

# Trip generation: comparison of neural networks and regression models

F. Tillema, K. M. van Zuilekom & M. F. A. M van Maarseveen  
*Centre for Transport Studies, Civil Engineering, University of Twente,  
The Netherlands*

## Abstract

Modelling the number of trips produced by the inhabitants of a zone, the *trip generation*, is complex and highly dependent on the quality *and* availability of data. It seems almost impossible to model/forecast the number of trips a person makes without adequate amounts of data. Transportation engineers are commonly faced with a question that is related to this topic; how to perform reliable trip generation with *scarce and expensive* field data. It is therefore interesting to find the method that gives the best results with the smallest data sets. This paper deals with trip generation and explores the performance of *neural networks* and commonly used regression models. This research tries to answer the question whether neural networks can out-perform traditional regression methods or not. The neural networks are tested in two situations with regards to the data availability; (i) data is scarce; and (ii) data is sufficiently at hand. Synthetic households, generated using travel diary data, are the basis for the research. These households are divided over a zone in varying complexities, from homogeneous without statistical deviation on the household characteristics to inhomogeneous with a deviation on the household characteristics. The use of synthetic data, without unknown noise, gives the opportunity to clearly determine the impact of complexity on the forecasting results. The question of whether neural networks can be used in trip generation modelling is answered positively. However, neural networks do not overall out-perform classical regression models in situations where data is scarce. The advantages over regression models are negligible.

*Keywords: trip generation, neural networks, regression model, synthetic data.*



## 1 Introduction

Transport planning, described in the classical 4-step model, is characterized by the dependence on data. The trip generation, that is number of trips that a person or household makes in a certain time period, is highly complex and dependent on many different measurable and non-measurable factors. The trips made by one person for example, can be dependent on measurable factors like characteristics of employment workplace and housing. In addition personal, non measurable, factors like preferences and habits are also to a large extent of influence on the trip generation. These factors are difficult to measure, contrary to the first set of characteristics. The challenge faced by engineers, is that they need information on as many factors as possible in amounts as large as possible for good forecasting keeping in mind the expense of collecting large amounts of data. Therefore it is important to find ways to optimise the use of data, since there is a large interest in accurate trip generation estimations.

Errors generated during the trip generation phase propagate through till the level of evaluating transport policies. This makes reliable transport planning dependent on accurate trip generation modelling. Currently used techniques, like amongst others classical regression models, try to use limited amounts of expensive data in order to forecast the trip generation. Since the beginning of the nineties, neural network models were introduced as alternatives for traditional (statistical) modelling approaches. Recent literature gives an insight into the opportunities of using neural networks to model spatial interactions. Openshaw and Openshaw [1] for example give their opinion on the advantages of using neural networks in geographical/transportation analysis. An eye-catching conclusion is the better performance of these models compared to more traditional models.

Several studies have explored the usefulness of neural networks in the context of trip generation modelling or strongly related topics and subscribe the conclusions of Openshaw and Openshaw [1]. Al-Deek [2] gives an example of the use of neural networks in truck trip generation in a harbour. Huisken and Coffa [3] conduct an extensive research into trip generation. Dantas *et al* [4] present a strategic planning model for urban transportation analysis. Conclusions are clear: neural networks are suitable models that are able to outperform the classical models.

So, literature survey shows that artificial neural networks are successfully used as data analyzing techniques in a trip generation context and the conclusions seem quite clear: neural networks are able to out-perform more traditional regression models. However neither of the mentioned studies gives conclusions on performance of trip generation on a household trip level. Al-Deek [2] shows the possibilities of trip generation in a harbor transport setting. Huisken and Coffa [3] only work with aggregated variables on a zonal base. They use known aggregated zone based data to calibrate the neural network. The calibrated model is then used to forecast the number of trips in other zones. The results of these studies are therefore not to a large extent generalisable. This complicates drawing conclusions on the performance of neural networks in trip



generation in general. This research tries to answer the question whether neural networks can really out-perform traditional methods.

The approach differs from the other research in several respects. In its simplest form a neural network is nothing more than a self-calibrating regression model. Therefore, the underlying hypothesis is that regression models cannot out-perform neural networks. Secondly, the evaluation is done based on a synthetic data set based on a real world data set. Thirdly, synthetic data on household level is used to explore neural network performances under circumstances of increasing complexity. The well-defined differences between the datasets increase the controllability of the test. Finally, the neural networks and regression models are calibrated using different percentages of hold out data, between 0.1 and 90%. As a comparison, Huisken and Coffa [3] assume that large data sets are available for the full 100%.

## 2 Organization of the test

To set up the test determine, a number of steps are performed. Firstly, the variables that characterise a household are determined. Secondly, synthetic input data is generated: synthetic households/individuals with different characteristics. Finally, a performance indicator is introduced.

### 2.1 Forecasting variables

The choice of variables used to predict trip generation rates has long been an area of concern (Ortuzar and Willumsen [5]). Income, car ownership, household structure, family size, value of land, residential density, accessibility, median income, total employment and the number of dwelling units are examples of trip generation variables. The variables used in this research to characterise different household types are based on the Dutch Regional Model, NRM (AVV [6]):

- number of employees in agriculture, industries, retail and other sectors;
- number of cars;
- number of students;
- total number of man/women working;
- number of people aged -14, 15-35, 35-65, 65-.

### 2.2 Synthetic data: building synthetic households

A set of 20 synthetic households classes define the major inputs in this test. The synthetic households are built using the Dutch national travel diaries (OVG), which holds aggregated data on trip frequencies. This data is used to produce trip generation factors. These factors are used to set up a data set with 20 representative household classes.

The test case is a zone/city with a population of 10000 households, approximately 30000 inhabitants. In order to test the capabilities of the three methods different complexities are defined to fill the test zone. Table 1 shows the complexity definitions.



Table 1: Complexity definitions.

Data structure	Homogeneous: 10000 households, 20 household classes, equally distributed	Inhomogeneous: 10000 households, 20 household classes, randomly distributed
Deviation on trips	1	2
Deviation on socio data	3	4

The first difference in complexity is brought about by the definition of both homogeneous and inhomogeneous distributions. Homogeneous zones are built around 20 household classes that are evenly distributed in the zone. The inhomogeneous zones are built around 20 randomly distributed household classes. Furthermore complexity varies in the way the data is presented. In the first two cases a statistical deviation is used on the total trips made per household. This results in household classes having the same socio data, but different trip productions. In cases 3 and 4 not only the trips are subject to a statistical deviation, but also the socio data is. This results in household classes with statistically deviated socio economic data and trips.

The synthetic data is used to calibrate both neural networks and regression models. Therefore, the data set is split up into a training/calibration set and a test/validation set. The test set is used to test the performance of both calibrated models. Out of the total set of 10000 households a training set corresponding with the training set percentage is randomly chosen. The test set is the remaining part of the 10000 households. This makes it easy to determine the influence of the training set percentage on the performance of both neural networks and regression models. The training set percentage is divided into two categories: low and high. The low percentages run from 0.1 to 0.9. The high percentages run from 1- 80% and 1-9% in respectively the less complex cases and the most complex cases. During the tests of complexity cases 1 and 2 it showed that the test percentages higher than 10% were not the most interesting. Therefore in cases 3 and 4, 9 % is the highest test set percentage.

### 2.3 Comparison measure

To compare the performances the error definition that was used is the Root Mean Square Error (RMSE). The RMSE is mathematically described by;

$$RMSE = \sqrt{\left(\frac{1}{N} \sum_{i=1}^N (x_i^{observed} - x_i^{predicted})^2\right)} \quad (1)$$

with:

$N$  = number of samples per matrix (10000, number of households);



A modelling method is said to out-perform the other if its goodness-of-fit is superior, as measured by the RMSE and the standard deviation. A good fit on the trip production and attraction on individual household levels is no guarantee for good estimates. The most important outcome of the trip generation process is the trip total per zone. So, extra analyses have to reveal information on the fit on the total number of trips.

### 3 Neural networks

#### 3.1 Short description of neural networks

Artificial neural networks (ANN's), or short neural networks, are based upon biological neural networks (like the human brain) by mimicking their architectural structure and information processing in a simplified manner. They both consist of building blocks or processing elements called neurons that are highly interconnected, making the networks parallel information processing systems. Although the artificial neural networks are a rudimentary imitation of biological ones, they are to some extent capable of tasks such as pattern recognition, perception and motor control which are considered poorly performed and highly processor time inefficient by conventional linear processing, whereas they seem to be done with ease by e.g. the human brain. These parallel systems are also known to be robust and to have the capability to capture highly non-linear mappings between input and output.

#### 3.2 Pre-processing

Neural networks have to be configured in order to work well; parameters in the neural network model formulation need to be determined. Questions answered relate to the size of the neural network (number of neurons) and parameters that determine the ability to 'learn'. Selecting a NN configuration and parameters is difficult and often based upon only a limited number of criteria. It is common practice to proceed by trial and error to for example select the number of neurons, and to test networks with layers of varying size. Pre-processing streamlines the process of working with neural networks.

Pre-processing is conducted for several reasons. Firstly, both neural network performance and computer time are strongly related to the network configuration. Pre-processing gives an indication of what the best number of neurons is. So computer and analysis time can be decreased. Secondly, computer time is dependent on the necessary 'learning' time of the neurons. Pre-processing gives an indication of the minimum necessary training time. Finally, pre-processing gives preliminary results on performances of neural networks and gravity models and helps to draw conclusions in the end.

Pre-processing showed that numerous variables and parameters can be altered in the search for a perfect neural network set-up. Finding this perfect set up is very difficult, if not impossible. It should be mentioned that a good instead of the perfect set-up was used. Therefore it cannot be guaranteed that the neural



network setup is the most suitable one for every complexity case. We will stress the problem of network configuration again in the next sections.

## 4 Comparison of model performance

### 4.1 Complexity cases 1 and 2

After pre-processing the final study was conducted with in mind the following hypothesis: in its simplest form, a neural network is nothing more than a self-calibrating regression model. In this sense it is impossible for a regression model to outperform the neural network. However, as mentioned, the setup of a neural network is very determining. The performances of both methods are presented in Figure 1. A distinction is drawn between the RMSE and the trip totals.

Neural networks mildly out-perform the regression models in both homogeneous and inhomogeneous configurations. Both RMSE and trip total results are in general better than the results of the regression model. The peak for regression models at 50% in the homogeneous configuration catches the eye. This seems to be a coincidence when looked at the results on 20 and 80 percent.

In real situations a data set percentage of 1 % is already high. Often data sets of less than 1 percent of data are used to perform model calibration. For more realistic results percentages from 0.1 to 0.9% are tested. Research into the results at lower percentages is conducted as presented in Figure 2. Looking at the RMSE values the same conclusions *cannot* be drawn.

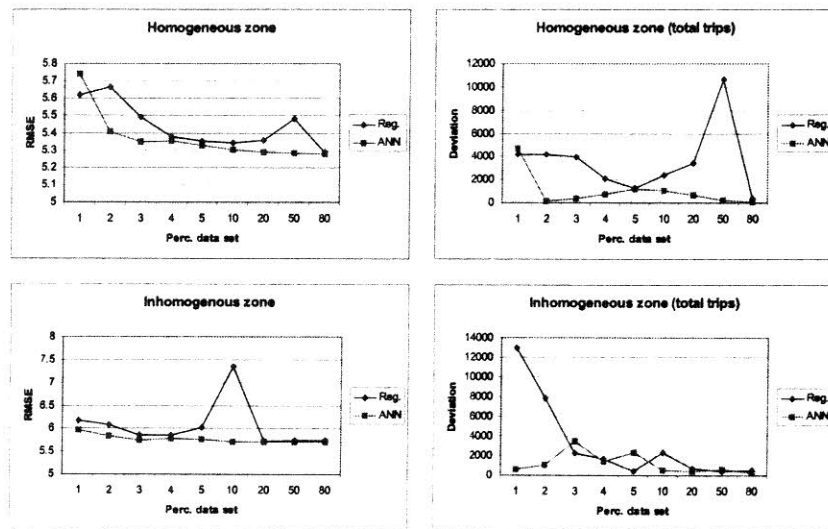


Figure 1: Performance of both methods (ANNs and Regression models (Reg.)) in complexity cases 1 and 2.



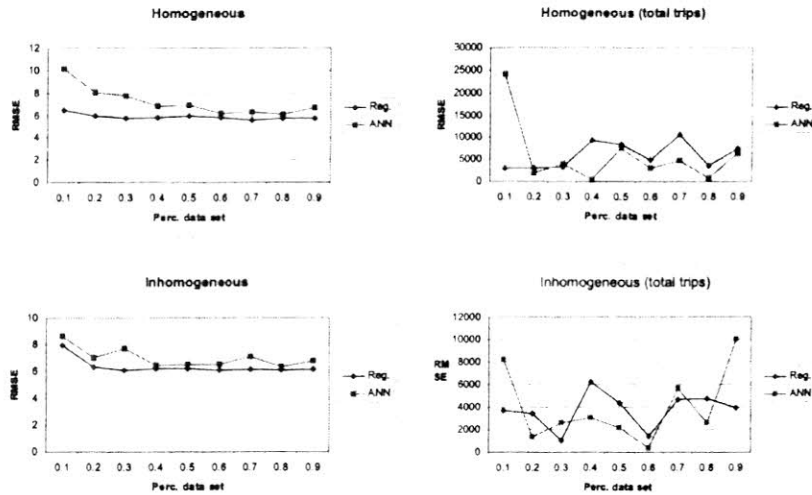


Figure 2: Performance of both methods in complexity cases 1 and 2 with scarce data.

The neural networks do not out-perform the regression models in the homogeneous problem. On the contrary, regression models seem to out-perform neural networks especially with data set percentages between 0.1 and 0.4 percent. The results of the trip totals show a somewhat different view. At some points the (in) homogeneous problem shows that the neural network model outperforms the regression model and the other way around. So no clear conclusion can be drawn based on the total trip values. This raises questions whether the right neural network configuration is chosen, however till so far no better suitable neural network configuration has been found. It is interesting to what extent the results in more complex situations are conform these results.

#### 4.2 Complexity cases 3 and 4

The previous figures showed that neural networks cannot significantly out-perform regression models. Figure 3 and Figure 4 show the results of case 3 respectively case 4.

The results show that regression models outperform the neural networks when looked at the RMSE. The trip totals show that neural networks in general score equally well or better than the regression model. How is this possible? Neural networks obviously give bad RMSE-results when the data percentage is low. The calibration process is difficult when data percentages are lower than 0.4. This can easily be explained by the necessary number of data records to train the networks. The currently used neural network configuration needs approximately 40 records to train (= 0.4%). The neural network system cannot be solved using less than 40 records. The system will be underdetermined.



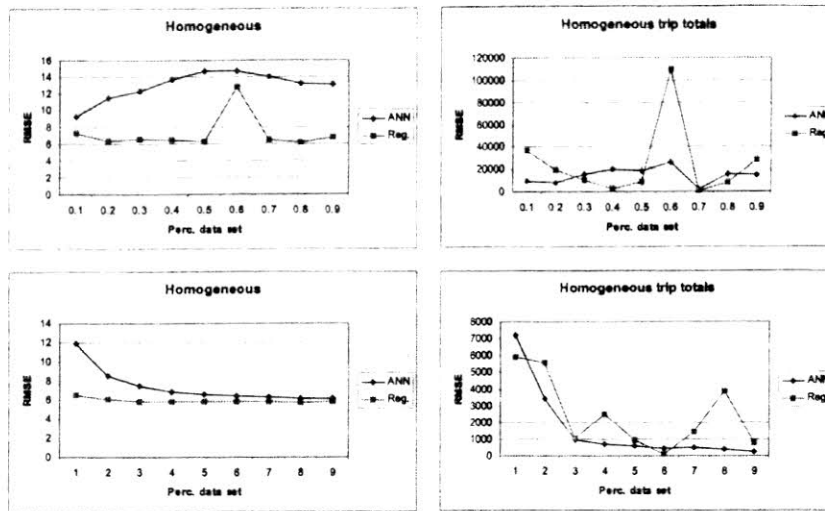


Figure 3: Performance of both methods in complexity case 3.

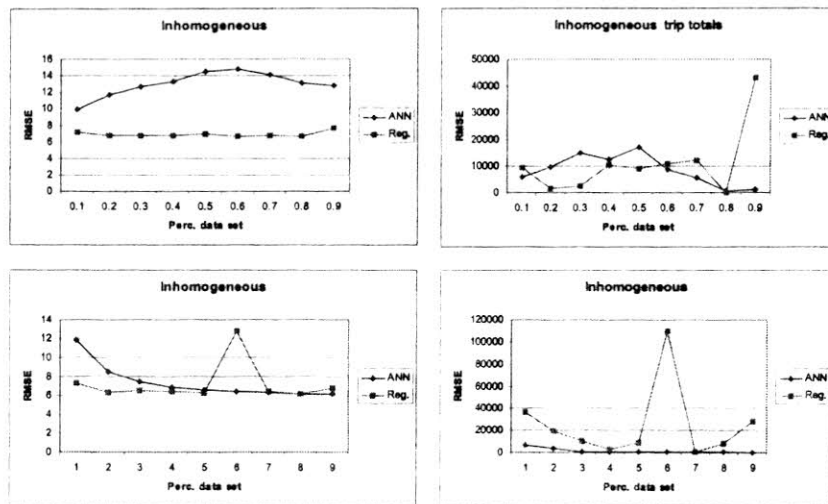


Figure 4: Performance of both methods in complexity case 4.

When neural networks have to perform better, the number of hidden nodes can be varied; in this case lowered to prevent the system from being underdetermined. Therefore different neural network configurations are tried. This resulted in better results (in favour of the neural networks) than presented in figures 3 and 4. However, there is no significant difference.





## 5 Discussion and conclusions

This research shows that neural networks do not overall out-perform classical regression models in situations when data is scarce. The total trip results are better for neural networks, but this can be a coincidence because the RMSE values are overall the same as those of regression models. These results are somewhat disappointing because the initial hypothesis was that neural networks were able to out-perform regression models.

A first interesting conclusion can be drawn on the relationship between the comparison measure, RMSE, and the total number of estimated trips in a zone. This however was not a conclusion that was looked for in first instance. The research shows that a good score on the RMSE, that means the number of trips of individual households is estimated good, does not always result in a good score on the total number of trips on a zonal level. Adding up household results can obviously result in better or worse (aggregated) results than indicated by the RMSE value. Therefore it can be concluded that neither of the comparison measures gives a good and thorough view on the results. Both measures therefore should be used.

In the least complex situation, homogeneous zone with a large available data set data, the neural network RMSE results are overall better. As expected the overall results in the inhomogeneous situation are worse than in the homogeneous situation. The results on both RMSE and trip totals with a calibration test set of 80% are comparable for both methods. Overall the score on the total number of trips is better for neural networks. Neural networks are better capable of abstracting the pattern in the dataset when enough data is available. Training percentages of over 1% can be quite large in a real world context. Performances of both methods are less good when calibrating is based on less than 1 % of the data. The stable results of the regression model are catching the eye. The neural network results are not better than the regression results. This seems to be a result of the underdetermined system in cases where less than 0.4% of the data is available. Regression models are less prone to being underdetermined.

In first instance the results of complexity cases 3 and 4 show different results. The results on the RMSE are better for regression models on both low and high percentages. However, the total trips results are better for neural networks. The observed results for neural networks indicate that the chosen network configuration from the pre-processing is not in every situation a good configuration. Research into the better neural network structures revealed that other structures give the better neural network results. However, these results are not significant.

So, there is not one best neural network topology and configuration for all proportions of available data. In addition to the network configuration, neural network performance is dependable on factors like the activation function, learning method and corresponding momentum, stop criteria for learning and software to run neural networks. During this research no extensive extra work has been conducted to investigate the influence of these parameters on neural



network performance. Future research has to be pointed in the direction of determining the impact of configuration parameters. This research emphasizes one of the difficulties of working with neural networks. In the complexity cases 1 to 4 the differences in optimal structure seem to be quite large, illustrating the difficulties in modelling neural networks. Regression models are easier to construct. A general rule of thumb to configure a neural network lacks. Trial and error is therefore the only solution.

So, can neural networks be used in trip generation modelling? Yes neural networks can. But there are hardly any advantages compared to regression models, at least in this setting. The performance of both regression and neural network models is good and are not significantly different. The differences in results of the RMSE-indicator and the indicator on total trips are eye-catching. Looking at the desired outcome, the total trips, neural networks have an advantage. But this is not significant either. Furthermore, the influence of a neural network configuration is present. The differences in RMSE when different neural network setups are used are not big. However, the necessity stays to do good pre-processing in order to find the best suitable network structure.

## References

- [1] Openshaw, S. & Openshaw, C., *Artificial intelligence in Geography*, John Wiley and Sons, Chichester, 1997.
- [2] Al-Deek, H.M., Which method is Better for Developing Freight Planning Models at Seaports – neural Networks or Multiple Regression, *Transportation research record*, 1763, pp. 90-97, 2001.
- [3] Huisken, G. & Coffa, A., Neural Networks and Fuzzy Logic to improve Trip Generation Modelling, *Proceedings of the 9<sup>th</sup> International Association for Travel Behaviour Research Conference*, Institute of Transport Studies, IATBR, Gold Coast, Queensland, Australië, 2 – 7 juli, 2000.
- [4] Dantas, A., Yamamoto, K., Lamar, M.V. & Yamashita, Y., Neural network for travel demand forecast using GIS and Remote sensing. <http://www.eletrica.ufpr.br/lamar/public/dantasitalia2000.pdf>, 2000.
- [5] Ortuzar, J. de D. & Willumsen, L.G., *Modelling Transport*, John Wiley & Sons, Ltd, Chichester, UK, 2001.
- [6] AVV, Adviesdienst Verkeer en Vervoer, *Nieuw Regionaal Model, Handboek*, 1997.

