

**VERGELIJKING VAN NEURALE NETWERKEN EN
ZWAARTEKRACHTMODELLEN VOOR RITDISTRIBUTIE**

Databeschikbaarheid versus de kwaliteit van de schattingen

Frans Tillema

f.tillema@utwente.nl

Kasper van Zuilekom

k.m.vanzuilekom@utwente.nl

en

Martin van Maarseveen

m.f.a.m.vanmaarseveen@utwente.nl

Civiele Techniek
Verkeer, Vervoer en Ruimte
Universiteit Twente
Postbus 217
7500 AE Enschede

Bijdrage voor het Colloquium Vervoersplanologisch Speurwerk 2003 te Antwerpen.
Enschede, 2-7-2003

Inhoudsopgave

1. Inleiding	4
2. Neurale netwerken	6
3. Onderzoeksopzet	8
4. Beoordelingsmaat	11
5. Resultaten	11
6. Verdieping van de resultaten	15
7. Discussie en Conclusies	16
Literatuur	19

Samenvatting

Vergelijking van neurale netwerken en zwaartekrachtmodellen voor ritdistributie 'Databeschikbaarheid versus de kwaliteit van de schattingen'

Transportplanners worden regelmatig geconfronteerd met de vraag hoe om te gaan met schaarse data/waarnemingen verkregen uit dure praktijk metingen. Ruimtelijke interactie patronen, bijvoorbeeld personen ritten tussen zones, de ritdistributie, zijn complex en daarom moeilijk te modelleren zonder adequate hoeveelheden data. Het doel van deze studie is om de kwaliteit van ritdistributie schattingen op basis van tweevoudig beperkte zwaartekrachtmodellen te vergelijken met schattingen op basis van neurale netwerken. Het onderzoek onderscheidt zich van ander onderzoek op een aantal punten, o.a. door het gebruik van zowel synthetische data alsmede praktijk resultaten. Hierdoor kan het grootste voordeel van neurale netwerken, extrapolatie van verbanden en ontbrekende data, worden onderzocht. Een synthetisch netwerk van 15 zones gecombineerd met bijbehorende skimmatrix, productie- en attractiewaarden vormt de basis voor deze studie. Door de onderbouwde verschillen in complexiteit van de data, kunnen latere verschillen in resultaten bijna uitsluitend worden toegeschreven aan de opbouw van de data. Data uit de Rotterdam Rijnmond regio is gebruikt om de modelprestaties van praktijkdata te vergelijken met synthetische. De resultaten laten zien dat neurale netwerken onder de meeste omstandigheden beter presteren dan klassieke zwaartekrachtmodellen. Met name wanneer data schaars is, schatten neurale netwerken significant beter. Een tweede voordeel van neurale netwerken is hun geringe steekproefgrootte om te komen tot statistisch significante uitspraken. De resultaten van beide methoden zijn sterk gecorreleerd met de beschikbaarheid en de kwaliteit van data.

Summary

Comparison between neural networks and gravity models in trip distribution 'Data availability vs. the performance in trip distribution'

Transportation engineers are commonly faced with the question of how to extract information from sparse field data. For example, modelling the distribution of trips between regions, trip distribution, is highly complex and dependent on data. The aim of this study is to explore the performances of neural networks in trip distribution modelling and to compare the results with more commonly used doubly constrained gravity models. The approach differs from other research in several respects; a.o. the use of both synthetic data as well as real world data. A synthetic network combined with synthesised impedances (skim matrix) and synthesised trip attraction and production values are trip distribution modelling inputs. Well-defined differences between origin-destination data complexity increase the controllability of the test; differences in results can easily be attributed to the built up of the matrices. Data of the Rotterdam Rijnmond region is used for comparison of the model performances of real world and synthetic cases. Neural networks out-perform gravity models when data is sparse in both synthesised as well as real world cases. Sample size for statistical significant results is much lower for neural networks. In general, the results show the strong dependence of performance on the availability and quality of data.

1. Inleiding

Transportplanners worden regelmatig geconfronteerd met de vraag hoe om te gaan met schaarse data/waarnemingen verkregen uit dure praktijk metingen. Een veel gebruikte aanpak is het ontwikkelen van een statistisch model waarin deze vaak schaarse data wordt gebruikt als calibratie/validatie invoer. Idealiter leidt een dergelijke aanpak op basis van een schaars aantal waarnemingen tot een model met de gewenste hoge nauwkeurigheid. Helaas bestaat er in de praktijk vaak een discrepantie tussen de gewenste en de verkregen nauwkeurigheid; het schatten op basis van schaarse data is geen eenvoudige klus en kan leiden tot grote spreiding in de resultaten. Verder bemoeilijken factoren als non-lineariteit en collineariteit de zoektocht naar een passend model (Huisken en Coffa, 2000).

De afhankelijkheid van data kenmerkt de klassieke transport planning, die beschreven wordt in het klassieke 4 traps model (Ortuzar en Willumsen, 2001). Ruimtelijke interactie patronen, bijvoorbeeld personen-ritten tussen zones, de ritdistributie, zijn sterk complex en moeilijk te modelleren zonder adequate hoeveelheden data. Fouten die gegenereerd worden tijdens het schatten van de ritdistributie propageren door tot op het niveau van de toedeling met kwalijke gevolgen voor goede planning. Huidige methodieken, zoals klassieke zwaartekrachtmodellen, proberen met beperkte hoeveelheden data toch tot zo goed mogelijke ritdistributie schattingen te komen. De vraag is echter in hoeverre deze methodieken leiden tot goede resultaten.

Recente literatuur geeft inzichten in het gebruik van andere methodieken om het ritdistributie probleem te modelleren. Parallele oplossingsmethoden en het gebruik van neurale netwerken passeren regelmatig de revue. Openshaw en Openshaw (1997) geven aan dat de voordelen van het gebruik van neurale netwerken in dit vakgebied groot kunnen zijn. Een in het oog springende conclusie is dat zij verwachten dat neurale netwerken betere resultaten zullen geven dan traditionele, statistische, modellen.

Illustratief voor deze conclusie is het werk van onder andere Collins *et al.* (2001), Pijanowski (in press), Padmakumarie (1999), Raju *et al.* (1998), Miller *et al.* (1995), Huisken en Coffa (2000), Currit (2002) en Faghri and Sandeep (1998). Deze onderzoeken laten duidelijk de mogelijkheden en voordelen van het gebruik van parallele oplossingsmethodieken. Voortschrijdend hierop geven Rodrigue (1997) en Tillema *et al.* (2002) inzicht in de

mogelijkheden om neurale netwerken te gebruiken voor integrale landgebruik en transport planning.

Verschillende studies hebben de mogelijkheden van neurale netwerken in een ritdistributie context onderzocht. Hoewel in het algemeen wordt geconcludeerd dat neurale netwerken goed presteren, is niet eenduidig vastgesteld dat neurale netwerken traditionele modellen kunnen overtreffen. Black (1995) stelt de vraag in hoeverre de basiskwaliteiten van neurale netwerken, patroonherkenning, kunnen worden uitgebuit voor ritdistributie. In zijn onderzoek concludeert hij dat het zeer goed mogelijk is om neurale netwerken te gebruiken. Hij benadrukt echter dat zijn geringe netwerk grootte (3x3 zones) de resultaten kunnen vertroebelen. Ook Fischer en Gopal (1994), Gopal en Fischer (1996) en Fischer (1998) vergelijken de schattingscapaciteiten van neurale netwerken met die van zwaartekrachtmodellen. Hoewel de context van hun onderzoek verschilt, namelijk distributie van telefoonverkeer tussen zones, zijn hun conclusies bemoedigend; neurale netwerken leveren betere resultaten dan traditionele modellen.

Mozolin *et al.* (2000) en Thill en Mozolin (2000) vergelijken in een uitgebreide studies dezelfde twee methodieken. Het onderzoek is echter belangrijk anders van opzet dan de eerder genoemde onderzoeken: (i) de modellen worden gebruikt voor voorspellingen; (ii) de modellen worden getest bij 2 verschillende netwerk groottes; (iii) het neurale netwerk model wordt voorzien van randvoorwaarden, iets wat eerder nog niet gedaan was en (iv) er wordt gebruik gemaakt van een dubbelzijdig beperkt zwaartekrachtmodel. De conclusies zijn minder positief dan in de eerder genoemde studies. Hoewel neurale netwerken prima presteren, blijven hun totale resultaten achter bij die van zwaartekrachtmodellen. Een mogelijke verklaring is dat de neurale netwerken niet in staat zijn extreme waarden te reproduceren.

Geconcludeerd kan worden dat neurale netwerken meer en meer gebruikt worden als tools voor data analyse en specifiek voor ritdistributie. Echter, in hoeverre neurale netwerken beter presteren dan zwaartekrachtmodellen in het laatste geval is onduidelijk. Het doel van deze studie is het vergelijken van de kwaliteit van ritdistributie schattingen van twee methodieken: (i) tweevoudige beperkte zwaartekrachtmodellen en (ii) neurale netwerken. Doel is om vanuit een beperkt aantal waargenomen relaties in een HB matrix, gegeven

bekende randtotalen, de volledige HB-matrix te schatten. In dit onderzoek zijn de onbekendheid met de werkelijke HB-matrix, ruis en meetfouten uitgesloten door de ‘volledige HB-matrix’ met een zwaartekrachtmodel op te stellen. Het onderzoek is opgesplitst in twee delen. De eerste test gaat uit van synthetische data. In de tweede test wordt gebruikt gemaakt van NRM data van Rotterdam Rijnmond.

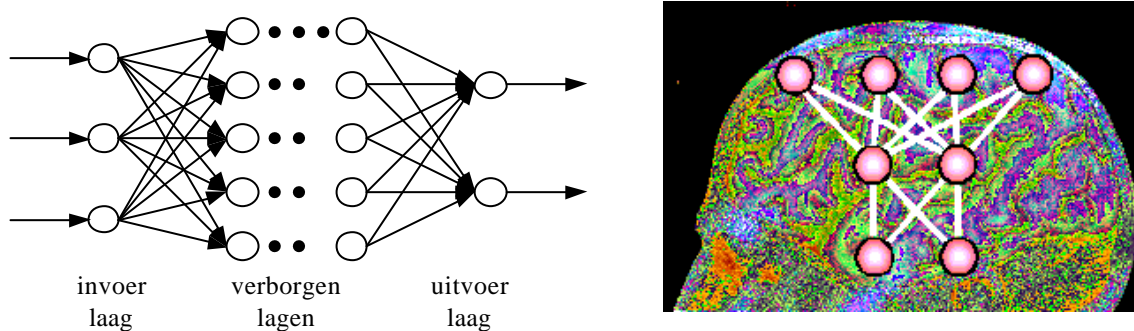
Het onderzoek onderscheidt zich van andere onderzoek op een aantal gebieden. Ten eerste wordt gebruik gemaakt van zowel synthetische alsmede echte datasets. Ten tweede zijn de synthetische datasets opgezet met een grote variatie aan complexiteit. Hierdoor kan duidelijk onderzocht worden in hoeverre complexiteit van de data de resultaten beïnvloedt. Ten derde zijn minimale steekproefgroottes vastgesteld om te komen tot statistisch significante resultaten. Ten vierde, overeenkomstig de studie van Thill en Mozolin (2000), worden randvoorwaarden gebruikt om het neurale netwerk te trainen. Tot slot worden beide methodes gecalibreerd op basis van grote variatie aan databeschikbaarheid. Met name de situaties waarin gecalibreerd wordt met weinig data, schaarse data, komen overeen met de werkelijkheid.

Dit artikel zal allereerst een korte introductie geven op neurale netwerken. Daarna zal achtereenvolgens ingegaan worden op de onderzoeksopzet, de resultaten en een verdieping op de resultaten. Afgesloten wordt met conclusies en aanbevelingen.

2. Neurale Netwerken

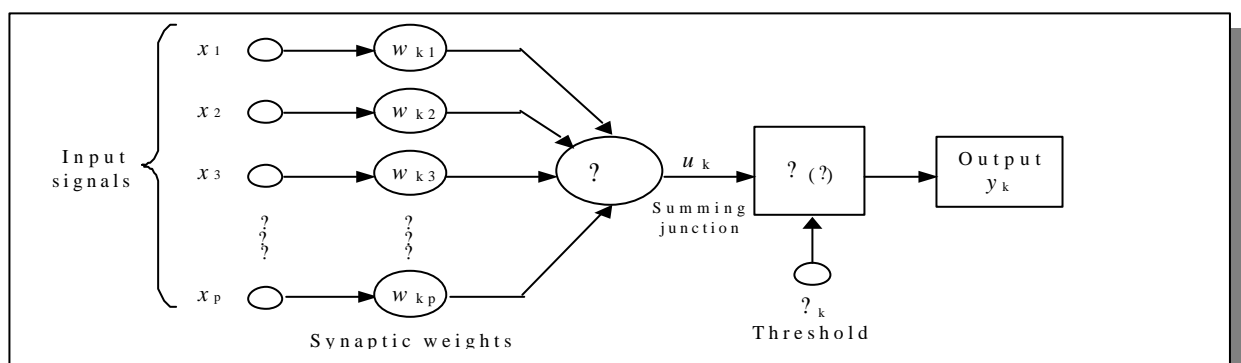
Artificiële Neurale Netwerken (ANN's) zijn parallele systemen (geënt op hersenen) en daardoor net als hersenen geschikt voor taken als patroonherkenning, perceptie en controle; taken waarmee lineaire systemen (computers) erg veel moeite hebben. Verder voordeel is dat ze robuust zijn: data voorzien van ruis wordt, tot op zekere hoogte, op een correcte manier verwerkt. De ANN's imiteren naast de hersen-netwerkarchitectuur (beide bestaan uit eenheden die neuronen worden genoemd) ook hersenprocessen: signalen die een eenheid bereiken na een non-lineaire bewerking te zijn ondergaan worden doorgegeven aan alle voorwaarts gekoppelde neuronen. De werking van ANN's in het kort (zie ook Huisken en van Maarseveen, 1998): signalen komen een netwerk binnen via een invoerlaag bestaande uit bijvoorbeeld n neuronen, worden non-lineair bewerkt en doorgegeven aan alle voorwaarts

verbonden eenheden in de volgende laag: dit kan een verborgen laag zijn (het aantal verborgen lagen is variabel), maar ook de uitvoerlaag. Deze bewerkingen herhalen zich totdat de uitvoerlaag met m neuronen is bereikt. Hierna geeft het ANN zijn signalen weer af (figuur 1).



Figuur 1: Schematische weergave van een ANN architectuur

Uit het bovenstaande blijkt dat een bewerking van een ANN een non-lineaire ‘mapping’ van een n -dimensionale invoervector (‘patroon’) op een m -dimensionale uitvoervector is. De manier waarop deze ‘mapping’ tot stand komt wordt bepaald door de netwerkarchitectuur, de overdrachtsfunctie en de leerregel die van kracht is tijdens het leerproces. De architectuur is afhankelijk van het aantal artificiële neuronen per laag en de verbindingen van neuronen tussen lagen en onderling in een laag. Ter verduidelijking is een schematisch overzicht van een artificeel neuron gegeven in figuur 2.



Figuur 2: Non-lineair model van een artificeel neuron k

Nog een overeenkomst naast de structurele analogie tussen ANN's en hersenen is het vermogen om te ‘leren’. Tijdens het leerproces wordt het ANN signalen aangeboden en die

bewerkt worden tot een uitvoersignaal. De vele ANN types behoren over het algemeen tot een der volgende netwerken:

- ?? netwerken die onder toezicht leren; de *supervised learning* netwerken. Het verschil tussen de doeluitkomst en de door het ANN gegenereerde uitkomst is een foutterm die een maat is voor het bijstellen van de synaptische gewichten en wel zó dat de cumulatieve foutterm van alle aangeboden trainingspatronen geminimaliseerd wordt.
- ?? netwerken die zonder toezicht leren; de *unsupervised learning* netwerken. Hierbij worden statistische regelmatigheden in de invoerdata gezocht tijdens de trainingsfase; er zijn dus geen doeluitkomsten behorende bij invoerpatronen. Deze ANN's bezitten na de trainingsfase een soort van interne classificatie.

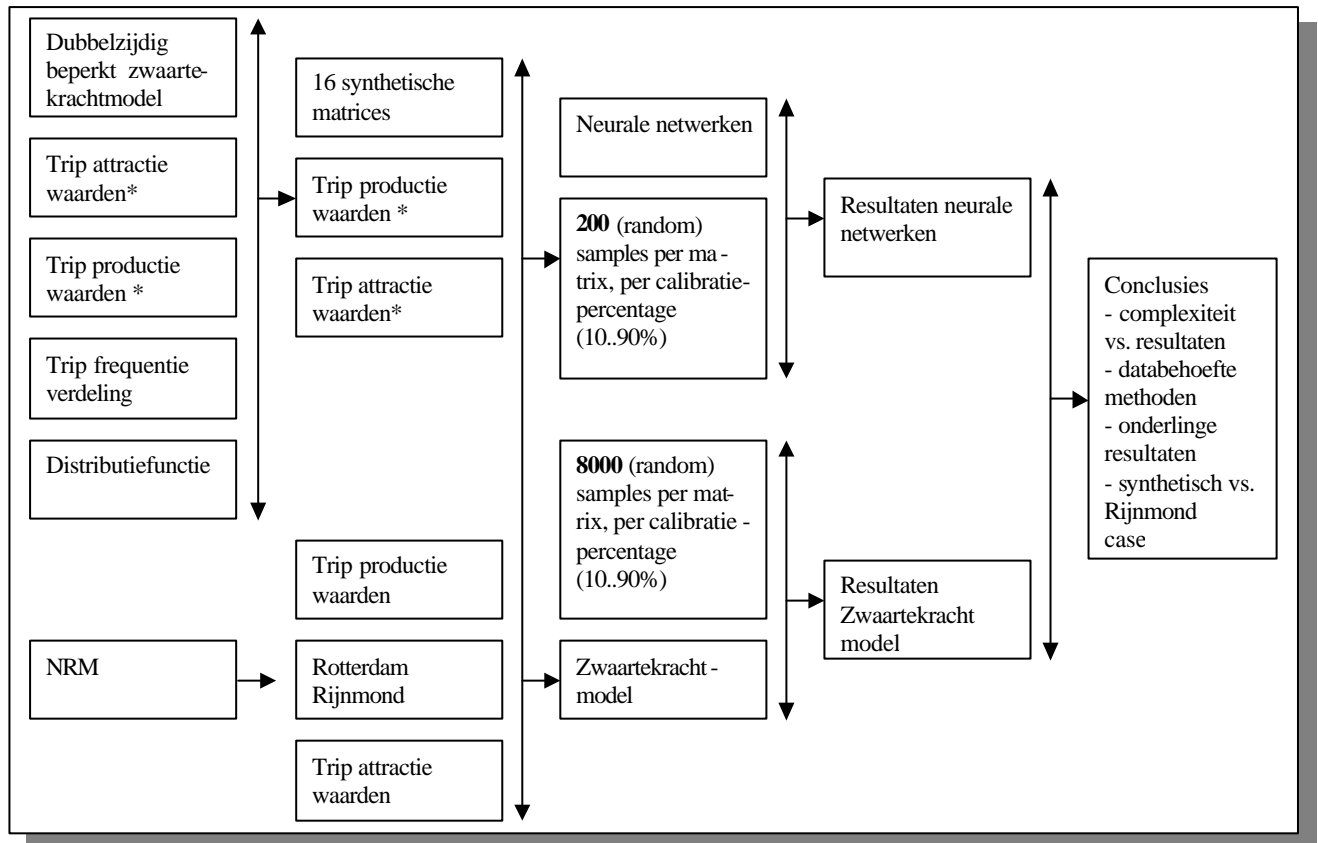
Het gaat te ver om hier allerlei verschillende ANN's te noemen. De voordelen van de techniek bestaan uit; (i) de korte ontwikkelingsfase; (ii) de relatieve ongevoeligheid voor ruis; (iii) de snelle processortijd (tijdens de operationele fase); (iv) de aanpassingmogelijkheden bij een veranderende omgeving; en (v) het vermogen om willekeurige functies tussen in- en uitvoer te leren. De nadelen van de techniek zijn; (i) de techniek is niet-deterministisch (black-box); (ii) de techniek heeft veel trainingsdata nodig; (iii) de optimale netwerkarchitectuur moet door middel van 'trial and error' gezocht worden.

Bishop (1995) en Haykin (1994) geven een uitgebreider overzicht op het gebied van neurale netwerken. Literatuuroverzichten op het gebied van verkeer en vervoer waarin neurale netwerken een rol spelen staan beschreven in Dougherty (1995), Himanen *et al.* (1998) en in Huiskens (1998).

3. Onderzoeksopzet

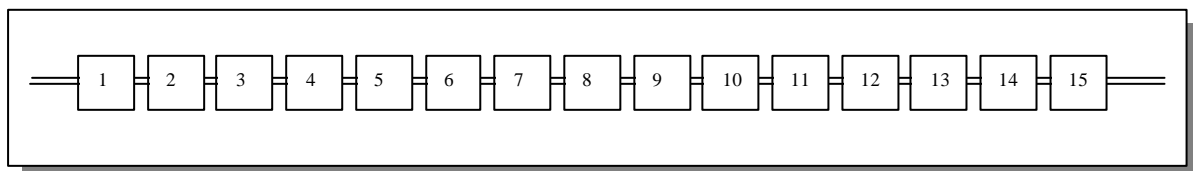
De focus van dit onderzoek ligt op het schatten van de ritdistributie met neurale netwerken en zwaartekrachtmodellen bij een grote variatie aan complexiteit: 16 synthetische en 1 werkelijke situatie. Het creëren van synthetische data geeft de mogelijkheid om te spelen met complexiteit. Deze aanpak geeft inzicht in de werkelijk te bestuderen complexiteit, zonder bang te zijn voor het modelleren van ruis of onduidelijke verbanden in de data. Afgeleid hiervan wordt onderzocht in welke mate beide methodieken goede resultaten leveren wanneer de databeschikbaarheid voor calibratie wordt gevarieerd. Voor de synthetische cases zijn

zowel een synthetisch netwerk alsmede een synthetische skimmatrix op basis van de euclidische afstand gecreëerd.



Figuur 3: Onderzoeksopzet











Het gebruik van de euclidische afstand verbetert de resultaten van zwaartekrachtmodellen (Thill en Mozolin, 2000). Het gebruikte netwerk bestaat uit 15 zones, gelegen op uniforme afstanden van elkaar op een streng (zie Figuur 4). In totaal is uitgegaan van 15000 ritten tussen de zones.



Figuur 4: Synthetisch netwerk

Op basis van het netwerk, de skimmatrix en de 15000 ritten zijn met behulp van een **zwaartekrachtmodel** 16 synthetische HB matrices (225 cellen) van toenemende complexiteit

gegenereerd. Het verschil in complexiteit tussen matrices wordt gecreëerd door variatie aan te brengen in de verdelingen van de ritten over de rand- en kolomtotalen (Figuur 5). De combinaties die weergegeven worden door streepjes zijn niet getest. Deze vormen een herhaling van een andere opgenomen combinatie.

1 (2)	2 (3)	3 (3)	4 (5)	5 (5)	6 (6)		Productie
	7 (4)	8 (5)	9 (6)	10 (6)			
			11 (6)	12 (6)			
			13 (8)	14 (8)			
				15 (8)			
					16 (10)	RND	
					RND		
Attractie							

1..16 geteste matrices
 ||||| niet getest: deze configuratie is al getest
 (..) matrix complexiteit

Figuur 5: Synthetische HB matrices

Vervolgens worden neurale netwerken en zwaartekrachtmodellen gecalibreerd voor alle matrices. Om de afhankelijkheid van en behoefte aan data te onderzoeken wordt de calibratie uitgevoerd op basis van random trekkingen van 10 tot 90% van de cellen van de HB matrices. Deze data, aantal ritten tussen zones, wordt gecombineerd met de bekend veronderstelde productie en attractie totalen die ook zijn gebruikt bij de generatie van de synthetische HB matrices (Figuur 5). Dit heeft als consequentie dat wanneer het percentage calibratiedata de 100% nadert, het zwaartekrachtmodel niets anders dan de eigen creatie reproduceert; immers, de data is gegenereerd met behulp van een zwaartekrachtmodel.

Uit uitvoerige statistische analyses vooraf is gebleken dat het benodigde aantal random trekkingen voor statistisch significante resultaten van dezelfde betrouwbaarheid, per methodiek sterk verschilt; 200 en 8000 trekkingen voor respectievelijk neurale netwerken en

zwaartekrachtmodellen. De zwaartekrachtmodellen worden gecalibreerd met behulp van het algoritme van Van Zuilekom (2003). Neurale netwerken worden getraind en getest met Matlab 5.3 (Math Works inc. 1996, 1998).

Naast de synthetische matrices is ook NRM data van Rotterdam Rijnmond gebruikt. In totaal bevat deze invoer 1,9 miljoen ritten, verdeeld over 225 HB-relaties. De weerstand tussen de zones is gedefinieerd als de reistijd.

4. Beoordelingsmaat

Om de resultaten van beide methodes te kunnen vergelijken is de RMSE (Root Mean Square Error) gebruikt. De RMSE wordt wiskundig als volgt beschreven:

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

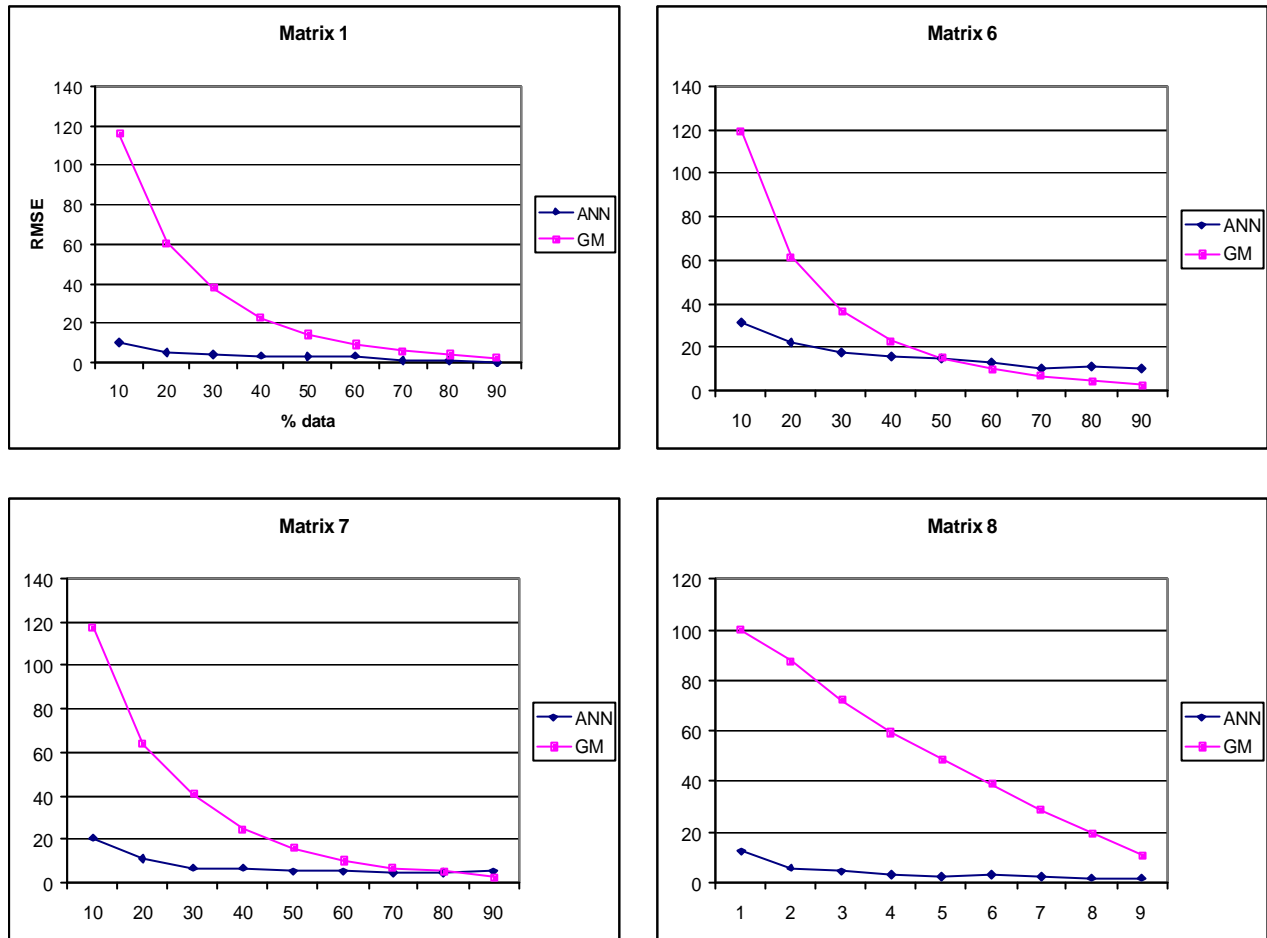
met:

N = het aantal cellen per matrix ($15 \cdot 15 = 225$)

Hoe lager de RMSE-waarde, hoe beter het resultaat. Verder is er nog gekeken naar de gemiddelde rit lengte verdeling en de standaard deviatie van de resultaten.

5. Resultaten

De resultaten van de calibratie van neurale netwerken en zwaartekrachtmodellen op basis van matrices 1, 6, 7 en 8 worden weergegeven in figuur 6. De overige matrices vertonen sterk overeenkomstige resultaten met matrices 1, 6 en 7. Matrix 8 geeft enigszins afwijkende resultaten. In het algemeen valt op dat het calibratiepercentage grote invloed heeft op het resultaat van beide methoden. Echter de figuren laten zien dat het zwaartekrachtmodel gevoeliger is voor de beschikbaarheid van data bij lage percentages. Bij lagere percentages, tot 50%, geven neurale netwerken daarom ook betere resultaten. Verder kan geconcludeerd worden dat geen van de neurale netwerken overall beter scoort dan zwaartekrachtmodellen. In gevallen waarbij voldoende data voorhanden is, >80-90%, lijken zwaartekrachtmodellen betere resultaten te geven. Dit is niet verwonderlijk. De zwaartekrachtmodellen reproduceren bij grote beschikbaarheid van data immers in hoge mate hun eigen creaties.

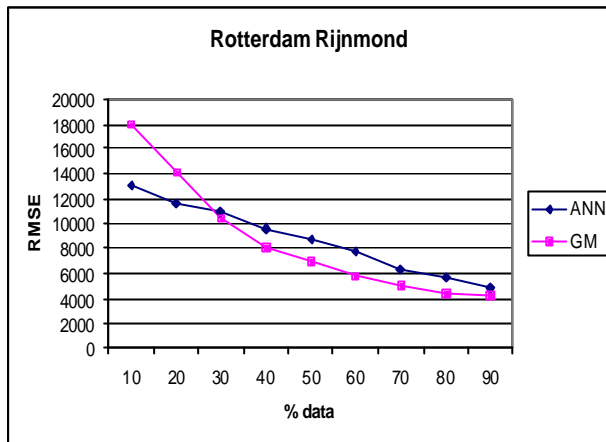


Figuur 6: RMSE van neurale netwerken (ANN) en zwaartekrachtmodellen (GM)

Matrix 8 vertoont afwijkende resultaten. De opbouw van de matrix, zie Figuur 5, bepaalt sterk het slechte resultaat voor het zwaartekrachtmodel. Doordat de randen van de matrix een distributiefunctie afdwingen die sterk afwijkt van een theoretische standaard functie, met op de korte afstanden relatief veel trips, heeft het model moeite met schatten. Hier valt het voordeel op voor neurale netwerken. Zonder voor-verondersteld verband tussen trips en afstand is het neurale netwerk in staat een goede schatting te bewerkstelligen.

De resultaten van de Rotterdam Rijnmond data leiden tot dezelfde conclusies. Benadrukt moet worden dat de resultaten van neurale netwerken fractioneel slechter zijn dan bij de synthetische matrices. Echter, het patroon is hetzelfde. Ten tweede moet opgemerkt worden dat de RMSE waarden sterk hoger liggen. Dit komt vanwege het feit dat het aantal ritten ongeveer 130 maal hoger is dan in de synthetische matrices. Wordt de RMSE uitgezet tegen

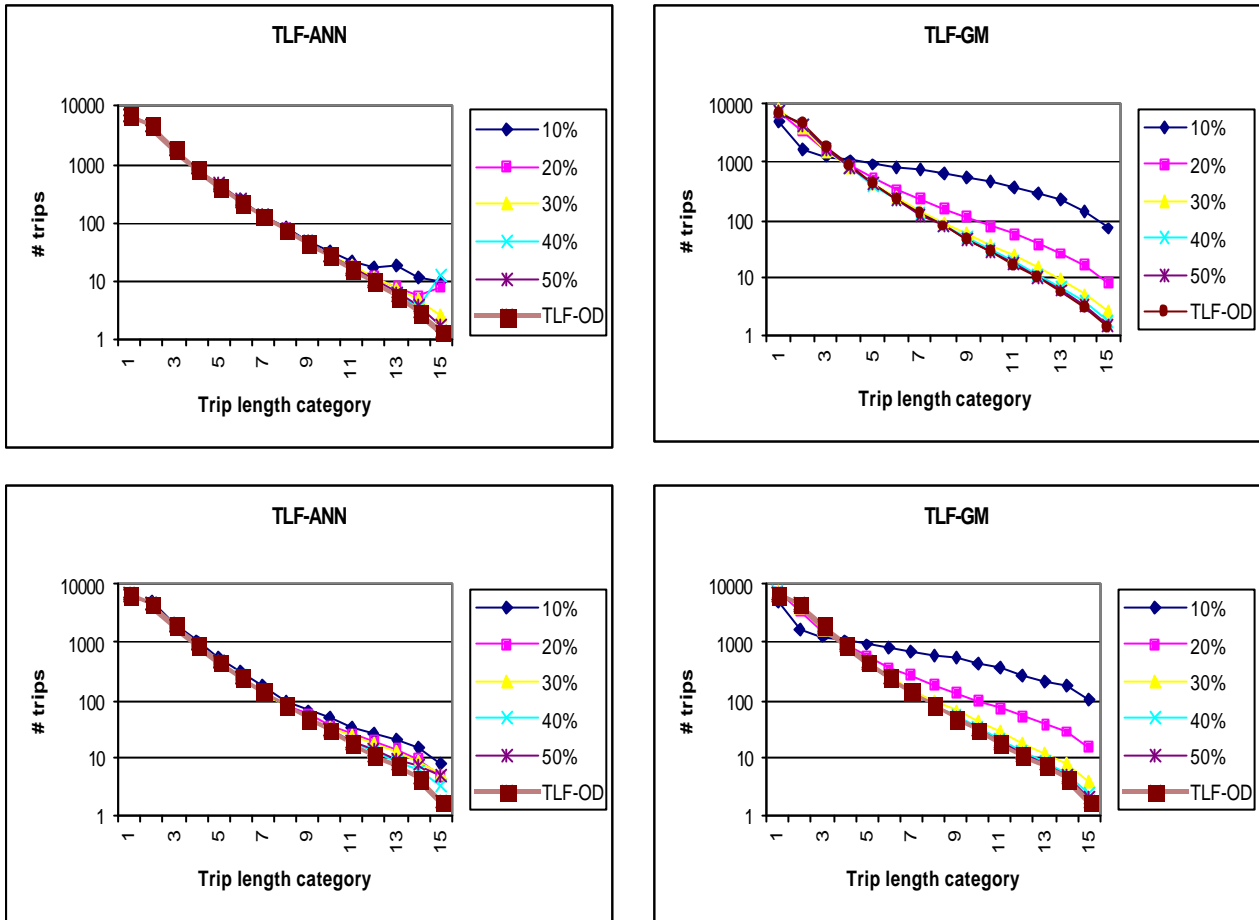
het aantal ritten dan blijkt dat de RMSE waarden 2 maal hoger zijn dan in het geval van de synthetische waarden.



Figuur 7: RMSE van neurale netwerken en zwaartekrachtmodellen in de Rijnmond case

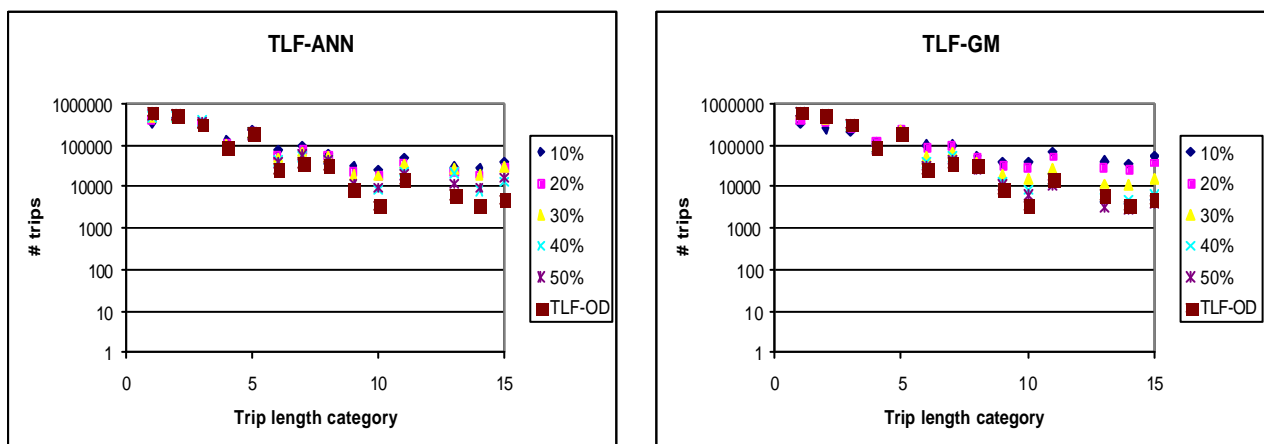
De resultaten op basis van de synthetische en werkelijke data laten zien dat neurale netwerken beter presteren als er een schaarste aan data is. In de werkelijkheid is er een gebrek aan data. Ortuzar en Willumsen (2001) concluderen dat waargenomen rit matrices altijd schaars gevuld zijn; het aantal lege cellen is groot wat leidt tot ijle matrices. *‘Als een steekproef bestaat uit 20% van de relaties, dan is de kans groot dat er op een specifieke HB relatie geen waarnemingen worden gedaan. Matrix expansie technieken kunnen dan gebruik worden om lege cellen alsnog te vullen. Echter, benadrukt moet worden dat expansieprocessen de fouten in de waargenomen rit matrices (kunnen) propageren.’*

Een goede fit op randtotalen van een HB matrix, gecombineerd met een lage RMSE waarde hoeft niet eenduidig te resulteren in een goed geschatte matrix op individueel cel niveau. Extra analyses zijn dus gewenst. Naast de RMSE waarden zijn daarom ook de ritlengte frequentie verdelingen bestudeerd. De onderstaande figuren tonen de frequentie verdelingen van beide methodes bij calibraties op basis van 10-50% van de beschikbare data. Alleen data van matrix 1, 6 en Rotterdam Rijnmond wordt getoond. De overigen matrices hebben vergelijkbare resultaten.



Figuur 8: Rit lengte frequenties voor neurale netwerken (ANN) en zwaartekrachtmodel (GM) vergeleken met de frequentieverdeling van de volledige data set (TLF-OD), matrix 1 en 6

Figuur 8 en 9 laten zien dat voor beide methoden de resultaten verbeteren naarmate de dataset voor calibratie groter wordt. Ten tweede lijken zwaartekrachtmodellen moeite te hebben met het schatten van zowel hoge als lage aantallen ritten.

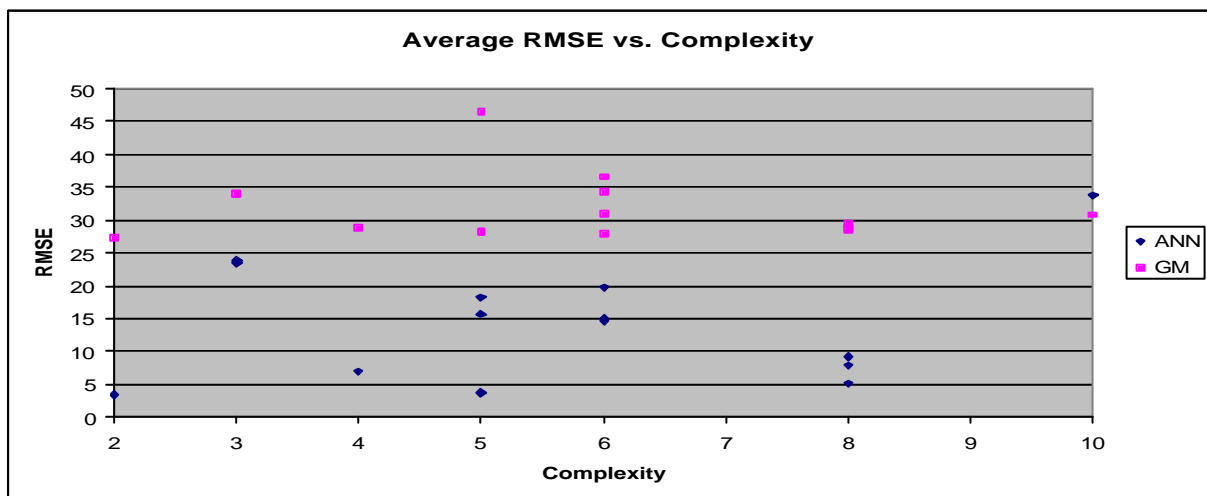


Figuur 9: Ritlengte frequenties voor neurale netwerken en zwaartekrachtmodel (GM) vergeleken met de frequentieverdeling van de volledige data set (TLF-OD) voor de Rijnmond case

Het totale resultaat bij lage percentages is beter als neurale netwerken worden gebruikt. In het algemeen kan geconcludeerd worden dat de rit lengte frequentieverdelingen goed benaderd worden.

6. Verdieping van de resultaten

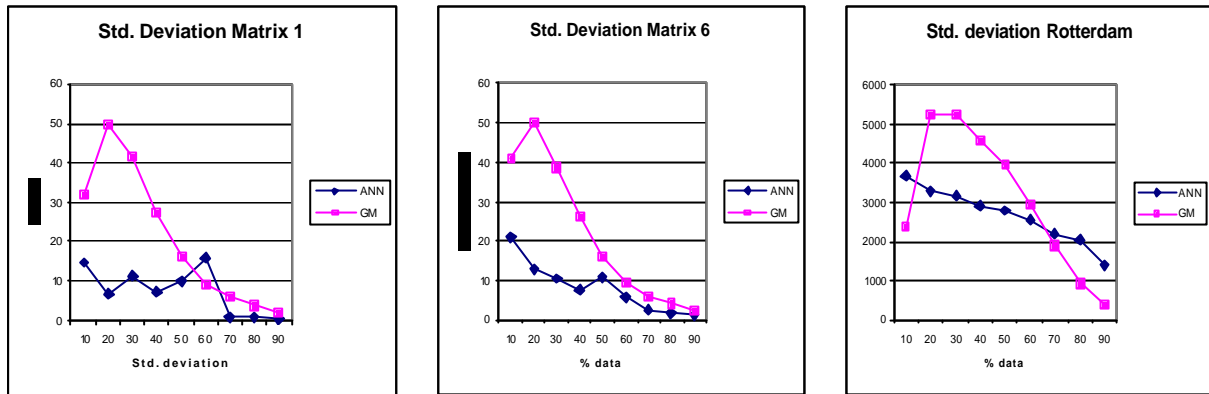
Is er een verklaring voor deze resultaten? Ten eerste laten alle matrices ruwweg hetzelfde patroon zien; bij lage percentages zijn de resultaten van neurale netwerken beter, bij hoge percentages daarentegen scoren zwaartekrachtmodellen beter. Neurale netwerken profiteren optimaal van hun karakteristiek om missende data te extrapoleren. De betere resultaten van zwaartekrachtmodellen komt voort uit de opbouw van de matrices, zoals beschreven in de onderzoeksopzet; de input matrices zijn ‘gebouwd’ gebruik makend van een zwaartekrachtmodel. Dus, wanneer de calibratie dataset 100% nadert zal het model haar eigen creatie opnieuw schatten. Dit verzwakt de conclusie dat bij hogere percentages van de dataset de zwaartekrachtmodellen betere resultaten geven. Ten tweede is er een duidelijk zichtbaar verschil wanneer de RMSE waarden worden uitgezet tegen de complexiteit van de matrices.



Figuur 10: Gemiddelde RMSE versus de complexiteit van de matrices

Figuur 10 geeft de resultaten van de gemiddelde RMSE per matrix, uitgezet tegen de complexiteit. De volgende conclusies kunnen getrokken worden op basis van dit figuur: (i) bij een minimale complexiteit, 2, is de data het meest gestructureerd en zodoende zijn de neurale netwerk resultaten het best; patroon herkenning; (ii) bij een maximale complexiteit zijn de neurale netwerk resultaten het slechts. De complexe matrix herbergt de minste patronen, waardoor de RMSE toeneemt; (iii) de resultaten binnen complexiteitsgroepen zijn grotendeels

gegroepeerd. Dit toont dat de opbouw complexiteit consistent is.; (iv) en de verschillen tussen de zwaartekracht modellen en neurale netwerken zijn niet stabiel. De standaard deviaties van neurale netwerk resultaten zijn lager dan de deviaties voor zwaartekrachtmodellen; de lage afwijking maakt kleinere steekproefgroottes mogelijk. Dit versterkt de voorkeur voor het gebruik van neurale netwerken andermaal.



Figuur 11: Standaard deviatie matrix 1, 6 en Rotterdam Rijnmond

7. Discussie en Conclusies

Dit onderzoek laat zien dat neurale netwerken beter presteren dan zwaartekrachtmodellen in zowel synthetische situaties als ook in de werkelijkheid. Met name in de situaties overeenkomstig de realiteit, waar data schaars is, zijn de resultaten voor neurale netwerken positief te noemen. Dit resultaat kan veelbelovend zijn voor het modelleren van ritdistributie, wat een belangrijke stap is voor goede transport planning. De controleerbaarheid van de invoer data versterkt de resultaten.

De resultaten moeten echter voorzichtig behandeld worden. Ten eerste geven de resultaten geen absolute antwoorden over de schattingen van de neurale netwerken en zwaartekrachtmodellen. De resultaten zijn gebaseerd op verschillende random trekkingen van calibratiedata van een van tevoren vastgelegd percentage. Deze resultaten zijn gemiddeld en kunnen dus uiteindelijk alleen voor vergelijkingsdoeleinden gebruikt worden.

De synthetische data is gegenereerd door middel van een zwaartekrachtmodel. Daarmee is een situatie gecreeerd die bij uitstek geschikt zou moeten zijn voor het zwaartekrachtmodel. Desondanks presteert het zwaartekrachtmodel alleen beter bij hogere databeschikbaarheid. In de meer praktijkrelevante situaties met een lagere databeschikbaarheid overtuigt het neurale netwerk. Dit versterkt in zekere mate de kwaliteit van de neurale netwerk resultaten.

Het onderzoek naar de rit lengte frequentieverdelingen toont dat beide methoden op cel niveau redelijke schattingen kunnen maken. De standaard afwijkingen laat zien dat er weliswaar veel variatie zit in de resultaten, maar dat de standaard afwijkingen voor neurale netwerken significant lager liggen.

Wat verder is benadrukt is het gedrag van beide methoden als de complexiteit toeneemt. De neurale netwerken geven de beste resultaten wanneer de data het meest gestructureerd is. Echter, neurale netwerken geven ook goede resultaten bij de random matrices (matrix 6 en 16) en bij de Rotterdam Rijnmond matrix. Deze zijn in principe minder gestructureerd. Door het grote aantal ritten in de Rijnmond case zijn de RMSE waarden hoger dan in het geval van de synthetische datasets.

Dit onderzoek toont dat het ritdistributie probleem ingewikkeld is. De tot op heden gebruikte zwaartekrachtmodellen slaan in sommige gevallen de plank geheel mis. Doordat in deze studie alle data voorhanden is kunnen de misschattingen van beide methodieken in kaart gebracht worden. In werkelijkheid is hiertoe geen mogelijkheid. De fouten die gegenereerd worden tijdens de distributiefase propageren verder in het klassieke 4-traps model.

Zoals reeds in de inleiding staat geschreven kan het schatten van modellen op beperkte hoeveelheden data leiden tot resultaten met een grote spreiding. Het feit dat er in de werkelijkheid maar één dataset beschikbaar is, doet twijfelen aan de nauwkeurigheid van het resultaat. De mate waarin de beschikbare data ‘past’ bij het gebruikte model, de modelfit, bepaalt uiteindelijk het succes van de calibratiecyclus. Ortuzar en Willumsen (2001) concluderen dat waargenomen matrices altijd ijl zijn. Als maar 20% van de te schatten data aanwezig is voor calibratie kan soms gesteld worden dat het schatten van een model niet meer is dan een grote gok. En omdat er geen referentiematrix beschikbaar is zal in werkelijkheid nooit een RMSE berekend kunnen worden.

Een belangrijke conclusie komt uit de statistische analyses. Bij lage percentages is het verschil in het aantal trekkingen, die noodzakelijke zijn voor statistisch significante resultaten, voor neurale netwerken 40x kleiner dan voor zwaartekrachtmodellen; 200 respectievelijk

8000. Dit geeft inzicht in enerzijds de totale calibratieduur voor zwaartekrachtmodellen. Anderzijds toont dit aan dat het gebruik van zwaartekrachtmodellen gevaarlijk kan zijn. De kans dat de beschikbare praktijkdata ‘past’ bij het zwaartekrachtmodel is vele malen lager dan voor een neuraal netwerk.

Uiteindelijk pleiten de resultaten voor meer onderzoek naar het gedrag van zowel zwaartekrachtmodellen als neurale netwerken onder verschillende omstandigheden. Dit onderzoek geeft aan dat nieuwe methoden een forse stap voorwaarts kunnen zijn in de richting van een betrouwbare ritdistributie.

Literatuur

- Bishop, C.M. (1995) *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford, UK.
- Black, W.R. (1995) Spatial interaction modelling using artificial neural networks, *Journal of Transport Geography*, Vol. 3, No. 3. pp. 159-166.
- Collins, M.G., Steiner, F.R. en Rushman (2001) M.J. Land-Use Suitability Analysis in the United States: Historical Development and Promising Technological Achievements, *Environmental Management*, Vol. 28, No. 5, pp. 611–621.
- Currit, N. (2002) Inductive regression: overcoming OLS limitations with the general regression neural network, *Computers, Environment and Urban Systems*, Vol. 26, pp. 335–353.
- Dougherty, M. (1995) A review of Neural Networks applied to Transport, *Transportation research-C*, Vol. 3. No.4 pp 247-260.
- Faghri, A. and Sandeep, A. (1998), Analysis of Performance of Backpropagation, ANN with Different Training Parameters, *Neural Networks in Transport Applications*, eds. Himanen, Nijkamp and Reggiani, Ashgate.
- Fischer, M.M. (1998) Computational Neural Networks: An Attractive Class of Mathematical Models for Transportation Research, *Neural Networks in Transport Applications*, eds. Himanen, Nijkamp and Reggiani, Ashgate.
- Fischer, M.M., en Gopal, S. (1994) Artificial Neural Networks. A new approach to Modelling Interregional Telecommunications Flows, *Journal of Regional Science*, Vol. 34, No. 4.
- Fotheringham, A.S. and O’Kelly, M.E. (1989). *Spatial Interaction Models: Formulations and Applications*, Kluwer Academic Publishers, London.
- 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*, Vol. 28. No. 1, pp. 38-55.
- Haykin, S. (1994) *Neural Networks: A Comprehensive Foundation*. Prentice Hall International, Inc., Upper Saddle River, New Jersey.
- Himanen, V. Nijkamp, P. and Reggiani, A. (1998) *Neural Networks in Transport Applications*, Ashgate.
- Huisken, G. en Coffa, A. (2000) Neural Networks and Fuzzy Logic to improve Rit 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.
- Huisken, G. en Maarseveen, M.F.A.M. van. (1998) Congestievoorspelling gebaseerd op neurale netwerken: een krachtig hulpmiddel voor sturing van verkeersstromen, *Colloquium Vervoersplanologisch Speurwerk 1998: sturen met structuren*, deel 4, pp. 1793 - 1812, Delft, 12-13 november 1998.
- Miller, D.M., Kaminsky, E.J. en Rana, S. (1995) Neural network classification of remote-sensing data, *Computers & Geosciences*, Vol. 21, No. 3, pp. 377–386.
- Mozolin, M., Thill, J.C. and Lynn Usery, E.L. (2000) Rit distribution forecasting with multiplayer perceptron neural networks: A critical evaluation, *Transportation Research C*, Vol. 34, No. 1, pp. 53.
- Openshaw, S. and Openshaw, C. (1997) *Artificial intelligence in Geography*, John Wiley and Sons, Chichester.
- Ortuzar, J. de D. en Willumsen, L.G. (2001) *Modelling Transport*, John Wiley & Sons, Ltd, Chichester, UK.
- Pijanowski, B.C., Brown, D.G. Shellito, B.A. en Manik, G.A. (in press) Using neural networks and GIS to forecast land use changes: a Land Transformation Model, *Computers, Environment and Urban Systems*.
- 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*, Vol. 22, No. 3, pp 203-218.
- Rodrigue, J.-P. (1997) Parallel modelling and neural networks: an overview for transportation/land use systems, *Transportation Research C*, Vol. 5, No. 5, pp. 259–271.
- The Math Works Inc. (1996) *Matlab: Using Matlab*, version 5.
- The Math Works Inc. (1998) *Neural Network Toolbox: for Use with Matlab*, eds. Demuth, H. and Beale, M., version 3.
- Thill, J.C. and Mozolin, M. (2000) Feedforward Neural Networks for Spatial Interaction: Are They Trustworthy Forecasting Tools?, in *dsljsal;djf;alsjdf;la*,
- Tillema, F. Huisken, G. and Maarseveen, M.F.A.M. (2002) Neurale netwerk technieken ten behoeve van landgebruik/transport modellen, *Presented at the Dutch Colloquium Vervoersplanologisch Speurwerk*.
- Zuilekom, van K.M. (2003), *Calibrating Gravity models: algorithm*, working document, Centre for Transport studies, University of Twente.