more redundant components that consume more and more electricity, influenced by the increasing demand for storage capacity and data processing. The concept of green data center is related to electricity consumption and CO₂ emissions, which depends on the utility power source adopted. For example, in Brazil, 73% of electrical power is derived from clean electricity generation [8], whereas in USA 82.1% of generated electricity comes from petroleum, coal or gas [9].

The main contributions of this work are as follows: considering the energy mix of the data centers to estimate the emission of carbon dioxide in the atmosphere through the energy consumption of the centers mentioned above; cost evaluation, availability and sustainability for the electric infrastructures of the data centers; and use of an artificial neural network (ANN) along with the energy flow model (EFM), to predict the consumption of energy in the next few months, based on the environment history.

In the present work, the power subsystem electrical flow is represented by an energy flow model (EFM) [8,10]. Both the proposed strategy and the EFM are supported by the Mercury tool [11]. In addition to the EFM, Mercury offers support models for reliability block diagrams (RBD) [12], Markov chains and stochastic Petri nets (SPN) [13].

The paper is organized as follows: Section 2 presents studies along the same line of research, highlighting the key differences. Section 3 introduces the basic concepts of sustainability, ANN and

ARIMA. Section 4 describes the adopted methodology. Section 5 describes the EFM. Section 6 illustrates how to consider different energy mixes in the EFM. Section 7 presents the artificial neural networks in EFM. Section 8 presents a case study. Section 9 concludes the paper.

2. Related Works

Over the last few years, considerable research has been conducted into energy consumption in data centers. This section presents studies related to this research field. Some papers have used neural networks and others have adopted strategies to reduce the energy consumption of data centers. To demonstrate the importance of the present study, we analyzed the environmental impact caused by the energy consumption of data centers and have proposed neural networks to forecast the impact over the following months based on information from previous data.

Zeng [14] proposed the use of a hybrid model of ANN of back propagations with an evolutionary algorithm adapted to predict the consumption of natural gas. Zeng applied the evolutionary algorithm to find initial values of the synaptic weights and the hit values (limits) of the activation functions of the neural network to improve the prediction performance. The simulation results prove that the weight and bias optimization technique improves the accuracy of the prediction. Our study is different from Zeng because it uniquely considered the electric consumption history to establish the prediction, whereas Zeng considered multiple variables as the input of the prediction model, such as gross domestic product, population and import and export data.

Wang [15] suggested using an effective and stable model of consumption prediction based on echo state network (ESN) by using a differential evolutionary algorithm to optimize the three essential parameters of an ESN, such as scale of the reservoir (N), connectivity rate (a) and spectral radius (p). In an ESN, the concealed layers are replaced by a dynamic reservoir of neurons. Besides, the prediction model of the electric power consumption is simulated in three situations: monthly consumption in the northeast of China between the period of January 2004 and April 2009, the yearly consumption of Taiwan from 1945 to 2003 and the monthly consumption in Zhengzhou from January 2012 to February 2017. The results of the experiments show that the proposed prediction model obtains superior results to the other models compared. Different from our proposal, this one includes the electric consumption of cities and not data centers, and it does not compare the environmental impact caused by energy consumption.

Wang [16] combined a novel sparse adaboost with echo state network (ESN) and fruit fly optimization algorithm (FOA) to enhance the forecasting accuracy of mid-term electricity consumption demand in China. The sparse adaboost (adaboostsp) is designed to overcome the instability and limited generalization ability of individual ESN, and reduce the computational complexity of adaboost through its sparse ensemble strategy. According to Wang [16], the proposed adaboostsp-ESN approach presents high performance in two IEC forecasting applications in China. The ensemble computation cost is slashed by the well-designed sparse adaboost structure. Time lag effects of influence factors on IEC are explored. Different from our proposal, this work does not relate energy consumption of data centers with environmental impact.

Yaoyao [17] presented a prediction method of consumption that combines the LASSO regressions with the Quantile Regression Neural Networks (LASSO-QRNN). The LASSO regression is used to produce high-quality attributes and to reduce the data dimensionality effectively. The suggested method is compared to state-of-the-art methods, including Radial Bias Function (RBF), Back-propagation (BP), QR and NLQR. Case studies are carried out using electricity consumption data from the Guangdong, China and California, USA. From the analysis of the results, the suggested method shows superior accuracy to the other methods and reduced data dimensionality. This work also refers neither to data centers nor issues related to sustainability.

The next works are related to the issues of sustainability and energy consumption in data centers, however they also do not fill the gaps mentioned above.

The goal is to provide an integrated workload management system for data centers that takes advantage of the efficiency gains possible by shifting demand in a way that exploits time variations in electricity price, the availability of renewable energy, and the efficiency of cooling. However, the focus is on the low level of the environment (load schedulers). Our proposed method enables conducting a global analysis of the energy consumption.

Reddy [18] presented various metrics relating to a data center and classification based on the different core dimensions of data center operations. They defined the core dimensions of data center operations as follows: energy efficiency, cooling, greenness, performance, thermal and air management, network, security, storage, and financial impact. They presented a taxonomy of state-of-the-art metrics used in the data center industry, which is useful for the researchers and practitioners working on monitoring and improving the energy efficiency of data centers. Our proposal method uses one of the metrics pointed out by Reddy (efficiency of energy use) and proposes a new metric, considering an integrated analysis of the energy cost and CO₂ emissions.

Dandres [19] affirmed that cloud computing technology enables real-time load migration to a data center in the region where the greenhouse gas (GHG) emissions per kWh are the lowest. He proposed a novel approach to minimize GHG emissions cloud computing relying on distributed data centers. In the case of the GJG, the GHG emission factor for the electric grid is the same as that of the electric grid. Results show that load migrations make it possible to minimize marginal GHG emissions from the cloud computing service. Our proposed method may be used as support for decision making regarding the need for service migration, thus complementing this research.

3. Basic Concepts

This section presents the basic concepts for a better understanding of this work. It starts by reviewing sustainability, followed by artificial neural networks and ARIMA model.

3.1. Sustainability

Sustainability is a term used to define human actions and activities that aim to meet the present needs of humans, without compromising the future of the next generations. Sustainability is directly related to economic and material development without harming the environment, using natural resources intelligently to preserve the future [20]. Following such principles, humanity can guarantee sustainable development.

Electricity generation may affect the environment in a number of ways. Electrical systems cause impacts on ecosystem functions including nutrient cycling, water distribution, soil dynamics, natural population dynamics and climate change [21]. Global climate change has been increasing and reaching the entire planet due to the gasses emitted that rapidly disperse into the atmosphere. Therefore, there is no difference related to the location where such gasses are emitted, causing global changes in the circulation of water and air [20].

3.2. Artificial Neural Networks

ANNs are computer systems that aim to work in a similar manner to the human brain. ANNs should be able to learn, and then make decisions based on the knowledge learned. Neural networks represent an attempt to overcome the limitations of conventional computing using certain advantages of the human brain, such as a high degree of parallelism, fault tolerance, robustness, adaptability and self-organization [22].

McCulloch and Pitts [23] presented a model combining neurophysiology and mathematical logic, using the property of the all-or-none firing of neurons, which is a binary element of discrete time. They demonstrated how excitation, inhibition, and threshold may be used to construct a wide variety of “neurons”. This represents the first model that tied the study of neural networks directly to the idea of computing in its modern sense.

The output of the neuron is the result of the activation function, responsible for firing the threshold of the neuron, in the sum of the input signals, weighted by their respective synaptic weights. This process may be described by the following equations:

$$\mu_k \sum_{i=1}^m X_i W_{ki}, \quad (1)$$

$$Y_k = f(\mu_k + b_k), \quad (2)$$

where x_1, x_2, \dots, x_m are the input signals; $W_{k1}, W_{k2}, \dots, W_{km}$ are the synaptic weights of neuron k ; μ_k is the result of the sum function; b_k is the bias, which has the effect of increasing or decreasing polarization in the activation function; $f(\cdot)$ is the activation function of the neuron; and Y is the output of neuron k .

Perceptron

The perceptron, created by Frank Rosenblatt, is the simplest configuration form of an artificial neural network, since it consists of a single neural layer and a single neuron. Through the training process, the algorithm learns to classify inputs into two different groups. The perceptron is only capable of solving problems that are linearly separable. However, nonlinearities are inherent to most real situations and problems, therefore it is necessary to use structures with non-linear characteristics in order to solve more complex problems [24].

A multi-layer perceptron (MLP) consists of three or more layers (an input layer, an output layer, and N hidden layers). The output layer receives the stimulus from the middle layer and constructs the response. MLPs use a backpropagation algorithm, standard solution for any supervised learning [24]. This algorithm updates not only the weights of the last layer of neurons (output), but also the weights of the intermediate layers. Therefore, the perceptron may be enlarged into layers, allowing more complex problems to be solved.

The following is a list of advantages of using the perceptron:

- Works fine in the case of incomplete information
- Does not require knowledge of the algorithm solving the problem (automatic learning)
- Processes information in a highly parallel way
- Can generalize to unknown cases
- Resistant to partial damage
- Performs associative memory (associative—similar to working memory in humans) as opposed to addressable memory (typical for classical computers).

3.3. ARIMA

In time-series analysis, an integrated autoregressive moving average model (ARIMA) is the generalization of a self-regression model of moving averages (ARMA) [25]. Both models are fitted to the time-series data to better understand the data or to forecast future points in the series. ARIMA models are applied in some cases where the data present evidence of non-stationarity, in which an initial differentiation step (corresponding to the “integrated” part of the model) may be applied one or more times to eliminate non-stationarity [25].

The autoregressive (AR) part of the ARIMA model indicates that the evolutionary variable of interest is returned in its own lagged values. The moving average (MA) part indicates that the regression error is actually a linear combination of the error terms, whose values occurred contemporaneously and at various times in the past. The integrated part (I) indicates that the data values have been replaced by the difference between their values and the previous values and this differentiating process may have been performed more than once. The purpose of each of these characteristics is to fit the data model in the best possible way [26].

Non-seasonal ARIMA models are generally denoted ARIMA (p, d, q), where the parameters p, d and q are nonnegative integers. “p” is the order (number of lags) of the autoregressive model, “d” is the degree of differentiation (the number of times past values were subtracted from the data) and “q” is the order of the moving average model. For more information regarding ARIMA, please see [25,26].

4. Methodology

Figure 1 depicts an overview of the proposed methodology for evaluating sustainability, and forecasting cost and energy issues in data center infrastructures. The first step of the methodology is concerned with understanding the system, its components, their interfaces and interactions. This phase should also provide (as a product) the set of metrics that need to be evaluated.

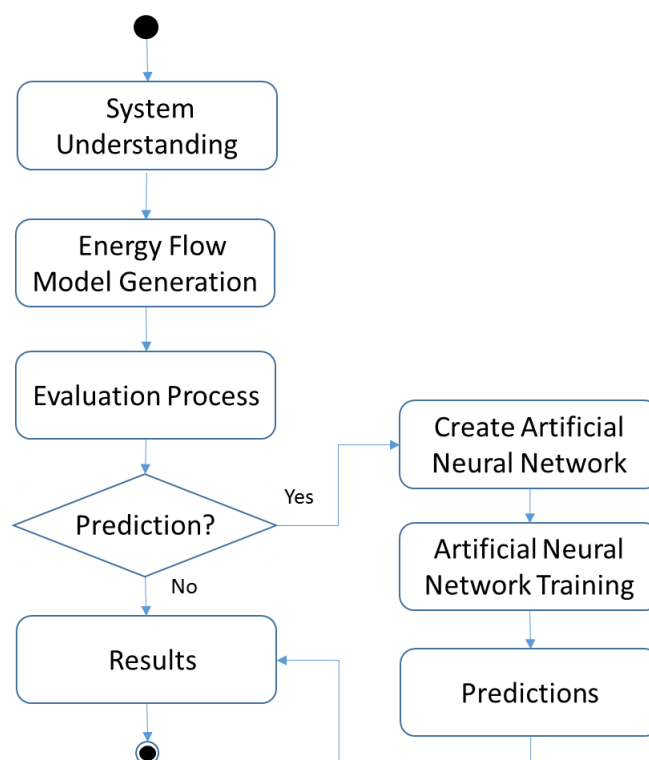


Figure 1. Methodology.

The next broad phase aims to create the high-level models that represent the data center architecture. The high-level models allow data center designers to specify power, cooling and IT systems following the standard adopted by engineers. These models may be converted into dependability models (e.g., fault tree, continuous time Markov chains (CTMC), SPN or RBD). It is important to state that sub-models may be generated to mitigate the complexity of the final model. The evaluation of each sub-model provides the results of the system.

The next broad phase aims to create the energy flow model. These models allow the integrated evaluation of dependability, cost and sustainability. Additionally, the EFM verifies that the energy flow does not exceed the maximum power capacity that each component is able to provide (considering electrical devices) or extract (assuming cooling equipment).

An evaluation process is directly conducted to provide the estimate results (e.g., operational costs, availability, downtime, exergy consumption, and CO₂ emissions). If the designer has no energy history for the environment, the results are displayed and the data may be analyzed. Otherwise, the forecasting option will be chosen and the data center designer should inform the energy history.

The next stage involves the steps for forecasting. The designer may choose the number of neurons in the input, output and hidden layers. The neural network is trained, considering 70% of the values

for training and 30% for validation. A forecast for the next 12 months is made and the results may be analyzed by the designer. Afterwards, the achieved results can be compared to a well-known prediction method called ARIMA to conduct the validation. More details about the validation process is presented in Section 8.3.

5. Energy Flow Model (EFM)

The EFM represents the energy flow between the components of a cooling or power architecture, considering the respective efficiency and energy that each component is able to support (cooling) or provide (power). The EFM is represented by a directed acyclic graph in which components of the architecture are modeled as vertices and the respective connections correspond to edges [8].

The following defines the EFM: $G = (N, A, w, f_d, f_c, f_p, f_\eta)$, where:

- $N = N_s \cup N_i \cup N_t$ represents the set of nodes (i.e., the components), in which N_s is the set of source nodes, N_t is the set of target nodes and N_i denotes the set of internal nodes, $N_s \cap N_i = N_s \cap N_t = N_i \cap N_t = \emptyset$.
- $A \subseteq (N_s \times N_i) \cup (N_i \times N_t) \cup (N_i \times N_i) = \{(a,b) \mid a \neq b\}$ denotes the set of edges (i.e., the component connections).
- $w : A \rightarrow \mathbf{R}^+$ is a function that assigns weights to the edges (the value assigned to the edge (j, k) is adopted for distributing the energy assigned to the node, j , to the node, k , according to the ratio, $w(j,k)/\sum_{i \in j^\bullet} w(j, i)$, where j^\bullet is the set of output nodes of j).
- $f_d : N \rightarrow \begin{cases} \mathbf{R}^+ & \text{if } n \in N_s \cup N_t, \\ 0 & \text{otherwise;} \end{cases}$ is a function that assigns to each node the heat to be extracted (considering cooling models) or the energy to be supplied (regarding power models).
- $f_c : N \rightarrow \begin{cases} 0 & \text{if } n \in N_s \cup N_t, \\ \mathbf{R}^+ & \text{otherwise;} \end{cases}$ is a function that assigns each node with the respective maximum energy capacity.
- $f_p : N \rightarrow \begin{cases} 0 & \text{if } n \in N_s \cup N_t, \\ \mathbf{R}^+ & \text{otherwise;} \end{cases}$ is a function that assigns each node (a node represents a component) with its retail price.
- $f_\eta : N \rightarrow \begin{cases} 1 & \text{if } n \in N_s \cup N_t, \\ 0 \leq k \leq 1, k \in \mathbf{R} & \text{otherwise;} \end{cases}$ is a function that assigns each node with the energetic efficiency.

Mercury engine provides support for EFM and an example is depicted in Figure 2. The rounded rectangles are the type of equipment, and the labels represent each item. The edges have weights that are used to direct the energy that flows through the components. For the sake of simplicity, the graphical representation of EFM hides the default weight 1.

TargetPoint1 and SourcePoint1 represent the IT power demanded and the power supply, respectively. The weights present on the edges, (0.7 and 0.3) are used to direct the energy flow through the components. In other words, UPS1 (Uninterrupted Power Supply) is responsible for providing 70% and UPS2 for 30% of the energy demanded by the IT system.

The EFM is employed to compute the overall energy required to provide the necessary energy at the target point. Assuming 100 kW as the demanded energy for the data center computer room, this value is thus associated to the TargetPoint1. Considering the efficiency of STS1 (Static Transfer Switch) is 95%, the electrical power that the STS component receives is 105.26 kW.

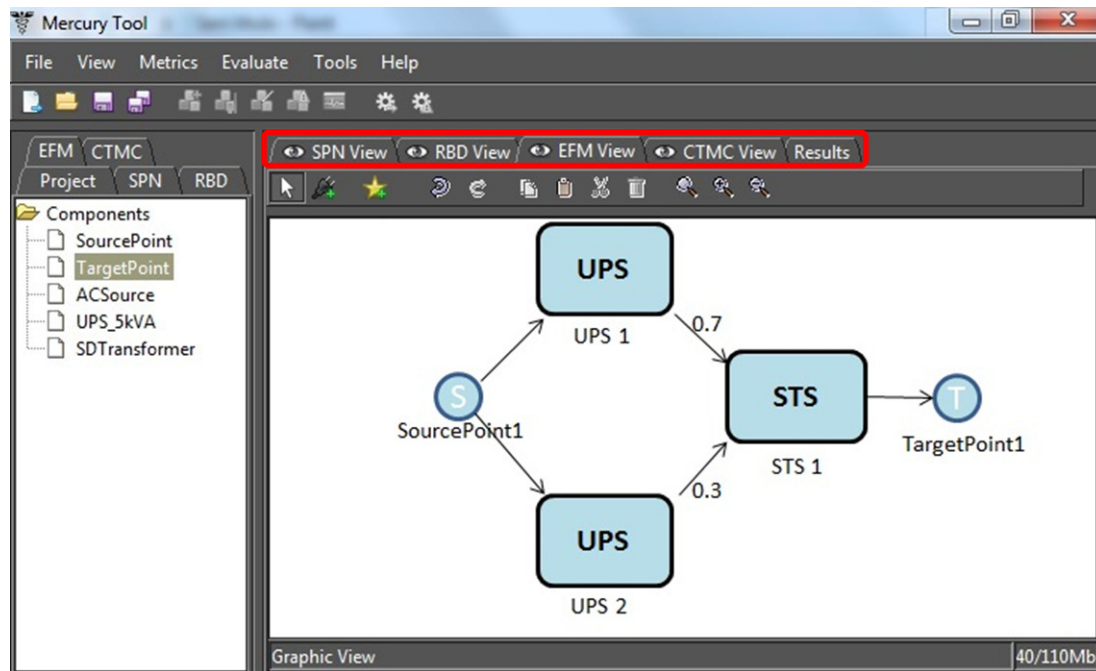


Figure 2. EFM Example in Mercury Tool.

A similar strategy is adopted for components UPS1 and UPS2, however, now, dividing the flow accord to the associated edge weights, 70% (73.68 kW) for UPS1 and 30% (31.27 kW) for UPS2. Thus, the UPS1 needs 77.55 kW, considering 95% efficiency, and UPS2 needs 34.74, considering 90% electrical efficiency. The *Source Point 1* accumulates the total flow (112.29 kW).

It is important to stress that the edge weights are defined by the model designer, and there is no guarantee that designers allocate the best values for the distribution, and the outcome may increase power consumption.

For more details about EFM, the reader is redirected to [27].

6. Considering Energy Mix in the EFM

Nowadays, there are different energy sources, such as solar, geothermal, thermoelectric, biomass, hydrogen fuel, tidal, ethanol, blue or melanin. However, in this study, we only considered the most frequently used: wind, coal, hydroelectric, nuclear and oil. Furthermore, this study also considered the amount of CO_2 emissions according to the energy source used.

The inclusion of the energy mix in the EFM is proposed for a more detailed analysis of the operational costs and the estimation of CO_2 emissions in the atmosphere, according to the energy consumed. By considering the energy mix in the EFM, it is possible to represent more details of the electrical infrastructures of a real-world data center. Data center designers may consider more than one energy source, which represents the energy mix of the utility. Additionally, this EFM extension allows operational costs and the environmental impacts of the electricity consumption to be calculated, as well as the CO_2 emissions from the adopted energy mix.

Table 1 presents the relation between the source used to produce the energy and the amount of CO_2 emitted by each. These values were obtained from [8]. This new feature is supported by the Mercury tool, in which the designer only needs to inform the amount of energy consumed by each source and the corresponding CO_2 emissions in the atmosphere will be calculated.

Table 1. Material used in energy generation x CO₂ emissions.

Energy Source	CO ₂ (g/kWh)
Wind	10
Coal	950
Hydroelectric	20
Nuclear	150
Oil	510

To compute the amount of CO₂ emissions, the percentage of each energy source is multiplied by its factor of aggression (see Table 1). This factor shows the amount of CO₂ that is provided by each energy source. This process is described by the following equation:

$$CO_2Emissions = \sum_{i=1}^n (P_i \times F_i) \quad (3)$$

where i is the energy source (wind, coal, hydroelectric, nuclear or oil), P_i is the percentage of energy source and F_i is the factor of aggression to the environment.

In this study, the operational cost was calculated based on the data center operation period, the energy consumed, the cost of energy and the data center availability. Equation (4) denotes the operational cost:

$$OperationalCost = \left(\sum_{i=1}^n (P_{Input(i)} \times C_{Energy(i)}) \right) \times T \times (A + \alpha(1 - A)) \quad (4)$$

where i is the energy source (wind, coal, hydroelectric, nuclear or oil), $P_{Input(i)}$ is the percentage of power supply for the source i , $C_{Energy(i)}$ is the energy cost of power energy unit, T is the considered time period, and A is the system availability. α is the energy percentage that continues to be consumed when the system fails.

7. Applying ANNs to the EFM

Several forecasting methods have been developed over the last decades. Various methods, e.g., regression models, neural networks, fuzzy logic, expert systems, and statistical learning algorithms, are commonly used for forecasting [28]. The development, improvement, and investigation of appropriate tools have led to the development of more accurate forecasting techniques. In this work, we integrated an ANN into the energy flow model.

An EFM with artificial neural networks (ANN) could expand the horizons of the modeling strategy previously adopted in [8,10,27]. In addition to computing exergy consumption, operational costs, input power and PUE, this new approach enables the values of these metrics to be forecast, based on a historical series. The multi-layer perceptron (MLP) [29] is adopted in this method. The MLP basically consists of a layer of nodes (input sources), one or more layers of hidden processor or computational nodes (neurons) and an output layer also composed of computational nodes. The layer composed of hidden neurons is called the hidden layer, as no access to the input or output of this layer may be reached.

In Figure 3, it is possible to identify three basic elements of the neural model:

- A set of synapses, each characterized by a weigh. Specifically, a signal X_j in the input of the synapse j connected to the neuron k is multiplied by the synaptic weigh W_{kj} . It is important to notice the way how the indexes of the synaptic weigh W_{kj} are written. The first index refers to the neuron under analysis and the second one refers to the terminal input of the synapse which the weight refers to.
- An adder to add two input signals, weighed by the respective synapses of the neuron. These operations were implemented in the Mercury tool and constitute a linear combiner.

- An activation function to limit the output range of a neuron. Typically, the normalized interval of the output range of a neuron is written as a closed unitary interval [0, 1] or alternatively [−1, 1].

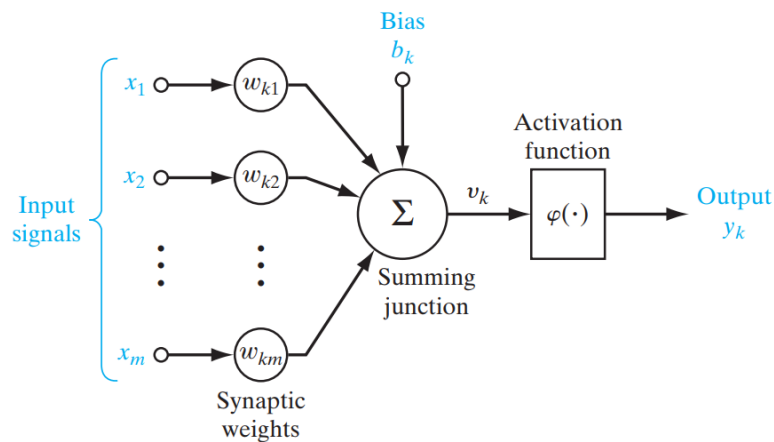


Figure 3. Nonlinear model of a neuron k [30].

The bias, represented by b_k , has the effect of increasing or reducing the input value of the activation function, depending if it is positive or negative, respectively.

In terms of mathematical concepts, it is possible to describe a neuron k by writing the following pair of equations:

$$\mu_k = \sum_{j=1}^m W_{kj} X_j, \tag{5}$$

$$Y_k = \varphi(\mu_k + b_k), \tag{6}$$

where x_1, x_2, \dots, x_m are the input signals; $W_{k1}, W_{k2}, \dots, W_{km}$ are the synaptic weights of the neuron k; μ_k (not shown in Figure 3) is the output of the linear combiner due to input signals; b_k is the bias, which has the effect of increasing or reducing the polarization in the activation function; $\varphi(\cdot)$ is the activation function of the neuron; and Y_k is the output signal of the neuron. The use of the bias b_k has the function of applying some transformation to the output μ_k of the linear combiner, as shown in:

$$v_k = (\mu_k + b_k), \tag{7}$$

The activation function, represented by $\varphi(v)$, defines the output of a neuron in terms of the induced local field v . There are three basic types of activation functions: threshold function, linear part function and sigmoid function [30]. In this study, a sigmoidal function was used, being the most common in the artificial neural network construction. Its graph is similar to the letter s. It is defined as a strictly rising function, which displays an adequate balance between linear and non-linear behavior. The sigmoidal function implemented in the Mercury was the logistics, defined by Equation (8).

$$\varphi(v) = \frac{1}{1 + \exp(-\alpha v)} \tag{8}$$

where α is the inclination parameter of the sigmoidal function. Figure 4 represents the sigmoidal function with a variation of α .

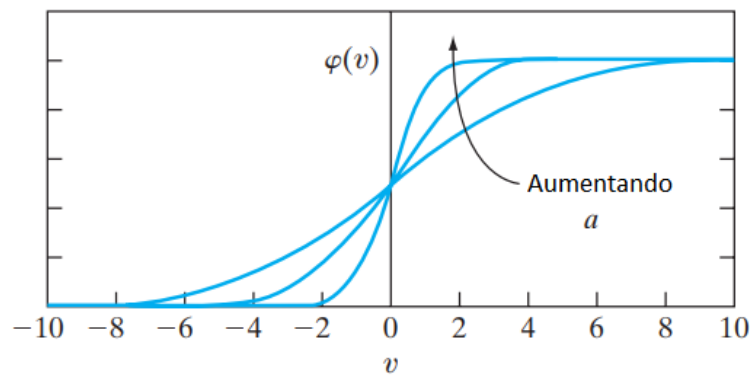


Figure 4. Sigmoidal function for the inclination parameter α variable [30].

The number of neurons, layers, degree of connectivity and the presence or absence of retro propagation connections define the topology of an artificial neural network [29]. This step is very important because it directly defines the processing power of the network. In light of modern knowledge, it is not possible to determine the exact number of neurons and the number of layers for general problems. It is imperative that the network does not suffer overfitting or underfitting, depending on the excess or lack of layers/neurons [29]. Thus, the number of neurons and layers must be extensively tested, with several configurations of neural networks, containing the same training tables. The final values used are defined in the validation phase. For example, a network is created and trained with a set of parameters. If the results are good, it is used as the final model, otherwise the network is trained again with other values.

ANN was implemented in the Mercury tool and is composed of five phases: load and normalize data, create an ANN, ANN training, forecasting and graph. Figure 5 presents the user view, where it is possible to create, train, forecast and visualize ANN in the Mercury tool.

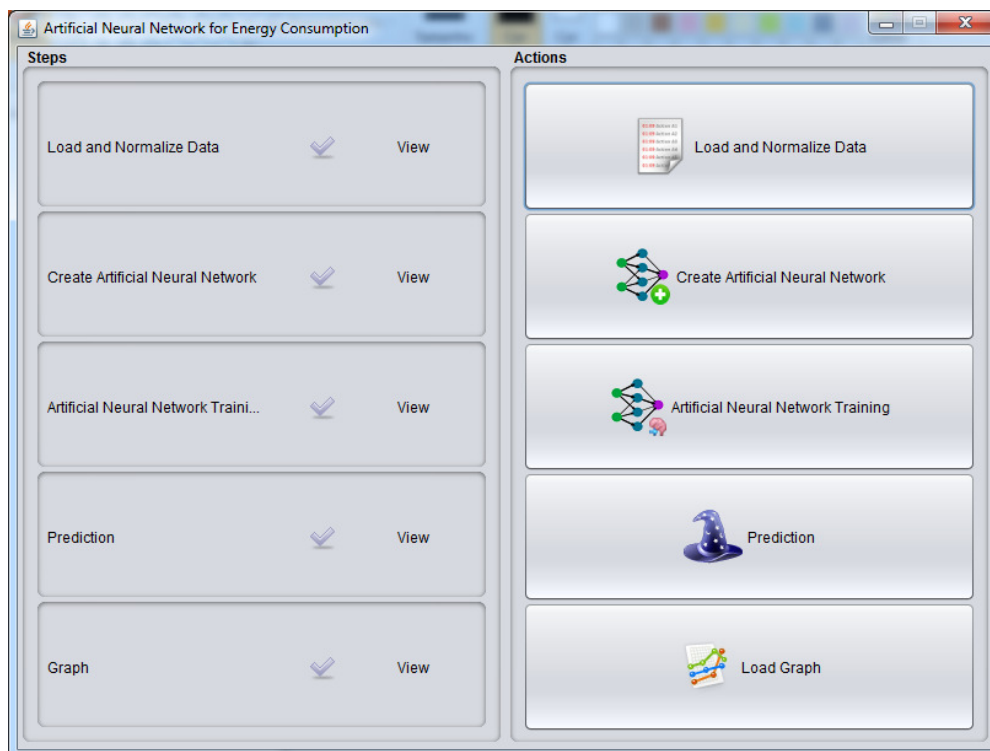


Figure 5. ANN view in Mercury Tool.

Load and normalize data: Mercury has been configured to accept spreadsheets in *odt*, *xls* and *csv* formats, with three columns (year, month and power consumption). Thus, users may upload a file with the monthly levels of power consumption from the previous years of a data center. These data are read and stored by the engine to use during the following phases.

Create an artificial neural network: In this phase, the basic parameters for creating the artificial neural network are set (e.g., number of neurons in the input layer, number of neurons in the first hidden layer and the number of neurons in the output layer). Empirical testing with the MLP backpropagation neural network [31] does not demonstrate a significant advantage in the use of two hidden layers rather than one for small problems. Therefore, most problems consider only one hidden layer.

ANN training: The most important property of neural networks is the ability to learn in their own environment, and thereby to improve their performance. This is done through an iterative process of adjustments applied to their weights, and training. The backpropagation training algorithm is the most popular algorithm for training multi-layer ANNs. The algorithm consists of two steps: propagation and backpropagation. In the first step, an input vector is applied to the input layer and its effect propagates across the network producing a set of outputs. The response obtained by the network is subtracted from the desired response to produce an error signal. The second step propagates this error signal in the opposite direction to the synaptic connections, adjusting them in order to approximate to the network outputs. Additionally, using the EFM with RNA, it is possible to set the training stop criteria for a specific error rate or a fixed number of iterations.

Prediction: This option produces forecasts related to the energy consumption of the environment over the next twelve months. At the end of the forecasts, the mean absolute percentage error is displayed.

Graph: This button graphically displays a comparison between the measurements, with a blue line for the actual data and a red line for the expected monthly consumption values.

8. Case Study

This case study aimed to estimate the operational costs and CO₂ emissions of a data center in the US, taking into account the energy mix adopted over a period of 15 years. Moreover, this paper provides the EFM models for a Tier III data center, besides estimating the environmental impact of the energy consumption of this data center using locations in different countries (with different energy mixes), namely Germany, China and Brazil. A neural network (multi-layer perceptron (MLP)) was used to forecast the energy consumption for the following 12 months, according to the consumption history of the data center. In the context, this work applied ANNs to approach the functions and to forecast future energy consumption based on the previous history.

8.1. Models

A data center infrastructure may be classified based on the redundancy features and fault tolerance [32]. This classification provides metrics for data center designers that identify the performance of the electricity and strategies adopted. This subsection presents an analysis of the proposed models to represent configurations of the Tier III data center, since, quantitatively, this is more representative of the world.

Data Center Tier III (Simultaneous Maintenance and Operation)

A Tier III data center does not require shutdowns for equipment replacement or maintenance. Figure 6 depicts an example of data center Tier III configuration. Each power component may be shutdown for maintenance without impacting the IT system operation. Similarly, a redundant cooling subsystem is also provided. These data centers are susceptible to downtime for planned activities and accidental causes. Planned maintenance activities may be carried out using the redundant components and capabilities of the reference distribution to ensure the safe operation of the remaining components.

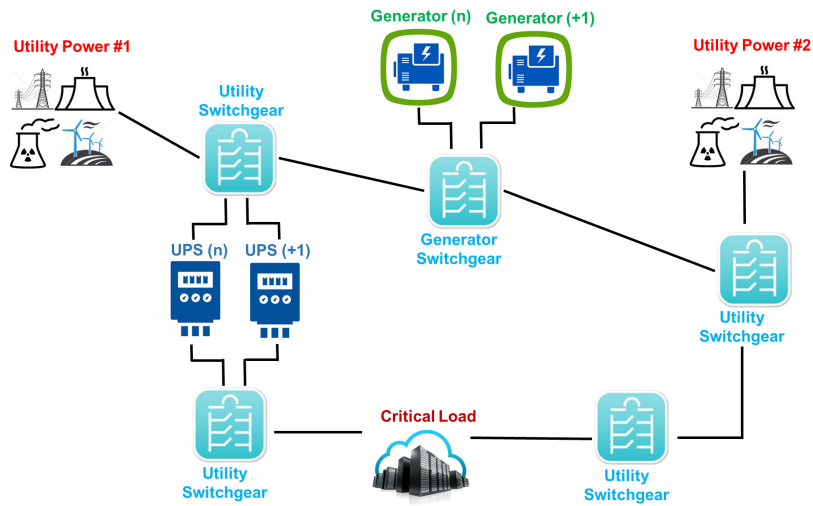


Figure 6. Tier III power system from utility to IT equipment.

Figure 7 represents the complete RBD model of Tier III. There are two redundant electric flow paths. One is composed of Utility Power 1 (UP1), Generator System and Subsystem X (Serial arrangement of ATS1, UPS System, ATS2, SDT1, Subpanel1 and JunctionBox1). The other path consists of a Utility Power 2 (UP 2) and Subsystem Y (Serial arrangement of SDT2, Subpanel2 and JunctionBox2). Both routes provide power to subsystem P (set of line filters).

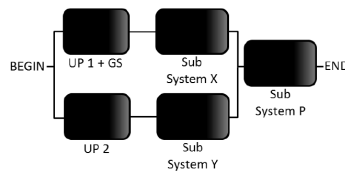


Figure 7. RBD model of Tier III.

Once availability was obtained (Figure 7), the EFM model could be analyzed to provide cost, operational exergy and CO₂ emissions, as well as to ensure that the power restrictions of each device are respected. Figure 8 presents the EFM model adopted for Tier III.

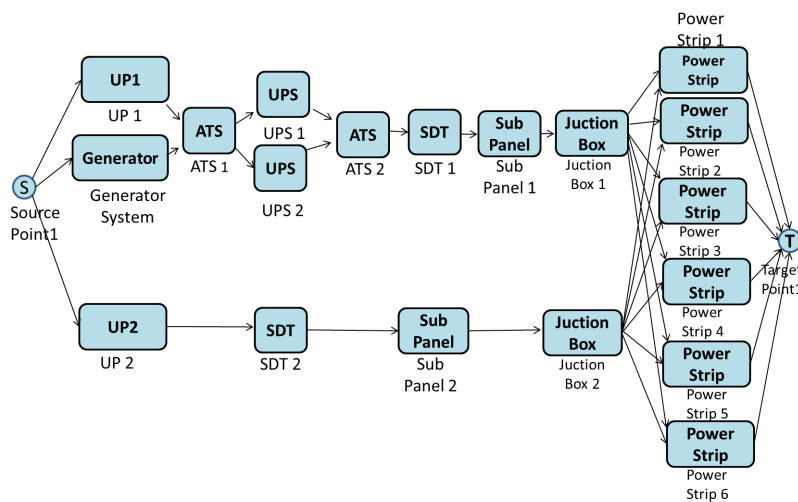


Figure 8. EFM model of Tier III.

8.2. Energy Mixes, Energy Cost and CO₂ Emissions

The energy consumption levels were obtained from [5]. Figure 9 depicts the relationship between the type of material used (in percentage) for power generation in Brazil, China, Germany and the US. We considered the five types of energy sources most frequently used in the world: wind, coal, hydroelectric, nuclear and oil (the “others” bar represents a combination of all the energy sources adopted in each country).

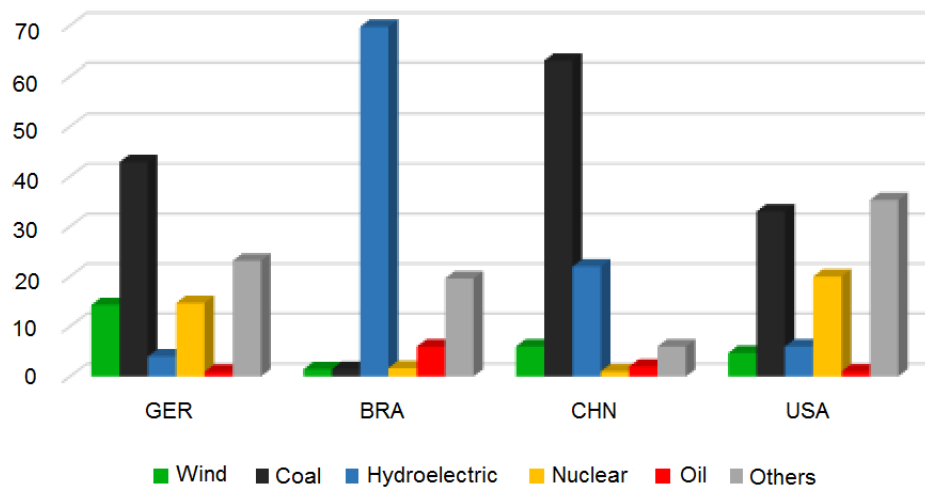


Figure 9. Germany, Brazil, China and USA energy mixes.

Table 2 presents the energy mixes used to estimate CO₂ emissions and cost, obtained from [5–7,33].

Table 2. Energy mixes, Energy cost and CO₂ emissions.

Energy Source	GER	BRA	CHN	USA	CO ₂ (g/kWh)
Cost kWh (USD)	0.25	0.18	0.43	0.12	-
Wind (%)	14.3	1.44	6	4.7	10
Coal (%)	42.9	1.5	63	33	950
Hydroelectric (%)	4	69.76	22	6	20
Nuclear (%)	14.7	1.68	1	20	150
Oil (%)	0.94	6	2	1	510
Others (%)	23.16	19.62	6	35.3	-

Figure 10 depicts a comparison of CO₂ emissions according to the energy mix adopted by each country (see Equation (3)), considering the same demand per year. China presented the highest CO₂ emissions, with 41,445 tons per year in 2014, followed by the US and Germany, with 37,177 and 35,883, respectively. Brazil, with 8459 tons, proved to be the cleanest. Due to the number of rivers and topology, Brazil is outstanding in its generation of clean energy, which may represent an interesting option for building a data center when considering only CO₂ emissions.

A great majority of people are still concerned however about the levels of CO₂ emissions produced by the US, which continue to be high despite these concerns. Currently, the US government shows little concern regarding such emissions, and a large increase in these levels was thus forecasted.

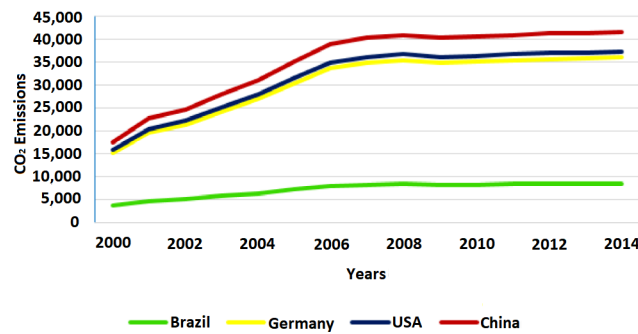


Figure 10. CO₂ emissions for one year in tons.

Figure 11 illustrates the operational costs (in USD, according to Equation (4)) of a data center during a period of one year. Additionally, this study also estimated the operational costs assuming that the same data center consumed energy as if it were in China, Germany and Brazil. According to these results, China presented the highest CO₂ emissions levels, as well as the highest kWh/year. Therefore, considering only power consumption in China, the corresponding cost was 3.5 times higher than if the data center were located in the US. Thus, the best choice according to operational costs, considering the price of energy per kWh, is the US and the worst is China. Considering both operational costs and CO₂ emissions, Brazil would be the best option.

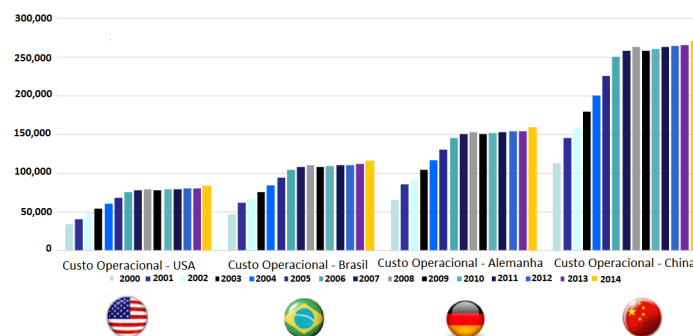


Figure 11. Data center operational cost by years: US, Brazil, Germany and China.

8.3. ANN Forecast

The main goal of this subsection is to present the forecasts made using the Mercury tool with the multi-layer perceptron (MLP), and the R tool using autoregressive integrated moving average (ARIMA). To accomplish this, an example is presented to demonstrate the applicability of the proposed method with the energy consumption history of a data center between 2000 and 2014.

Analyzing the energy consumption, we identified a significant movement trend for the data, which denotes a non-stationary series. Thus, we applied the non-parametric correlation test (spearman coefficient [25]) to confirm this fact. The autocorrelation option of the R tool was used and the filters chosen were the ARIMA (3,2,2), where 3 represents the size of the regressive factor, 2 a quantity of difference for the objective and 2 the moving average. Figure 12 presents the energy consumption of the data center from 2000 to 2014 and provides a forecast for the next 12 months, highlighted in blue with the confidence interval in gray.

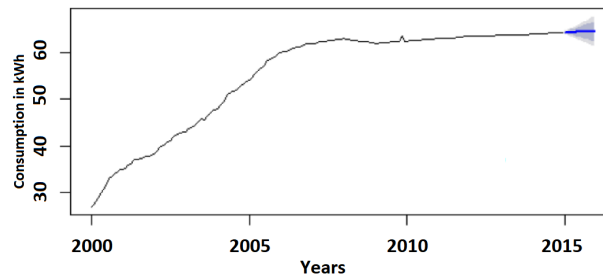


Figure 12. Forecast from ARIMA (3,2,2).

Figure 13 depicts a comparison between the real consumption and that forecast by the learning of ANN. In this case, ANN performed a learning in 100,000 interactions, reaching an error of 1.58×10^{-4} . This achieved error was verified through the graph presented in the figure, which corresponds to a very small difference and, therefore, it may be considered that the ANN demonstrated a good learning.

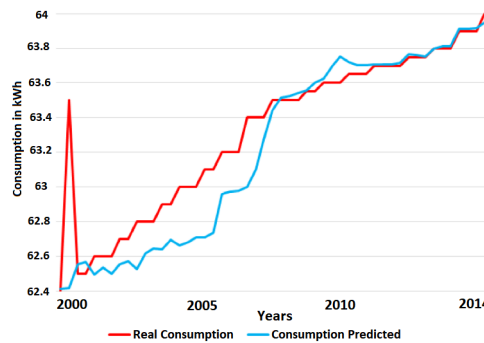


Figure 13. Real consumption \times consumption predicted.

With the aforementioned energy consumption values, we forecasted the energy consumption for the coming months using the MLP (Figure 14). This MLP was configured to use 70% of the measurements for training and 30% for testing. The MLP was configured with an input layer, a hidden layer and an output layer.

Considering a 90% confidence interval for ARIMA, the energy consumption might decrease by 1.32% or increase by 2.21%. Using the standard mean error in MLP, the changes in the energy consumption would be between a decrease of 1.25% and an increase of 1.04%. Therefore, we estimated an increase in CO₂ emissions, which might cause health problems for the world population in a few decades. Table 3 presents the ARIMA and MLP forecasts, for the following 12 months starting from January, 2015.

Table 3. ARIMA and MLP predictions for next 12 months (2015).

	ARIMA		MLP	
	Lower	Upper	Lower	Upper
January	64.07	64.59	63.63	63.77
February	63.95	64.75	63.23	63.37
March	63.82	64.92	63.43	63.57
April	63.69	65.10	63.20	63.34
May	63.56	65.28	63.20	63.34
June	63.42	65.47	63.33	63.47
July	63.27	65.66	63.03	63.17
August	63.12	65.85	63.53	63.67
September	62.97	66.05	63.13	63.27
October	62.81	66.26	63.53	63.67
November	62.65	66.47	63.23	63.37
December	62.49	66.68	63.83	63.97

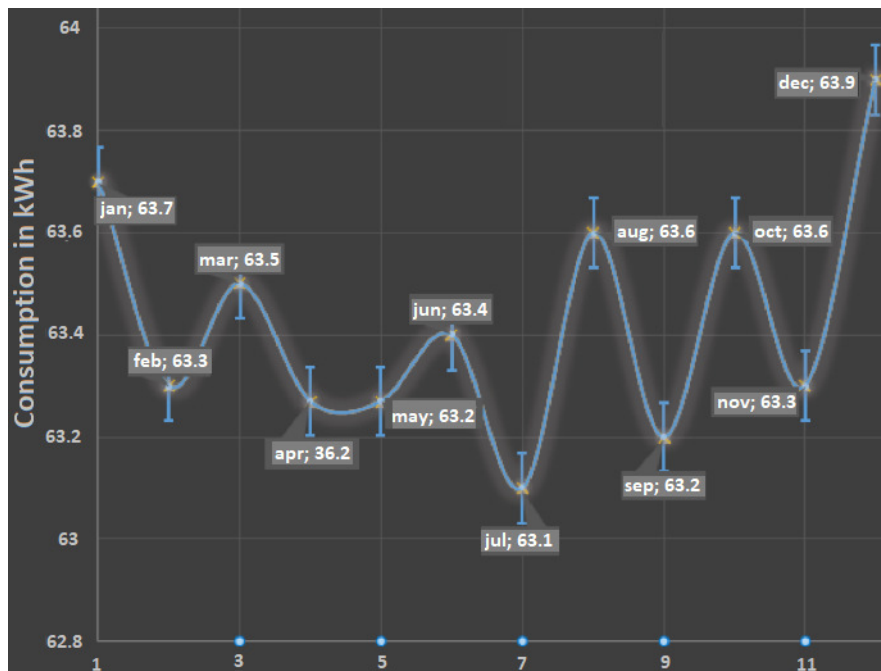


Figure 14. Predictions to 12 months using MLP.

8.4. Considerations

An artificial neural network (MLP) was used to estimate the emissions of carbon dioxide emitted into the atmosphere through the energy consumption of the data centers. This network was validated through a comparison with ARIMA. The Tier III modeling was used and the variation of the energy sources considered. Finally, a comparison with the energy mix of Brazil, Germany, China and the US was made.

9. Conclusions

This paper proposes an extension to the energy flow model (EFM) that may consider different energy sources in the system under analysis. This new feature on the EFM, in addition to estimating energy consumption, cost and the environmental impact, is also able to estimate consumption over the coming months, based on an artificial neural network. The formal definition of EFM and its extension are presented.

We conducted experiments with this strategy and achieved interesting results. For instance, we estimated CO₂ emissions and data center energy consumption considering the variation of the generating source, as well as forecasting the energy consumption over the next months.

A comparative study of data center power architectures, considering the operational costs and CO₂ emissions associated with Brazilian, Chinese, German and US energy mixes was performed, including the proposal of formal models for a Tier III data center. In addition, this approach enabled the use of an artificial neural network to forecast the values of these metrics. We evaluated the power consumption of data centers located in the US over the past 15 years, considering the energy mix (wind, coal, hydroelectric, nuclear and oil). The results reveal that China's energy mix presents the highest levels of CO₂ emissions and Brazil the lowest. The US presented the lowest operating costs per kWh/year, while China presented the highest. Furthermore, we adopted the MLP to forecast energy consumption over the next 12 months and the increase was around 1.25–1.04%. These features are available for academic use through the Mercury tool.

Thus, this study demonstrated that our proposal is effective and is useful for academic use worldwide, since it provides a new integrated strategy for evaluating the electrical infrastructure of a

data center, considering energy consumption and operational costs, besides having the possibility of forecasting consumption over the coming months.

Main Contributions and Future Work

The following contributions from this work can be highlighted:

- Global benefits by reducing environmental impact through reduced energy consumption
- Energy efficient architectures
- Accurate modeling of the electrical infrastructure of data centers that use variation of energy sources
- Possibility of predictions based on artificial neural network for cost and environmental impact of electric flow models

The following items summarize some possibilities for future work, allowing the continuity of this research:

- **Consider the LCA of the equipment:** The sustainability impact was estimated considering the exegetical consumption during the operational phase of the data center. One possible extension is to consider the impact of sustainability throughout the cycle (life cycle assessment (LCA)) of the equipment.
- **Consider cooling infrastructures:** This study considered only the components of the electrical infrastructure; it would be interesting to consider the cooling infrastructure, responsible for almost 50% of data centers' energy consumption.

Author Contributions: J.F. conceived of the presented idea, developed the theory, implemented the algorithms, proposed the formal models and performed the computations. G.C. and A.J. verified the analytical models and algorithms, and revised the paper. P.M. encouraged J.F. to investigate maximum and minimum flow and to propose a new solution to data centers' electrical power. P.M. and D.T. supervised and revised the findings of this work. All authors discussed the results and contributed to the final manuscript.

Funding: This study was financed in part by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Ciência e Tecnologia de PE (FACEPE).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hallahan, R. Technical information from Cummins Power Generation. In *Data Center Design Decisions and Their Impact on Power System Infrastructure*; Power Generator: Plymouth, MN, USA, 2011. Available online: <https://power.cummins.com/system/files/literature/brochures/PT-9020-Data-Ctr-Design-Decisions.pdf> (accessed on 15 July 2017).
2. Environmental Protection Agency. Report to Congress on Server and Data Center Energy Efficiency. Available online: http://www.energystar.gov/ia/partners/prod_development/downloads/EPA_Datacenter_Report_Congress_Final1.pdf (accessed on 23 May 2016).
3. Delforge, P. America's Data Centers Consuming and Wasting—Growing Amounts of Energy. Available online: <http://switchboard.nrdc.org> (accessed on 1 May 2017).
4. Matt McGrath. Climate Change: CO₂ Emissions Rising for First Time in Four Years. Available online: <https://www.bbc.com/news/science-environment-46347453> (accessed on 24 December 2018).
5. Shehabi, A.; Smith, S.; Sartor, D.; Brown, R.; Herrlin, M.; Koomey, J.; Masanet, E.; Horner, N.; Azevedo, I.; Lintner, W. *United States Data Center Energy Usage Report*; Lawrence Berkeley National Laboratory: Berkeley, CA, USA, 2016.
6. Willkommen bei den Energy Charts. Available online: <https://www.energy-charts.de/> (accessed on 23 April 2017).

7. U.S. Energy Information Administration (EIA). Available online: <https://www.eia.gov/beta/international/analysis.cfm?iso=CHN> (accessed on 13 February 2017).
8. Callou, G.; Maciel, P.; Tutsch, D.; Ferreira, J.; Araújo, J.; Souza, R. *Estimating Sustainability Impact of High Dependable Data Centers: A Comparative Study between Brazilian and US Energy Mixes*; Springer: Vienna, Austria, 2013.
9. Institute for Energy Research. Energy Encyclopedia. Available online: <http://instituteenergyresearch.org/topics/encyclopedia/> (accessed on 29 June 2016).
10. Ferreira, J.; Callou, G.; Maciel, P. A power load distribution algorithm to optimize data center electrical flow. *Energies* **2013**, *6*, 3422–3443. [[CrossRef](#)]
11. Silva, B.; Matos, R.; Callou, G.; Figueiredo, J.; Oliveira, D.; Ferreira, J.; Dantas, J.; Junior, A.L.; Alves, V.; Maciel, P. Mercury: An Integrated Environment for Performance and Dependability Evaluation of General Systems. In Proceedings of the IEEE 45th Dependable Systems and Networks Conference (DSN-2015), Rio de Janeiro, Brazil, 22–25 June 2015.
12. Kuo, W.; Zuo, M.J. *Optimal Reliability Modeling—Principles and Applications*; John Wiley and Sons: New York, NY, USA, 2003.
13. Molloy, M.K. Performance analysis using stochastic Petri nets. *IEEE Trans. Comput.* **1982**, *9*, 913–1007. [[CrossRef](#)]
14. Zeng, Y.-R.; Zeng, Y.; Choi, B.; Wang, L. Multifactor-influenced energy consumption forecasting using enhanced back-propagation neural network. *Energy* **2017**, *127*, 381–396. [[CrossRef](#)]
15. Wang, L.; Hu, H.; Ai, X.-Y.; Liu, H. Effective electricity energy consumption forecasting using echo state network improved by differential evolution algorithm. *Energy* **2018**, *153*, 801–815. [[CrossRef](#)]
16. Wang, L.; Lv, S.-X.; Zeng, Y.-R. Effective sparse adaboost method with ESN and FOA for industrial electricity consumption forecasting in China. *Energy* **2018**, *155*, 1013–1031. [[CrossRef](#)]
17. He, Y.; Qin, Y.; Wang, S.; Wang, X.; Wang, C. Electricity consumption probability density forecasting method based on LASSO-Quantile Regression Neural Network. *Appl. Energy* **2019**, *233*, 565–575. [[CrossRef](#)]
18. Reddy, V.D.; Setz, B.; Rao, G.S.V.; Gangadharan, G.R.; Aiello, M. Metrics for sustainable data centers. *IEEE Trans. Sustain. Comput.* **2017**, *2*, 290–303. [[CrossRef](#)]
19. Dandres, T.; Moghaddam, R.F.; Nguyen, K.K.; Lemieux, Y.; Samson, R.; Cheriet, M. Consideration of marginal electricity in real-time minimization of distributed data centre emissions. *J. Clean. Prod.* **2017**, *143*, 116–124. [[CrossRef](#)]
20. Helm, J.L. *Energy: Production, Consumption, and Consequences*; National Academy Press: Washington, DC, USA, 1990.
21. Kammen, D.M.; Pacca, S. Assessing the costs of electricity. *Annu. Rev. Environ. Resour.* **2004**, *29*, 301–344. [[CrossRef](#)]
22. Haykin, S. *Neural Networks: A Comprehensive Foundation*; Prentice Hall PTR: Upper Saddle River, NJ, USA, 1994.
23. McCulloch, W.S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **1943**, *5*, 115–133. [[CrossRef](#)]
24. De Pádua Braga, A.; de Leon Ferreira, A.C.P.; Ludermir, T.B. *Redes Neurais Artificiais: Teoria e Aplicações*; LTC Editora: Rio de Janeiro, Brazil, 2007.
25. Box, G.E.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. *Time Series Analysis: Forecasting and Control*; John Wiley and Sons: Hoboken, NJ, USA, 2015.
26. Pindyck, R.S.; Rubinfeld, D.L. *Econometria: Modelos & Previsões*; Elsevier: Amsterdam, The Netherlands, 2004.
27. Ferreira, J.; Callou, G.; Dantas, J.; Souza, R.; Maciel, P. An algorithm to optimize electrical flows. In Proceedings of the 2013 IEEE International Conference on Systems, Man, and Cybernetics, Manchester, UK, 13–16 October 2013; pp. 109–114.
28. De Gooijer, J.G.; Hyndman, R.J. 25 years of time series forecasting. *Int. J. Forecast.* **2006**, *22*, 443–473. [[CrossRef](#)]
29. Gardner, M.W.; Dorling, S.R. Artificial neural networks (the multilayer perceptron)—A review of applications in the atmospheric sciences. *Atmos. Environ.* **1998**, *32*, 2627–2636. [[CrossRef](#)]
30. Haykin, S. *Neural Networks and Learning Machines*; Prentice Hall: Pearson Upper Saddle River, NJ, USA, 2009.
31. Rumelhart, D.E.; Hinton, G.E.; McClelland, J.L. A general framework for parallel distributed processing. *Parallel Distrib. Process. Explor. Microstruc. Cognit.* **1986**, *1*, 45–76.

32. The Up Time Institute. Available online: <https://uptimeinstitute.com/> (accessed on 9 July 2016).
33. Neoenergia Group. Available online: <http://www.neoenergia.com/> (accessed on 28 April 2017).



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).