Application of Neural Networks’ Models to Predict Energy Consumption

Rita Sereno Serrão da Veiga Alves
rita_veigaalves@hotmail.com
Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisbon, Portugal
June, 2016

Artificial neural networks (ANN) can be used in the modelling and prediction of dynamic complex problems, which cannot be treated using conventional solutions. This study is innovative, because it uses useful exergy to predict energy consumption. It also makes uses of ANN to model and predict the energy consumption in Portugal and compares this value with others obtained from different modelling techniques, namely, multiple linear regression and multivariate linear regression. Available data from 1960 to 2009 of Gross Value Added (GVA) and useful exergy, namely High Temperature Heat (HTH), Medium Temperature Heat (MTH), Low Temperature Heat (LTH), Mechanical Drive (MD) and Other Electric Uses + Light (Elec) was analysed and used to build the dynamic prediction model. Prior to the construction of the dynamic model, the variables with the greatest impact on the process were selected using correlation analysis and principal component analysis (PCA) and then the nonlinear ANN model was tested. For this study, the Neural Network Toolbox™ from MATLAB was used. The dynamic neural network has been trained through the backpropagation algorithm using the Levenberg-Marquardt optimization method with various combinations of input data, in order to search for the model that best fits the data. The training, testing and validation of the model construction were performed using data between the years 1960-1999 while the data from decade 2000-2009 was only used to verify the ANN prediction capabilities. Compared to multiple linear regression and multivariate linear regression, the ANN demonstrated a far superior approach capacity. This supports the adaptability of ANN for modelling and predicting complex dynamical problems compared to conventional methods.

Keywords: Gross Value Added, Exergy, Prediction, Modelling, Multivariate Data Analysis, Artificial Neural Networks

I. INTRODUCTION

In this thesis is presented a new study that combines the artificial neural networks (ANN) approach with the temporal evolution of useful exergy in order to make predictions of energy consumption. There are already few studies that have used ANN to make forecasts of energy consumption using final energy data, such as Greece, South Korea and a study that uses useful exergy China to make forecasts. There is a study by Percebois [5], supported by Serrenho [7], where it is suggested that energy intensity (energy consumption by GDP) is better calculated at the last level of the energy chain when aiming to calculate the needs to meet different end-users. This work is the first one to combine the use of ANN with useful exergy to predict the useful exergy in its different uses. In this study the value of exergy instead of energy is used, since exergy emphasizes the quantitative aspect of energy but it also emphasizes the quality of the different forms of energy. This quality is measured as the capacity of transforming a certain form of energy in work. However, this study typically examines primary or final energy data, rather than useful work values obtained using an exergy analysis-based technique. Exergy analysis takes a broader, whole-system approach to energy analysis, giving “a measure of the thermodynamic quality of an energy carrier”, thereby enabling a robust view of useful work consumed in provision of energy services. Energy analysis also has the benefit of taking into account more aspects of the energy supply chain and in a more consistent way than traditional energy analysis [1]. It is also important to refer that there are three stages of exergy: - primary, final and useful. The primary energy intents to quantify the energetic content of the used resources while the useful energy (exergy) quantifies the needs of energy (or exergy) of final consumers. Thus, it represents the real necessities of a country and it gives a perception of the direct relationship with the productivity and economic level. The aim of this study is the:
1. Analysis of the evolution curves of the useful exergy in Portugal between 1960-2009;
2. Creation of a computational tool that, based in models of artificial neural networks, is able to project the useful exergy levels per type of use;
3. Verification of the importance of correctly defining input and output data of the neural net, by comparing the results of this study with the results of similar studies;
4. Conclusion that sensibility analysis studies would help identify the most important parameters in the forecast.

II. METHODOLOGY AND DATA

A. Operational Data

The sample of data used in the construction of the prediction model is forty years long (between 1960
ANN, being the model, we want to eliminate. The existence of outliers and empirical distribution of variables provided the first step was to represent a good organization as well as to help introduce it in the MATLAB computational tool developed.

The available data is the following:

**Heat (TJ)** [8]: heat process, where HTH and MTH are mainly used in industry (ex: glass) and LTH is used in all sectors but mainly in the services and the residential sectors for the heating of spaces.

<table>
<thead>
<tr>
<th>High Temperature Heat</th>
<th>Medium Temperature Heat</th>
<th>Low Temperature Heat</th>
</tr>
</thead>
<tbody>
<tr>
<td>500°C</td>
<td>150°C</td>
<td>50-120°C</td>
</tr>
</tbody>
</table>

**Mechanical Drive (TJ)** [8]: transportation sector and stationary industry, being the work of the gas and diesel that move motors, steam engines, diesel-electric engines, aircraft motors, ships motors and electric motors.

**Other electric uses** [8]: residential and services sectors such as communication, electronics, electric devices and electrochemical industry.

**Light (TJ)** [8]: refers to the use of electricity in all sectors.

Other electric uses and light were grouped in one variable denominated Ele.

**Gross Added Value (Mrd EURO_PTE)** [3] e [6]: final result of the productive activity during a certain period. It is the difference between the value of the production and the value of the intermediate consumption.

### B. Steps of Data Processing and the Development of Prediction Models

The main steps of data processing and the major steps that were made in the development of the prediction models are described below.

The first step was to represent in boxplots all the variables provided. With the aim of analyzing the empirical distribution of the data it was verified the existence of outliers and, when necessary, they were eliminated.

With the objective of better understanding the relationship between each of the variables, a map of correlations with all the variables was made. The Pearson Correlation was used, which is obtained by dividing the covariance of two variables by the product of their standard deviation.

In PCA the number of principal components to be used in the construction of the model, the score plot and the loading plot were analysed. With PCA analysis a matrix of correlations was obtained which allows the selection of the variables that most contribute to the description of each component. The information thus gained leads us to select variables more related with the variables that we want to predict.

For each of the variables both linear regressions (multiple and multivariate) were performed to determine the weight of each variable in the definition of the remaining variables and also to obtain a comparison with the results obtained through the ANN.

With these variables the prediction models, using both Multiple Linear Regression and Multivariate Linear Regression and ANN methods, were constructed for useful exergy prediction, using data from 1960 to 1999. With the best obtained model from the ANN method, the prediction for the period between 2000 and 2009 was made.

### ANN model

**Training:** The dynamic neural network has been trained using the mechanism of backpropagation with the Levenberg-Marquardt optimization algorithm and a supervised learning, since it had fifty years of consecutive data. The network performance was evaluated based on the mean squared error (MSE), which results from the difference between the real output and the one calculated by ANN, according to Equation (2.1):

\[
MSE = \frac{1}{N} \sum_{i=0}^{i=N} (y_i - \hat{y}_i)^2
\]  


### 1-step-ahead-prediction

\[
y = \left[ \frac{\bar{y}_{1961} \bar{y}_{1962} - \bar{y}_{1991}}{\bar{e}_{1990} - \bar{e}_{1991}} \right]_{\text{ANN}} = \hat{y} - [\bar{e}_{1992} \bar{e}_{1993} - \bar{e}_{1994}]
\]
Where U is the real value of the supplied data table and \( \hat{Y} \) the values estimated by ANN.

To the training phase, the configuration of Picture 1 was used, where the aim was to compare the estimated values with the actual values. In a first analysis was selected the neural network that best reproduces the data for each exergy. Each variable or group of variables were taken into account in the construction of the final network to forecast all exergies simultaneously.

Simulation: In order to find a network simulation that best fits the data and produces the best forecasts, two types of prediction models were used: 1-step-ahead-prediction and 2-step-ahead-prediction.

Furthermore, it was considered that only knew forty years (1960-1999) of the fifty available. This assumption has forced the use of a third network type, the closed loop in which the last ten years are fed back to the network.

The 2-step-ahead-prediction’s structure is depicted in Picture 2 as follows:

Models with more than 2 steps ahead often do not obtain good results [4]. So, we tried to carry out the analysis of such networks before testing higher order prediction models. In this case it is first estimated the \( \hat{U}_n(k + 1) \) and returned to the network, followed by the estimation of \( \hat{U}_n(k + 2) \), which is then compared to the real \( U_n(k + 2) \) and so on for the year interval tested.

\[
0 = \begin{bmatrix} \hat{P}_{n_1} & \ldots & \hat{P}_{n_q} \\ \hat{E}_{r_1} & \ldots & \hat{E}_{r_q} \end{bmatrix} \begin{bmatrix} \hat{E}_n(k) \\ \hat{E}_n(k - 1) \end{bmatrix} + \begin{bmatrix} \hat{E}_{r_1} \\ \hat{E}_{r_2} \end{bmatrix} = \gamma = \begin{bmatrix} U_{r_1} \\ U_{r_2} \end{bmatrix}.
\]

Closed-loop

During the learning phase the corresponding data for the last ten years (2000-2009) has not been used. These were used to test the ANN ability to forecast these data. Thus, we used the closed-loop simulation presented in Picture 3.

Neural network used

After being achieved satisfactory forecast results for the years 2000 to 2009, evaluated by observing the graphs generated by MATLAB and the evolution of the MSE, the desired neural network was obtained. In the present study the dynamic neural network used was nonlinear autoregressive with exogenous input (NARX), defined be Equation (2.2).

\[
y(t) = f(y(t - 1), y(t - 2), \ldots, y(t - n_y), u(t), u(t - 1), \ldots, u(t - n_u))
\]

(2.2)

Where the next value of the output \( y(t) \) is calculated through the regression of previous output values and an independent input variable (exogenous), \( u(t) \). These are characterized by feedback connections that span across multiple network layers.

After the construction of a code that would allow the systematic train of networks with different numbers of layers and different number of years of delays, forty networks have been trained in each of settings (for a total of 640 trainings), presented in order to search the network that best fits the data. This systematic search for the best network is required since, due to the allocation of the values of the network weights being random, it’s not guaranteed a priori that the network that best approximates the data is obtained.

Next, we tried to identify the network that had the best performance in the dynamic modelling of the
problem. For that, the network with smaller MSE in bringing data was identified. Nevertheless, the analysis was accompanied by analysis of a performance graph showing the MSE value versus the number of iterations (Picture 4) and performance of training and testing validation. This analysis is needed because if the curve increased significantly before the increase of the validation curve, then it was possible that some overtraining had occurred.

C. Algorithms and Software libraries

This work was performed with the help of MATLAB which implements various algorithms and methods commonly used in machine learning. To find the optimal number of iterations and neurons, an exhaustive search was performed across the space containing all possible combinations of iterations number and nodes number.

III. RESULTS AND DISCUSSION

This chapter presents the main results. Initially, it will be made an analysis of the data to identify outliers and thus separate the disparate data from the remaining data. Afterwards, the data correlation will be analyzed. After this analysis we will try to understand the relationship between the variables and the contribution of each one for the evolution description of useful exergy per each type of use in Portugal. As this analysis is limited to the correlation between pairs of variables, it is completed with the analysis of Principal Components in order to find the variables with major influence in the explanation of exergies. Based on this conclusions it is then possible to find a prediction model for 10 years.

A. Data Pre-Processing

Boxplots: In order to work the data with similar dimensions we first worked in the standardization so that the data would have null average and unitary standard deviation. Without loss of generality, with \( x \) aleatory variable with average \( \mu \) and standard deviation \( \sigma \), the standardization is given by the Equation (3.1):

\[
    z = \frac{(x - \mu)}{\sigma} \quad (3.1)
\]

As each variable has fairly distinctive dimensions it is important to make that normalization in order to minimize subsequent numerical problems. The following graph (Picture 5) shows the absence of outliers.

B. Data Processing

Correlations analysis: One of the most used methods to investigate the relationship between each pair of variables is the Cartesian scatterplots. In Table 1 and Picture 6, the correlation between each variable after Pearson can be seen. We can observe that a Pearson correlation efficiency of -1 means that the variables are not correlated, a coefficient of 1 means that the variables are perfectly correlated and a coefficient of 0 means that the variables are uncorrelated.

<table>
<thead>
<tr>
<th></th>
<th>HTH</th>
<th>MTH</th>
<th>LTH</th>
<th>MD</th>
<th>Ele</th>
<th>GVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTH</td>
<td>1.00</td>
<td>0.81</td>
<td>0.90</td>
<td>0.84</td>
<td>0.91</td>
<td>0.90</td>
</tr>
<tr>
<td>MTH</td>
<td>0.81</td>
<td>1.00</td>
<td>0.94</td>
<td>0.95</td>
<td>0.93</td>
<td>0.94</td>
</tr>
<tr>
<td>LTH</td>
<td>0.90</td>
<td>0.94</td>
<td>1.00</td>
<td>0.98</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>MD</td>
<td>0.84</td>
<td>0.95</td>
<td>0.98</td>
<td>1.00</td>
<td>0.98</td>
<td>0.99</td>
</tr>
<tr>
<td>Ele</td>
<td>0.91</td>
<td>0.93</td>
<td>0.99</td>
<td>0.98</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>GVA</td>
<td>0.90</td>
<td>0.94</td>
<td>0.99</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

picture 4 - Performance graph obtained through the MATLAB

picture 5 - Representation boxplot of variables, namely, 1-HTH, 2-MTH, 3-LTH, 4-MD, 5-Ele, 6-GVA
Analyzing the graphic above, the existence of strong correlations between variables is identified, once the major part shows correlations values near the unit. However there are variables such as the HTH and the MTH that, in spite having the Pearson’s coefficient relatively lower, have significant dispersions with all variables when compared to the remaining. This shows that these variables will be preponderant in the description of the problem, since they have non-linear behavior and therefore will add additional information to the forecast.

**PCA analysis:** After this, the principal components were analysed and the significance of each component variable was observed. Table 2 demonstrates the covariance matrix which explains these relationships.

### Table 2 – Principal Components of PCA

<table>
<thead>
<tr>
<th>Components</th>
<th>1ºComp</th>
<th>2ºComp</th>
<th>3ºComp</th>
<th>4ºComp</th>
<th>5ºComp</th>
<th>6ºComp</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTH</td>
<td>0.3840</td>
<td>0.8418</td>
<td>0.2741</td>
<td>-0.0753</td>
<td>-0.2400</td>
<td>-0.0740</td>
</tr>
<tr>
<td>MTH</td>
<td>0.3999</td>
<td>-0.4364</td>
<td>0.8023</td>
<td>-0.0386</td>
<td>0.0507</td>
<td>-0.0423</td>
</tr>
<tr>
<td>LTH</td>
<td>0.4168</td>
<td>-0.0173</td>
<td>-0.1960</td>
<td>0.8536</td>
<td>0.1683</td>
<td>-0.1747</td>
</tr>
<tr>
<td>MD</td>
<td>0.4122</td>
<td>-0.3098</td>
<td>-0.3604</td>
<td>-0.2271</td>
<td>-0.6693</td>
<td>-0.3318</td>
</tr>
<tr>
<td>Ele</td>
<td>0.4171</td>
<td>0.0286</td>
<td>-0.2719</td>
<td>-0.4563</td>
<td>0.6844</td>
<td>-0.2730</td>
</tr>
<tr>
<td>GVA</td>
<td>0.4183</td>
<td>-0.0614</td>
<td>-0.1971</td>
<td>-0.0658</td>
<td>-0.0228</td>
<td>0.8818</td>
</tr>
</tbody>
</table>

As we can see in the eigenvalues of the shown covariance matrix in Table 3, it is verified that the first component has an eigenvalue of 5.683 and so it describes 94.75% of the sample. The second component only shows an eigenvalue of 0.2197, inferior to the unit, the reason why Elizabeth Reis [2] could not use it in the analysis.

To select the variables that describe the first component, the major and/or minor element must be analysed. In this case, as the column shows very close values, the major one will be selected better to represent the corresponding variable GVA.

In the second column three situations are identified:

- **HTH** has a strong positive correlation;
- Negative correlations – MTH and MD. The most negative will be chosen, MTH;
- The remaining have a low correlation.

Therefore, for the second component the variables **HTH and MTH** will be selected.

In the third component the more negative variable is selected, i.e., **MD**. The major element, because it is related with the variable MTH and since it has already been selected, will not be taken into account in this component. Thus, as all the variables have been around MD, they will be not selected.

In the fourth component, the major variable is again selected: **LTH**. Finally, the last choice is **Ele** with the most negative value.

In terms of priority in the variables selection for the description of the problem, we have respectively **GVA, HTH, MTH, MD, LTH and Ele**.

### Table 3 – Eigenvalues of each principal component

<table>
<thead>
<tr>
<th>Components</th>
<th>Eigenvalues</th>
<th>Explation of variance (%)</th>
<th>Selected variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1º</td>
<td>5.6853</td>
<td>94.7551</td>
<td>GVA</td>
</tr>
<tr>
<td>2º</td>
<td>0.2197</td>
<td>3.6621</td>
<td>HTH e MTH</td>
</tr>
<tr>
<td>3º</td>
<td>0.0758</td>
<td>1.2636</td>
<td>MD</td>
</tr>
<tr>
<td>4º</td>
<td>0.0123</td>
<td>0.2042</td>
<td>LTH e Ele</td>
</tr>
<tr>
<td>5º</td>
<td>0.0049</td>
<td>0.0814</td>
<td>-</td>
</tr>
<tr>
<td>6º</td>
<td>0.0020</td>
<td>0.0336</td>
<td>-</td>
</tr>
</tbody>
</table>
Similar to what has been done in the selection through the correlation matrix, the biplot graphic was analyzed. It has the components in a system of axis and the vectors of each of the described variables in these components are presented. As an example the first two components in Picture 7 are presented. Thus, one can visualize the dimension and the orientation of the variables described in those principal components, besides reinforcing the selection of GVA, HTH, MTH and MD.

![Image of biplot graphic](image)

**Picture 7 – Component 1 versus Component 2**

### C. Development of Prediction Models

**ANN:** Finally, we identified a model based on dynamic neural networks. When performing the training of the network the weight values are randomly defined. So, the quality of the network for forecasting also has an inherent randomness. We proceeded to the training of 640 networks where various combinations of entry variables were very often tested with one or two years of delay for the variables of exergy. After that analysis the structure (Picture 8) of the network was identified which gave the best MSE results and so it was considered to best model and describe the problem. The entry variables of the network are: GVA, HTH, MTH, LTH, MD and Ele and also the latter five delayed one year.

In the network structure shown it is observed that the network receives 6 input variables (x(t)) where 5 of these variables (y(t)) also come with a year prior to the prediction: x(t) = [matrix with variables GVA, HTH, MTH, LTH, MD and Ele] without delay (represented by 0 of the “clock” in hidden) and y(t) = [matrix with variables HTH, MTH, LTH, MD and Ele]. This variable corresponds to the variables that are one year late, as these data remains in a cycle (represented by 1 of the “clock” in hidden) waiting to enter the network. The network consists of six hidden layers (number 6 in hidden) in which the neurons are modeled by sigmoid functions and the five required variables are provided (number 5 in output).

In Picture 9, Picture 10, Picture 11, Picture 12, Picture 13 the results of the net with 6 hidden layers and 85 neurons in the hidden layer are presented. In black we can see the prediction of the first forty years using the training, validation and testing of the nets (1-step-ahead-prediction), in green the prediction to ten years closed loop process and the red points represent the known data we try to model.

![Image of network structure](image)

**Picture 8 – Network structure removed from MATLAB**

![Image of second ANN](image)

**Picture 9 - Second ANN, representation of HTH**

![Image of second ANN](image)

**Picture 10 - Second ANN, representation of MTH**
On the pictures above we verify that good approaches were achieved in the data used to train and validate the network (first 40 years), having a satisfactory prediction for the next 10 years. It is observed that the largest absolute prediction errors do not exceed 5% in some years (e.g. LTH forecast in 2007), a perfectly acceptable value. Since the first forty years of training were reached and because the forty years were used in the network training and validation it was expected that the model was very near the mentioned years. However, the last ten years will hardly reach that approach, since the network had no information about them during the stage 1. Besides, in these years because each prediction is found based in years that had already been estimated, there is an increased number of mistakes every year. Picture 14 consolidates the previous conclusions.

Another difficulty closely linked to the prevision years is the abrupt variation of the variables HTH and MTH which may introduce desirable oscillations on the remaining variables since they are not constant.

For the graphic analysis it was concluded that the error in absolute value displays a random behavior and the variables that most contribute to the error in the prediction years are the MTH, MD and Ele. However, the behavior of Picture 10, Picture 12, Picture 13 follows the trend of the variables data.

Other models useful in the prediction of the exergy are the multiple linear regression and the multivariate linear regression. We observe the multivariate analysis in order to determine if this model shows better results. Because it is a multivariate model it is expected that there is a coupling between the variables.

D. Comparison between the three prediction models

From the analysis of the different methods it can be verified that the neural network was able to reach values with inferior errors. This fact is essential due to the nonlinear description given by the sigmoid function of neurons. Besides, the fact that the net has an over-definition of the parameters in its structure provides some flexibility to ANN, being thus possible to obtain predictions with an order of magnitude three times inferior in the case of ANN seen on Table 4.
Table 4 – Error comparison of ten years forecast

<table>
<thead>
<tr>
<th>Method</th>
<th>HTH</th>
<th>MTH</th>
<th>LTH</th>
<th>MD</th>
<th>Ele</th>
<th>Global Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multivariate</td>
<td>$8.17 \times 10^{-1}$</td>
<td>$2.48 \times 10^{-1}$</td>
<td>$1.99 \times 10^{-1}$</td>
<td>$4.53 \times 10^{-2}$</td>
<td>$1.11 \times 10^{-1}$</td>
<td>$1.57 \times 10^{-1}$</td>
</tr>
<tr>
<td>ANO</td>
<td>$2.74 \times 10^{-1}$</td>
<td>$3.17 \times 10^{-1}$</td>
<td>$1.58 \times 10^{-1}$</td>
<td>$0.65 \times 10^{-2}$</td>
<td>$0.57 \times 10^{-2}$</td>
<td>$0.40 \times 10^{-1}$</td>
</tr>
<tr>
<td>ANN</td>
<td>$4.10 \times 10^{-4}$</td>
<td>$1.46 \times 10^{-4}$</td>
<td>$3.86 \times 10^{-4}$</td>
<td>$5.12 \times 10^{-4}$</td>
<td>$8.25 \times 10^{-4}$</td>
<td>$1.65 \times 10^{-4}$</td>
</tr>
</tbody>
</table>

No matter what the adopted method is, the training data is few descriptive of the behavior to predict. This leads to increased difficulties when extrapolating. In the case of linear regressions this made the functions more rigid to present variations in the last ten years. On the other hand, in case of ANN this factor caused some overtraining situations, creating nets which interpolated significantly well the training years. This led sometimes to unacceptable responses in years of prediction, or just the contrary, meaning predictions relatively acceptable but with the training years having great alterations and amplitude.

Finally, the graphs (Picture 15, Picture 16, Picture 17, Picture 18, Picture 19) of the errors of the ten-year forecast for ANN analysis are shown. It was concluded that in general terms looking at the sum of the errors, although there are some years better than others, the graphics individually approach to the real data. It is noted that the larger the “web” the greater the error in the predicted year. If the “web” is small the error is smaller.

It was concluded, by graphical observation, the neural network was the forecast model that best adapted to the data provided, providing very useful information on the estimates of the variables individually.

Picture 15 – Error evolution of variables in 2000 and 2001 for the ANN predictive analysis

Picture 16 - Error evolution of variables in 2002 and 2003 for the ANN predictive analysis

Picture 17 - Error evolution of variables in 2004 and 2005 for the ANN predictive analysis

Picture 18 - Error evolution of variables in 2006 and 2007 for the ANN predictive analysis
In order to verify the quality of the net the results were compared with models obtained by the multiple and multivariate linear regression. They have demonstrated to have an inferior quality when compared to ANN.

In every case, the difficulty in the extrapolation due to the behavior discrepancy of the training data (first forty years) and the prediction of the last ten years was verified. Those difficulties appear as the lack of flexibility to the sharp variations in the case of linear regressions.

As a future study it is important to evaluate the relationship with other input variables and to test a new network structure. The neural networks require a lot of data in order to obtain a model that can well adjust to each case study, therefore it needs a lot of inputs to run properly. Hence, a suggestion for future development is to analyze these variables for a higher period of time because if the data is larger the models could have better performances, since the growth in the data can contribute to obtain better results in exergy prediction.

This study shows that dynamic neural networks have good results for exergy prediction and can eventually be adopted by other countries. Another proposal is to adapt this dynamic neural network in predicting other types of consumption, besides energy.

As a final note and based on the article developed for China [1] it would be interesting to develop a forecasting study, for e.g. 2030 - 2050, with GVA scenarios for Portugal and compare it with existing scenarios.

REFERENCES