Modelling the energy consumption of trains applying neural networks

January, 2016.

ABSTRACT

This paper presents the training of a neural network using consumption data measured in the underground network of Valencia (Spain), with the objective of estimating the energy consumption of the systems. After calibration and validation of the neural network using part of the consumption data gathered, the results obtained show that the neural network is capable of predicting power consumption with high accuracy. Once fully trained, the network can be used to study the energy consumption of a metro system and for testing hypothetical operation scenarios.

Keywords

Gradient; energy consumption; artificial neural networks; Metro; railway; track layout.

1. INTRODUCTION

The transport sector contributes greatly to global energy consumption. According to the International Energy Agency [1], overall energy consumption in 2013 was 2 563.52 Million Tonnes of Oil Equivalent (Mtoe), with the transport sector being responsible of up to 27.6%.

Railways are generally much more efficient than road transport in terms of energy consumption for both freight and passengers [2], [3], [4]. Despite this, it is still necessary to reduce their energy consumption in order to improve their competitiveness and contribute to a more sustainable world. For this reason, many strategies are implemented to reduce energy consumption in railways. There are strategies proposed concerning line design, rolling stock and operation [5].

Traditionally, energy consumption of an electric train is monitored at the substations. This provides information about the total energy consumed in an instant, or during a given period of time. However, substations do not give information on how the energy is consumed by each element and subsystem of the railway system, and thus it is not possible to know in detail the impact of any action taken to reduce energy consumption.

The current energy consumption in railways depends on many factors such as gradients, maximum speeds, loads, patterns of stops, electrical efficiency of train and power supply system, running resistance, driving style, etc.

Researchers have estimated the energy consumption and explored improvements in rail transport through track layout optimisation by means of Geographic Information Systems (GIS) [6], [7]. Other authors have used genetic algorithms to optimise different aspects such as track alignments and operator and user costs for rail operation [6], [7], [8] or crew scheduling [9], [10]. There are methods that aim to optimise travel time and coasting points by using models based on artificial neural networks and genetic algorithms [11]. But these methods do not include gradient or real time measured energy consumption as data.

This paper aims to develop, train and validate a neural network to simulate the energy consumption of a metropolitan line using measured empirical data, and use the neural network to predict the energy consumption at each instant. This network will be then used as a tool to study and optimise different variables with an important impact on energy consumption of the train, like speed, acceleration or gradient.

2. METHODOLOGY

2.1. Data gathering and processing

In order to check the energy consumption of the train in a global way, three MSAVDC meters devices, manufactured by Mors-Smitt, were installed in the front car of the train: one in the pantograph (circuit breaker), another one in the auxiliary converter input, and the last one in the braking resistors. These devices allow measuring not only the overall train energy consumption in real time, but also the energy consumed by each subsystem: traction, auxiliary devices and rheostatic brake. Figure 1~~Figure 1~~ shows a diagram of the MSAVDC meters devices installed in the train.

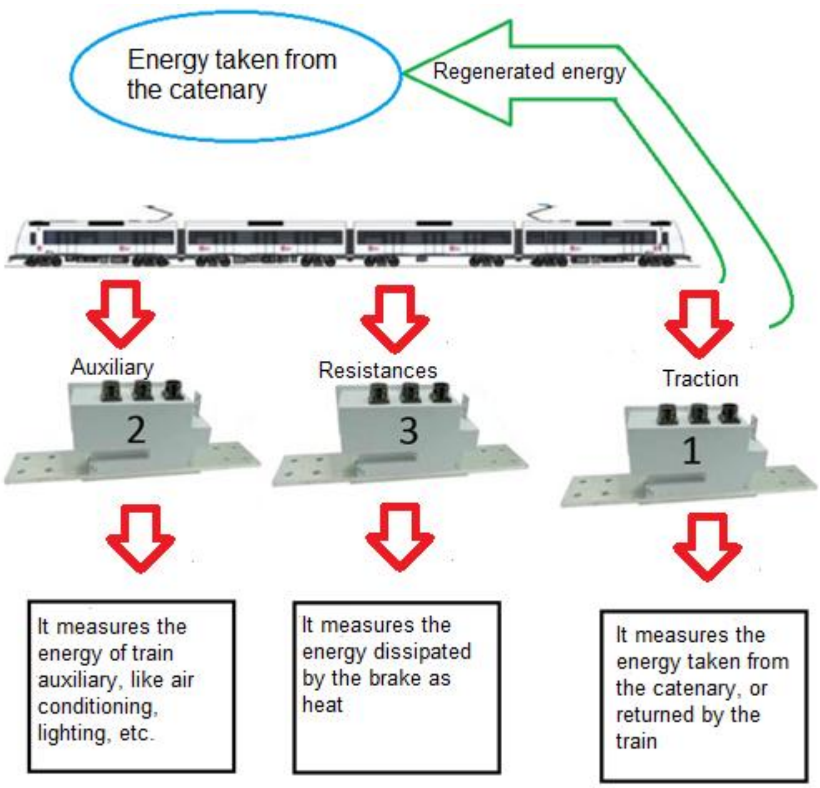


Figure 1. Diagram of the MSAVDC meters devices installed in the train

The speed was measured by a Knorr sensor model BB0457681100, fed by a phonic wheel on one axis of another car of the train.

After verifying the correct operation of all devices, measurements were made with passengers on board, on August 4th, 2014. Twelve trips were measured in line 5 of

MetroValencia between Marítim-Serrería and Alameda stations, six trips towards Alameda and six trips towards Marítim-Serrería.

2.2. Neural networks

A Neural network is a computational model inspired by the structure of biological nervous systems. It has several elements called neurons which operate in parallel and can be trained to yield a target output data when supplied with specific input data.

In this case, neural networks were implemented to evaluate variables with an important impact on energy consumption. In order to define the framework of the neural network, an approach adjustment function has been chosen. The neural network, supplied with *input* data to be defined (such as speed, acceleration, etc.) and *target* data (in this case, measured energy consumption) is trained, which means that the *output* provided by the network (simulated energy consumption) is compared with the available *target* data, and the network parameters are adjusted through an iterative process until a good agreement between model and reality is achieved.

The structure chosen to accomplish this objective is a two layer feed-forward neural network (Figure 2~~Figure 2~~), because it is a common and tested scheme, and with a great ability to adjust functions [12], [13].

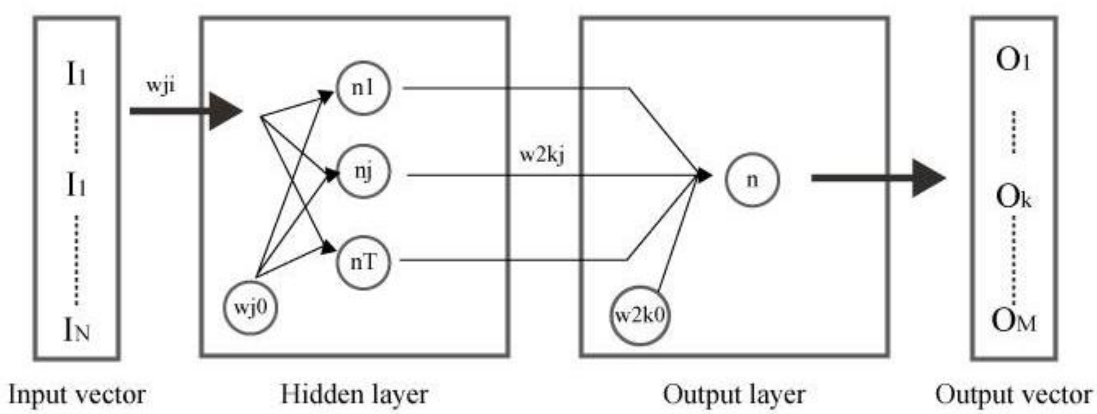


Figure 2. Neural network framework

The first layer, called *hidden layer*, has a number of neurons to define. The second layer (*output layer*) has a single neuron with a linear transfer function. Eq. 1 shows the formulation of the neural network:

$$O_k = \tilde{g}\left(\sum_{j=0}^{M} w2_{kj} \cdot g\left(\sum_{i=0}^{N} w_{ji} \cdot I_i\right)\right) \tag{1}$$

Where O_k is the network output, M is the number of *output* elements, I_i is the *input* data, N is the number of input variables, w_{ji} is the synaptic weight of the first layer and $w2_{kj}$ is the synaptic weight of the second layer. The synaptic weight w_{ji}, for example, defines the strength of a synaptic connection between two neurons, the presynaptic neuron j and the postsynaptic neuron i. This structure can identify non-linear relations between *input* and *output* data [12] using the Log-Simoid function as a transfer function between the hidden layer and the output layer.

The training method used is called *Back-Propagation*: The network is evaluated, the results are checked based on certain criteria, and the synaptic weights are changed in an iterative loop [14]. The chosen calibration criterion is the minimization of the Mean Square Error (MSE) between the network *output* and the *target* data, which is verified by deriving the MSE with respect to the network synaptic weights. The specific training algorithm used is called *Levenberg-Marquardt* algorithm, very efficient and widely checked [12].

When a neural network is evaluated, it is not only important to assess whether it agrees with the training data. A well trained network must encompass the subjacent patterns of the data, a feature called generalisation. In order to assess this aspect of the network, after the training phase the network is once again checked using previously unused data [15]. To accomplish this, the data available is divided randomly in three subsets, one for training (70%), one for validation (15%) and one for testing (15%). The network is trained with the first data block, and after each iteration a check-up with the second block (validation) is performed. When the validation MSE begins to increase (while the training MSE continues to drop), the network is starting to adjust the data error (*overfitting*) and the training is stopped. At this point, the third block (testing) is used to perform a final check-up to the validity of the neural network.

In order to avoid o*verfitting*, the *early stopping* training method is employed, which stops the training process when the training criteria (i.e. minimising the MSE) has not been fulfilled completely. This method helps ensuring that the network is capable of generalisation, i.e. it is not conditioned by the specific error of the training data and has learned the subjacent pattern of the modelled phenomenon [16].

The use of an *early-stopping* method limits the possibility of *overfitting*, and therefore there is not a theoretical limitation for the size of the network (i.e. the number of neurons in the hidden layer) [17]. However, this number determines the degrees of freedom of the neural network, and certainly influences its generalisation ability. In order to determine the optimal size, different network sizes were tested and the training and validation MSE were compared. The first one will always tend to drop as the number of neurons grows, while the validations MSE will start to grow up when a certain size is reached. This specific size is considered the optimum, as further increasing the number of neurons will cause *overfitting* [18].

The neural network presented in this paper uses speed, acceleration and gradient as *input* data and measured empirical energy consumption as *target* data. Once trained, the neural network may predict the train energy consumption *(output data)* with high accuracy as proven when compared to the measured empirical data (*target* data).

3. CASE STUDY

3.1. Introduction

MetroValencia has six subway lines (lines 1, 2, 3, 5, 7 and 9) and three tram lines (lines 4, 6 and 8). MetroValencia has 132 stations with a total length of 146.8 km, and 121 trains [19]. Trips along line 5 were analysed, between stations Marítim-Serrería and Alameda.

Regarding the traction and power systems of the network, there is only a single input voltage to the substations with a magnitude of 20 kV AC. However, there are two different output voltages: 1 500 V DC, (used in all six subway lines) and 750 V DC (used in all three tram lines), with annual energy consumption around 64.4 GWh and 18.1 GWh, respectively (78% for the subway lines and 22% for the tram lines). This energy is consumed by all elements and systems of MetroValencia. If the energy consumption of each network component is considered, 70% of the overall energy goes to traction (53 GWh) and 24% goes to stations, other power consumptions being negligible. [20].

3.2. Input data

3.2.1. Gradient profile and stations along the line

Focusing on Line 5 of MetroValencia, particularly between the stations of Marítim-Serrería and Alameda, there are three stations in between with a total length of 2 720 m and the route has four stops. Figure 3~~Figure 3~~ shows a diagram of the vertical track layout of the studied route, indicating the stations (Marítim-Serrería, Ayora, Amistat, Aragón and Alameda) and the gradient profile along the line. The maximum gradient is 2%.

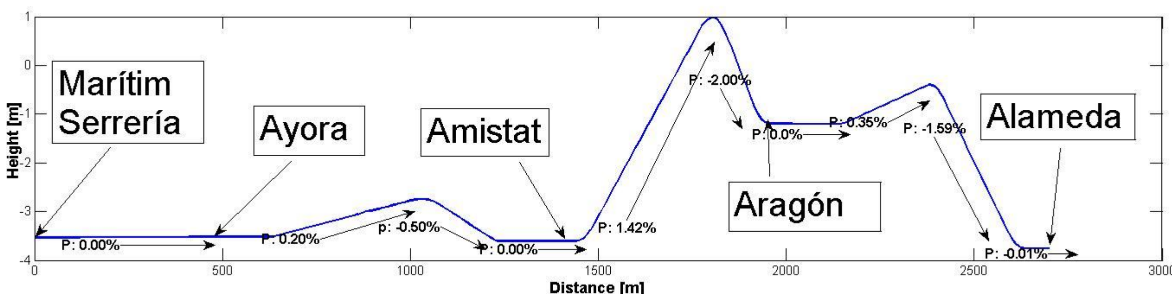


Figure 3. Vertical layout between Marítim Serrería and Alameda and the stations (stops) in between

3.2.2. Speed and acceleration

Speed was measured using an odometer placed in one of the wheels of the monitored train. Acceleration is directly derived from the speed. Figure 4~~Figure 4~~ shows these two variables as measured during the first trip.

Three variables were chosen as *input* data (gradient, speed and acceleration). During the neural network training, all *input* variables and their combinations were tested until the one that provided a better fit with the *target* data was chosen.

Figure 4. Speed (blue) and acceleration (orange) of the first trip between Marítim-Serrería and Alameda

3.3. Target data

The energy consumption in the pantograph (measured in the circuit breaker) was monitored in real time while the train performed conventional services with passengers on board. This measured energy consumption was used as *target* data.

The monitored train was a Metro Series 4300 (Vossloh) with 4 cars, a maximum speed of 80 km/h, a nominal tension of 1500 V DC and a power of 1480 kW.

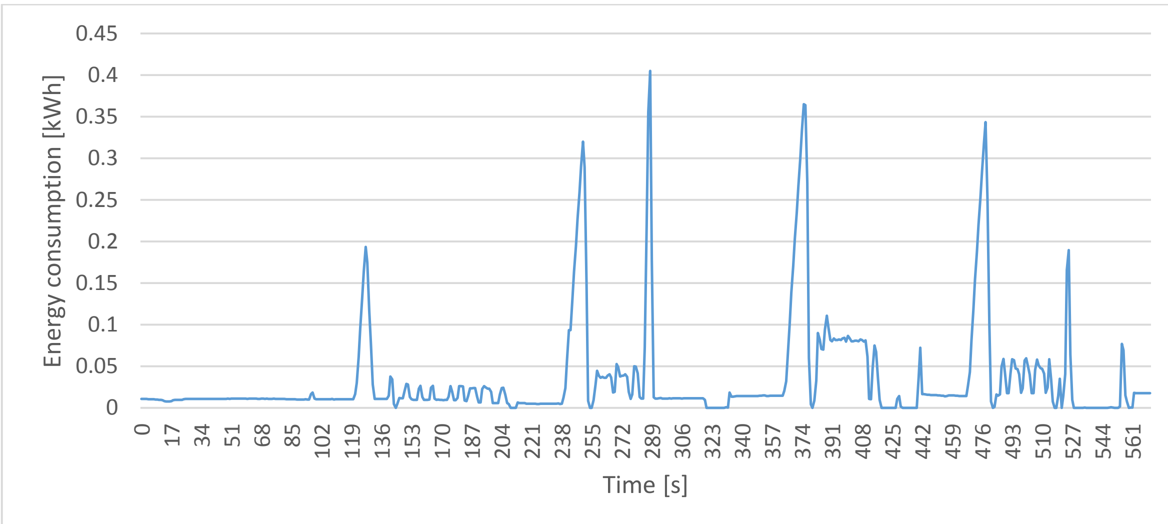
Energy consumption data obtained for each trip is shown in Table 1~~Table 1~~. Every trip is the same, with four stops between the first station and the last one. The stopping time in every station was not measured.

Table 1. Global consumption of the registered trips

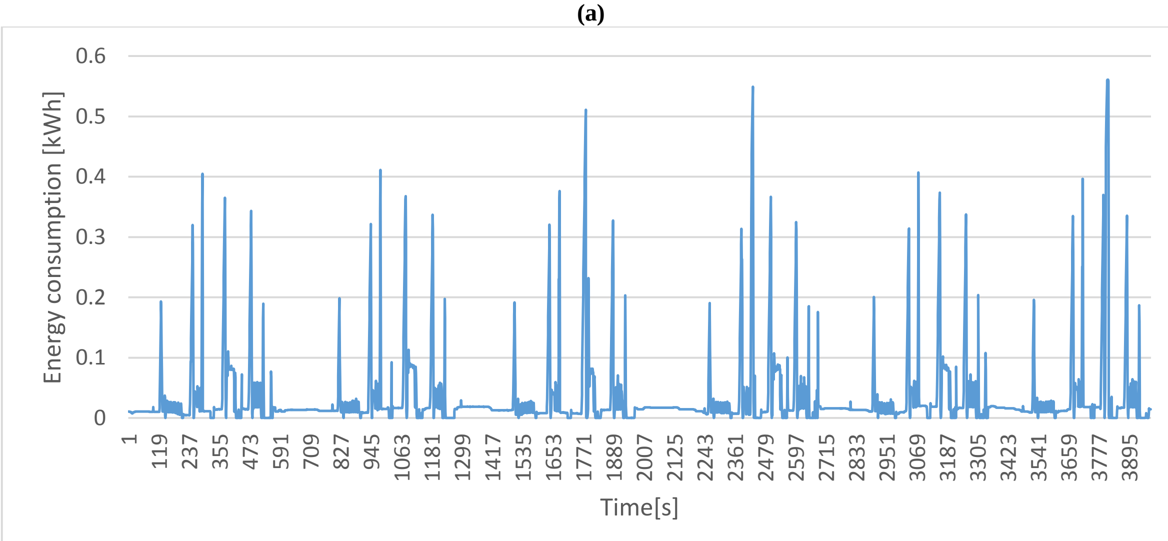
	Date	Travel time (min)	Energy consumption measured in the circuit breaker of the train (kWh/km)
Marítim-Serrería - Alameda	04/08/2014	9.52	7.18
	04/08/2014	11.78	8.11
	04/08/2014	11.68	8.33
	04/08/2014	11.87	9.25
	04/08/2014	11	8.16
	04/08/2014	10.38	9.40
Alameda – Marítim-Serrería	04/08/2014	10.58	8.18
	04/08/2014	11.1	7.89
	04/08/2014	10.83	8.84
	04/08/2014	10.08	9.56
	04/08/2014	9.82	9.00
	04/08/2014	10	9.96

Table 1~~Table 1~~ shows that there is not a clear correlation between traction energy and travel time. This points out that the differences in the total travel time are mainly due to the variation of the stopping time at each station, while the energy consumption depends on external factors such as the geometric track layout, temperature, degree of occupation of the train and driving style of each driver, among others.

The measuring devices provide the energy consumption measured in the circuit breaker of the train every second as shown in Figure 5~~Figure 5~~. Data presented in Figure 5~~Figure 5~~ represents the training *target* of the neural network.



(a)



(b)

Figure 5. Energy consumption measured in the circuit breaker: (a) the first trip between Marítim-Serrería - Alameda, (b) the six trips made between Marítim-Serrería – Alameda towards Alameda

4. RESULTS

Two different criteria were used to assess the performance of the network and to decide whether the training was successful or not. The first one was the Pearson correlation coefficient (R) between the neural network *output* (modelled energy consumption) and the *target* data (measured energy consumption), which has to be equal or greater than 90% for all the three subsets (training, validation and testing).

The second criterion was the relative Mean Square Error (rMSE) defined as follows in eq. 2:

$$rMSE = \frac{MSE}{Var(Q)} \leq 0.2 \tag{2}$$

Where MSE is the mean square error, and Var(Q) is the variance of the measured consumption data (*target*). This rMSE has to be lower than 20% of the variance of the data for all three subsets (training, validation and testing) [21], in order to control the dependence of the neural network with respect to the specific data used for training.

The process of creation, training, and validation of the neural network was performed using the Neural Fitting Tool, from MATLAB R2014a (The MathWorks, Inc.).

Different tests with combinations of the three input variables previously defined (speed, acceleration and gradient) were performed, in order to identify which of them fits better the energy consumption data. Table 2~~Table 2~~ shows the results for each combination studied considering the criteria defined previously (R coefficient and rMSE) using 12 trips.

Table 2. Results of network training with different variables

	20 NUMBER OF NEURONS						
Model	**Input variables**	**Training**		**Validation**		**Test**	
		R	**rMSE**	**R**	**rMSE**	**R**	**rMSE**
1	Speed	0.49	0.64	0.45	0.65	0.43	0.70
2	Gradient	0.27	0.79	0.26	0.81	0.28	0.72
3	Speed Acceleration	0.86	0.21	0.87	0.20	0.84	0.25
4	Speed Gradient	0.63	0.52	0.60	0.50	0.55	0.58
5	Speed Acceleration Gradient	0.91	0.16	0.90	0.17	0.90	0.16

Table 2~~Table 2~~ shows that if speed is the only input variable, the results are clearly negative, with a rMSE well over 60%. However, if all three variables are used the network satisfies all the criteria; with a rMSE lower than 20% and the R coefficient greater than 90%. Therefore those were the input variables chosen for the analysis because they have an important impact on energy consumption.

Once the input variables were finally defined, the neural network size was determined studying the rMSE values for training and validation, varying the number of neurons in the hidden layer. Figure 6~~Figure 6~~ shows the results, where each value is the average of 20 simulations.

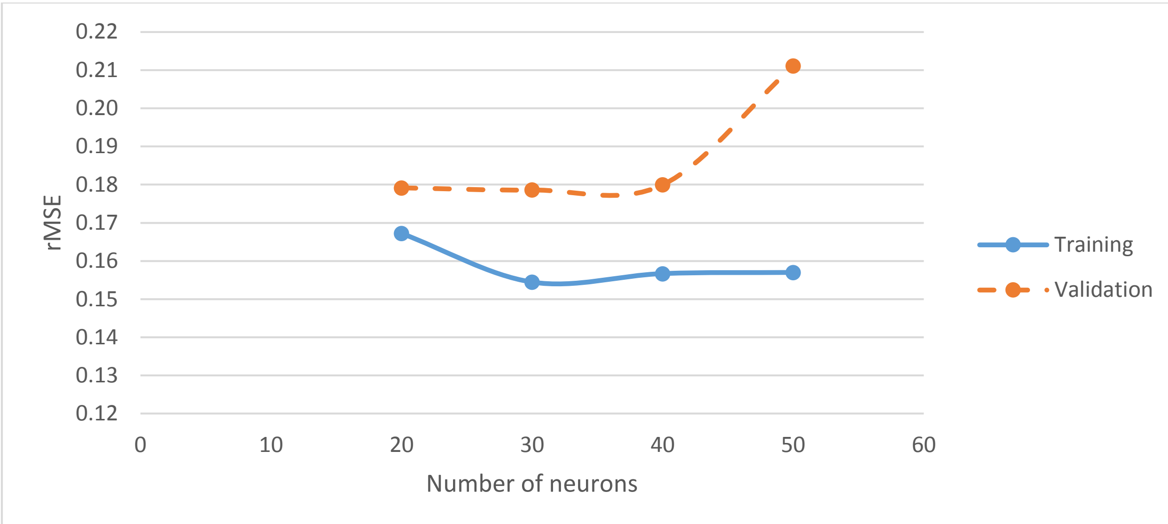
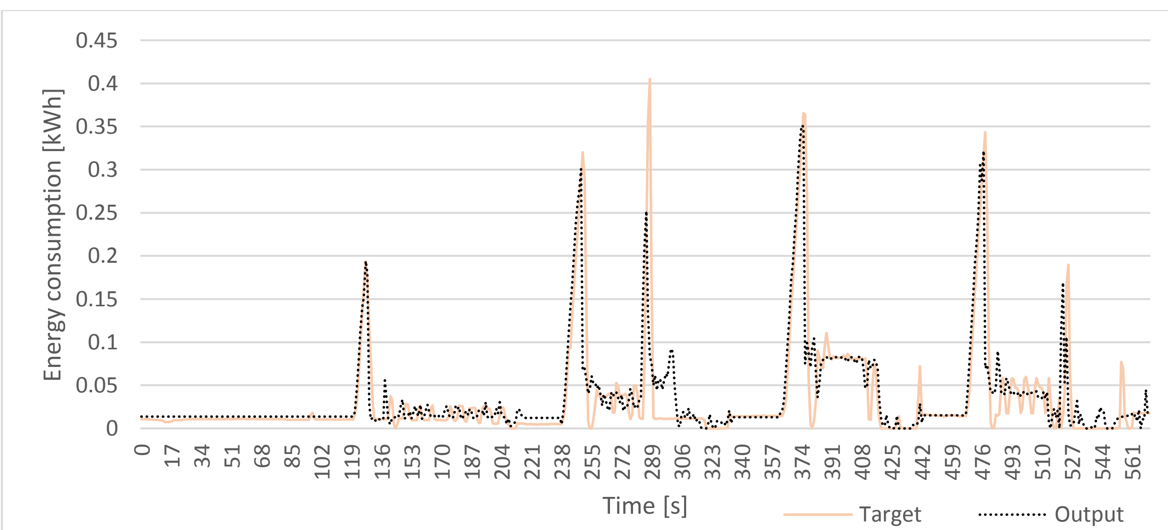


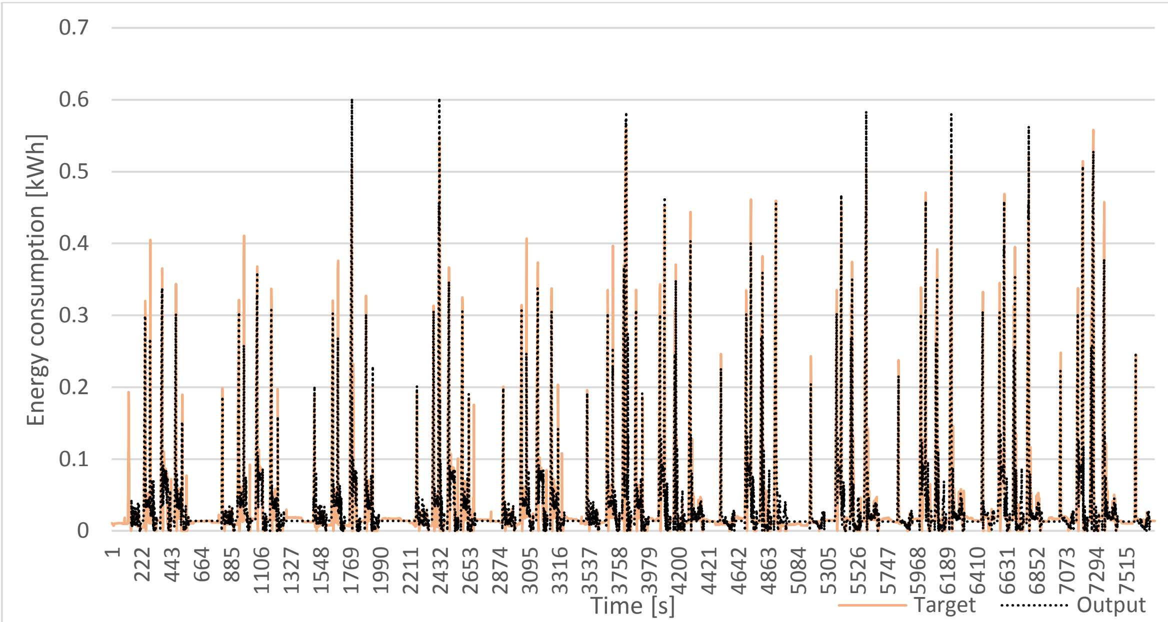
Figure 6. rMSE value by number of neurons in the hidden layer for training (blue) and validation (orange)

Figure 6~~Figure 6~~ shows that the training rMSE decreases while the network size increases due to the higher capacity of the network. The validation rMSE has the same behaviour until reaching a size of 40 neurons. At that point the validation rMSE increases, indicating that the network is experiencing *overfitting*. Therefore, the optimum size for the hidden layer is 30 neurons.

Figure 7~~Figure 7~~ shows the comparison between the measured energy consumption (*target*) and modelled energy consumption (*output*) yielded by the neural network for twelve trips.



(a)



(b)

Figure 7. Comparison between the measured energy consumption (*target*) and modelled energy consumption (*output*) by the neural network: (a) in the first trip between Marítim-Serrería and Alameda, (b) in the twelve trips

5. DISCUSSION

Figure 7~~Figure 7~~ shows that the neural network adjusts reasonably well the energy consumption measured in the circuit breaker, reproducing the peaks due to traction and valleys where the train is coasting. However, the neural network omitted small oscillations in energy consumption, and indeed shows small oscillations and negative peaks that do not correspond to the measurement. This shows that it is possible to refine the training of the neural network, possibly with a post-processing of the *output* data.

In any case, the trained neural network shown, with a size of 30 neurons and three *input* variables (speed, acceleration and gradient), provides a good estimation of the energy consumption of the train, always within the range considered for every variable, so the neural network could be used as a tool to test alternatives in track layout, train operation and driving style, aiming to reduce energy consumption and improve efficiency.

When analysing the global consumption by trip (Figure 8~~Figure 8~~), the average measured energy consumption is found to be 8.66 kWh/km, while the average modelled energy consumption is 8.51 kWh/km, a small deviation of 0.14 kWh/km (1.64%).

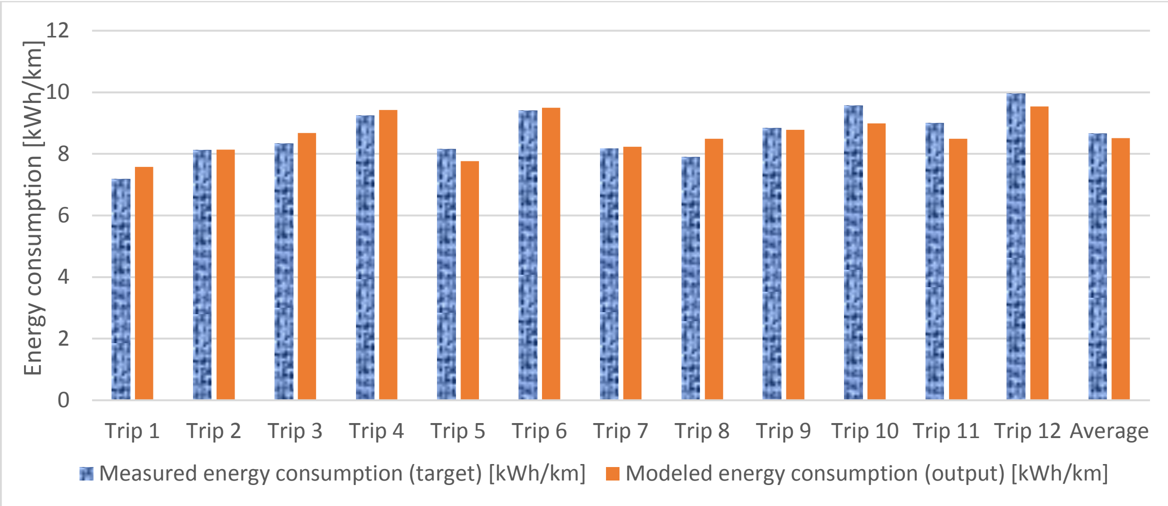


Figure 8. Global consumption by trip

6. CONCLUSIONS

This paper describes the training and validation of a neural network to model the energy consumption of a metro line in the Valencia Metro Network operated by Metro Valencia. In order to do so, real energy consumption was measured using a monitored train operating normally along line 5 of the metro network. This data was analysed and used to train and validate the neural network. Three *input* variables were chosen: speed, acceleration and gradient. These *input* variables combined predict the energy consumed with high accuracy, proving that just one variable cannot explain this phenomenon by itself.

The global consumption shows an average measured energy consumption of 8.66 kWh/km, while the trained neural network estimates an energy consumption of 8.51 kWh/km, a small deviation of 0.14 kWh/km (just 1.64%). The neural network yields a good estimation of the real time energy consumed by the train, including traction peaks and coasting.

The neural network fully trained is a useful tool for studying the energy consumption of the metro system. The advantages of this method lie in its adjustment speed and simulation, and, specially, in the fact that the neural network may function as a virtual laboratory where it is possible to test hypothetical scenarios, modifying variables such as track layout and train driving style in order to reduce the train energy consumption.

The next step of research will be to analyse how the model may be improved if energy recuperation is included, and then to use it to test hypothetical operation and construction scenarios, seeking to minimise the energy consumption of the system.

7. REFERENCES

[1] International Energy Agency, "Key World Energy STATISTICS," 2015. [Online]. Available: http://www.iea.org/publications/freepublications/publication/key-world-energy-statistics-2015.html. [Accessed 13 Junio 2016].

[2] A. García Álvarez, "High speed, energy consumption and emissions," Study and Research Group for Railway Energy and emissions, Madrid, 2010.

[3] K. Kim and S. I. Chien, "Optimal train operation for minimum energy consumption considering schedule adherence," *TRB Anual Meeting Compendium. Transportation Research Board,* 2010.

[4] A. Kosinski, L. Schipper and E. Deakin, "Analysis of high-speed rail's potential to reduce CO2 emissions from transportation in the United States," *TRB Annual Meeting Compendium. Transportation Research Board,* 2011.

[5] A. García Álvarez, "High speed, energy consumption and emissions," Study and Research Group for Railway Energy and emissions, Madrid, 2010.

[6] X. Lai and P. Schonfeld, "Optimizing rail transit alignment connecting several major stations," *TRB Annual Meeting Compendium. Transportation Research Board,* 2010.

[7] M. K. Jha, P. Schonfeld and S. Samanta, "Optimizing rail transit routes with genetic algorithms and geographic information systems," *J. Urban Plann. Dev., 133(3),* pp. 161-171, 2007.

[8] S. Samanta and M. K. Jha, "Modeling a rail transit alignment considering different objectives," *Transportation Research: Part A, 45(1),* pp. 31-45, 2011.

[9] S. Dundar and I. Sahin, "A genetic algorithm solution for train scheduling," *TRB Annual Meeting Compendium. Transportation Research Board,* 2011.

[10] M. Liu, A. Haghani and S. Toobaie, "A genetic Algorithm-based column generation approach to passenger rail crew scheduling," *TRB Annual Meeting Compendium. Transportation Research Board,* 2010.

[11] S. Acýkbas and M. Soylemez, "Coasting point optimisation for mass rail transit lines using artificial neural networks and genetic algorithms," *IET Electric Power Applications,* vol. 2, no. 3, pp. 172-182, 2008.

[12] C. Bishop, Neural networks for pattern recognition, Oxford, UK: Clarendon Press, 1995.

[13] R. Insa Franco, P. Salvador Zuriaga, I. Villalba Sanchis, P. Martínez Fernández, C. García Román, J. D. Pineda-Jaramillo and R. Sánchez Alandí, Monitorización, modelización y simulación del consumo energético en el ferrocarril, Valencia, Valencia: Universitat Politècnica de València, 2016.

[14] J. Fritisch, "Modular neural networks for speech recognition," 1996.

[15] T. Sanger, "Optimal unsupervised learning in feedforward neural networks," *En neural networks,* vol. 2, pp. 459-473, 1989.

[16] C. Chinrungrueng, "Analysis of simple neural networks," Berkeley, 1988.

[17] I. V. Tetko, D. J. Livingstone and A. I. Luik, "Neural network studies, 1. Comparison of Overfitting and Overtraining," *Journal of Chemical Information and Computer Sciences,* no. 35, pp. 826-833, 1995.

[18] M. A. Van Gent, H. P. Van den Boogard, B. Pozueta and J. R. Medina, "Neural network modelling of wave overtopping at coastal structures," *Coastal engineering,* no. 54, pp. 586-593, 2007.

[19] MetroValencia, "www.metrovalencia.es," [Online]. Available: http://www.metrovalencia.es/descargas/pdf/Plano%20de%20red_2011_1200x1760_ok.pdf. [Accessed 15 Octubre 2015].

[20] I. Villalba Sanchis, "Estudio y optimización de la eficiencia energética en la conducción de trenes metropolitanos de la ciudad de Valencia," Valencia, 2015.

[21] J. Molines, "Stability of crown walls of cube and cubipod armoured mound breakwaters," *PIANC E-Magazine*, no. 144, pp. 29-44, 2011.