

XII Conference on Transport Engineering, CIT 2016, 7-9 June 2016, Valencia, Spain

Modelling electric trains energy consumption using Neural Networks

Pablo Martínez Fernández^a, Carla García Román^a, Ricardo Insa Franco^a *

^a*Department of Transport Engineering and Infrastructure, Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia, Spain*

Abstract

Nowadays there is an evident concern regarding the efficiency and sustainability of the transport sector due to both the threat of climate change and the current financial crisis. This concern explains the growth of railways over the last years as they present an inherent efficiency compared to other transport means. However, in order to further expand their role, it is necessary to optimise their energy consumption so as to increase their competitiveness.

Improving railways energy efficiency requires both reliable data and modelling tools that will allow the study of different variables and alternatives. With this need in mind, this paper presents the development of consumption models based on neural networks that calculate the energy consumption of electric trains. These networks have been trained based on an extensive set of consumption data measured in line 1 of the Valencia Metro Network.

Once trained, the neural networks provide a reliable estimation of the vehicles consumption along a specific route when fed with input data such as train speed, acceleration or track longitudinal slope. These networks represent a useful modelling tool that may allow a deeper study of railway lines in terms of energy expenditure with the objective of reducing the costs and environmental impact associated to railways.

© 2016 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of the organizing committee of CIT 2016

Keywords: Railways; Neural Networks; Energy efficiency; Modelling; Metro.

* Corresponding author. Tel.: +34-96-387-7007; fax: +34-96-387-7379.
E-mail address: rinsa@tra.upv.es

1. Introduction

The current context of climatic change and financial crisis points out the need for more efficiency in all aspects of our economy. This is particularly important with regards to the transport sector, which is a key element of our society and greatly influences the global carbon footprint. Railways are inherently efficient compared to other transport means (Barkan, 2007; Garcia, 2007) but there is still ample room for improvement in terms of reducing their energy expenditure in order to make them more competitive and environmentally friendly. However, measuring and evaluating the energy consumption of trains tends to be problematic and usually only average values and estimations are available for railway operators, both in diesel trains (Baumel, 2011) and electric (García and Martín, 2008).

This context points out the need for more comprehensive data and reliable modelling tools, as these two elements are essential to achieve a proper assessment of the energy consumption of trains and develop new solutions and alternatives in order to improve their efficiency. Taking this into account, the paper aims to develop a modelling tool based on neural networks capable of estimating the energy consumed by metro trains during normal operation.

The development, training and validation of the network are described in detail, paying particular attention to the network size and the input variables. The data used for training has been gathered during an extensive monitoring campaign carried out in the Valencia Metro Network (Spain).

2. Materials and methods

2.1. Energy consumption data

The consumption data used to train and validate the neural network was obtained from a comprehensive monitoring campaign carried out in the Valencia Metro Network (Spain) operated by FGV. A Metro Series 4300 train (Vossloh) with four carriages was equipped with three DC voltage and current measuring systems model MSAV-DC (Fig. 1) developed by Mors Smitt in accordance with EN 50463. These devices provide real time values of voltage, current and power.



Fig. 1. DC voltage and current measuring device. Source: Mors Smitt.

One device was installed in the catenary in order to measure the overall energy consumed by the train (including traction and auxiliary systems). Another device measured the energy used by said auxiliary systems (i.e. heating, lights, automatic doors, etc.) and the last one measured the energy consumed by the rheostatic braking system. The energy solely used for traction was then obtained from the difference between the total energy and the energy consumed by auxiliary equipment and braking. Additionally, train speed was also measured through an odometer. The sampling frequency of the DC equipment was 1 Hz while the speed was sampled at 100 Hz.

Once fully equipped with the measuring devices, the train operated normally along lines 1, 2, 3, 5 and 7 of the

Valencia Metro Network between July and October 2014. The data was measured and stored on a laptop placed on board the train cabin. Up to 229 train services were monitored, which accounts for more than 230 hours of data. Data thus gathered was then processed in MATLAB 7.12.0 (The MathWorks, Inc.).



Fig. 2. Line 1 of the Valencia Metro Network. Source: FGV.

For the purpose of this paper, a subset of the whole data measured was chosen; corresponding to train services along line 1 (Fig. 2). This line consist of a central underground stretch (‘Empalme-Sant Isidre’) operated automatically with ATO, and two surface stretches (‘Bétera-Empalme’ and ‘Sant Isidre-Castelló de la Ribera’) operated manually with ATP and FAP. The neural network has been trained only for the subterranean part of the line.

2.2. Neural network development and training

Neural networks are computational models based on the underlying structure of biological nervous systems. A neural network is a made of components (i.e. the neurons) which work in parallel to provide an output value when fed with certain input values. The process of training the network consists on an iterative modification of the parameters which regulate the connections between the neurons until an error function is minimised according to pre-established criteria.

Although there are several possible network structures, one of the most common and widely used to fit functions is the two layer feed-forward network (Bishop, 1995) shown in Fig. 3.

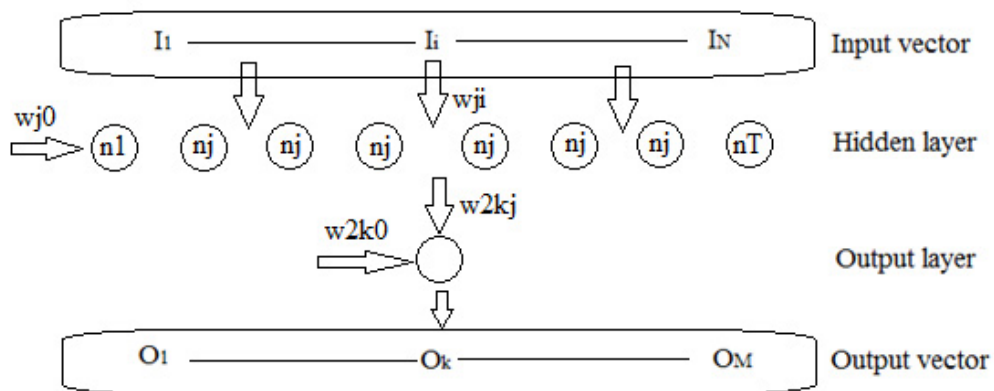


Fig. 3. Two layer feed-forward network.

This is the network structure chosen for this study, whose overall equation (1) is:

$$O_k = \tilde{g}(\sum_{j=0}^M w_{2kj} \cdot g(\sum_{i=0}^N w_{ji} \cdot I_i)) \quad (1)$$

Where O_k is the network output, M is the number of output elements, I_i is the input, N is the number of input variables, T is the number of neurons in the hidden layer, w_{ij} are the weights of the first layer and w_{2kj} are the weights of the second layer (bias values correspond to w_{j0} and w_{2k0}). Using this configuration it is possible to fit non-linear functions to an arbitrary degree of accuracy (Bishop, 1995) and thus it has been extensively used in many applications.

In order to train the network, a Back-Propagation method is used, which aims to minimise the Mean Square Error (MSE) between the network output and the target data, as defined in equation (2).

$$MSE = \frac{1}{N} \sum_{i=1}^N (t_i - o_i)^2 \quad (2)$$

Where N is the number of data, t_i is the target output and o_i is the network output. The Levenberg-Marquardt Algorithm of minimisation, combined with an Early Stopping Method, was used to effectively carry out the training. Early stopping consists, as its name suggests, on stopping the training before the MSE is completely minimised. This is done by dividing randomly the data in three subsets: a larger one used for Training and two smaller (Validation and Test) used to control the process. This training methodology is efficient and reliable (Bishop, 1995) and ensures generalisation while avoiding overfitting i.e. an excessive learning of the network, which becomes affected by the errors and noises of the specific data used for training.

Two criteria were defined to assess the goodness of the training process. The first one was to require the Pearson correlation coefficient (R) between target and output to be 0.9 or greater. The second one was that the relative MSE (i.e. the ratio between the MSE and the variance of the target data) must be lower than 0.2. This means that less than 20% of the network variance estimated by the MSE is due to the training data variance (Molines, 2011).

An element of the network required special attention during training: the network size i.e. the number of neurons in the hidden layer. A larger network is more powerful and capable of modelling complex trends, but too large a network may experience overfitting even after training with early stopping. Therefore, a balance between complexity and capacity must be achieved. This was done by testing different networks sizes and comparing the MSE obtained for the training and the validation datasets. When the latter started to rise with the addition of more neurones (while the former kept decreasing), an optimum size was reached (Van Gent et al., 2007).

3. Results and discussion

Table 1 shows the results obtained during the training process for different combinations of input variables, namely train speed and acceleration and the track longitudinal slope and radius of curvature, taking into account the two criteria previously defined (R coefficient and $rMSE$).

Table 1. Training, validation and test results for different input combinations.

Input variables	Training		Validation		Test	
	R	rMSE	R	rMSE	R	rMSE
Speed	0.35	0.872	0.33	0.883	0.42	0.874
Acceleration	0.74	0.460	0.76	0.441	0.69	0.488
Speed & acceleration	0.88	0.230	0.91	0.185	0.89	0.201
Speed, acceleration & slope	0.89	0.208	0.88	0.232	0.89	0.217
Speed, acceleration & radius	0.89	0.220	0.88	0.236	0.88	0.192
Speed, acceleration, radius & slope	0.90	0.190	0.90	0.200	0.90	0.159

As the table shows, using the speed or the acceleration as a single input variable gives a rather poor result, although the latter seems to provide more information to the network. Combining both variables gives a result that almost fulfils both criteria, but there is still room for improvement. By adding information regarding the track geometry (i.e. the longitudinal track slope and the radius of curvature) the network finally reaches a 0.9 value of the R coefficient and an rMSE below 20% for the three stages of development. Therefore, the combination of the four variables is chosen as input for the network.

Regarding the network size, Fig. 4 shows the evolution of the MSE for training and validation as the number of neurons in the hidden layer increases. Notice that each marker in the figure represents the average MSE after twenty training processes. As expected, the training MSE tends to decrease as the network complexity increases, while the validation MSE decreases at first and then tends to rise, thus pointing out the size after which the network is overfitting. Therefore, the optimum size network is set to 15 neurons in the hidden layer.

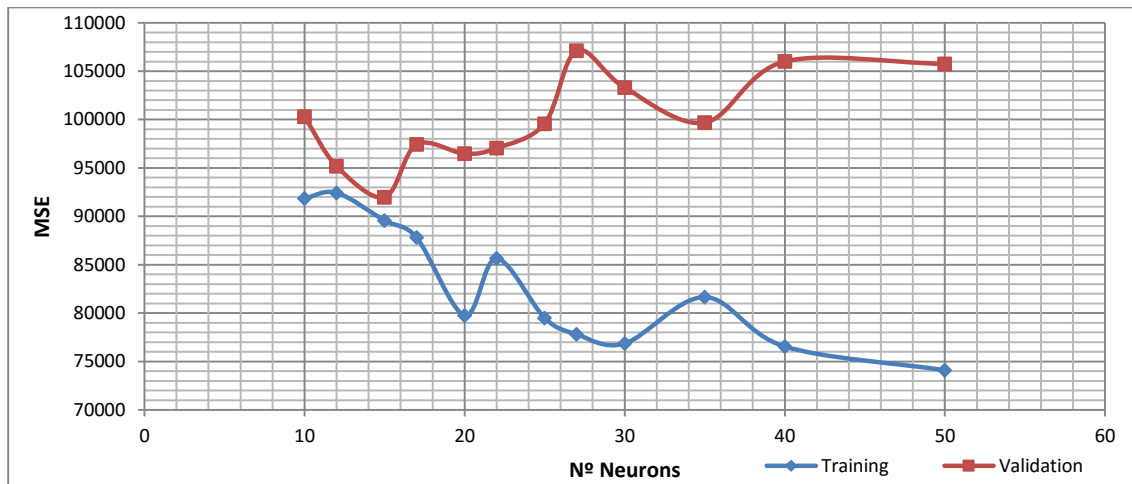


Fig. 4. Training and validation MSE vs number of neurons.

Once the network main features have been settled, the training process has been completed. Fig. 5 shows the comparison between network output and target data.

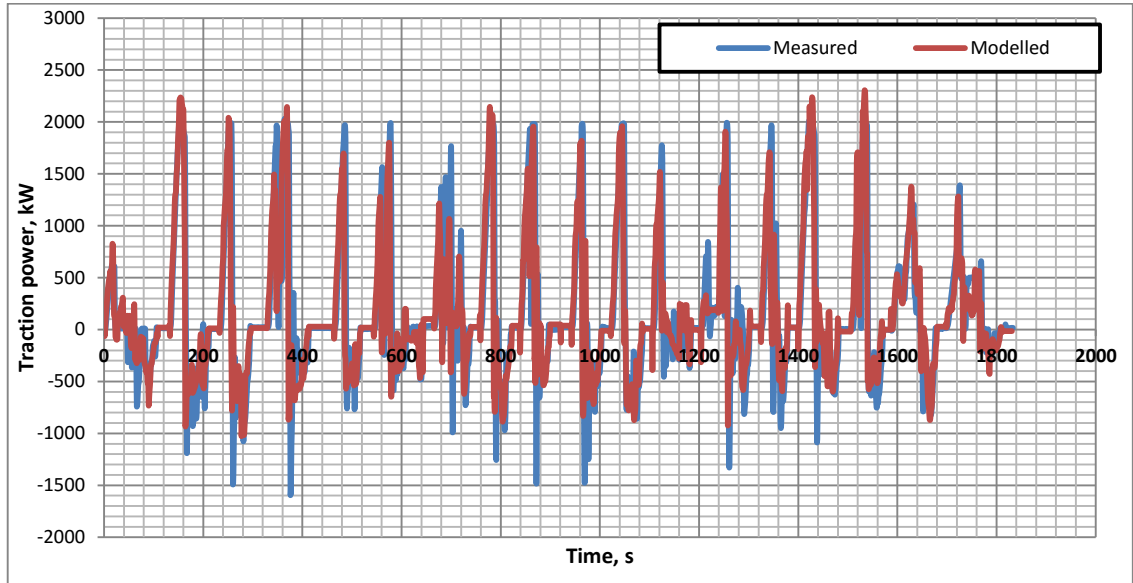


Fig. 5. Network output vs target data.

As Fig. 5 shows, there is a good agreement between the network and the data. The overall energy consumption measured for that service is 83.852 kWh, while the network yields a value of 83.131 kWh.

As a further test, the trained network was used to model the energy consumption of a different service along the same line (but in the opposite direction), and the output was compared to the measured energy. This comparison is shown in Fig. 6.

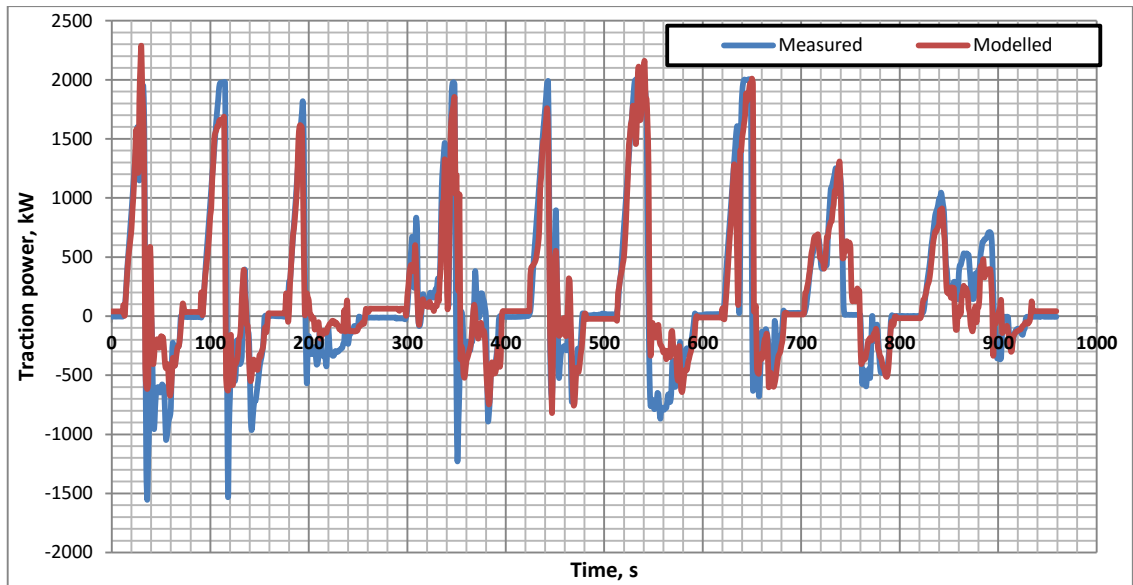


Fig. 6. Energy consumption modelled vs measured.

In this case the agreement is not as good, as the network apparently underestimates the negative peaks of power, which correspond to braking events and energy generated back to the catenary. For this reason, the network yields a total energy consumed equal to 44.579 kWh while the energy actually consumed was 38.974 kWh. Nevertheless, both criteria are still fulfilled as the R coefficient in this case is 0.9 and the rMSE is 0.192. This points out that, although the network has been fully trained and provides a good estimation of the energy consumption, there is still room for improvement.

4. Conclusions

In order to develop a useful tool to estimate the train energy consumption of electric trains, a neural network has been built, trained and validated using real consumption data measured in line 1 of the Valencia Metro Network (Spain). Different input variables were considered, including train speed and acceleration and track longitudinal slope and radius of curvature.

The training process was carried out by means of a combination of the Levenberg-Marquardt Algorithm of minimisation and the Early Stopping Method. The optimum network size was also assessed and found to be 15 neurons in the hidden layer.

The network was properly trained according to predefined criteria, and shows a good agreement with the target data. As an additional test, the network output was compared with another subset of measured data and once again it provided a good estimation of the energy consumption, although a slight underestimation of negative energy peaks was observed, pointing out the need for further refinement of the network with additional data.

Acknowledgements

The authors wish to thank Ferrocarrils de la Generalitat Valenciana (FGV) for their permission and help during the monitoring campaign. Project funded by the Spanish Ministry of Economy and Competitiveness (Grant Number TRA2011-26602).

References

- Barkan, C., 2007. Railroad transportation energy efficiency. University of Illinois, Chicago, USA.
- Baumel, P., 2011. Measuring bulk product transport fuel efficiency. *Journal of Transportation Research Forum* 50(1), 79-88.
- Bishop, C. M., 1995. *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford, UK.
- García, A., 2007. Consumo de energía y emisiones del tren de alta velocidad en comparación con otros modos (energy consumption and emissions from high-speed trains compared to other transport means). *Anales de Mecánica y Electricidad* 84, 26-34.
- García, A., Martín, M. P., 2008. Cuantificación del consumo de energía eléctrica del ferrocarril español (Quantification of the electric consumption of Spanish Railways). *Fundación de los Ferrocarriles Españoles*, Madrid, Spain.
- Molines, J., 2011. Stability of crown walls of cube and cubipod armoured mound breakwater. *PIANC E-Magazine* 144, 29-44.
- Van Gent, M. R. A., Van den Boogaard, H. F. P., Pozueta, B., Medina, J. R., 2007. Neural network modelling of wave overtopping at coastal structures. *Coastal Engineering* 54, 586-593.