Prediction of Wheel-Rail Forces, Derailment and Passenger Comfort using Artificial Neural Networks

Gualano L., Iwnicki S.D., Ponnapalli P.V.S. and Allen P.D.
Rail Technology Unit
Manchester Metropolitan University

Summary
After briefly introducing the conventional techniques used for modelling wheel-rail interactions, mostly based on the construction of the system's equations of motion, Artificial Neural Network (ANN) modelling techniques are introduced, followed by a discussion of how these have been applied to model wheel-rail interactions. An analysis of ANN models efficiencies is then presented, aimed to assess aspects of the ANN models most relevant in the context of wheel-rail interaction modelling. Finally results and conclusions from the application of this efficiency analysis to some of the obtained ANN models is discussed.

Abstract
Over the past two decades, Artificial Neural Network (ANN) techniques have been used in many fields of research, due to their high speed and improved robustness and failure tolerance capabilities compared with conventional modelling approaches. In applications such as railway vehicle dynamics, discrete scenarios of vehicle/track interactions are currently modelled using computer packages such as Simpack, ADAMS/Rail, Vampire or Medyna, which use multi-body techniques to accurately model different aspects of rail vehicles and tracks such as derailment or passenger comfort.

The authors have developed new ANN techniques which make it possible to achieve ANN model accuracies comparable to those of multi-body techniques. A particular ANN structure has been designed with the aim of simplifying the training of ANNs with long training data sets. This is a Recurrent Neural Network (RNN) structure characterised by an optimised feedback technique which requires very little computational power. This novel structure and other more conventional RNN structures have been trained and tested and the processing times compared. The efficiency of the novel ANN structure in modelling a number of vehicle types, from passenger to friction damped freight vehicles, has been validated against commonly used techniques and also with a newly designed method, which consists of a combination of statistical functions applied to assess different aspects of the ANN models responses.

The novel ANN structure appears to be much faster than other structures commonly used for non-linear system modelling and adequate for the purposes of rail vehicle modelling. Compared to conventional ANN validation techniques, such as the mean square error and the cross-correlation function analysis, the novel assessment technique results in a more accurate quantification of the error terms and therefore, in a safer assessment which may be focused on aspects of the ANN model responses which are relevant in the context of railway engineering. It is possible that this novel approach to designing efficient ANN's could be applied to a wide variety of scientific fields involved in the application of ANN techniques.

Introduction

In the field of railway engineering, vehicle models are currently implemented in software packages for design and assessment purposes. These vehicle models are mostly based on the use of the multi-body technique which relies on the derivation of complex equations of motion to describe the dynamics of the considered system. Current multibody models can reach high levels of accuracy, but they require lot of computational power and modelling time can be high. In this context, ANNs [1 – 4] represent a valid alternative in modelling vehicle dynamics and track interactions. ANN techniques could contribute to railway technology by speeding up the modelling process and also allowing more variables to be included into the models, improving track measurement and maintenance techniques or constituting a more cost effective means of measurement for implicit aspects of vehicle-rail dynamics.

Attention in this paper will be focused on ANN modelling of wheel-rail interactions. Some of these interactions have already been modelled with acceptable accuracy [5 – 8], such as the vertical forces present at the wheel-rail contact patch. Other more complex aspects, such as the lateral forces or the derailment coefficient are more difficult to model with a level of accuracy sufficient for ANNs to be used as stand-alone tools in safety-critical applications.

The main cause of this problem lies in the *training sets* which the ANNs use to *learn* the dynamics of the target systems: the learning process of an ANN consists of the application of a *training algorithm* which uses a subset (called a training set) of possible target system input-output combinations. This subset needs to contain most of the dynamics (in terms of amplitude and frequency) of the target system. To achieve such a training set, the availability of another non-ANN model of the same target system is often necessary. Such a model would be given a set of inputs designed to excite the model across most of its dynamic range, so that a training set can be constructed from its inputs and outputs.

This reduces the usefulness of ANNs as substitute methods for conventional modelling techniques. An alternative is that of using a set of data measured from a real system, but again, this has to include most of the dynamics of the system. This usually leads to very large datasets, resulting in an excessive amount of computational power being required to train the ANNs. An ANN structure is presented in this work, aimed to overcome this limitation.

Another problem with the application of ANNs to railway engineering is the difficulty of accurately assessing the ANN performance. Currently, performance assessment techniques for ANNs are based on general probability methods which do not contribute to an increase in confidence of ANN applications in railway engineering, mainly because of their deficiency in assessing specific aspects of the ANN performance. A technique for improving these performance validation methods, to assess more specific aspects of ANN performance, is also presented.

Modelling Rail Vehicle Dynamics

When constructing a mathematical model of a railway vehicle, it is usually considered as a multibody system. Once the multibody system has been defined, the equations of motion for the system can be derived. These are usually differential equations, expressed in matrix form.
Once the equations of motion for the system have been obtained, these can be analysed and the system behaviour can be simulated. This is done by looking at the solutions of the equations for given inputs and desired outputs. Performance criteria are used as guidelines to prevent vehicle derailment, vehicle and track damage or passengers discomfort.

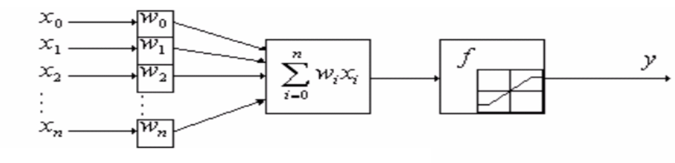
In this work, derailment prediction and ride quality have been the main issues studied. The risk of derailment can be predicted by calculating an L/V ratio, where L and V are the lateral and vertical wheel-rail contact forces respectively. When this ratio exceeds a certain maximum (often set at around 1), it is considered that there is high risk of derailment. The ride quality analysis is usually performed on the basis of two ranges of horizontal and lateral vibrations respectively, to which humans are most sensitive ([0.2 – 10]Hz for vertical vibrations and [0.2 – 2]Hz for lateral vibrations).

Software packages are available for the simulation and analysis of the railway vehicle system and, in most of the cases, these use the methods described above. The drawback however is, the computational time that these techniques require. Usually many approximations are made to reduce the simulation times. In this context, ANN techniques represent a valuable alternative due to their speed and also the fact that the increase in required computational power which comes when adding more variables to the models is not significant when using ANNs.

Neural Network Models

An *Artificial Neural Network* (ANN) is a system that uses a finite number of *neurons* (processing elements which apply transfer functions on their inputs) connected to each other in different ways in order to perform operations.

A neuron can be defined as an operative cell (also named by mathematicians as a *Processing Element - PE*) whose main task is to execute an **activation function** on its inputs.
A neuron accepts a certain number of inputs and to each input it applies a *weighting* factor which indicates the contribution of each input to the neuron output. Once the inputs have been weighted, a *net*-input is calculated (normally by adding or multiplying all the weighted inputs together) and the result is forwarded to an output function (*activation function*) from which the output of the neuron is calculated. This process can be described mathematically by the block diagram shown in Figure 1.



$x_0, x_1, x_2, ..., x_n = inputs$
$w_0, w_1, w_2, ..., w_n = weights$
$y = output$
$f = activation\ function$

Figure1

An additional component of a neuron which must be considered before a working artificial neuron can be achieved is an input, often referred to as a *threshold* or *bias*. The demonstration of the operation of this input is beyond the scope of this paper and here it will only be said that its value is assumed to be a constant input +1 and therefore, only the weight has to be determined.

Because of its ability to learn and generalise, an ANN can theoretically approximate any system. However, there is still the need to choose the most appropriate ANN structure (the way the different neurons are connected to each other and the different transfer functions that these neurons implement) which is dependant on the type of operation which has to be performed by the ANN.

For the purpose of modelling wheel-rail interactions, *Recurrent Neural Networks* (RNNs) have been used, which are ANNs with feedback connections between the outputs and the inputs of the ANN neurons. This allows data from previous time steps and thus dynamic characteristics to be included.

Training an ANN

The training algorithms used in this work are based on minimisation of the error between the ANN outputs and the desired outputs. Every time a set of inputs is presented to the ANN to train, this will produce a set of outputs accordingly. The error between the ANN outputs and the desired outputs is calculated and the weights of the neurons of the ANN are updated to minimise the error. The process is then repeated until the ANN gives outputs within a tolerable range of error. The number of updates which the training algorithm performs on the weights of an ANN is defined as number of *epochs*.

Normally, only the desired output for the output neurons of an ANN are known, but this problem is overcome by using techniques which will approximate the error of each neuron of the ANN on the basis of the error of the output neurons, activation functions, and the weights of the ANN.

Different training algorithms update the ANN weights on the basis of different formulas. The training algorithm used in this work is the Levenberg-Marquadt, which is an approximation of Newton's method used to find the minima of the error curve at the ANN output with respect to the ANN weights.

Neural Network models of Wheel-Rail Interactions

It has already been demonstrated that (RNNs) are suitable structures to model wheel-rail dynamics [5 – 8]. RNNs are a class of ANNs and are characterised by backward connections (also called *feedback connections*) between outputs and inputs of neurons in the hidden and output layers. The feedback connections allow RNNs to model time-series systems.

However, the drawback with RNNs is that the feedback connections significantly increase the amount of data that an RNN has to process. This limitation, in conjunction with the problem of using large training sets from measured input-output data, can make the RNN modelling of the target system very tedious.

However, in this work it was found that the amount of feedback connections in a RNN can be optimised to a single set of connections between the output of the output neurons and the inputs of the neurons of the input layer, as shown in Figure 2.

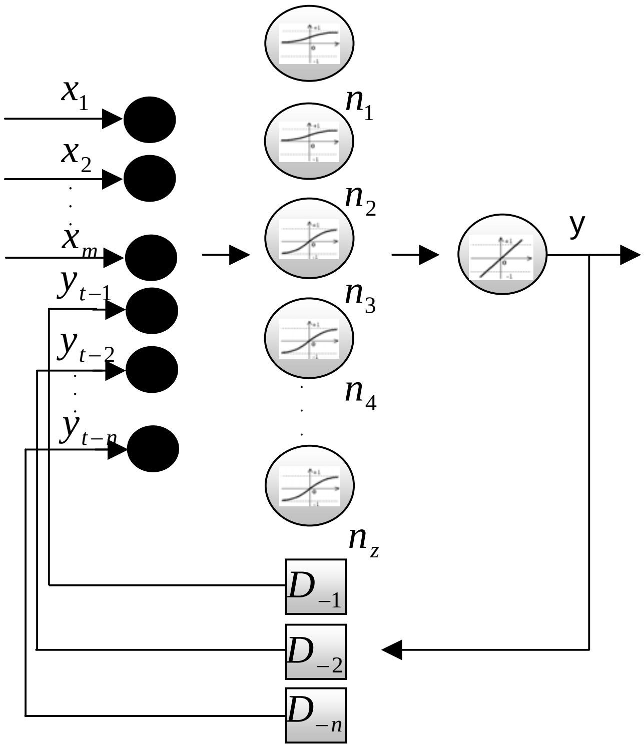


Figure 2

For many vehicle-track interactions the network of Figure 2 can be used and a small number of neurons in the hidden layer is sufficient. However, if the target model is highly non-linear, small modifications can significantly improve performance, that is, including neurons with different non-linear activation functions [4] in the hidden layer (as shown in Figure 2) or adding an extra hidden layer to the ANN.

The external inputs supplied to the model were track vertical and lateral irregularities, cross-level irregularities, curvature irregularities and gauge variations.

Evaluation of the Neural Network Performance

When comparing the response of an ANN model with that of the target system, the most commonly used function is the *mean-square-error* (MSE) function, defined as shown in equation 1.

$$mse = \frac{1}{Q}\sum_{k=1}^{Q}(t(k) - a(k))^2 \qquad 1$$

where Q is the number of samples of the target response (equal to the number of samples of the corresponding ANN response), $t(k)$ is the output of the target model to a certain input $u(k - 1)$, and $a(k)$ is the output of the ANN to the same input $u(k - 1)$.

However, there are aspects of the ANN behaviour which cannot be assessed by using the MSE function (unwanted peaks, phase errors, lack of frequency content, etc...). Hence, other test functions have been implemented which can analyse other factors of the ANN responses.

Power Spectral Density analysis
A Power Spectral Density (PSD) describes how the power (or variance) of a time series is distributed with frequency. A PSD analysis is performed by comparing the PSD of the ANN response to a certain input with the PSD of the response of the modelled system to the same input. One way of comparing the PSD of the two signals is by using the MSE. This has the advantage of minimising the amount of data when several ANNs are assessed.

In the context of prediction of ride quality [9, 10], it was found more efficient to compare the ANN output PSD and target system output PSD within a the range of frequencies to which humans are most sensitive.

Cross-Correlation (XCORR) Functions Analysis
The Cross-Correlation function is a measure of similarity of two signals, commonly used to find features in an unknown signal by comparing it to a known one.

In the context of the work here presented, ANN and the target system responses to the same input dataset have been compared.

First, it was necessary to modify the two responses in a way that the function autocorrelation (the cross-correlation of a signal with itself) of each of them would be defined within the interval [0 – 1]. An example is given in Figure 3.

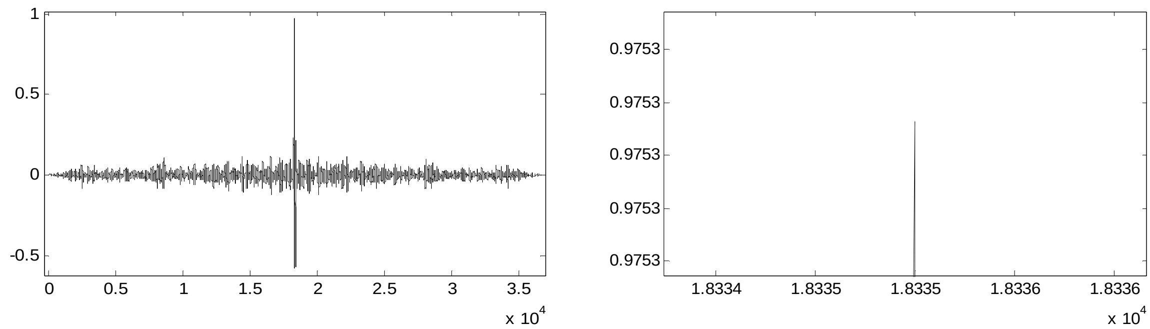


Figure 3

If the autocorrelation of each of the cross-correlated signals is defined in the interval [0 – 1], it can be shown that the nearer the peak of the cross-correlation function is to 1, the more similar the two compared responses are in terms of their amplitudes.

This normally gives a more practical value compared to the MSE as results from different tests can be compared.

The cross-correlation also analysis can also assess the phase error between the two compared responses. This is done by analysing the position of the peak of the cross-correlation graph with respect to the x-axis. If N is the number of samples of the ANN response (and the target response), if the x-axis value x_p, corresponding to the peak of the XCORR curve is the same as N, then there is no phase error between the two signals. When this is not the case, the phase error is proportional to the distance on the x-axis of x_p from N.

In the case of Figure 3, each of the compared signals is composed of 18334 samples. As The peak value falls on the 18335th sample this means that phase error of 1 sample occurs.

Analysis of Variance (ANOVA)
The ANOVA test determines the probability that the similarities between two sets of data are not due to chance. This is done by analysing the variances of two datasets in order to decide if their means are significantly different This is also to say that, if there is a significant error between two datasets, the ANOVA test would detect it.

The theory of the ANOVA technique is beyond the scope of this paper and it is well documented in [11]. It is only stated here, that if the ANOVA value is below 0.05, a common assumption is that the two compared datasets are considered significantly different.

Maximum Error (ME)
The purpose of the maximum-error function is to detect peak errors between the ANN and the target system responses, which may be outside a tolerance range defined by the user. A software routine was used to compare the ANN and target system output samples one by one and store the maximum error found.

A maximum error found is in many cases irrelevant (for example, due to a very short duration, its energy is not sufficient to affect the behaviour of the model).

Therefore, an analysis of the maximum error needs to be performed by looking at the area surrounding the error in order to assess the impact of the detected error on the overall performance of the model.

In this work, the Analysis of Variance (ANOVA) and the cross-correlation functions have been used as a means of assessment of the area surrounding the maximum error.

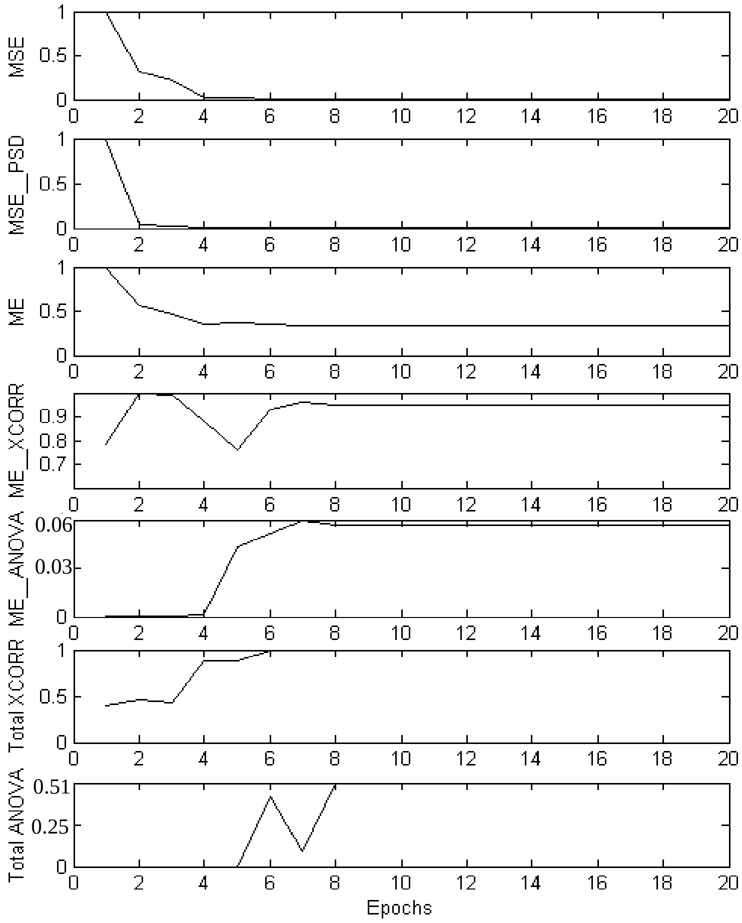
Results
The ANN validation methods presented above have been applied to assess the performance of vehicles modelled with the RNN structure, also introduced in this paper. Figure 4 shows an example of the results obtained for an ANN modelling the lateral forces of a 4-axle passenger vehicle with Linear Stiffness and Damping.

The data inputs which can be presented to an ANN can be classified as **seen** or **unseen** data. Seen data is the data with which the ANN has been trained, while unseen data is data that the ANN has not seen during the training phase.

Predictably, ANNs perform much better in the prediction of system response to seen data than to unseen data. For this reason, all results presented here are obtained by performing validation on **unseen** data.

Figure 4 was obtained by validating an ANN after each training epoch and recording the validation results. It can be seen from the figure, that at the 7^{th} epoch, the ANN already shows results which indicate a good performance (Table 1 shows the results recorded at the 7^{th} and 8^{th} epochs). The maximum error detected is large (about 0.3), considering that the ANN has been trained with data normalised in the interval [-1 1], but both the ANOVA and the

XCORR indicate that the error is not significant and does not noticeably affect the overall behaviour of the model.



	7 Epochs	8 Epochs
MSE	0.001153	0.0011511
MSE_PSD	6.86E-05	9.76E-05
ME	0.337	0.341
ME_XCORR	0.95	0.94
ME_ANOVA	0.060	0.056
TOT_XCORR	0.987	0.999
TOT_ANOVA	0.093	0.51

Table 1

Figure 4

Figure 5 shows the target system and the ANN response to a set of unseen data.

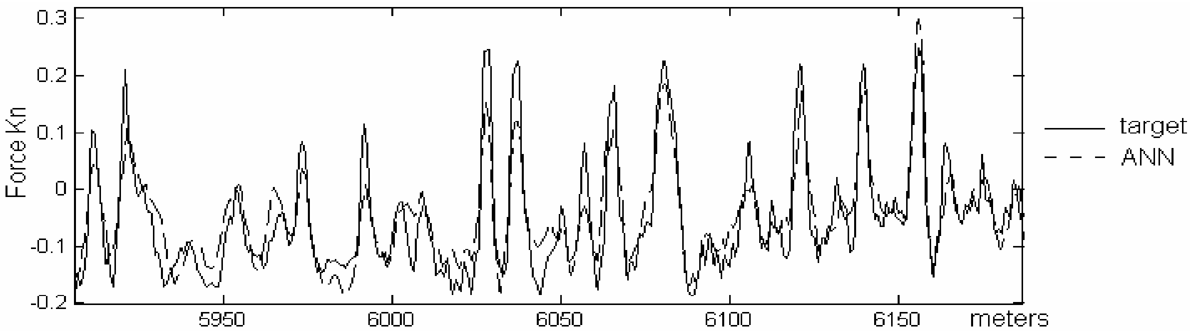


Figure 5 – Target and ANN responses after 7 training epochs

The ANN was composed of 15 hidden neurons (with a tangent sigmoid transfer function) and using a feedback signal describing 6 samples of past history, each sample being recorded at a distance of 0.4 meters.

Similar results have been obtained for various other wheel-rail interaction systems modelled with similar ANNs.

Conclusions

Figures 4, 5 and Table 1 provide a more complete description of the level of accuracy of the ANN model than that of the single MSE function. The MSE is still used, together with the MSE of the PSDs of the compared responses and their XCORR, to assess the overall performance of the ANN. However, the ANOVA test is also used to determine the probability that the similarities between the two responses are not due to chance, that is, an analysis of the reliability of the ANN model. The test results show that the ANN can be considered reliable after 7 training epochs.

Also, the ME between the two responses has been determined and assessed (by performing XCORR and ANOVA tests using a 2m window surrounding the point where the maximum error was detected) for its impact on the overall behaviour of the system. The test results show that the ME can be considered irrelevant after 7 training epochs and the 2m window surrounding it reaches an acceptable level of correlation with the target response after the 6th training epoch.

It can also be seen that the MSE and the MSE of the PSD indicate that the ANN has an acceptable performance after the 4th training epoch. However, Figure 6 clearly shows that only 4 training epochs are insufficient in achieving an efficient ANN model. The XCORR, ANOVA and ME analysis indicate that an optimal number of training epochs is approximately 7 or 8, which produces the ANN model whose response is shown in Figure 5.

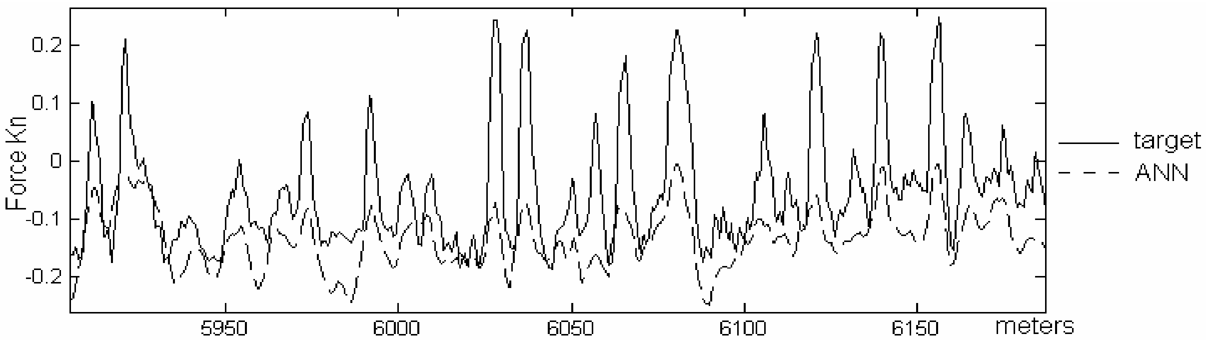


Figure 6 – Target and ANN responses after 4 training epochs

It is therefore concluded that the error analysis proposed here can better describe the ANN behaviour and it also constitutes a more accurate assessment than the MSE function.

With respect to the ANN models developed here, this error analysis cannot determine with certainty if an ANN is safe to use for wheel-rail interaction modelling, as the assessment is based on statistical results.

However, the same error analysis could be performed by comparing conventional wheel-rail models (such as those based on multibody techniques) with real measured data. In this context, if similar results to those of the ANN models are achieved, then it may be possible that ANNs could be considered safe to use. Therefore, the proposed error analysis constitutes a means of validation for the suitability of ANNs for applications like wheel-rail interactions modelling and perhaps many other applications of ANNs.

Finally, a most common RNN (more precisely Elmann ANN) has been used to model the same target system. Apart for the number of feedback connections (one set of feedback from the output to the input of each neuron), all the other specifications of the RNN were the same as for the previous ANN. This was carried out with the aim of comparing the speed of training and testing of the two ANNs, so that the speed improvement of the ANN structure proposed here could be quantified. The test could not be completed as the Elmann RNN was causing a computational error after an extreme length of time taken for completion of the first training epoch.

It is therefore concluded that the feedback technique proposed in this paper is more efficient than other most common RNN feedback techniques.

Future Work

The proposed ANN validation method, can also be applied to assess the accuracy of conventional models (such as those based on multibody techniques), so that a record of results for different vehicles could be constructed. This record would then be useful as a benchmark to finally validate the suitability of ANN techniques for applications such as wheel-rail interactions modelling.

This method could also be adapted for application to other fields of research, especially where the suitability of ANN techniques to a certain application has to be assessed.

References

1. W. McCulloch, W. H. Pitts, “A Logical Calculus of the Ideas Immanent in Neural Nets”, Bullettin of Mathematical Biophysics, Volume 5, 1943
2. K. Hornik, M. Stinchcombe, H. White, “Multilayer Feedforward Networks are Universal Approximators”, Neural Networks Vol. 2, Pergamon Press plc. 1989
3. L. Fausett, “Fundamentals of Neural Networks”, Prentice Hall, 1994
4. K.C. Ho, P.V.S. Ponnapalli, M. Thomson, “A novel MULTI-TYPE architecture for FANNs”, IEE 5th International Conference on Artificial Neural Networks, Cambridge, England, 7-9th, July 1997
5. H. Parkinson, S. Iwincki, “An Intelligent Track Monitoring System” AIC conference on Infrastructure Management, London, February 1999
6. S. Iwnicki, J. Stow, H. J. Parkinson, “Assessing railway Vehicle derailment potential using Neural Networks”, I.Mech.E. Conference Transactions, Fault free Infrastructure. Pub Professional Engineering Publications Ltd, 1999
7. S. Nefti, M. Oussalah, “A Neural Network approach for Railway Safety Prediction”, Proceedings og the IEEE International Conference on Systems, Man and Cybernetics, pages 3915 – 3920, Vol. 4, 2004
8. M. El-Sibaie, “Computer Model Developed to Predict Rail Passenger Car Response to Track Geometry, US Department of Transportation, Federal Railroad Administration, RR000-04, October 2000
9. International Standards Organization, ISO 2631, “Guide for the evaluation of human exposure to whole body vibration”, 1974
10. M. G. Pollard, N. A. Simmons, “Passenger comfort, the role of active suspensions”, Proc. IMechE Vol. 198 No. 135, 1984
11. Hogg, R. V., J. Ledolter, “Engineering Statistics”, MacMillan Publishing Company, 1987