TRIP GENERATION AND TRIP DISTRIBUTION:
COMPARISON OF NEURAL NETWORKS AND TRADITIONAL METHODS

F. Tillema, K.M. van Zuilekom, M.F.A.M van Maarseveen, G.Huisken
Centre for Transport Studies, Civil Eng., Univ. of Twente
PO Box 217, 7500AE, Enschede, The Netherlands
f.tillema@utwente.nl

Abstract

The aim of this study is to explore the performances of neural networks in both trip generation and trip distribution modelling and to compare the results with more commonly used models, respectively regression models and doubly constrained gravity models.

Trip generation and trip distribution are complex and highly dependent on the quality *and* availability of data. Transportation engineers are commonly faced with a question that is related to this topic; how to perform reliable trip generation and -distribution 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 research tries to answer the question whether neural networks can be a better alternative for traditional methods in trip generation and -distribution.

The research design relies on the use of synthetic data. The use of synthetic data, without unknown noise, gives the opportunity to clearly determine the impact of data complexity on the forecasting results.

In trip generation, neural networks do not overall out-perform classical regression models. The advantages over regression models are negligible. In trip distribution, neural networks out-perform gravity models when data is scarce. Gravity models perform slightly better than neural networks when sufficient data is available.

Keywords: Trip generation; Trip distribution; Neural networks; Regression model; Gravity model
Topic Area: D1 Passenger Transport Demand Modeling

1. Introduction

Transportation engineers are commonly faced with the question of how to extract information from expensive and scarce field or survey data. A common approach is to create a model that describes the behaviour of the phenomenon observed in which the data is used for calibration/validation. Ideally, such an approach leads to a model with the desirable high accuracy. Unfortunately, there is often a discrepancy between the desired and the obtained accuracy; estimating a model, based on scarce data is not an easy job and can lead to results with high deviations. Furthermore, it is not always easy to construct a statistical model from data, due to the fact that many phenomena are non-linear, and/or collinear (Huisken and Coffa, 2000).

Classical transport planning, described in the classical 4-step model (Ortuzar and Willumsen, 2001), is characterized by the dependency on data. Spatial interaction patterns, for example the trips generated by a zone or the trips between zones, respectively the trip generation and the -distribution, are highly complex and difficult to model without adequate amounts of data. Errors that are generated during the trip generation and distribution estimation process propagate through till the assignment phase. This causes difficulties for a good transport planning. Currently used techniques try to use limited amounts of data. The question rises, whether these techniques are able to give good trip generation and distribution estimations.

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 in classical transport planning. Openshaw and Openshaw (1997) 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. Research conducted by Miller *et al.* (1995), Dougherty (1995), Collins *et al.* (2001), Pijanowski (in press), Padmakumarie (1999), Raju *et al.* (1998), Huisken and Coffa (2000), Currit (2002) and Faghri and Sandeep (1998) support this notion. These studies carefully reveal the opportunities of applying neural networks in transport planning context.

This research tries to answer the question whether neural networks can out-perform traditional methods for trip generation and distribution. Neural networks are compared to regression models in a trip generation context. In addition, the neural networks are compared to doubly constrained gravity models in a trip distribution context. Both regression models and gravity models are commonly used models in these contexts.

The paper is organized as follows. The first section gives an introduction into neural networks. The second section of this paper goes deeper into the subject of trip generation. The third section deals with the trip distribution. In the second and third part we will focus on the organisation of both tests and the performances of neural networks compared to the commonly used methods. Finally, conclusions are drawn on the capabilities of neural networks for classical transport planning in general.

2. 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. Research study 1: Performance of neural networks in trip generation

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 (1997). Al-Deek *et al.* (2001) gives an example of the use of neural networks in truck trip generation in a harbour. Huisken and Coffa (2000) conduct an extensive research into trip generation. Dantas *et al.* (2000) present a strategic planning model for urban transportation analysis.

Literature survey shows that artificial neural networks are successfully used as data analysing 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. And 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. Firstly, 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 created with 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%.

3.1. Organisation of the test

To set up the test 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 and the results are determined.

3.2. Forecasting variables

The choice of variables used to predict trip generation rates has long been an area of concern (Ortuzar and Willumsen, 2001). 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, 1997):

- 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-.

3.3. 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 two methods different complexities are defined to fill the test zone. Table 1 shows the complexity definitions.

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.

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 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.

3.4. 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;

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

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.

3.5. Comparison of model performance

3.5.1. Complexity cases 1 and 2

After a pre-processing stage, 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 set-up 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.

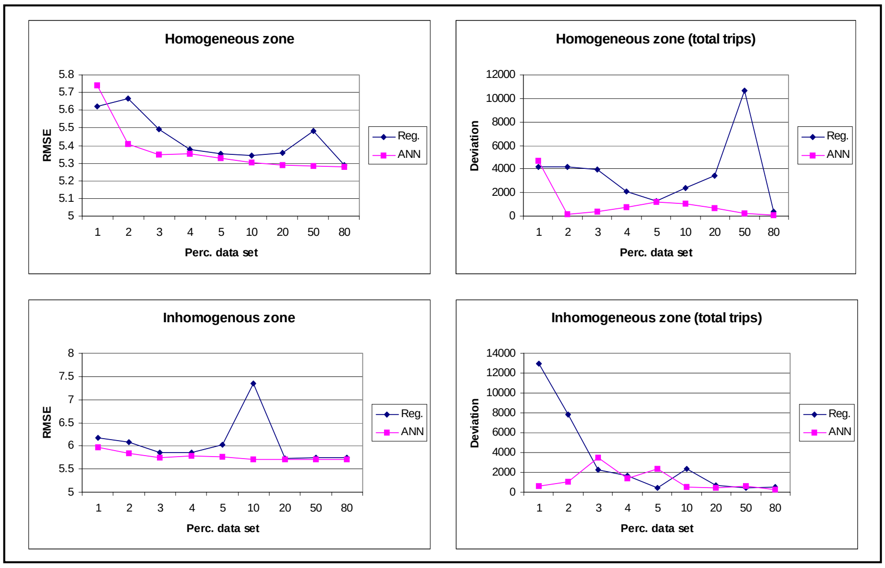

Figure 1: Performance of both methods (ANNs and Regression models (Reg.)) in complexity cases 1 and 2.

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.

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 %. 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.

3.5.2. 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 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.

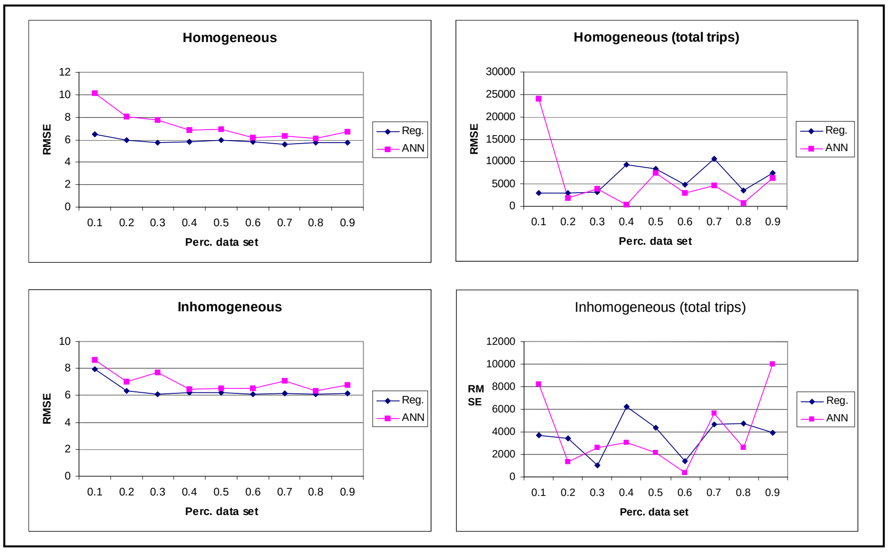

Figure 2: Performance of both methods in complexity cases 1 & 2

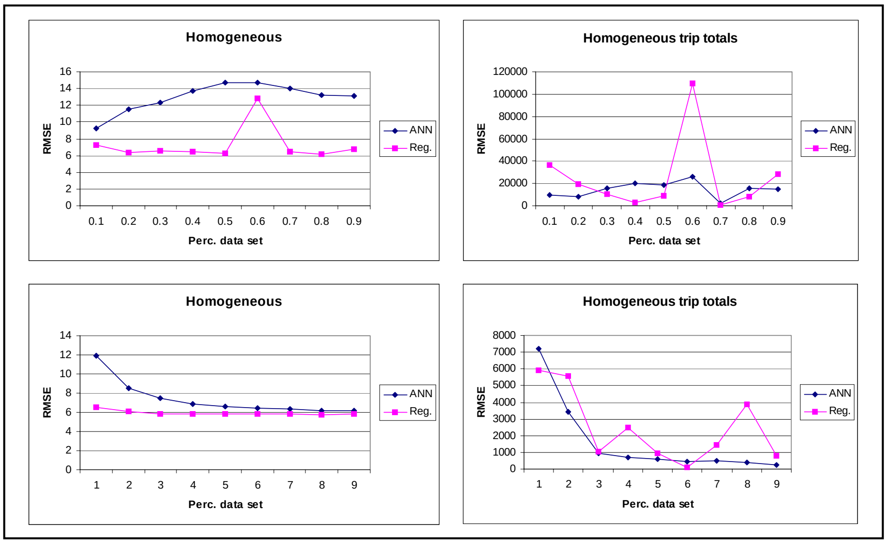

Figure 3: Performance of both methods in complexity case 3.

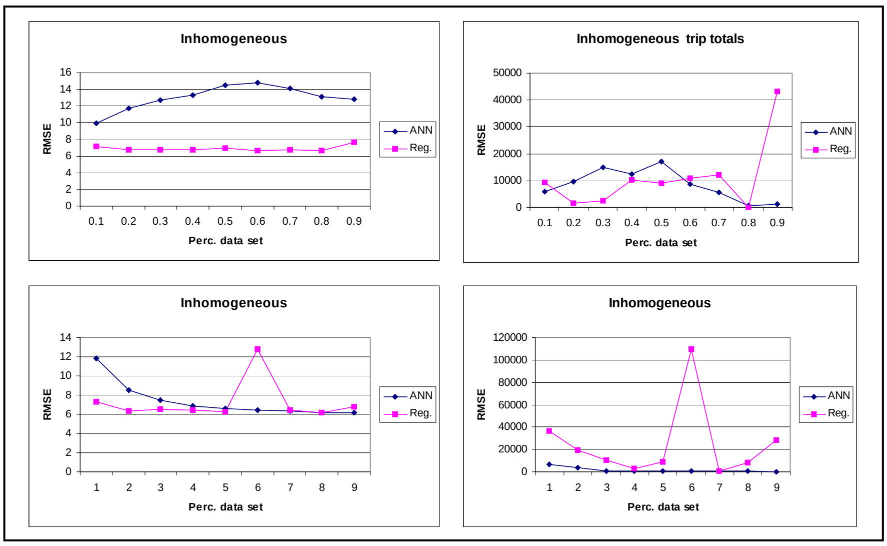

Figure 4: Performance of both methods in complexity case 4.

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.

3.6. 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 calibration 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. Research into the better neural network structures revealed that other structures give better neural network results. However, these results are not significant.

4. Research case 2: Trip distribution

The previous section showed the results of neural networks in trip generation. It is interesting to see whether these results are the same as for the trip distribution problem. Several studies have explored the usefulness of neural networks in the context of trip distribution modelling. However, the empirical results leave questions open whether neural networks give better results than traditional trip distribution methods. Black (1995) asks the question if the basic purpose of neural networks, identifying patterns in data and to replicate those patterns for new data, can be utilized in a spatial context. He makes a comparison between a gravity model and neural networks. Black uses two case studies: (i) a three-region flow problem; and (ii) a commodity flow problem. The first problem is a very simple three-region flow problem. Both doubly constrained gravity models as well as a neural network model give excellent results. Black emphasizes that one should not lose sight of the fact that the matrices have only nine flow values. The commodity flow problem gives similar results, ranking the scores of artificial neural networks above the scores of gravity models. Finally, Black concludes that neural networks are capable of high levels of accuracy based on their use in other fields and are suitable for future flow forecasting.

Fischer and Gopal (1994), Gopal and Fischer (1996) and Fischer (1998) compare the forecasting results of neural networks to those of a traditional gravity model. Research is done into the distribution of interregional telecommunication flows. Although the test case is not traffic and transport related problem, the problem is to large extent comparable with a trip distribution problem. The basic conclusion is that the neural network models out-perform the conventional gravity model.

Mozolin *et al.* (2000) compare the performances of neural networks and maximum likelihood doubly constrained models for commuter trip distribution. The authors state that their approach differs drastically from others in several respects: (i) the models are used in a predictive mode and calibration is done on observed data, while testing is conducted on data for a projection year; (ii) the baseline problem is a doubly constrained model estimated by maximum likelihood; (iii) the models are evaluated on origin-destination matrices of different sizes to be able to test the sensitivity of the conclusions to the size of the interaction system; and (iv) the model applies an adjustment factor to flows predicted by the neural network output to satisfy constraints. It is concluded that neural networks exhibit good to very good ability to predict future commuter flows. Yet, none of the tested neural networks outperforms the corresponding doubly constrained model. The authors find this fact puzzling and unexpected. After further data analysis the following results are formulated: (i) due to over-fitting the ability to generalize is rather poor and the prediction accuracy is low particularly where training data are scarce; (ii) networks fail to extrapolate around and beyond the limits of the training sample; (iii) networks with less hidden nodes are less prone to over-fitting; (iv) the ability to approximate data structures with great accuracy is also their weakness.

So, artificial neural networks are increasingly used as data analysing techniques in a spatial interaction, trip distribution context. Yet, the conclusions whether neural networks out-perform more traditional models are still under discussion. The aim of this part of the study is to explore the performance of neural networks in trip distribution modelling and to compare the results to more commonly used doubly constrained gravity models.

The approach differs from other research in several respects. Firstly, the evaluation is done based on large synthetic datasets, as well as a real world data set. Secondly, synthetic data (OD-matrices) are used to explore neural network performances under circumstances of increasing complexity. The well-defined differences between OD matrices increase the controllability of the test; differences in results can easily be attributed to the built up of the data. Thirdly, statistical analysis is conducted in order to find minimum necessary sample sizes for both models. Fourthly, like Mozolin *et al*., the neural network output is enforced on the production and attraction constraints; in this case by using the Furness method (Orthuzar and Willumsun, 2001). Finally, the neural networks and gravity models are calibrated using different percentages of hold out data. In this way one of the biggest advantages of neural networks, extrapolating/forecasting of missing data (patterns), can be examined.

A complete OD matrix is generated using a gravity model. This results in a completely known OD matrix. In addition, noise and measurement errors are prevented. The basic test is a synthetic spatial network of 15 regions, combined with synthetic impedances and attraction/production values. The second test is a comparison of different estimation methods on observed trip patterns in a real world network, Rotterdam Rijnmond (National Regional model, NRM). The known data from the generated OD matrix is split up into calibration and test data. The calibration percentage is varied between 10 and 90. So, a complete OD matrix will be estimated, using limited (observed) data and trip attraction and production totals.

4.1. Organisation of the test

The focus of this research is on comparing the performances of neural networks and gravity models in well-defined basic and real world cases. Firstly, synthetic input data is generated: (i) synthetic network; (ii) synthetic skim matrix (impedance); and (iii) synthetic input data of different complexities (OD matrices). As mentioned in section 3, synthetic data gives the opportunity to play with complexity. This approach gives an insight into the impact of complexity, without modelling noise or unclear relations between variables. Neural networks and gravity models are calibrated on different percentages of the input data. Finally conclusions are drawn on the performances and performances in relation to different percentages of hold out data.

4.2. Synthetic data: building synthetic OD matrices

4.2.1. Synthetic network

A synthetic network combined with synthesized impedances (skim matrix) and synthesized trip attraction and production values define trip distribution modelling inputs. The choice for 15 regions results in a 225 cells Origin-Destination (OD) dataset. The use of a simple synthetic 15-region network gives the opportunity to carefully explore neural networks usefulness in trip distribution modelling. The regions are located on a straight line and distances in between regions are equally distributed.

The logistics of spatial interaction modelling requires clearly defined regions with no, or small, flows across the borders. In the case of the synthetic network, this assumption is not violated. Setting intra zone distance to zero is known to generate systematic measurement errors. Therefore spatial separation within the regions, a inter zone distance greater than zero, is introduced within the network.

4.2.2. Synthetic OD-matrices

Trip distribution estimation requires input values for the distances between regions as well as trip generation and attraction values. Trip generation and attraction have been synthesized. A total of 15000 trips is distributed among the zones as schematised in Figure 5.

From 1 to 16 the matrices' complexity increases: e.g. matrix 1 is built with evenly distributed origins/destinations, matrix 3 is built with evenly distributed origins and a descending pattern for destinations. The well-defined differences between OD matrices increase the controllability of the test. Differences in results can easily be attributed to the built-up of the matrices. Matrices indicated by lines are not tested; configurations of these matrices are already tested in one of matrices 1-16. The complexity of the matrices is shown in between brackets. The complexity is based on the patterns for destination and origins and the interactions within the matrix.

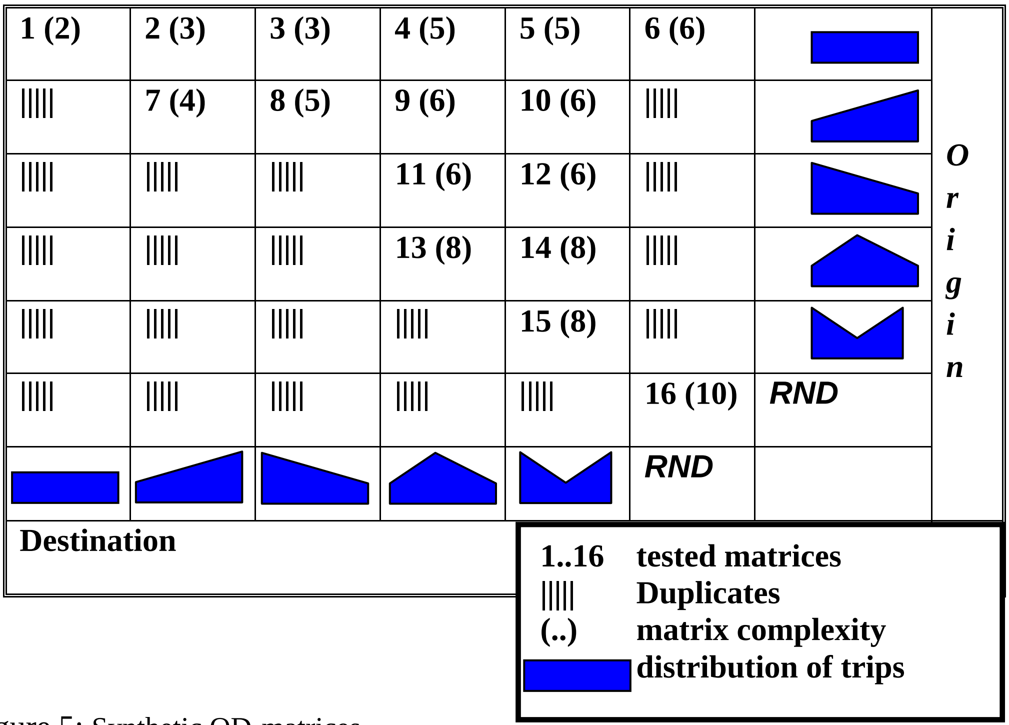

Figure 5: Synthetic OD-matrices

4.3. Real world matrix: Rotterdam Rijnmond Region

Rotterdam, famous for its harbour, is the second city of The Netherlands with a population of 0.6 million inhabitants. Rotterdam Rijnmond is the whole area of Rotterdam including the harbour and suburbs. Using the NRM (National Regional Model) zoning method, Rijnmond is divided into 15 zones. A total number of nearly 1.9 million car trips per 24 hours is made, calculated over all motives. Spatial impedance between counties is simply measured as time between the zone centroids.

4.4. Comparison of model performances

For comparison of the results the same comparison measure is used as in the case of trip generation. A model 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 totals and a low RMSE are no guarantee for good estimates. So, extra analyses have to reveal new information on the fit on OD-cell level. Therefore comparisons are made between both methods on the average trip length en the trip length distribution. Trip length frequencies give insights into the results of both methods on all trip length categories. The performances of calibrating both ANN and GM on the data of a representative sample of the 16 matrices are presented in Figure 6. None of the neural network models outperforms calibrated gravity models for all percentages. Gravity models outperform neural network models when sufficient data is at hand to perform a good calibration. Figure 6 shows that most gravity models start outperforming the neural networks when the total percentage of data exceeds 50%.

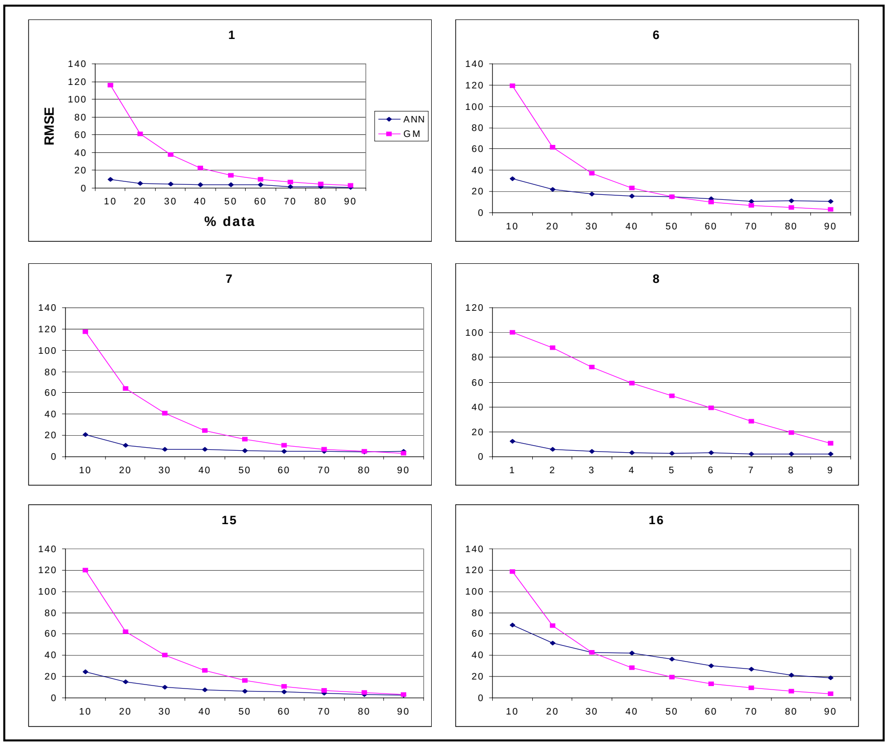

Figure 6: Model performances

In general the RMSE results are strongly influenced by the percentage of calibration data. At low percentages the results are worse than at high percentages, as expected. Furthermore, the gravity model results are far more influenced by data percentage as the neural networks. Especially at low percentages is the performance of neural networks better. Therefore, at lower percentages, up to 50%, neural networks give better results.

In addition, neural networks do not outperform gravity models on the whole scale. When the calibration dataset percentage is higher than 80-90%, gravity models give better results. This is not surprising, because of the fact that gravity models estimate to a high extent their own creations. At 100% gravity models always replicate the complete matrix that was created before the result comparison. Matrix 8 shows strange results. The built up of the matrix strongly determines the bad results of the gravity model. The trip attraction and production values differ strongly from the theoretical standard distribution function. This standard function gives high trip rates at low distances and low trip rates at high distances. Matrix 8 shows quite the opposite. At low distances people make far less trips than at high distances. Here the advantages of neural networks are shown. Without pre-assuming a certain function, the neural network is better capable of estimating the trip distribution.

At the other end of the scale, percentages 10 up to 50 %, neural networks *strongly* outperform classical gravity models. Looking at the Rotterdam Rijnmond matrix, the same observations can be made (Figure 7).

Two facts have to be stressed. Firstly, the ANN results are fractionally worse than in case of the synthetic matrices. Yet, the same pattern is still visible. Secondly, the RMSE values are higher due to the fact that the total number of trips is approximately 130 times higher. However, when the number of trips is related to the RMSE values, the RMSE values are still 2 times higher. RMSE value of the best-guess (dividing 1.9 million trips over 225 cells) results in a RMSE of 79.000.

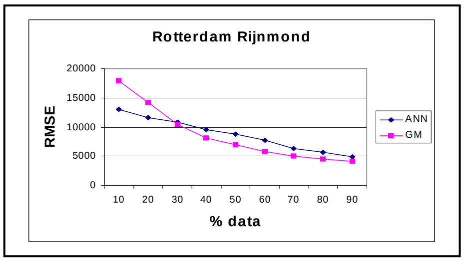

Figure 7: Model performances in the Rijnmond case

A good fit on the trip production and –attraction totals and a low RMSE are no guarantee for good estimates. So, extra analyses have to reveal new information on the fit on OD-cell level. Therefore the results of neural network and gravity model trip length frequencies are studied. The results give an insight into three facts. Firstly, for both gravity models and artificial neural networks, the performance on the trip length frequency goes up when data percentage goes up. Secondly, gravity models seem to have difficulties estimating both high and low number of trips. Neural networks only seem to have problems with low number of trips. Finally, the performance at low percentages is much better when neural networks are used. The results perfectly illustrate the conclusions based on Figure 6; at low percentages, the ANN outperforms the GM. When the input data increases, differences in performance decrease. The results of the Rotterdam Rijnmond case show roughly the same pattern; at low percentages, the ANN out-performs the GM. When the input data increases, differences in performance decrease. Neural networks have more difficulties in estimating extreme values.

4.5. Explaining the results

Is there an explanation for these results? Firstly, all matrices show roughly the same pattern; at low percentages neural networks outperform gravity models, at high percentages gravity models outperform neural networks. Neural networks show their ability of extrapolation of data; they can very well cope with small data sets. The performances of gravity models, when calibration data percentage is high, seems to be related to the built-up of the matrices; the matrices were built using gravity models.

So, when the calibration dataset nears 100%, the only matrix the gravity model estimates is its own creation. Therefore, the conclusion that gravity models outperform neural networks, when high percentages of data are used, is not very strong. This could favour the use of neural networks, even when datasets are large. Due to the data management after training, neural networks were able to estimate both high and low trip values, also beyond the limits of the training sample.

In first instance it seems that no general conclusion can be drawn upon the relationship between complexity and results; no clear relationship is shown for either of the models. The gravity models appear to be less sensitive to complexity and more stable in results whereas neural networks show strongly varying results when complexity increases. The following four conclusions can be drawn: (i) when complexity is minimal, 2, data is most structured and neural network performance is best. This stresses one of the qualities of neural networks: pattern recognition; (ii) contrary to this point, when complexity is at its maximum, 10, neural network performance is worst. The complex matrix reveals the least order and therefore the fewest patterns. This results in an RMSE increase; (iii) results within complexity groups are mostly grouped together. This fact shows that the complexity built up of the test sets is consistent; (iv) and the differences in results between neural network models and gravity models are not stable. Especially the difference between both models in matrix 16 is small. This stresses the fact that neural networks perform best in situations where data is most structured; matrix 1, complexity 2, shows the largest difference in (average) RMSE values.

4.6. Discussion and conclusions

This research part shows that neural networks out-perform gravity models in both synthetic and real situations; when data is scarce. These results are promising for future trip distribution modelling, which is an important step for good transport planning. The results were obtained using both synthetic and real world datasets. This gives the opportunity to control the test.

As seen in the different figures, neural network performance, compared to gravity models, is best when data is scarce. As stated before, the synthetic matrix data was generated using a gravity model. This creates a situation in which gravity models should give good results. However, the gravity model only gives the best results when calibration percentage is high; gravity models only reproduce their own results. In situations close to reality, with only limited amounts of data, neural networks show their abilities. This strengthens the results of neural networks.

The investigation into the trip length frequencies gives an insight into the absolute performances of both methods. Neural networks show better performances on trip length frequencies when data is scarce. Due to the data management after training, neural networks were able to estimate both high and low trip values, also beyond the limits of the training sample.

The behaviour of both methods changes when complexity increases. The datasets are complex enough, especially the random matrices, to come close to reality. Results show that neural networks perform better under conditions in which data is structured. But, results show also that even performances of the random matrix and real world matrix are good. Due to the large number of trips in the Rijnmond case, the RMSE values were higher than in the synthetic cases.

It is difficult to obtain a good estimation for the total number of samples necessary to be sure about the results. In addition, it can be concluded that large sample sizes are necessary due to amongst others the random initialisation process of neural networks. Furthermore, the used calibration process for gravity models needs *40 times* more samples than the neural network.

5. Discussion and conclusions

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 set-ups are used are not big. However, the necessity stays to do good pre-processing in order to find the best suitable network structure.

Can neural networks be used in trip distribution modelling? The study shows that neural networks outperform gravity models when data is scarce. The conclusion that gravity models out-perform neural networks when more than 50% data is available seems less certain, due to the research method and the generation of the synthetic data.

So, the performance of neural networks is promising. The research shows that the trip generation and -distribution problems are complicated ones, but also two important steps in transport planning. Many errors generated during these two phases are passed on to the next steps. Often, real world problems have only limited data. And contrary to this research, real world problems have only one sample of that data. Scarce data can give difficulties during calibration of models and results have high standard deviations. The extent to which the available data suits a calibration process determines the performance. However, in the case of trip distribution if only 20 percent or less data is available, the calibration process can lead to a large number of different matrices. In this study, a total dataset is available as a reference for determining the quality of the estimated data; the RMSE.

This paper adds new inputs to the discussion of trip generation and -distribution modelling with neural networks. New methods, like neural networks, can be a leap into the direction of good and accurate results.

References

Al-Deek, H.M., 2001. Which method is Better for Developing Freight Planning Models at Seaports – neural Networks or Multiple Regression, Transportation research record, 1763 90-97.

AVV (Adviesdienst Verkeer en Vervoer), 1997. Nieuw Regionaal Model, Handboek, in Dutch.

Black, W.R., 1995. Spatial interaction modelling using artificial neural networks, Journal of Transport Geography 3 (3) 159-166.

Collins, M.G., Steiner, F.R., and Rushman, M.J., 2001. Land-Use Suitability Analysis in the United States: Historical Development and Promising Technological Achievements, Environmental Management 28 (5) 611–621.

Currit, N., 2002. Inductive regression: overcoming OLS limitations with the general regression neural network, Computers, Environment and Urban Systems 26 335–353.

Dantas, A., Yamamoto, K., Lamar, M.V., and Yamashita, Y., 2000. Neural network for travel demand forecast using GIS and Remote sensing. Downloadable from website http://www.eletrica.ufpr.br/lamar/public/dantasitalia2000.pdf.

Dougherty, M., 1995. A review of Neural Networks applied to Transport, Transportation research-C 3. (4) 247-260.

Faghri, A. and Sandeep, A., 1998. Analysis of Performance of Backpropagation, ANN with Different Training Parameters, In: Himanen, Nijkamp and Reggiani (Eds.), Neural Networks in Transport Applications, Ashgate.

Fischer, M.M., 1998. Computational Neural Networks: An Attractive Class of Mathematical Models for Transportation Research, In: Himanen, Nijkamp and Reggiani (Eds.), Neural Networks in Transport Applications, Ashgate.

Fischer, M.M., and Gopal, S., 1994. Artificial Neural Networks. A new approach to Modelling Interregional Telecommunications Flows, Journal of Regional Science 34 (4).

Gopal, S. and Fischer, M.M., 1996. Learning in Single Hidden-Layer feed forward network models: Backpropagation in a Spatial Interaction Modelling Context, Geographical Analysis 28 (1) 38-55.

Huisken, G., and Coffa, A., 2000. Neural Networks and Fuzzy Logic to improve Trip Generation Modelling, Proceedings of the 9th International Association for Travel Behaviour Research Conference, Institute of Transport Studies, IATBR, Gold Coast, Queensland, Australië, 2 – 7 juli.

Miller, D.M., Kaminsky, E.J., and Rana, S., 1995. Neural network classification of remote-sensing data, Computers & Geosciences 21 (3) 377–386.

Mozolin, M., Thill, J.C. and Lynn Usery, E.L., 2000. Trip distribution forecasting with multiplayer perceptron neural networks: A critical evaluation, Transportation Research C 34 (1).

Openshaw, S. and Openshaw, C., 1997. Artificial intelligence in Geography, John Wiley and Sons, Chichester.

Ortuzar, J. de D., and Willumsen, L.G., 2001. Modelling Transport, John Wiley & Sons, Ltd, Chichester, UK.

Pijanowski, B.C., Brown, D.G. Shellito, B.A., and Manik, G.A., in press. Using neural networks and GIS to forecast land use changes: a Land Transformation Model, Computers, Environment and Urban System.

Raju, K.A., Sikdar, P.K., and Dhingra, S.L., 1998. Micro-simulation of residential location choice and its variation, Comput., Environ, and Urban Systems 22 (3) 203-218.