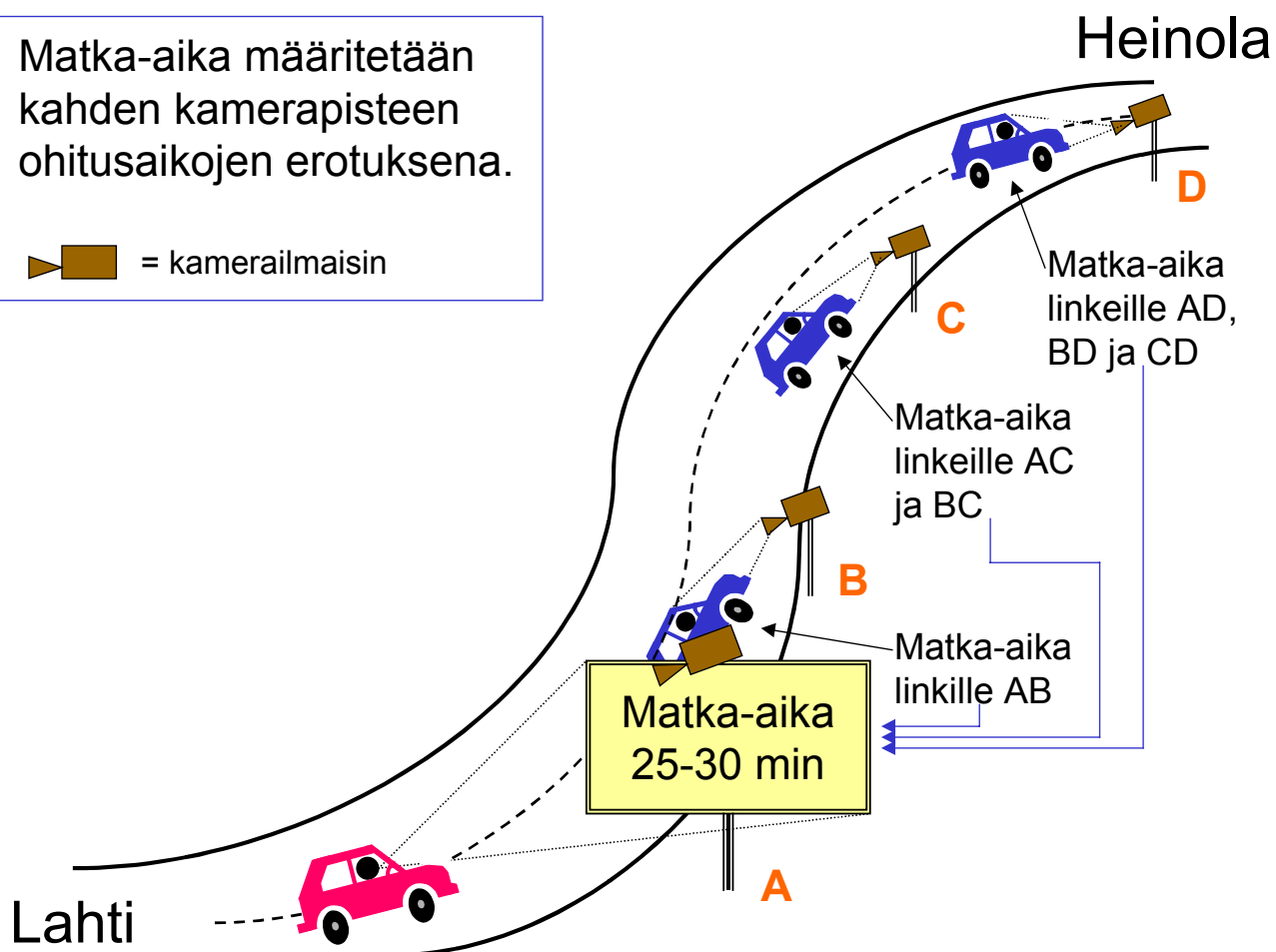


Pääteiden lyhyen aikavälin matka-aikaennusteet

Tiehallinnon selvityksiä 5/2002

Matka-aika määritetään kahden kamerapisteen ohitusaikojen erotuksena.

 = kamerailmaisim



Satu Innamaa, Laura Lanne, Kerkko Vanhanen ja Matti Pursula

Pääteiden lyhyen aikavälin matka-aikaennusteet

Tiehallinnon selvityksiä 5/2002

Tiehallinto

Helsinki 2002

ISSN 1457-9871
ISBN 951-726-863-7
TIEH 3200739

Oy Edita Ab
Helsinki 2002

Julkaisua myy/saatavana:
Tiehallinto, julkaisumyynti
Telefaksi 0204 22 2652
S-posti julkaisumyynti@Tiehallinto.fi
www.Tiehallinto.fi/julk2.htm

Tiehallinto
TIEHALLINTO
Liikenteen palvelut
Opastinsilta 12 A
PL 33
00521 HELSINKI
Puhelinvaihte 0204 2211

Satu Innamaa, Laura Lanne, Kerkko Vanhanen ja Matti Pursula: Pääteiden lyhyen aikavälin matka-aikaennusteet. Helsinki 2002. Tiehallinto, Liikenteen palvelut. Tiehallinnon selvityksiä 5/2002. 81 s. + liitt. 16 s. ISSN 1457-9871, ISBN 951-726-863-7, TIEH 3200739.

Asiasanat: liikennetiedotus, ennusteet, matka-aika
Aiheluokka: 11, 20

TIIVISTELMÄ

Tämän tutkimuksen tarkoituksena oli selvittää lyhyen aikavälin matka-aikaennusteiden tekemistä sekä sitä, kuinka hyvin matka-aika saadaan estimoitua pistemittausten avulla. Tutkimuskohteina olivat valtatie 4 Lahden ja Heinolan välillä sekä Kehä I. Koejaksot on varustettu automaattisilla matka-ajan seurantajärjestelmillä.

Ennustemalleja tehtiin kahdella eri menetelmällä: neuroverkkoihin ja sumeaan logiikkaan perustuen. Mallien hyvyttä verrattiin olemassa olevan järjestelmän antaman tiedon tarkkuuteen. Ennusteiden hyvyttä tarkasteltiin tilastollisesti ja matka-aikatiedotuksen näkökulmasta.

Valtatien 4 ennustemallit ennustivat tarkasteluhetkeä seuraavan minuutin aikana linkeille lähtevien ajoneuvojen keskimatka-aikaa viimeksi mitattujen liikennetietojen perusteella. Oikein ennustettujen matka-aikojen osuus vaihteli voimakkaasti linkiltä toiselle.

Kehä I:lle tehtiin sekä estimaatti- että ennustemallit. Estimaattimalleilla pyrittiin estimoimaan tietyn linkin matka-aika muiden mittaustietojen perusteella tilanteessa, jossa kyseisen linkin matka-ajan mittaaminen ei onnistu. Tuloksena oli, että ennustemallit antoivat selvästi parempia tuloksia kuin estimaattimallit.

Kaikki sumean mallin tuloksille lasketut tilastolliset virhetermit olivat neuromallin vastaavia virheitä suurempia. Neuromalli toimi sumeaa mallia paremmin myös matka-aikatiedotuksen näkökulmasta tarkasteltuna. Sumeita malleja keskenään vertailtaessa nähtiin, että Gödelin algebran mukainen malli oli hiukan similaarista mallia parempi. Nykyjärjestelmä estimoit ruuhkaliikennettä sumeita malleja huonommin.

Heinola–Kymijärvi- ja Kymijärvi–Heinola-linkeille tehtiin ruuhkamallit kokoomalla opetusjoukko ainoastaan ruuhkaisten päivien aineistoista. Molemmilla linkeillä päästiin näiden mallien avulla parempaan tulokseen kuin parhaalla koko päivän aineistolla opetetulla mallilla. Paras ruuhkamalli ennusti ruuhkassa 84 prosenttia ajasta oikein.

Tutkimuksessa haluttiin lopuksi selvittää, kuinka mittausjärjestelmän rakenne vaikuttaa ennusteen laatuun. Havaittiin muun muassa, että matka-aikaennuste oli sitä parempi, mitä pidempi linkki oli, ja että linkillä oleva ylimääräinen kamera paransi tuloksia enemmän kuin sen ulkopuolella oleva mittauspiste. Ennen Lahtea sijainneen LAM-aseman liikennetiedot näyttivät parantavan tulosta enemmän kuin Vierumäen LAM-aseman tiedot.

Valtatiellä 4 saavutetut tulokset olivat suhteellisen tyydyttäviä. Kehä I:n malleja sitä vastoin voi pitää korkeintaan välttävinä. Kehä I:llä matka-aikojen seurantajärjestelmän yksityiskohtaisuus ei selvästikään riitä ennusteiden tekoon ja kamerapisteitä pitäisi olla enemmän. Molempien kohteiden ennusteet paranisivat varmasti, mikäli liikennetilannekuva täydennettäisiin matka-aikatietojen lisäksi kattavammilla liikennemäärätiedoilla. Havaituista puutteista huolimatta tulokset olivat sen verran lupaavia, että mallista kannattaisi tehdä kokeiltavaksi tuotantoversio.

SAMMANDRAG

Avsikten med denna undersökning var att klargöra hur man gör kortsiktiga restidsprognoser samt hur bra man kan estimerastiden av punktmätningar. Studieobjekten var riksväg 4 mellan Lahtis och Heinola och Ring I. Experimentsträckorna är utrustade med automatiska system för restidsuppföljning.

För prediktionsmodellerna användes två olika metoder: en baserad på neuronnät och en på oskarp (fuzzy) logik. Modellernas godhet jämfördes med noggrannheten för informationen från det existerande systemet. Prognosernas godhet testades statistiskt och ur restidsinformationens synvinkel.

Prediktionsmodellerna för riksväg 4 förutspådde den genomsnittliga restiden för de fordon som körde in på länken under den nästa minut med stöd av den senaste uppmätta trafikinformation. Andelen rätt predikterade restider varierade kraftigt från en länk till en annan.

För Ring I gjordes både estimations- och prediktionsmodeller. Med estimationsmodellerna försökte man estimerastiden för en viss länk, då man inte lyckas mäta restiden för ifrågasvarande länk. Prediktionsmodellerna gav klart bättre resultat än estimationsmodellerna.

Alla statistiska fel som beräknades för resultaten med den oskarpa modellen var större än motsvarande fel för neuronnätmodellen. Neuronnätmodellen fungerar bättre än den oskarpa modellen även ur restidsinformationens synvinkel. Då man jämför de oskarpa modellerna sinsemellan, ser man att modellen enligt Gödels algebra var en aning bättre än den similära modellen. Det nuvarande systemet förutspådde rusningstrafik sämre än de oskarpa modellerna.

För länkarna Heinola–Kymijärvi och Kymijärvi–Heinola gjordes rusningsmodeller genom att bygga upp inlärningsmängden med mätdata enbart från rusningsdagarna på dessa länkar. På båda länkarna uppnåddes bättre resultat med dessa modeller än med modellerna som var inlärd med data från hela dagen. För den bästa modellen var andelen rätt predikterade tider 84 procent av tiden i rusningstrafik.

Slutligen ville man klargöra hur mätsystemets uppbyggnad inverkar på prognosernas kvalitet. Man observerade bland annat att ju längre länken var desto bättre var restidsprognosen och att en extra kamera på länken förbättrade resultaten mer än en extra mätpunkt utanför länken. Mätdata från den av Vägförvaltningens automatiska mätpunkter som fanns före Lahtis såg ut att förbättra resultatet mer än den mätpunkt som fanns i Vierumäki.

De resultat som uppnåddes på riksväg 4 var relativt tillfredsställande. Modellerna för Ring I kan däremot högst anses vara försvarliga. För Ring I räcker utförligheten i uppföljningssystemet för mätning av restiderna helt klart inte till för att prognosticera och antalet kamerapunkter borde vara fler. Prognoserna för bägge studieobjekten skulle säkert förbättras om trafiksituationen utöver restidsinformation skulle kompletteras med mera täckande information om trafikmängder. Trots de observerade bristerna är resultaten så lovande att det skulle löna sig att genomföra ett pilotprojekt, som utnyttjar modellen.

Satu Innamaa, Laura Lanne, Kerkko Vanhanen and Matti Pursula: Pääteiden lyhyen aikavälin matla-aikaennusteet. [Short-Term Prediction of Travel Time on Main Roads] Helsinki 2002. Finnish Road Administration. Finnra Reports 5/2002. 81 p. + app. 16 p. ISSN 1457-9871, ISBN 951-726-863-7, TIEH 3200739.

Keywords: road user information, predictions, travel time

SUMMARY

The purpose of this research work was to study the short-term prediction of travel time and how well travel times can be estimated from point measurements. The study sections were located on highway 4 between Lahti and Heinola and Ring Road I and they are equipped with an automatic travel time monitoring system.

Prediction models were made with two different methods: based on neural networks and on fuzzy logic. The goodness of the models was compared to the accuracy of the information given by the current system. The goodness of the forecasts was measured statistically and from the point of view of the travel time information system.

The models made for highway 4 predicted the travel time for vehicles entering the section within the following minute based on the latest measurements. The proportion that was predicted correctly varied from one link to another.

Both estimation and prediction models were made for the Ring Road I. The purpose of the estimation models was to estimate the travel time of a particular link in situations when the measurement of the travel time itself is not possible. The result was that the prediction models gave clearly better results than the estimation models.

All the statistical error terms of the fuzzy models were inferior to the error terms calculated for the neural models. Neural models were also better when the goodness was measured from the point of view of the travel time information. When the fuzzy models were compared among themselves it could be seen that the model based on Gödel algebra was slightly better than the model based on fuzzy similarity. The current information system estimated travel time in congested conditions worse than the fuzzy models.

Specific congestion models were made for the links Heinola–Kymijärvi and Kymijärvi–Heinola based on data from congested days only. Better results were obtained for both links than with the models trained with the whole data. The best congestion model predicted correct travel times 84 percent of the time in congested conditions.

Finally it was studied how the structure of the measurement system affects the forecasts. It was found out that increasing link length improved the forecast and that an additional camera improved the results more if it was located within the link than if it was outside. The point measurement station before Lahti improved the results more than the station in Vierumäki.

The results obtained from highway 4 were rather satisfying. On the contrary, the models for the Ring Road I can be considered at most tolerable. The travel time monitoring system of the Ring Road I is not detailed enough for the prediction and more camera stations would be required. The forecasts of both study sections would improve if the traffic information were complemented with more covering flow data. Despite of the deficiencies, the results are promising and it would be worthwhile to pilot the model as a production version.

ESIPUHE

Tiehallinto sekä Liikenne- ja viestintäministeriö ovat tilanneet tutkimuksen automaattisista liikenteenohjaus- ja liikenteen tiedotusjärjestelmistä osana ministeriön TETRA ja FITS – Liikennetelematiikan rakenteiden ja palvelujen tutkimus- ja kehittämisohjelmia 1998–2000 ja 2001–2004. Tämä tutkimus on valmistunut osana ko. toimeksiantoa.

Tutkimuksesta vastasi diplomi-insinööri Satu Innamaa Teknillisen korkeakoulun liikennelaboratoriosta ohjaajanaan professori Matti Pursula. Sumeaan logiikkaan perustuvat mallit perustuvat tekniikan ylioppilas Laura Lanteen erikoistyöhön. Tutkimusapulaisena toimi tekniikan ylioppilas Kerkko Vanhanen. Tutkimuksen loppuvaiheessa Satu Innamaa siirtyi VTT Rakennus- ja yhdyskuntatekniikkaan, jossa työ viimeisteltiin.

Tutkimuksen etenemistä valvoi asiantuntijaryhmä, johon kuuluivat tekniikan tohtorit Jussi Sauna-aho ja Matti Roine Liikenne- ja viestintäministeriöstä, diplomi-insinöörit Kari Hiltunen Tiehallinnon tiestötiedot-yksiköstä, Sami Luoma liikenteen palvelut -yksiköstä ja Pekka Rajala Uudenmaan tiepiiristä sekä tekniikan tohtorit Risto Kulmala VTT:ltä ja lisäksi Kosonen TKK:lta.

Neuraalilaskennassa on hyödynnetty CSC - Tieteellinen Laskenta Oy:n myöntämiä resursseja.

Helsinki, helmikuussa 2002

Tiehallinto
Liikenteen palvelut

Sisältö

1	JOHDANTO	9
2	TUTKIMUKSEN TAUSTAA	11
2.1	Tutkimuskohteet	11
2.2	Aineistot	13
2.3	Tutkimusasetelma	15
3	KIRJALLISUUSKATSAUS	17
3.1	Piste- ja matkanopeuden välinen vertailu	17
3.2	Linkkiparametrien estimoiminen anturiajoneuvojen avulla	17
3.3	Matka-ajan määrittäminen induktioilmaisimien avulla	19
3.4	Suora matka-ajan mittaaminen	20
4	MALLIT	22
4.1	MLP-neuroverkkomallit	22
4.2	Sumea logiikka mallintamisvälineenä	23
4.2.1	Sumeat joukot	23
4.2.2	Sumeat järjestelmät ja sumea päättely Gödelin algebran mukaisesti	24
4.2.3	Lukasiewiczin algebra ja sumea similaarisuus	26
5	NYKYINEN TIEDOTUSJÄRJESTELMÄ VT 4:LLÄ	28
6	NEUROMALLIT	30
6.1	Syötteiden valinta	30
6.1.1	Valtatien 4 mallit	30
6.1.2	Kehä I:n mallit	40
6.2	Ennusteet ja estimaatit	42
6.2.1	Valtatie 4	42
6.2.2	Kehä I	51
7	SUMEAT MALLIT	57
7.1	Sumean ennustemallin rakentaminen	57
7.2	Sumea ennustemalli ja sen hyvyys	59
8	RUUHKAPÄIVÄMALLIT	65
9	MITTAUSJÄRJESTELMÄN RAKENNE	70
10	JOHTOPÄÄTÖKSIÄ JA SUOSITUKSIA	72
10.1	Neuroverkkomallit	72
10.2	Sumeat ennustemallit	72
10.3	Suosituksia	73
11	YHTEENVETO	76
	LÄHDELUETTELO	79
	LIITTEET	81

1 JOHDANTO

Kasvavat liikennemäärät ja ruuhkat luovat tarpeen entistä monipuolisemmalle liikenteenohjaukselle ja sitä kautta liikenteen seurannalle. Perinteisesti liikennettä on tarkasteltu pistemittausten avulla. Kun liikenteenohjausjärjestelmät kehittyvät pisteohjauksesta kohti väylä- tai verkko-ohjausta, seurattavan alueen koko kasvaa ja seurannan toteuttaminen pistemittausten avulla on vaikeaa ja tulee kalliiksi. Linkkikohtaisilla suureilla verkko voidaan kattaa helpommin, ja esimerkiksi keskimääräinen linkin matka-aika antaa hyvän käsityksen liikenteen sujuvuudesta tarkasteltavalla tieosuudella.

Dynaamiset liikenteen tiedotus- ja ohjausjärjestelmät eivät voi toimia reaktiivisesti, vaan niiden täytyy osata ennakoida ja sopeuttaa strategiansa ajoissa muuttuviin liikennetilanteisiin. Myös nykytilanteen tuntemiseen tarvitaan ennuste, koska liikennetiedot saadaan järjestelmän käyttöön aina pienellä viiveellä.

Yksilötasolla tieosuuden tai muun liikennejärjestelmän osan liikenteen sujuvuutta arvioidaan sen perusteella, kuinka häiriöttömästi ja odotusten mukaisesti tienkäyttäjä kykenee kulkemaan sen läpi. Tärkein yksittäinen objektiivinen liikenteen sujuvuuden mittari on matka-aika. Odotusten mukaisuudella kuvataan matkan ennustettavuutta. Liikkujien kannalta olennaista näyttää olevan kokonaismatka-ajan ohella se, kuinka paljon joudutaan poikkeamaan oletetusta tai tavoitematka-ajasta. (Luoma 1998.)

Liikenteen sujumisella on tärkeä merkitys teollisuuden ja elinkeinoelämän kuljetuksille, joille sujuvuus ilmenee lähinnä kuljetusten varmuutena eli matka-aikabudjetin pitävyytenä. Matka-ajan minimoinnin sijaan kuljetuksissa on tärkeintä matka-ajan hajonnan minimointi, sillä kuljetusyrityksille on oletettavasti tärkeämpää tietää mahdollisimman luotettavasti kuljetuksen saapumisajankohta määränpäähän kuin pyrkiä matka-ajan minimointiin, minkä jälkeen tavara seisoo esimerkiksi terminaalissa useiden tuntien ajan. (Liikenneministeriö 1997.)

Kertomalla ajantasaisesti liikennetilanteesta yritetään auttaa tienkäyttäjää valitsemaan liikkumisen kannalta parhaimmat ajankohdat ja reitit. Näin tieverkon välityskykyä käytetään tehokkaasti hyväksi. Reitin valintaan vaikuttaa lähinnä ennen matkaa saatu tieto, kun taas matkan aikana saatu tieto vaikuttaa ajotavan muuttamiseen. (Kiljunen & Summala 1996.)

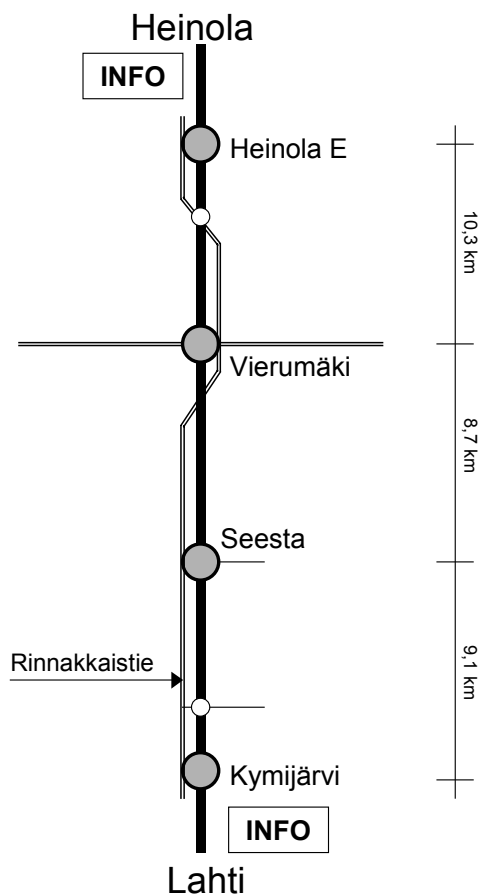
Ennustetun matka-ajan merkitys on liikenneoperaattorin näkökulmasta katsottuna erittäin merkittävä, sillä juuri se vaikuttaa päätöksen, mitä tietoa tienkäyttäjille tarjotaan vaihtoehtoisista, paremmista ajoreiteistä. Tiedotuksella pyritään vaikuttamaan tienkäyttäjien odotuksiin matka-ajan suhteen ja siten tavoitematka-aikoihin. Täten pyritään parantamaan koettua sujuvuutta. (Luoma 1998.)

Tämän tutkimuksen tarkoituksena on selvittää lyhyen aikavälin matka-aikaennusteiden tekemistä. Selvityksen luvussa 3 on luotu lyhyt katsaus kirjallisuuteen. Aiheesta on enemmän viitteessä Innamaa (1999). Luvussa 2 on esitelty tehdyn tutkimuksen tutkimuskohteet, aineistot ja tutkimusasetelma. Käytetyt mallit on esitelty luvussa 3 ja nykyinen tiedotusjärjestelmä luvussa 5. Luvussa 6 käsitellään neuroverkkoihin perustuvia matka-ajan ennuste- ja estimaattimalleja ja luvussa 7 sumeaa logiikkaan perustuvia ennustemalleja. Luvussa 8 on tehty neuroverkkoihin perustuvia malleja pelkkien ruuhkapäivien liikennetiedoilla ja luvussa 9 on selvitetty mittausjärjestelmän rakenteen vaikutusta ennusteen laatuun. Luvussa 10 on tehty johtopäätöksiä ja annettu suosituksia saatujen tulosten pohjalta. Yhteenveto koko raportista on esitetty luvussa 11.

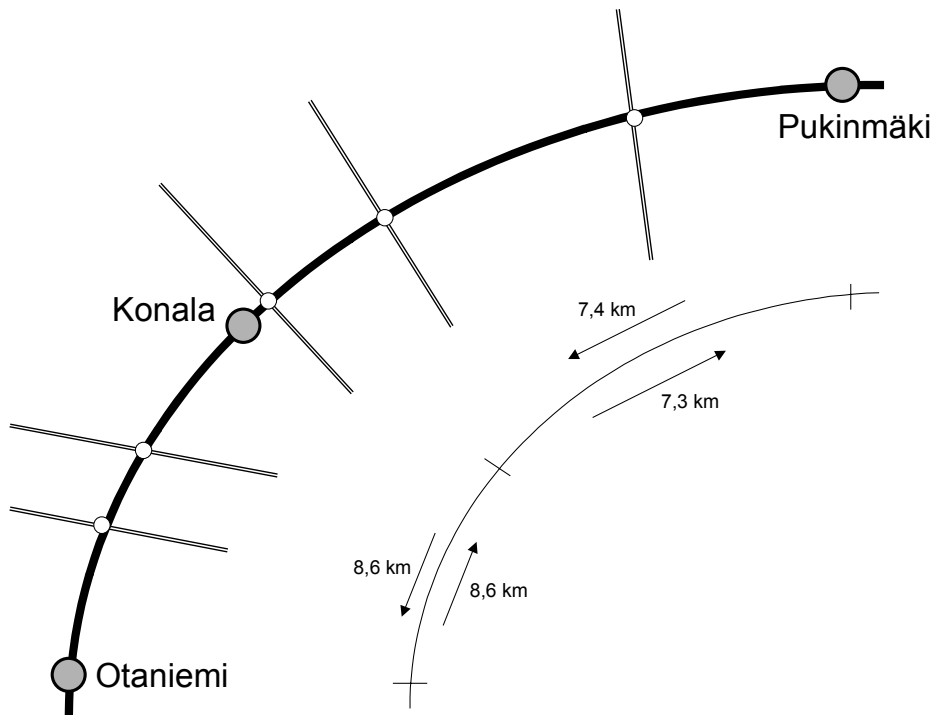
2 TUTKIMUKSEN TAUSTAA

2.1 Tutkimuskohteet

Tutkimuskohteina olivat valtatie 4 Lahden ja Heinolan välillä (kuva **Virhe. Viitteen lähde ei löytynyt.**) sekä Kehä I Otaniemen ja Pukinmäen välillä (kuva **Virhe. Viitteen lähde ei löytynyt.**). Molemmat koejaksot on varustettu automaattisilla matkanopeuden seurantarjestelmillä, joiden avulla matka-ajat mitataan rekisteritunnusten perusteella. Valtatien 4 kohteessa on molemmilla suunnilla muuttuvat opastetaulut, jotka kertovat odotettavissa olevan matka-ajan. Opaste sijaitsee molemmilla suunnilla ennen ensimmäistä matka-ajan mittauskameraa, noin 36 kilometriä ennen kohteena olevaa kaupunkia. Eteläpäässä opaste sijaitsee tarkemmin sanottuna noin neljä kilometriä ennen Joutjärven eritasoliittymää ja pohjoispäässä heti Heinolan tähden vesistö sillan jälkeen.



Kuva 1. Tutkimusalue ja mittauspisteet sekä niiden väliset etäisyydet valtiellä 4. Muuttuvat opastetaulut kertovat matka-ajasta tieosan kummassakin päässä. Kamerapisteet ovat Heinola E:ssä, Vierumäellä, Seestassa ja Kymijärvellä. Lisäksi Vierumäellä on LAM-piste.



Kuva 2. Tutkimusalue ja mittauspisteet sekä niiden väliset etäisyydet Kehä I:llä. Konalassa ja Pukinmäessä on sekä kamera- että LAM-piste ja Otaniemessä pelkkä kamerapiste.

Matkanopeuden seurantajärjestelmän toiminta perustuu ajoneuvojen tunnistamiseen liikennevirrasta ilman ajoneuvoihin asennettavia erillisiä tunnisteita tai lähettämiä. Ajoneuvo pyritään tunnistamaan yksilöllisesti rekisterikilven perusteella aina, kun se ohittaa järjestelmän mittauspisteen. Pisteen ohi kulkeneiden ajoneuvojen rekisterikilvet ja ohitusajankohdat tallennetaan ja peräkkäisten pisteiden tietoja verrataan keskenään. Yhdistämällä saman ajoneuvon rekisteritunnustiedot kahdessa pisteessä saadaan laskettua ajoneuvon matkaan käyttämä aika, eli ajoneuvon matka-aika. Järjestelmä perustuu Golden River Traffic Ltd:n toimittamiin laitteistoihin. (Eloranta 1999.)

Valtatiellä 4 Kymijärven ja Heinolan mittauspisteet sijaitsevat moottoritieellä, vajaa kilometri ennen ohituskaistatien alkua / ohituskaistatien loppumisen jälkeen. Näissä molemmissa pisteissä kaistoja on siis kaksi molempiin suuntiin. Seesta ja Vierumäki sijaitsevat ohituskaistatieellä, jossa on peruskaistojen lisäksi yksi ohituskaista. Tämä ohituskaista on molempien viimeksi mainittujen mittauspisteiden kohdalla etelään eli Lahden suuntaan. Matka-aikaseuranta kattaa yhden kaistan suuntaa kohti kussakin mittauspisteessä. Seestassa ja Vierumäellä seurataan siis kaikkia Heinolan suunnan kaistoja, mutta muuten seuranta kohdistuu ainoastaan peruskaistaan eikä ohituskaistaa seurata. Kehä I:llä liikennettä seurataan kaikilla kaistoilla.

Tutkimuksen tarkoituksena oli selvittää matka-aikaennusteiden tekemistä sekä matka-ajan arvioimista tilanteissa, joissa tietyltä linkiltä ei saada matka-aikamittausten tuloksia. Ennusteet ovat tärkeä osa liikenteen tiedotusjärjestelmiä, ja tutkimuksen jälkimmäistä osaa voidaan hyödyntää esimerkiksi tilanteissa, joissa yksi kamerapiste on poissa toiminnasta.

Matka-aikaennusteiden tutkimisessa pääpaino oli valtatie 4 malleissa, koska tiellä on Kehä I:stä enemmän koko koealueen päästä päähän ajavaa liikennettä ja alue on jaettu useampaan osalinkkiin kuin Kehä I:llä. Ennustemallit tehtiin kuitenkin myös Kehä I:lle. Matka-ajan estimoimista pistemittauksen avulla tutkittiin ainoastaan Kehä I:n aineiston avulla, koska Kehä I:llä on kahden peräkkäisen kamera-aseman kohdalla LAM-ilmaisoin toisin kuin valtatiellä 4, jossa LAM-asemia on ainoastaan yksi.

2.2 Aineistot

Valtatiellä 4 matka-aika-aineiston keruu aloitettiin 22.6.2000. Vierumäestä alkaville tai sinne päättyville linkeille keruu lopetettiin kamerapisteen rikkoutumisen takia 7.–15.7.2000. Muilla linkeillä keruu päättyi kesän 2000 aineiston osalta 15.8. Aineistoa kerättiin lisää 20.6.–17.9.2001. Ilmaisimissa oli aineiston keruujaksojen aikana lyhyitä toimintahäiriöitä, joiden takia yksittäiset ilmaisimet olivat satunnaisesti poissa päältä muutamasta tunnista pariin päivään kestäviä jaksoja. Ilmaisimissa tulkittiin olevan toimintahäiriö, jos kahden peräkkäisen havainnon väli oli yli puoli tuntia. Tästä syystä aikajaksot, joissa jollain linkillä peräkkäisten havaintojen väli oli yli puoli tuntia, jätettiin tutkimuksen ulkopuolelle.

Tarkasteltavalla tieosalla valtatiellä 4 kesäajan keskimääräinen vuorokausiliikennemäärä oli noin 15 000 ajon./vrk (Tielaitos 1999). Mitattujen matka-aikojen keskimääräinen osuus kaikista mitattavissa olevista matka-ajoista on esitetty linkeittäin taulukossa 1 vuoden 2000 aineiston osalta. Havainnointiajasta vähennettiin yli puolen tunnin mittaiset katkokset. Taulukossa on oletettu suuntajakauman olevan tasainen (50/50 prosenttia).

Taulukko 1. Kesällä 2000 kerättyjen havaintojen osuus kesän keskivuorokausiliikennemäärästä (KKVL) valtatie 4 aineistossa. Vuorokausihavaintomäärää laskettaessa havainnointiajasta vähennettiin yli puolen tunnin mittaiset katkot.

Linkki	Havain-		Katkot (vrk)	Havain- toja/vrk	KKVL	Osuus	
	toja	Havainnointi (vrk)					
Etelään	Heinola–Kymijärvi	39 469	53,7	8,3	870	7 500	12 %
	Heinola–Seesta	22 188	53,7	11,2	523	7 500	7 %
	Heinola–Vierumäki	5 995	14,9	7,8	856	7 500	11 %
	Vierumäki–Kymijärvi	8 070	14,9	7,9	1 165	7 500	16 %
	Vierumäki–Seesta	4 878	14,8	8,9	817	7 500	11 %
	Seesta–Kymijärvi	32 439	53,6	11,6	771	7 500	10 %
Pohjoiseen	Kymijärvi–Heinola	29 601	53,6	3,1	585	7 500	8 %
	Kymijärvi–Vierumäki	15 413	22,7	4,7	855	7 500	11 %
	Kymijärvi–Seesta	35 419	53,7	4,9	726	7 500	10 %
	Seesta–Heinola	39 061	53,7	8,6	868	7 500	12 %
	Seesta–Vierumäki	17 975	22,7	9,9	1 405	7 500	19 %
	Vierumäki–Heinola	22 816	22,7	8,8	1 641	7 500	22 %

Valtatiellä 4 linkin molemmissa päissä havaittujen ajoneuvojen osuus kaikista linkin ajoneuvoista vaihtelee kesän 2000 aineistossa 7 ja 22 prosenttiin. Alhaista prosenttia selittää osittain se, että suuressa osassa mittauspisteitä liikennettä seurataan yhdellä kaistalla suuntaa kohti (ohituskaistaa ei seurata). Kysymykseksi jää, ovatko matka-aikahavainnot satunnaisotos kaikista matka-ajoista vai ovatko tietyn tyyppiset matka-ajat yliedustettuina. Koska ohituskaistaa ajavien matka-aikaa ei saada mitattua, on luultavaa, että ly

himmät matka-ajat ovat aineistossa aliedustettuina. Jos ainoastaan joka kymmenes matka-aika tulee mitatuksi on esimerkiksi vaikea sanoa, kasvaako matka-aikojen hajonta liikennetilanteen muuttuessa.

Aineistossa on mukana myös sellaisia ajoneuvoja, jotka ovat poikenneet reitiltä (esimerkiksi huoltoasemalle) ja palanneet takaisin. Näiden ajoneuvojen matka-ajat eivät ole kiinnostavia eikä niitä myöskään pitäisi käyttää matka-aikaennusteen tekemiseen. Tästä syystä aineistosta karsittiin pois ne matka-aikahavainnot, jotka eroavat seitsemän havainnon liukuvasta keskiarvosta (tarkasteltava havainto on havainnoista keskimäinen) yli 33 prosenttia.

Karsintaa ei kuitenkaan tehty automaattisesti, vaan ehdon täyttävät havainnot käytiin vielä käsin läpi. Tällä haluttiin varmistaa esimerkiksi, ettei yksittäinen erittäin pitkä matka-aika pääse kasvattamaan liukuvaa keskiarvoa ja johda näin havaintojen turhaan karsintaan tai ettei havaintoja turhaan karsiudu kohdasta, jossa matka-aika muuttuu voimakkaasti.

Aineistosta haluttiin myös karsia rekisterilaattojen väärintulkinnoista tai eri pisteiden havaintojen väärinyhdistelystä aiheutuneet väärät havainnot (käytännössä: liian lyhyet matka-ajat). Taulukossa 2 on esitetty valtatiellä 4 mittattujen matka-aikahavaintojen kokonaismäärät ja edellä kuvatun karsinnan jälkeen jäljelle jääneiden havaintojen osuudet linkeittäin vuoden 2000 aineistossa.

Taulukko 2. Alkuperäiset havaintojen ja karsinnan jälkeen jäljelle jääneiden havaintojen lukumäärät valtatie 4 vuoden 2000 aineistossa.

Linkki	Alkuperäisten havaintojen lkm	Havaintojen lkm karsinnan jälkeen	Osuus	
Etelään	Heinola–Kymijärvi	39 469	38 621	98 %
	Heinola–Seesta	22 188	21 708	98 %
	Heinola–Vierumäki	5 995	5 990	100 %
	Vierumäki–Kymijärvi	8 070	7 859	97 %
	Vierumäki–Seesta	4 878	4 770	98 %
	Seesta–Kymijärvi	32 439	32 012	99 %
Pohjoiseen	Kymijärvi–Heinola	29 601	28 248	95 %
	Kymijärvi–Vierumäki	15 413	14 634	95 %
	Kymijärvi–Seesta	35 419	34 877	98 %
	Seesta–Heinola	39 061	37 614	96 %
	Seesta–Vierumäki	17 975	17 304	96 %
	Vierumäki–Heinola	22 816	22 795	100 %

Kehä I:n matka-aika-aineisto on ajalta 29.4.–25.9.1999. Kehä I:llä linkkien päästä päähän ajavien ajoneuvojen osuus mittauspisteen liikennemäärästä on todennäköisesti pienempi kuin valtatiellä 4. Tästä syystä matka-aikahavaintojen lukumäärää ei kannata verrata pisteen liikennemäärään. Eloranta (1999) selvitti tunnistusastetta Konalan pisteestä pienellä otoksella. Hän sai tulokseksi, että länteen päin menevällä linkillä rekisterikilpien tunnistusaste oli noin 70 prosenttia liikennemäärästä. Itään päin menevällä linkillä vastaava tunnistusaste putosi ruuhka-aikaan alle 40 prosenttiin. Tässä tutkimuksessa käytetyssä aineistossa itään päin menevällä suunnalla (Otaniemi–Konala–Pukimäki) havaintoja oli enemmän kuin päinvastaisella suunnalla (taulukko 3).

Taulukko 3. Havaintojen määrät ja havainnointiaika Kehä I:n aineistossa. Vuorokausihavaintomäärää laskettaessa havainnointiajasta vähennettiin yli puolen tunnin mittaiset katkot.

	Linkki	Havainnointiaika (vrk)	Katkot (vrk)	Havaintoja/vrk	
Itään	Otaniemi–Konala	212 612	149,68	25,47	1 712
	Konala–Pukinmäki	197 461	149,74	32,07	1 678
Länteen	Pukinmäki–Konala	86 607	149,49	59,77	965
	Konala–Otaniemi	144 456	149,74	43,80	1 364

Kehä I:n aineistolle tehtiin vastaavanlainen karsinta kuin valtatie 4 aineistolle. Karsinnassa kaikki havainnot, jotka poikkesivat liukuvasta keskiarvosta yli 33 prosenttia tulkittiin aineistoon kuulumattomiksi havainnoiksi, jolleivät ne kuuluneet voimakkaan muutoksen rintamaan. Samoin systemaattisesti selvästi väärin (liian lyhyiden) havaintojen jaksot karsittiin pois. Näihin jaksoihin on todennäköisesti ollut syynä kellojen huono kalibrointi.

Kun aineistosta karsittiin pois ne matka-aikahavainnot, jotka erosivat seitsemän havainnon liukuvasta keskiarvosta yli 33 prosenttia tai olivat muuten selvästi virheellisiä, jäljelle jäi 88–95 prosenttia alkuperäisistä havainnoista. Näitä muuten selvästi virheellisiä havaintoja olivat muun muassa sellaiset havainnot, jotka oli kerätty aikana, jolloin toisen kameran kello oli ollut väärässä, mikä johti epärealistisen pieniin matka-aikoihin.

Taulukko 4. Alkuperäiset ja karsinnan jälkeen jäljelle jääneiden havaintojen lukumäärät Kehä I:n aineistossa.

	Linkki	Alkuperäisten havaintojen lkm	Havaintojen lkm karsinnan jälkeen	Osuus
Itään	Otaniemi–Konala	212 612	192 489	91 %
	Konala–Pukinmäki	197 461	173 976	88 %
Länteen	Pukinmäki–Konala	86 607	81 874	95 %
	Konala–Otaniemi	144 456	133 058	92 %

2.3 Tutkimusasetelma

Malleja tehtiin kahdella eri menetelmällä: neuroverkkoihin ja sumeaan logiikkaan perustuen. Valtatie 4 mallien hyvyttä verrattiin olemassa olevan järjestelmän antaman tiedon tarkkuuteen.

Neuroverkkoihin perustuvien matka-ajan ennustemallien osalta tutkimuksen tarkoituksena oli selvittää, mitkä muuttujat korreloivat parhaiten koko linkin matka-ajan kanssa ja minkälaiset syötteet johtavat parhaaseen tulokseen. Tutkimuksen tarkoituksena oli myös selvittää, mitkä tekijät korreloivat matka-ajan kanssa parhaiten silloin, kun matka-aika estimoidaan neuroverkolla muiden linkkien matka-aikojen ja pistekohtaisten mittausten avulla.

Sumeiden mallien osalta oli tarkoitus selvittää syy-seuraussuhteita matka-ajan kehityksessä ja tutustua sumean logiikan antamiin mahdollisuuksiin lyhyen aikavälin ennusteiden tekemisessä.

Lopuksi oli tarkoitus selvittää, riittäisikö ennusteiden tarkkuus mallien käyttöönottoon. Tiehallinto on asettanut tavoitteeksi, että pääteille saataisiin liittymäväleittäin matka-aikaennusteet, joiden tarkkuus olisi ± 10 prosenttia ja jotka kykenevät ennustamaan ruuhkan korkeintaan kymmenen minuutin viipeellä (Tiehallinto 2001). Tässä sovittiin, että malli on riittävän hyvä, jos 90 prosenttia ennusteista jää kymmenen prosentin virhemarginaalin sisään.

Valtatielle 4 tehtiin matka-ajan ennustemalleja. Malleja tehtiin joka osalinkille ja osalinkiyhdistelmälle neuroverkkojen avulla. Lisäksi Kymijärvi–Heinola-linkille tehtiin sumeaan logiikkaan perustuvia malleja. Kehä I:lle tehtiin vastaavat matka-ajan neuroverkkoihin perustuvat ennustemallit kuin valtatielle 4.

Kehä I:n aineistojen avulla selvitettiin sitä, kuinka hyvin matka-ajan voi estimoida muiden linkkien matka-aikatietojen ja LAM-tietojen avulla. Estimointimallissa siis tarkasteltiin tilannetta, jossa yksittäiseltä linkiltä ei jostain syystä saada matka-aikatietoa. Kehä I soveltuu estimointimallien kokeilemiseen paremmin kuin valtatie 4, koska Kehä I:llä on kahden kamerapisteen kohdalla LAM-asema. Valtatien 4 tarkasteltavalla tieosuudella LAM-pisteitä on vain yksi.

3 KIRJALLISUUSKATSAUS

3.1 Piste- ja matkanopeuden välinen vertailu(Haugen 1996)

Piste- ja matkanopeuden välisiä eroja on selvitetty muun muassa Norjassa. Nopeuksien välisten erojen analyysi osoittaa, että matkanopeus on vakaampi suure kuin pistenopeus ja sopii paremmin liikenteenohjauksen tarpeisiin. Sellaisilla aikajaksoilla, jolloin liikenteessä tapahtui häiriöitä, matkanopeuksissa oli vain vähän vaihtelua ajoneuvojen välillä. Tästä syystä matkanopeus on hyödyllinen liikennevirran laadun mittari.

Pistenopeuksien ja linkin matkanopeuksien vaihtelua vertailtiin koelinkillä. Viiden minuutin keskimääräisen pistenopeuden lisäksi vertailussa käytettiin myös pistenopeuden 30 minuutin liukuvaa keskiarvoa. Pistenopeuksien mittauspiste sijaitsi noin puolivälissä matkanopeuden määrittäjäjaksoa.

Tavallisesti ruuhka ja viivytykset syntyivät pistenopeuksien mittauspisteeltä alavirtaan (eli koejakson loppupäässä). Liikenteen häiriöiden ensimmäiset merkit näkyivät koealueella matkanopeuden alentumisena. Matkanopeus laski ja jonkin ajan kuluttua pistenopeudet alkoivat heilahdella, mikä viittasi siihen, että epävakaa liikennevirta oli saavuttanut pistenopeuksien mittauspisteen. Pistenopeuksien liukuvan keskiarvon alapuolella oleva matkanopeus viittaa liikennevirran häiriöihin koealueen loppupäässä. Tämä on myös varoitus siitä, että epävakaa liikennevirta on saavuttamassa pistenopeuksien mittaustaikaa.

Pitkillä linkeillä pistenopeuksien ja matkanopeuksien välinen suhde riippuu mittauspisteiden sijainnista. Jos liikennevirta oletetaan vapaaksi, joissain pisteissä pistenopeus on korkeampi kuin matkanopeus ja joissain alhaisempi.

Vakaan ja epävakaa virran välillä pistenopeus liittyy korkeampaan tai matalampaan matkanopeuteen, kun mennään vakaalta alueelta epävakalle ja toisinpäin. Tämä voidaan selittää sillä, että tietyllä ajanjaksolla ne ajoneuvot, joilta pistenopeudet mitataan, ovat eri ajoneuvoja kuin ne, joita käytetään matkanopeuden määrittämiseen. Lisäksi sillä on oma vaikutuksensa, että liikenteen ongelmat liikkuvat alavirrasta ylävirtaan.

3.2 Linkkiparametrien estimointi anturiajoneuvojen avulla (Westerman 1995)

Westerman (1995) on tutkinut linkkikohtaisten parametrien (liikennemäärä ja -tiheys sekä matkanopeus ja -aika) estimointia anturiajoneuvojen avulla. Hänen tuloksensa osoittavat, että parametrit voidaan estimoida vain tietyllä tarkkuudella. Korkeallakin anturiajoneuvojen osuudella estimaatteihin jää epätarkkuutta. Tämä vahvistaa sen, että liikennemäärä ja -tiheys ovat sellaisia suureita, joiden tarkaksi määrittämiseksi pitää käytännössä mitata kaikki virran ajoneuvot. Anturiajoneuvot eivät siis ole hyvä menetelmä näiden parametrien määrittämiseen, koska otos jää liian pieneksi. Epätarkkuudesta huolimatta anturiajoneuvoja voidaan kuitenkin käyttää sen tiedon arvioimiseen, onko liikennetilanne "hiljainen", "ei niin hiljainen" vai "vilkas".

Liikenteen tiedotusjärjestelmät tarvitsevat ajantasaisen tiedon sekä tilaan etä aikaan liittyvistä liikenteen vaihteluista. Anturiajoneuvokonsepti täyttää liikenteen tiedotusjärjestelmien molemmat vaatimukset. Tielinkin keskinopeus saadaan laskemalla kaikkien linkillä olevien anturiajoneuvojen matkanopeuksien keskiarvo linkin pituudella ja mittausajalla painotettuna. Tielinkin keskinopeutta voidaan siis pitää eräänlaisena aika- ja tilakeskinopeuden yhdistelmänä.

Yksinkertaisin ja suoraviivaisin lähestymistapa tielinkin keskinopeuden estimoimiseen anturiajoneuvojen liikenneraporteista kootun nopeusotoksen perusteella on ottaa nopeuksista keskiarvo. Linkkiparametrien estimointi pelkästään yksinkertaisen keskiarvoestimaatin avulla on kuitenkin arveluttavaa etenkin silloin, kun otoskoko on pieni. Tielinkin keskinopeuden ja -matka-ajan estimoimiseksi voidaan määrittellä liikennevirtamalli, joka kuvaa yksittäisten anturiajoneuvojen nopeuksien ja liikennevirran todellisen keskinopeuden välistä yhteyttä. Malli voi perustua esimerkiksi liikennevirran peruskuvajaan (liikennemäärä-nopeuskuvaaja). Kuvaajasta voidaan erottaa kolme aluetta: vapaan virran alue, epävakaata alue ja ruuhka-alue. Koska epävakaata alue voidaan tulkita siirtymävaiheeksi kahden muun alueen välillä, liikennevirtamalliin on määritelty vain kaksi aluetta, joista molemmat pitävät sisällään osan epävakaasta alueesta. Nopeusjakauman varianssi riippuu nopeuden sijainnista peruskuvajalla. Tielinkin todellinen keskinopeus vaihtelee kyseisen ajankohdan, sää- ja kelityypin pitkäaikavälin nopeuskeskiarvon ympärillä. Pitkän aikavälin keskiarvot määrittellään erikseen vapaalle ja ruuhkaiselle liikenteelle.

Toinen tapa parantaa liikennevirtamallin avulla saatavia estimaatteja on verrata tielinkin keskinopeutta pitkän aikavälin keskinopeuden lisäksi edellisten mittaustulosten mukaisiin nopeuksiin. Tämä tekee jäykät pitkän aikavälin keskiarvot joustavammiksi ja ne sopeutuvat paremmin kulloiseenkin tilanteeseen. Myös erilaisia tasoituskeinoja (kuten eksponentiaalista tasoitusta) voidaan käyttää.

Estimaattien vakautta voidaan parantaa sekä vapaan että ruuhkaisen liikennevirran alueella käyttämällä tasoitusta. On selvää, että vapaan virran alueella otoksen anturiajoneuvojen nopeushavaintojen merkitys kasvaa otokseen kasvaessa ja pitkän aikavälin keskiarvon vastaavasti laskee. Sama pätee myös ruuhkaisella alueella.

Kaikissa olosuhteissa tilastollinen estimaatti, joka ottaa huomioon tietokantaan tallennetut tiedot muuttujan käyttäytymisestä, on vähintään yhtä hyvä kuin estimaatti, joka ei ota näitä huomioon olettaen, että tietokannan tiedot ovat luotettavia. Suoran keskiarvoestimaatin etuna on estimoinnin yksinkertaisuus, kun taas tietokannan tietojen mukaan tuominen mutkistaa asioita. Tietokannan tietojen mukaan ottaminen on kuitenkin vaivansa arvoista, koska tielinkin keskinopeuden estimaatti tarkentuu ja sen luotettavuus paranee. Tietokantaa hyväksi käytävää estimaattia kannattaa kuitenkin käyttää vain silloin, kun sen sisältämien tietojen voidaan olettaa olevan oikeita.

3.3 Matka-ajan määrittäminen induktioilmaisimien avulla

Matka-aika voidaan määrittää induktioilmaisinaistosta vähintään neljällä eri tavalla. Menetelmistä ensimmäinen perustuu kahteen mittauspisteeseen, toinen kumulatiivisiin jakaumiin, kolmas liikennemäärien väliseen korrelaatioon ja neljäs ajoneuvojen ”sormenjalkien” tunnistamiseen.

1) Kahden mittauspisteen liikennevirrat

Vapaassa liikennevirrassa keskimääräinen matka-aika voidaan arvioida karkeasti tiejakson alkupisteen ja loppupisteen ajoneuvojen nopeuksien perusteella sillä oletuksella, että ajoneuvot ajavat tiejakson alussa mitattua nopeutta jakson puoliväliin ja muuttavat siellä nopeutensa tiejakson lopussa mitatuksi nopeudeksi. Menetelmä antaa riittävän tarkan tuloksen myös ruuhkaliikenteessä, kunhan tiejakso on riittävän lyhyt (korkeintaan yksi kilometri) eikä laskentajakso ylitä kahta minuuttia. (Westerman 1995.)

Toinen menetelmä on käyttää liikennemäärän ja nopeuden osamäärää, jolloin saadaan määriteltyä paikallinen liikennetiheys. Liikenteen oletetaan olevan vapaata ja mitattujen parametrien ajatellaan edustavan koko tiejaksoa. Tiejakson alusta ja lopusta määriteltyjen liikennetiheyksien keskiarvo kerrotaan tiejakson pituudella, jolloin saadaan tulokseksi tiejaksolla olevien ajoneuvojen määrä. Saatua ajoneuvojen määrää päivitetään laskemalla liikennemäärien erotus laskentajakson kuluttua, kertomalla erotus laskentajakson pituudella ja lisäämällä tuloon aiemmin saatu ajoneuvojen määrä. Nyt matka-ajaksi saadaan päivitetty ajoneuvomäärä (ajon.) jaettuna jälkimmäisellä liikennemäärällä (ajon./h). (Westerman 1995.)

Molemmat menetelmät sopivat pienen mittakaavan mittauksiin. Jos tiejakson pituus ei ylitä 500 metriä, menetelmät toimivat hyvin. Ensimmäinen menetelmä antaa hieman parempia tuloksia kuin jälkimmäinen, joka on alttiimpi mittausvirheille. Kilometrin pituisella tiejaksolla tulokset ovat vielä tyydyttäviä, mutta tulosten laatu heikkenee nopeasti tiejakson pituuden kasvaessa. Kun tiejakson pituus on 5 - 10 km, molemmat menetelmät ovat huonoja. (Westerman 1995.)

2) Kumulatiiviset jakaumat

Matka-aika saadaan määriteltyä vertailemalla peräkkäisten mittauspisteiden kumulatiivisia ajoneuvojen saapumisista tehtyjä jakaumia toisiinsa ja määrittämällä aika, jolla jakaumat ovat mittauspisteiden välillä siirtyneet. Vertailu voidaan tehdä myös ajoneuvoluokittain, mikä parantaa menetelmän tarkkuutta erityisesti silloin, kun eri ajoneuvoluokat ajavat hyvin erilaisilla nopeuksilla. Menetelmän huono puoli on se, että jos tietty määrä ajoneuvoja jää havaitsematta molemmissa mittauspisteissä tai ne tulkitaan väärin, matka-aika-arvio on virheellinen. (Westerman 1995, Cassidy & Windover 1995.)

3) Liikennemäärien välinen korrelaatio

Menetelmä perustuu oletukseen, että liikenne on jatkuva virtausprosessi ja että tietyssä paikassa mitattu liikennemäärä voidaan mitata uudestaan toisessa paikassa alavirran suuntaan. Liikennemäärä mitataan kahdessa paikassa induktioilmaisimilla ja saatuja signaaleja verrataan keskenään. Kahden täsmäävän liikennemäärän välinen aikaviive on sama kuin keskimääräinen matka-aika. Menetelmä sopii ainoastaan vapaalle liikennevirralle, koska ruuhkassa korrelaatio häviää. (Westerman 1995.)

Liikennemäärien väliseen korrelaatioon perustuvasta menetelmästä on kehitetty parannettu kaksivaiheinen menetelmä. Ensimmäinen vaihe on määrittää edellä kuvatulla tavalla liikennemäärien välinen korrelaatio ja käyttää tätä ruuhkan ilmaisemiseen. Jos korrelaatio löytyy, liikennettä voidaan pitää suhteellisen vapaana, kun taas korreloimattomuus on merkki ruuhkasta. Jos ruuhkaa löytyy, käynnistetään menetelmän toinen vaihe, deterministinen jonomalli, jossa määritellään sen hetkinen ajoneuvomäärä, ”liikaa” olevien ajoneuvojen määrä ja tiejaksolle arvioitu viivytys. Viimeksi mainittu tehdään suhteuttamalla ajoneuvojen ylijäämä tiejaksolta poistuvien ajoneuvojen määrään. (Westerman 1995.)

4) Pulssimuoto

Liikenteestä voidaan määrittää linkkikohtaisia tietoja korrelaation avulla käyttämällä hyväksi ajoneuvojen ”sormenjälkiä” eli niiden synnyttämiä yksilöllisiä pulsseja induktioilmaisimen jännitteen muutoksissa. Yksittäisten ajoneuvojen sormenjäljet ovat hyvin samantapaisia muiden vastaavanlaisten ajoneuvojen sormenjälkien kanssa, mutta erot eri ajoneuvotyyppien välillä ovat sen verran merkittäviä, että ajoneuvoryhmät voidaan tunnistaa linkin lopussa ja linkkikohtaiset suuret määrittää. Tällä tavalla suhteellisen pienellä mittauspistemäärällä saadaan sama tieto liikenteestä, joka perinteisellä pistekohtaiseen aineistoon perustuvalla menetelmällä olisi vaatinut huomattavaa mitauspistetiheyttä. (Pfannerstill 1991.)

Tunnistamalla sama jono linkin alku- ja loppupäässä voidaan laskea jonon keskimääräinen matka-aika linkillä. Liikennetiheys taas voidaan laskea niiden ajoneuvojen lukumäärän avulla, jotka ovat menneet linkille juuri tunnistetun jonon jälkeen. Aineisto voidaan määritellä erikseen eri ajoneuvotyypeille tai kaistoitoille. Tämä johtaa suhteellisen tarkkaan kuvaan liikennetilanteesta. (Pfannerstill 1991.)

3.4 Suora matka-ajan mittaaminen (Haugen 1996)

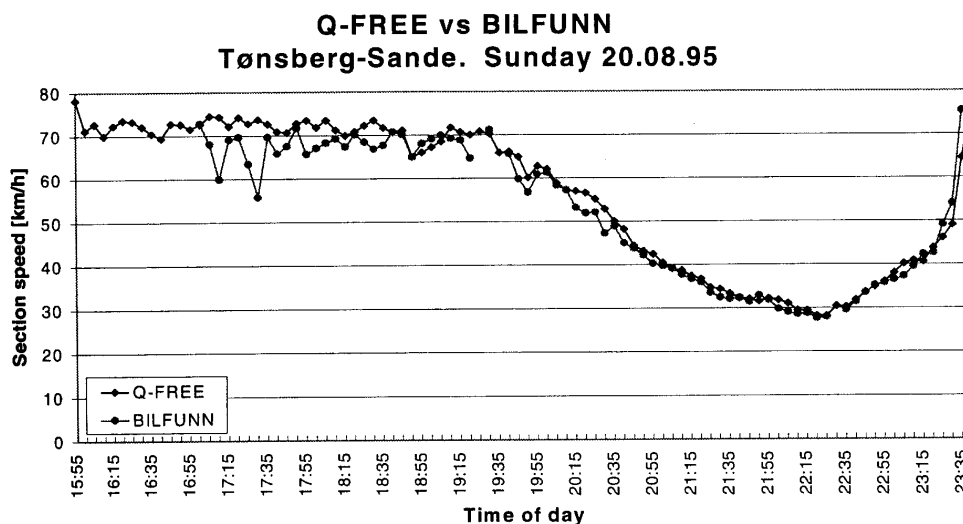
Anturiajoneuvoja ja muutamaa muuta menetelmää lukuun ottamatta suurin osa matka-ajan mittaamisen menetelmistä perustuu ajoneuvon tunnistamiseen kahdessa tai useammassa pisteessä tutkittavalla tiejaksolla. Kussakin mittauspisteessä ajoneuvon erityispiirteet tallennetaan yhdessä kellonajan kanssa. Ajoneuvojen tunnistaminen tehdään pääasiassa jommallakummalla kahdesta päämenetelmästä.

- **Absoluuttinen tunnistaminen** tarkoittaa sitä, että jokainen ajoneuvo tunnistetaan kussakin mittauspisteessä yksiselitteisesti, kuten rekisteritunnuksen tai yksilöidyn elektronisen tunnistimen avulla. Menetelmän etuna on se, että ajoneuvojen tunnistaminen tehdään varmasti, jolloin ei ole vaaraa, että ajoneuvot sekoitetaan keskenään ja näin syntyisi virheitä kerättävään aineistoon. Tämän tyyppinen tunnistaminen saattaa kuitenkin rikkoa yksilösuoja ja tietojen keräämiseen täytyy olla lupa.
- **Suhteellinen tunnistaminen** tarkoittaa sitä, ettei ajoneuvoja tunnisteta yksiselitteisesti. Tämä voidaan toteuttaa usealla eri tavalla, esimerkiksi ajoneuvon fyysisien ominaisuuksien kuten akselipohjan, akseleiden lukumäärän, renkaiden pohjan, akselipainojen tai ajoneuvon kokonaispainon perusteella. Suhteellinen tunnistaminen voidaan tehdä myös absoluuttisen tunnistamisen keinoin rekisteröimällä vain osa rekisteritunnuk

sesta tai elektronisesta tunnisteesta, jolloin tallennetta ei voi yksiselitteisesti yhdistää mihinkään tiettyyn ajoneuvoon. Suhteellinen tunnistaminen ei yksilöi ketään ja siksi siihen ei tarvita lupia tienkäyttäjiltä. Toisaalta ajoneuvon virheellisen tunnistamisen riski on olemassa ja tämän takia algoritmin, joka yhdistelee ajoneuvotiedot ja muuttaa ne matka-aikahavainnoiksi, täytyy olla pidemmälle kehitetty.

Se, onko absoluuttisella tunnistamisella merkittäviä etuja suhteelliseen tunnistamiseen verrattuna, riippuu useista erilaisista seikoista, kuten liikennemäärän suhteesta välityskykyyn, mittauspisteiden välimatkasta, ohitusten määrästä ja niiden ajoneuvojen osuudesta kokonaisliikenteestä, jotka ohittavat molemmat mittausasemat.

Haugen (1996) on vertaillut sähköisen tunnistimen ja fyysisten ominaisuuksien perusteella tapahtuvaa tunnistamista (kuva 3). Fyysisten ominaisuuksien perusteella tapahtuvan tunnistamisen havaittiin antavan hyviä tuloksia. Parhaat tulokset saatiin silloin, kun liikennevirta oli vakaa ja mittauspisteiden välinen etäisyys pieni. Kuvasta 3 voidaan havaita, että erot menetelmien välillä olivat suhteellisen pieniä.



Kuva 3. Sähköisen tunnistimen (Q-free) ja fyysisten ominaisuuksien (Bilfunn) perusteella tapahtuvan tunnistamisen välinen vertailu (Haugen 1996).

4 MALLIT

4.1 MLP-neuroverkkomallit

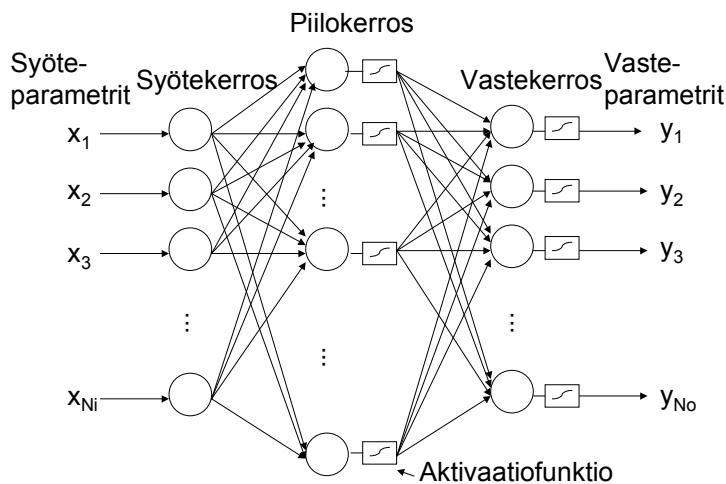
Ennustemalleina käytettiin monikerrosperseptronineuroverkkoja (MLP-verkot, kuva 4) niiden aikaisemmissa tutkimuksissa (mm. Innamaa ym. 2002, Innamaa & Pursula 2000, Lee ym. 1998, Smith & Demetsky 1994 ja 1997) antamien hyvien tulosten perusteella. Syötekerroksen neuronien määrä oli yhtä suuri kuin syöteparametrien määrä ja vastekerroksen neuronien määrä vastaavasti yhtä suuri kuin vasteparametrien määrä. Piilokerrosten määräksi valittiin yksi ja piiloneuronien määrä määriteltiin Widrow'n nyrkkisäännön avulla: opetusjoukon koon (T) pitäisi olla vähintään kymmenen kertaa estimoitavien parametrien (paino- ja harhakertoimet, W) lukumäärä. N_i on syöteuronien lukumäärä, N_h piiloneuronien lukumäärä ja N_o vastaneuronien lukumäärä (Oja 1998).

$$T_{\min} = 10 \cdot W$$

$$W = (N_i + 1)N_h + (N_h + 1)N_o$$

$$\Rightarrow N_{h_{\max}} = \frac{0,10 \cdot T - N_o}{N_i + N_o + 1}$$

Piiloneuronien määräksi valittiin näin laskettu maksimimäärä, mikäli tämä ei ylittänyt 20:tä. Muuten määräksi asetettiin 20, jottei opetusprosessi käynyt liian raskaaksi.



Kuva 4. Monikerrosperseptroniverkko.

Aineisto esikäsiteltiin normeeraamalla se siten, että keskiarvoksi tuli nolla ja keskihajonnaksi yksi. Piilokerroksen aktivaatiofunktioiksi valittiin hyperbolinen tangenti ja vastekerroksen aktivaatiofunktioiksi lineaarinen funktio.

Mallien opettamiseen käytettiin Fletcher-Reevesin päivitystä, joka kuuluu konjugaatti-gradienttimenetelmiin. Perusgradienttimenetelmissä neuroverkon painokertoimia päivitetään siihen suuntaan, jossa hyvyysfunktio (esimerkiksi virheen neliösumma) pienenee nopeimmin. Tämä ei kuitenkaan välttämättä takaa nopeinta konvergenssia. Konjugaatti-gradienttimenetelmässä suun

naksi valitaan konjugaatin suunta, jolloin päästään nopeampaan konvergenssiin kuin suurimman gradientin suunnassa. (Demuth & Beale 1998.)

Jos neuroverkko oppii opetusaineiston liian hyvin, se muistaa sen ulkoa eikä kykene yleistämään. Eräs tapa ehkäistä ulkoa oppimista on varmistaa, että opetusjoukon koko on riittävän suuri (piiloneuronien maksimimäärä perustuu opetusjoukon kokoon). Toinen tapa välttää ylioppiminen on asettaa opetusprosessille lopetuskriteerejä. Nyt tehdyssä tutkimuksessa näitä kriteerejä asetettiin useita. Opettaminen lopetettiin, kun saavutettiin opetuskierrosten maksimimäärä, kun gradientin arvo tai keskimääräinen neliövirhe laskivat riittävän pieniksi tai kun kalibrointiaineiston keskimääräinen neliövirhe lakkasi pienenemästä. Käytännössä opetus lopetettiin aina viimeisen lopetuskriteerin perusteella.

Viimeksi mainittua kalibrointiaineistokriteeriä varten alkuperäinen opetusaineisto jaettiin kolmeen osaan: opetus-, kalibrointi- ja validointiaineistoon. Opetusaineisto koostui niistä havainnoista, joiden perusteella neuroverkko yritti oppia mallinnettavaa ilmiötä. Kalibrointiaineiston perusteella yritettiin havaita, milloin opetus oli edennyt niin pitkälle, ettei neuroverkko enää oppinut yleisiä mallinnettavaan ilmiöön liittyviä ominaisuuksia, vaan alkoi opetella käytetyn opetusjoukon erityispiirteitä. Tämä ilmeni tilanteena, jossa opetusjoukon virhe pieneni, mutta kalibrointijoukon virhe alkoi kasvaa. Validointijoukko oli kolmas aineisto ja sitä ei käytetty opetusprosessissa millään tavalla. Sen tarkoituksena oli testata, kuinka hyvin opetuksessa onnistuttiin eli kuinka hyvin neuroverkko toimi.

Neuroverkkomalleista on kerrottu tarkemmin viitteessä Innamaa & Pursula (2000).

4.2 Sumea logiikka mallintamisvälineenä

4.2.1 Sumeat joukot

Teorian sumeasta joukko-opista perinteisen, kaksiarvoiseen logiikkaan perustuvan joukko-opin yleistykseenä esitteli L. A. Zadeh vuonna 1965 pyrki myksenään mallintaa ihmisen ajattelumeکانismia numeroiden sijasta sumeilla, lingvistikillillä arvoilla. (Yager, Filev 1994.)

Sumea joukko (*fuzzy set*) tarkoittaa joukkoa, jonka alkioit voivat kuulua joukkoon vain osittain. Alkion joukkoon kuulumisen asteen ilmoittaa jäsenyysfunktio (*membership function*). Täsmällinen joukko on sumeassa joukko-opissa erikoistapaus: alkio joko kuuluu täysin joukkoon, jolloin jäsenyysfunktio saa arvon 1, tai alkio ei kuulu joukkoon ollenkaan, jolloin jäsenyysfunktion arvo on 0. Sumean joukon ollessa kyseessä jäsenyysfunktio voi saada mitä tahansa arvoja väliltä $[0, 1]$. Tällöin jäsenyysfunktion arvon ollessa 0 alkio ei kuulu joukkoon ja jäsenyysfunktion arvolla 1 se kuuluu joukkoon kokonaan. Muilla jäsenyysfunktion arvoilla alkio kuuluu joukkoon vain osaksi. Näin ollen joukon ja sen ulkopuolisen universumin raja ei ole täsmällinen. (Pursula 1995, Yager, Filev 1994.)

Matemaattisesti määriteltyinä sumea joukko A on järjestettyjen parien joukko (X, μ_A) , missä $X (\neq \emptyset)$ on referenssijoukko ja $\mu_A: X \rightarrow L$ on jäsenyysaste-funktio. L puolestaan on jäsenyysasteiden joukko, reaali-lukuintervalli $[0, 1]$ varustettuna jollakin sopivalla algebrallisella rakenteella. (Turunen 2001.)

4.2.2 Sumeat järjestelmät ja sumea päättely Gödelin algebran mukaisesti

Sumeiden joukkojen käyttöön perustuva päättely mahdollistaa johtopäätösten tekemisen epätarkan tiedon perusteella. Sumea päättely (*approximate reasoning*) noudattaa perinteisen kaksiarvoisen logiikan sääntöjä huolimatta siitä, että lähtötiedot ja/tai relaatiot ovat sumeita.

Järjestelmää, jonka toimintaan liittyy sumeutta, kutsutaan sumeaksi järjestelmäksi (*fuzzy system*). Sumean järjestelmän kuvaus on joukko lingvistisiä sääntöjä, joista jokaisen voimassaolon aste arvioidaan samanaikaisesti sumean päättelyn avulla. Sen perusteella muodostetaan järjestelmän toiminnan määräävä kokonaisarvio eli johtopäätös, joka on kompromissi kaikista asiaan liittyvistä säännöistä. (Pursula 1995.)

Sumean järjestelmän rakentaminen eroaa perinteisestä järjestelmän mallintamisesta siinä, että matemaattinen malli korvataan sanallisella sääntömuotoisella kuvauksella. Järjestelmän toimintaan liittyvät kausaaliset syyseuraussuhteet muutetaan säännöiksi, joihin liittyvät ehdot ja riippuvaisuudet ovat luonteeltaan sumeita. Sumean päättelyn tuloksena järjestelmä antaa jäsenyysastefunktion, joka täsmällistetään valitulla menetelmällä, jotta päästään yksikäsitteiseen tulokseen. (Pursula 1995.)

Gödelin algebrassa loogista operaatiota \vee (tai) vastaa joukkojen unioni eli jäsenyysfunktioiden maksimiarvo. Tällöin väitteen "x kuuluu joukkoon A tai B" totuusarvo $\mu_{A \vee B}(x)$ voidaan määrittää seuraavasti:

$$\mu_{A \vee B}(x) = \mu_{A \cup B}(x) = \max\{\mu_A(x), \mu_B(x)\} \quad (1)$$

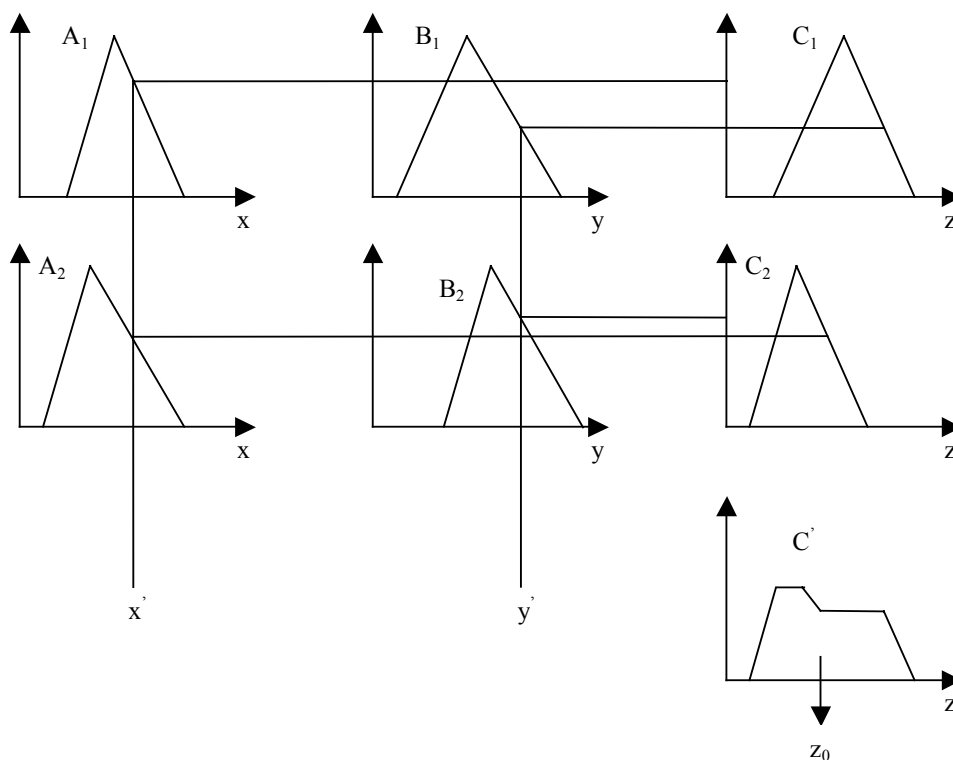
Yhtälössä $\mu_A(x)$ on jäsenyysfunktio, joka ilmoittaa alkion x joukkoon A kuuluminen asteen.

Samoin loogista operaatiota \wedge (ja) vastaa joukkojen leikkaus eli Gödelin algebran mukaisesti jäsenyysfunktioiden minimiarvo. Väitteen "x kuuluu joukkoihin A ja B" totuusarvo $\mu_{A \wedge B}(x)$ määritetään siis seuraavasti:

$$\mu_{A \wedge B}(x) = \mu_{A \cap B}(x) = \min\{\mu_A(x), \mu_B(x)\} \quad (2)$$

(Yager, Filev 1994, Pursula 1995.)

Kuvassa 5 esitetään periaatekaavio Gödelin algebran mukaisen sumean järjestelmän toiminnasta graafisesti.



Kuva 5. Gödelin algebran mukaisen sumean järjestelmän toiminnan graafinen kuvaus täsmällisillä syöttötiedoilla (Pursula 1995).

Kuvan 5 järjestelmä sisältää kaksi päättelysääntöä:

- Jos x on A_1 ja y on B_1 , niin z on C_1
- Jos x on A_2 ja y on B_2 , niin z on C_2

Säännöissä A , B ja C ovat sumeita. Järjestelmän saadessa täsmälliset syöttötiedot x' ja y' , määritetään niiden yhteensopivuus kaikkien päättelysääntöjen sisältämien ehtojen kanssa, eli jäsenyysaste sumeissa joukoissa A ja B . Vain niitä sääntöjä, joissa syöttötietojen jäsenyysaste sekä joukossa A että B on nollaa suurempi (*rules that fire*), käytetään päätöksenteossa (kuvan 5 tapauksessa molempia päättelysääntöjä).

Kaavan (2) perusteella näiden yksittäisten sääntöjen vaikutus lopputulokseen määräytyy syöttötietoja vastaavien pienimpien jäsenyysasteiden mukaan kuvan 5 mukaisesti. Yhdistämällä ne kaavan (1) mukaisesti valitsemalla kutakin z :n arvoa vastaava jäsenyysasteen maksimi saadaan johtopäätös, sumea joukko C' , joka kuvaa sääntöjen yhteistä toteutumista.

Yksikäsitteiseen lopputulokseen z_0 päästään esimerkiksi laskemalla jäsenyysastefunktion rajoittaman pinta-alan painopiste (*Center of Gravity, COG*) (Klir ym. 1997). Täsmällistäminen (*defuzzification*) voidaan tehdä myös esimerkiksi määrittämällä maksimien keskiarvo (*Mean of Maxima, MOM*) tai valitsemalla z_0 :ksi pienin jäsenyysastefunktion maksimia vastaava z eli ns. ensimmäinen maksimi riippuen siitä, millä menetelmällä katsotaan saatavan parhaiten sumeaa joukkoa C' kuvaava luku (Yager, Filev 1994).

4.2.3 Lukasiewiczin algebra ja sumea similaarisuus

Ekvivalenssirelaatio on yleinen tapa luokitella keskenään samanlaisia matemaattisia objekteja. Sumea similaarisuus on klassisen ekvivalenssirelaation moniarvologinen yleistys, joka kuvaa objektien samanlaisuuden astetta. Moniarvologiikassa ekvivalenssirelaation algebrallinen vastine on bi-residuaali \leftrightarrow , joka määritellään

$$\forall x, y \in [0,1]: x \leftrightarrow y = \min\{(x \rightarrow y), (y \rightarrow x)\} \quad (3)$$

Bi-residuaalilla on seuraavat ominaisuudet:

$$x \leftrightarrow 1 = x$$

$$x = y \text{ jos ja vain jos } x \leftrightarrow y = 1$$

$$x \leftrightarrow y = y \leftrightarrow x$$

$$(x \leftrightarrow y) \otimes (y \leftrightarrow z) \leq x \leftrightarrow z$$

Jokainen sumea joukko $A = (X, \mu_A)$ generoi sumean similaarisuuden S_A , joka määritellään seuraavasti:

$$\forall x, y \in X : S_A \langle x, y \rangle = \mu_A(x) \leftrightarrow \mu_A(y) \quad (4)$$

Bi-residuaalin ominaisuuksien nojalla S_A toteuttaa seuraavan sumean similaarisuuden määritelmän: *L-arvoinen X:ssä määritelty binäärinen relaatio S on sumea similaarisuus, jos se toteuttaa seuraavat ehdot:*

$$\forall x \in X : S \langle x, x \rangle = 1 \quad (\text{refleksiivisyys})$$

$$\forall x, y \in X : S \langle x, y \rangle = S \langle y, x \rangle \quad (\text{symmetrisyys})$$

$$\forall x, y, z \in X : S \langle x, y \rangle \otimes S \langle y, z \rangle \leq S \langle x, z \rangle \quad (\text{heikko transitiivisuus})$$

Sumea similaarisuus voidaan jännöksettömästi redusoida Lukasiewicz-Pavelka-logiikkaan. Lukasiewicz-algebrassa

$$x \rightarrow y = 1, \text{ jos } x \leq y \quad (5)$$

$$x \rightarrow y = 1 - x + y, \text{ jos } x > y$$

Kaavojen (3), (4) ja (5) nojalla saadaan Lukasiewicz-algebran mukainen laskentakaava sumealle similaarisuudelle

$$\forall x, y \in A : S_A \langle x, y \rangle = \mu_A(x) \leftrightarrow \mu_A(y) = 1 - |x - y|$$

Kokonaissimilaarisuus on osasimilaarisuuksien painotettu keskiarvo

$$S(x, y) = \frac{1}{n \sum_{i=1}^n S_i(x, y)}$$

Kokonaissimilaarisuuteen perustuvassa sumeassa päättelyalgoritmista pyrkimyksenä on etsiä syötemuuttujien kanssa samanlaisin väite, päättelysäännön JOS-osa, ja toteuttaa kyseisen säännön NIIN-osa. Jokaista sumeaa päättelysääntöä (JOS-osaa) siis verrataan erikseen syötemuuttujiin ja lasketaan vastaavat kokonaissimilaarisuudet. Sääntö, jota vastaava kokonaissimilaarisuus on suurin, valitaan perustaksi päätökselle. Säännön NIIN-osa toteutetaan katkaisemalla sen jäsenyysastefunktion kuvaaja saatua kokonaissimilaarisuutta vastaavan jäsenyysasteen kohdalta ja suorittamalla täsmällistä menetelmällä. (Haukilehto 2001, Turunen 2001.)

5 NYKYINEN TIEDOTUSJÄRJESTELMÄ VT 4:LLÄ

Tällä hetkellä tiedotusjärjestelmä kertoo tienkäyttäjille arvion Lahden ja Heinolan välisestä matka-ajasta noin 36 kilometrin matkalta ennen kyseistä kaupunkia. Matka-aikaestimaatti perustuu viimeisimpiin mittauksiin. Mittaus-tiedot välitetään mittauskojeilta liikennekeskukseen 20 minuutin välein, kesä-ruuhkissa tietoja välitettiin 5–10 minuutin välein (Hämeen tiepiiri 2000). Mit-tausten jälkeen järjestelmän Internet-yhteyksiä on parannettu ja päivitykset on mahdollista tehdä jatkuvasti viiden minuutin välein. Tätä parannusta ei kuitenkaan ole otettu huomioon, kun puhutaan nykyjärjestelmän antaman matka-aikatiedon hyvydestä.

Matka-aikaestimaatti lasketaan eri osalinkkien matka-aikojen summana. Osalinkin matka-aika määritetään viimeisten 10 minuutin matka-aikamediaa-nien perusteella, mikäli jaksolta on vähintään yksi havainto. Tieto saa kui-tenkin olla korkeintaan 60 minuuttia vanhaa. Jos jokin kamerapisteistä ei ole toiminnassa, kyseinen kamerapiste (tai -pisteet) voidaan ohittaa pidempien linkkien matka-ajoilla olettaen, että päätepisteiden kamerat toimivat. (Hä-meen tiepiiri 2000.)

Matka-aikaestimaatti (JTe) lasketaan seuraavan yhtälön avulla.

$$JTe = Vakio(ajoaika) + MA1 + MA2 + MA3 + Vakio(käyttäjä)$$

Yhtälössä vakio(ajoaika) on ajoaika seuranta-alueen ulkopuolella (vakio 191 sekuntia), MA1–MA3 ovat linkkien 1–3 mediaanimatka-ajat ja vakio(käyttäjä) on käsin syötettävä vakio (liikennekeskuksen päivystäjä voi lisätä matka-aikaan ruuhkan alkaessa 3–5 minuuttia). (Hämeen tiepiiri 2000.)

Näytöllä ei kuitenkaan esitetä yksittäistä matka-aikaestimaattia, vaan ajat, joiden välillä matka-aika kullakin hetkellä vaihtelee. Nämä matka-ajan ylä- ja alarajat määritellään yllä olevan yhtälön mukaisesti siten, että mediaanit on korvattu 25/75-prosenttipisteillä. Rajoille on kuitenkin asetettu ehdoksi, että niiden erotuksen täytyy olla vähintään viisi minuuttia ja että pienin alaraja on kesänopeuksilla 21 minuuttia ja talvinopeuksilla 22 minuuttia. Minimiarvot vastaavat tieosan nopeusrajoitusten mukaista ajoaikaa. Miniminäytön yläaraja perustuu raskaiden ajoneuvojen maksiminopeuteen. (Hämeen tiepiiri 2000.)

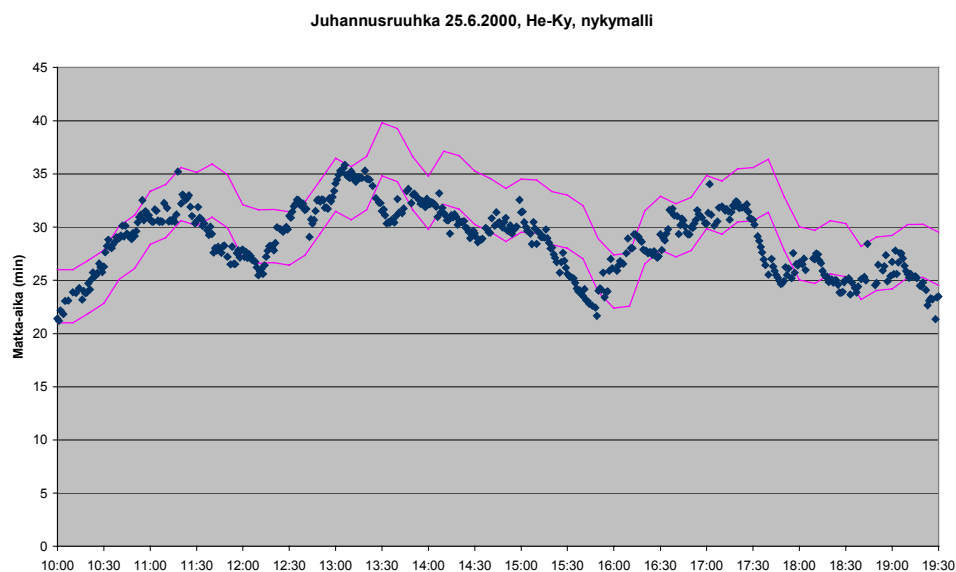
Oletetaan, ettei liikennekeskuksen päivystäjä korjaa matka-aikaestimaattia manuaalisesti. Oletetaan lisäksi, että estimaatti on oikeassa, jos mitattu mat-ka-aika mahtuu kymmenen prosentin virhemarginaalin sisään tai jos esti-maatti näyttää minimiarvoja ja mitattu matka-aika on lyhyempi kuin nopeus-rajoitusten sallima.

Heinola–Kymijärvi-suunnalla matka-aikatieto oli oikeassa 97,3 prosenttia ajasta, näytti liian suurina matka-aikoja 0,9 prosenttia ja liian pieniä matka-aikoja 1,8 prosenttia ajasta, kun tarkastellaan koko kesän 2000 liikennettä. Jos tarkastellaan ainoastaan ruuhkaliikennettä, tieto oli oikeassa ainoastaan 49,7 prosenttia, näytti liian suurina matka-aikoja 22,3 prosenttia ja liian pieniä matka-aikoja 18,8 prosenttia ajasta.

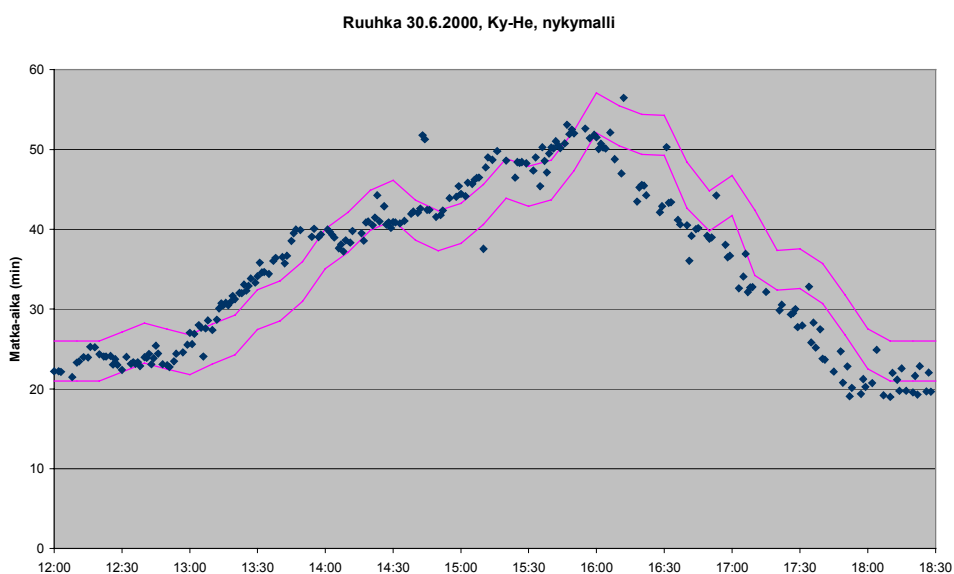
Päinvastaisella suunnalla Kymijärvi–Heinola matka-aikatieto oli oikeassa 95,4 prosenttia, näytti liian suurina matka-aikoja 2,4 prosenttia ja liian pieniä matka-aikoja 2,2 prosenttia ajasta, kun tarkastellaan koko kesän 2000 lii-kennettä. Jos tarkastellaan ainoastaan ruuhkaliikennettä, tieto oli oikeassa

32,9 prosenttia ajasta, näytti liian suuria matka-aikoja 49,3 prosenttia ajasta ja liian pieniä matka-aikoja 29,4 prosenttia ajasta.

Kuvissa 6 ja 7 on tehty esimerkit nykymallin mukaan määritetyistä matka-aikatiedotuksesta ja siitä, kuinka hyvin mitatut matka-ajat ovat sitä vastanneet. Molemmista kuvista voidaan helposti havaita, kuinka matka-aikatiedotus laa-
haa perässä todellisuudesta.



Kuva 6. Nykyisen tiedotusjärjestelmän mukaan ilmoitettu matka-aika (ylä- ja alaraja) sekä mitatut arvot (pisteet) juhannusruuhkassa 25.6.2000 Heinola–Kymijärvi-linkillä.



Kuva 7. Nykyisen tiedotusjärjestelmän mukaan ilmoitettu matka-aika (ylä- ja alaraja) sekä mitatut arvot (pisteet) perjantairuuhkassa 30.6.2000 Kymijärvi–Heinola-linkillä.

6 NEUROMALLIT

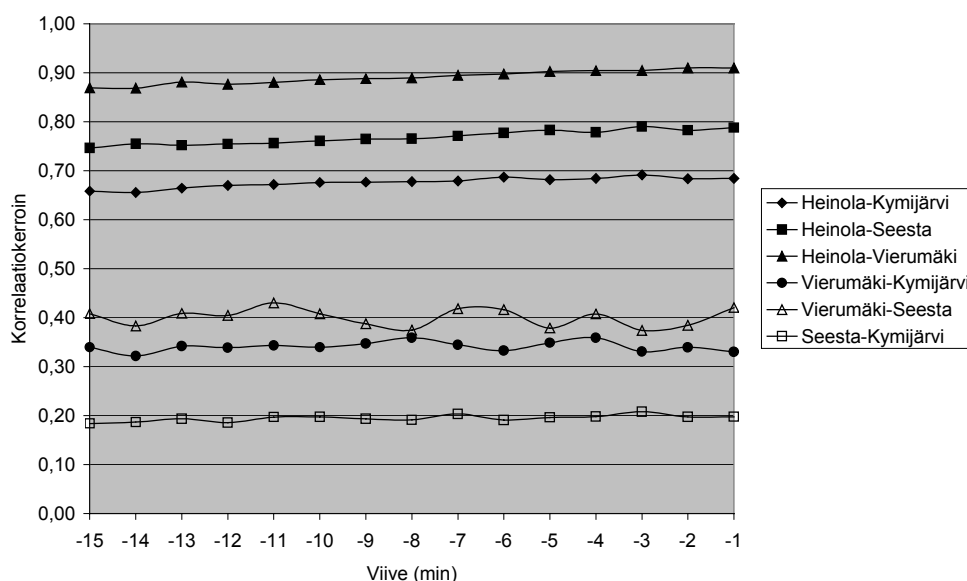
6.1 Syötöiden valinta

6.1.1 Valtatien 4 mallit

Tarkoituksena oli ennustaa valtatiellä 4 yksittäisille osalinkeille tai osalinkiyhdistelmille lähdössä olevien ajoneuvojen keskimatka-aikoja viimeisten matka-aikamittausten ja LAM-tietojen perusteella.

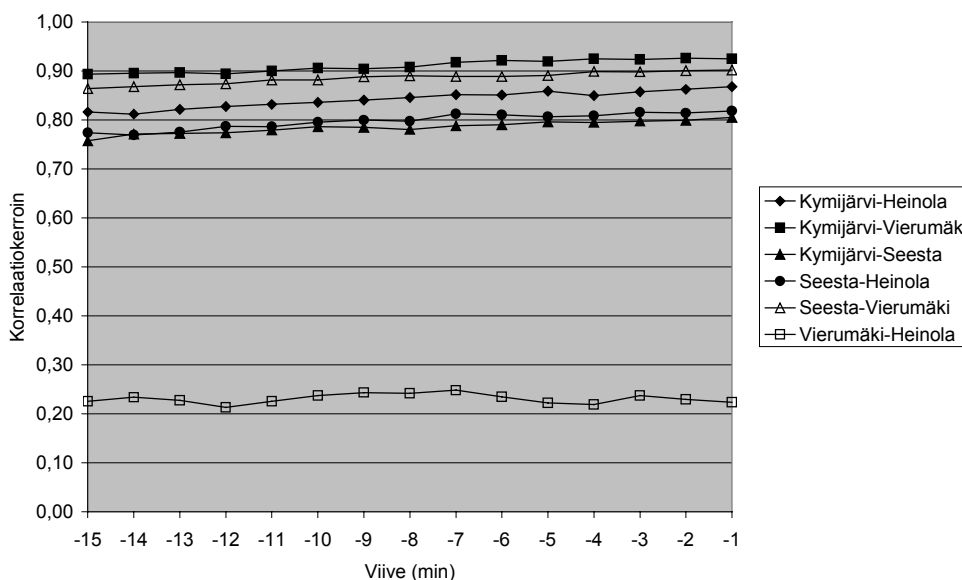
Syötemuuttujien valintaa varten kesän 2000 minuuttikohtaisesta aineistosta selvitettiin, kuinka eri muuttujat korreloivat keskenään. Ensimmäisenä selvitettiin, kuinka eri linkkien matka-ajat korreloivat ennustettavien Heinola–Kymijärvi- ja Kymijärvi–Heinola-linkeille lähtevien ajoneuvojen keskimatka-aikojen kanssa (kuvat 8 ja 9). Tarkastelussa käytettiin minuuttikohtaisia keskiarvoja erilaisilla viiveillä osalinkkien päätepisteet ohittaneiden ajoneuvojen matka-ajoista. Viiveellä tarkoitetaan aikaa mittausjakson alkamisen ja tiedon käyttämisen välillä. Esimerkiksi minuuttikeskiarvoissa viive -1 minuuttia tarkoittaa keskiarvoa, joka on laskettu ajalta 0–1 minuuttia ennen tarkasteluhetkeä.

Heinolasta lähtevien osalinkkien matka-ajat korreloivat ennustettavan Heinola–Kymijärvi-linkin matka-ajan kanssa parhaiten (kuva 8). Toiseksi parhaiten korreloivat Vierumäeltä lähtevät linkit ja Seesta–Kymijärvi-linkin matka-ajalla oli pienin korrelaatiokerroin. Korrelaatio oli sitä parempi, mitä lyhyempi linkki oli. Toisin sanoen korrelaatio oli siis sitä parempi, mitä lähempänä matka-aikatiedon näkevää ajoneuvoa matka-aikamittaukseen käytetyt ajoneuvot olivat tiellä kulkevassa ajoneuvojonossa. Tämä on varsin loogista. Heinolasta lähtevillä linkeillä korrelaatio huononi viiveen kasvaessa. Muilla linkeillä korrelaatioissa ei ollut havaittavissa laskevaa tai nousevaa trendiä.



Kuva 8. Eri osalinkkien keskimatka-ajan korrelointi Heinola–Kymijärvi-linkille lähtevien ajoneuvojen keskimatka-ajan kanssa erilaisilla viiveillä.

Vierumäki–Heinola-linkkiä lukuun ottamatta kaikkien Kymijärvi–Heinola-suunnan linkkien matka-ajat korreloivat varsin voimakkaasti ennustettavan Kymijärvi–Heinola-linkin keskimatka-ajan kanssa (kuva 9). Tällä suunnalla on huomattava, että se, mitä lähempänä matka-aikatiedon näkevää ajoneuvoa matka-aikamittaukseen käytetyt ajoneuvot olivat tiellä kulkevassa ajoneuvojonossa, ei selittänyt korrelaatiokerrointen suuruusjärjestystä. Pikemminkin korostuu, että kaikkien Seesta–Vierumäki-linkin sisällään pitävien linkkien matka-aika korreloi kokonaismatka-ajan kanssa parhaiten. Ainoastaan muiden linkkien matka-ajan korrelaation paremmuusjärjestyksen selittää tämä edellä mainittu vastakkaisella suunnalla pätenyt sääntö. Ilmeisesti ruuhka alkaa usein Seesta–Vierumäki-väliltä. Korrelaatiokerroin oli Vierumäki–Heinola-linkkiä lukuun ottamatta pienillä viiveillä yli 0,80 ja kerroin pieneni viiveen kasvaessa. Vierumäki–Heinola-linkin matka-ajan korrelaatiokerroin oli selvästi alhaisempi eikä siinä voitu havaita nousevaa eikä laskevaa trendiä.



Kuva 9. Eri osalinkkien keskimatka-ajan korrelointi Kymijärvi–Heinola-linkille lähtevien ajoneuvojen keskimatka-ajan kanssa erilaisilla viiveillä.

Matka-ajoista laskettiin keskiarvon lisäksi mediaaneja ja keskihajontoja. Yksittäisen minuutin aikana matka-aikahavaintoja oli vähän (esimerkiksi Kymijärvi–Heinola-linkillä 73 prosenttia ja Kymijärvi–Heinola-linkillä 76 prosenttia minuuttihavainnoista perustui yhteen matka-aikahavaintoon), eli viimeiset 10 tai 20 ajoneuvoa tarkoittavat usein huomattavasti minuuttia pidempää ajanjaksoa.

Mediaanien ja keskihajontojen korrelaatiot ennustettaviin matka-aikoihin määritettiin (taulukot 5 ja 6). Erilaiset keskihajonnat korreloivat selvästi huommin kuin mediaanit. Useammasta ajoneuvosta laskettu nopeuskeskihajonta korreloi paremmin kuin pienemmästä otoksesta laskettu. Matka-aikojen keskihajonnoilla korrelaation suuruus ja etumerkki kuitenkin vaihtelivat voimakkaasti linkiltä toiselle. Mediaanien korrelaatiokerrointen suuruus oli hyvin lähellä vastaavien keskiarvojen korrelaatiokertoimia, kuten saattoi odottaa.

Taulukko 5. Eri linkeiltä laskettujen suureiden korrelaatio ennustettavan Heinola–Kymijärvi-linkin matka-ajan kanssa. Suureissa on viiveenä yksi minuutti.

Etelään	He-Ky	He-Se	He-Vi	Vi-Ky	Vi-Se	Se-Ky
Viimeisen viiden minuutin matka-aikojen keskihajonta	-0,01	0,20	0,06	-0,29	-0,12	-0,04
Viimeisen 10 ajoneuvon matka-aikojen keskihajonta	-0,22	-0,05	-0,37	-0,48	-0,40	-0,18
Viimeisen 20 ajoneuvon matka-aikojen keskihajonta	-0,19	0,18	-0,12	-0,49	-0,48	-0,18
Viimeisen minuutin matka-aikojen mediaani	0,82	0,88	0,89	0,42	0,37	0,26
Viimeisen 10 ajoneuvon matka-aikojen mediaani	0,84	0,88	0,89	0,56	0,37	0,34
Viimeisen 20 ajoneuvon matka-aikojen mediaani	0,84	0,87	0,88	0,55	0,39	0,39

Taulukko 6. Eri linkeiltä laskettujen suureiden korrelaatio ennustettavan Kymijärvi–Heinola-linkin matka-ajan kanssa. Suureissa on viiveenä yksi minuutti.

Pohjoiseen	Ky-He	Ky-Vi	Ky-Se	Se-He	Se-Vi	Vi-He
Viimeisen viiden minuutin matka-aikojen keskihajonta	0,40	0,43	0,40	0,31	0,51	-0,07
Viimeisen 10 ajoneuvon matka-aikojen keskihajonta	-0,02	0,19	0,25	0,04	0,36	-0,26
Viimeisen 20 ajoneuvon matka-aikojen keskihajonta	0,10	0,35	0,46	0,17	0,39	-0,31
Viimeisen minuutin matka-aikojen mediaani	0,92	0,95	0,80	0,85	0,92	0,31
Viimeisen 10 ajoneuvon matka-aikojen mediaani	0,92	0,95	0,80	0,85	0,92	0,45
Viimeisen 20 ajoneuvon matka-aikojen mediaani	0,91	0,94	0,80	0,85	0,91	0,51

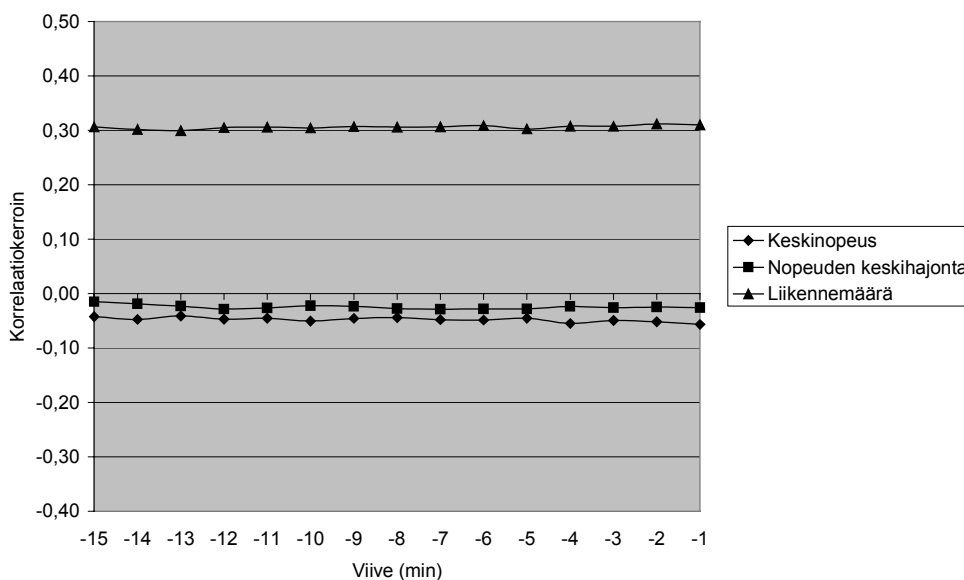
Se, että muuttuja korreloi voimakkaastikin ennustettavan matka-ajan kanssa, ei välttämättä tarkoita, että sitä kannattaa käyttää ennustemallin syötteenä. Jos muuttuja korreloi erittäin voimakkaasti jonkin toisen syötteen kanssa, se ei anna lisätietoa ja näin ollen sen arvo syötteenä on kyseenalainen.

Taulukossa 7 on esitetty matka-aikojen erilaisten mediaanien korrelaatiot yhden minuutin havainnoista lasketun keskimatka-ajan kanssa. Taulukosta havaitaan, että saman linkin keskimatka-aika ja minuutin matka-aikamediaani korreloivat erittäin voimakkaasti (korrelaatiokerroin 0,96 - 1,00). Tämä on seurausta havaintojen vähäisestä määrästä. Mediaani, joka on laskettu viimeisten 10 ajoneuvohavainnon matka-ajasta, korreloi hieman voimakkaammin kuin viimeisten 20 ajoneuvohavainnon perusteella laskettu mediaani. Nämä mediaanit korreloivat kuitenkin muutamissa tapauksissa selvästikin heikommin kuin minuuttimediaani. Se seikka, että Heinolaa kohti mentäessä Seestan ja Vierumäen mittauspisteissä ei ole kuin yksi kaista, eli kaikkia kaistoja seurataan, saattaa osaltaan selittää suuntien välistä eroa. Se voi olla selityksenä ainakin Seesta–Vierumäki-linkin hyvälle korrelaatiokertoimelle myös 10 ja 20 viimeisen havainnon osalta.

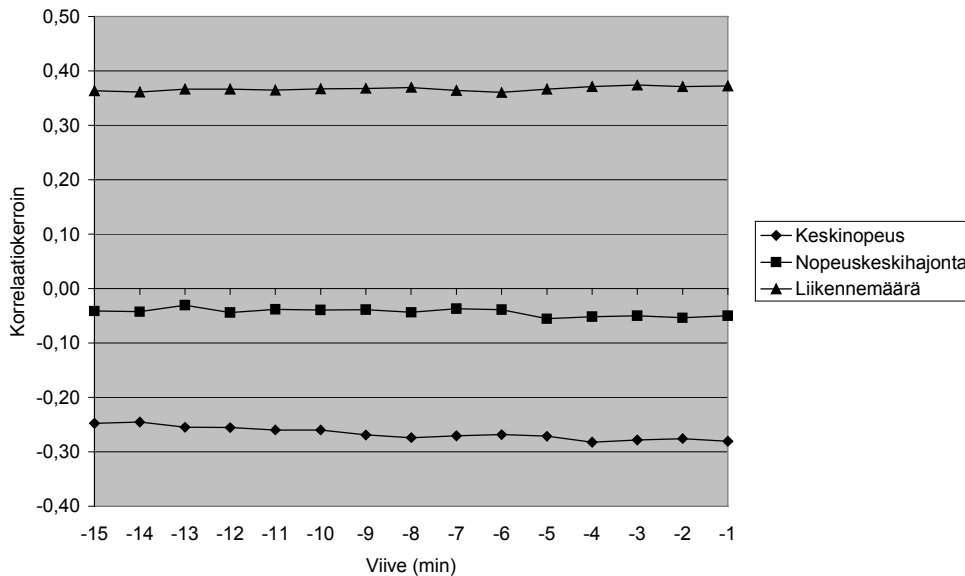
Taulukko 7. Eri linkeiltä laskettujen mediaanien korrelaatio saman linkin yhden minuutin keskimatka-ajan kanssa.

Linkki	Havainnot, joista mediaani laskettiin			
	1 min aikana saapuneet	10 viimeistä ajoneuvoa	20 viimeistä ajoneuvoa	
Etelään	He-Ky	0,98	0,80	0,78
	He-Se	0,99	0,90	0,88
	He-Vi	1,00	0,98	0,98
	Vi-Ky	0,96	0,51	0,46
	Vi-Se	0,97	0,55	0,50
	Se-Ky	0,97	0,53	0,48
Pohjoiseen	Ky-He	0,99	0,94	0,93
	Ky-Vi	1,00	0,98	0,98
	Ky-Se	1,00	0,97	0,96
	Se-He	0,99	0,94	0,93
	Se-Vi	1,00	0,98	0,98
	Vi-He	0,96	0,45	0,39

Vierumäen LAM-pistessä 601 mitatulle keskinopeudelle, nopeuden keskihajonnalle ja liikennemäärälle määriteltiin korrelaatiokertoimet (kuvat 10 ja 11). Nopeuden keskihajonta korreloi erittäin huonosti ennustettavien matka-aikojen kanssa. Liikennevirran keskinopeus korreloi Heinola–Kymijärvi-linkin matka-ajan kanssa lähes yhtä huonosti kuin nopeuskeskihajonta, mutta Kymijärvi–Heinola-linkin matka-ajan kanssa korrelaatio oli huomattavasti parempi. Liikennemäärän korrelaatio oli molemmilla suunnilla LAM-suureista paras. LAM-suureiden korrelaatio ei näyttänyt juurikaan muuttuvan ajan funktiona.



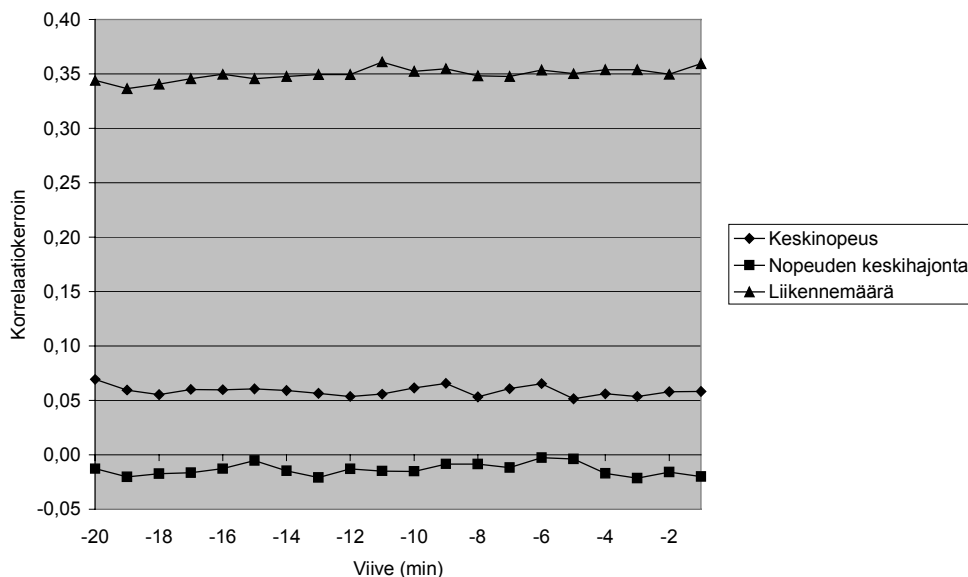
Kuva 10. Vierumäen LAM-pisteessä mitatun keskinopeuden, nopeuden keskihajonnan ja liikennemäärän korrelaatio Heinola–Kymijärvi-linkille lähtevien ajoneuvojen keskimatka-ajan kanssa erilaisilla viiveillä.



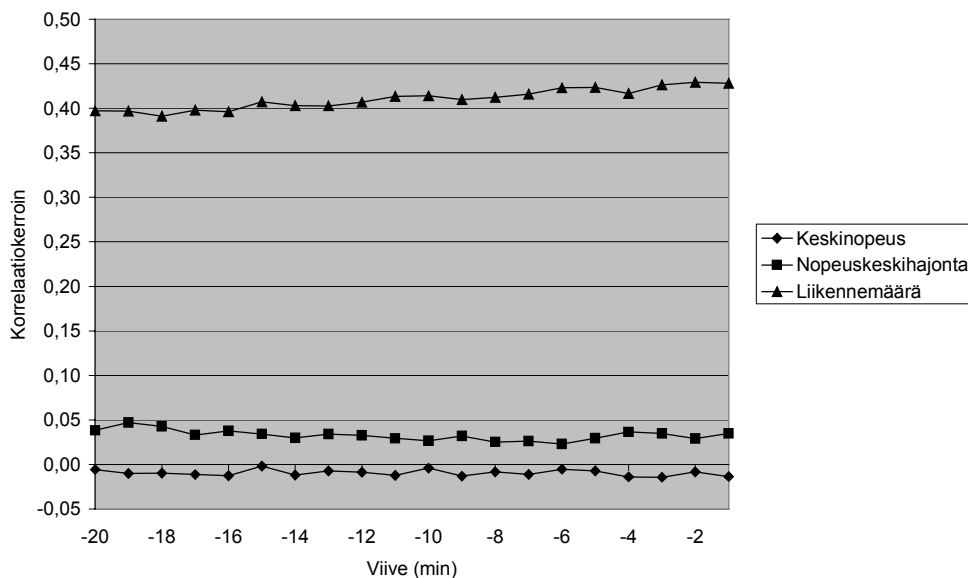
Kuva 11. Vierumäen LAM-pisteessä mitatun keskinopeuden, nopeuden keskihajonnan ja liikennemäärän korrelaatio Kymijärvi–Heinola-linkille lähtevien ajoneuvojen keskimatka-ajan kanssa erilaisilla viiveillä.

Tieto linkille tulevan liikenteen määrän kehityksestä voi olla mallissa tärkeä syöte. Lopuksi selvitetiinkin, kuinka ennustettavia linkejä ennen olevien LAM-pisteiden liikennetiedot korreloivat ennustettavien matka-aikojen kanssa. Heinola–Kymijärvi-suunnalla ennen tarkasteltavaa aluetta sijaitseva viimeinen LAM-piste on Murhamäessä (LAM 628), 10,0 kilometriä Heinolan kamerapisteen pohjoispuolella. Kymijärvi–Heinola-suunnalla viimeinen LAM-piste taas sijaitsee Hollolassa Lahden eteläpuolella (LAM 424), 11,9 kilometriä Kymijärven kamerapisteen eteläpuolella.

Molemmilla suunnilla LAM-pisteen liikennemäärätiedot korreloivat jossain määrin (korrelaatiokerroin 0,35–0,43) ennustettavan matka-ajan kanssa (kuvat 12 ja 13). Keskinopeuden ja nopeuden keskihajonnan korrelaatio oli sitä vastoin erittäin vähäinen.



Kuva 12. Murhamäen LAM-pisteessä mitatun keskinopeuden, nopeuden keskihajonnan ja liikennemäärän korrelaatio Heinola–Kymijärvi-linkille lähtevien ajoneuvojen keskimatka-ajan kanssa erilaisilla viiveillä.



Kuva 13. Hollolan LAM-pisteessä mitatun keskinopeuden, nopeuden keskihajonnan ja liikennemäärän korrelaatio Kymijärvi–Heinola-linkille lähtevien ajoneuvojen keskimatka-ajan kanssa erilaisilla viiveillä.

Myös Heinola–Kymijärvi-suunnan osalinkkien matka-aikojen väliset korrelaatiot selvitettiin (taulukko 8). Havaittiin, että Heinola-alkuisten linkkien matka-ajat korreloivat voimakkaasti keskenään (korrelaatiokerroin vähintään 0,83). Seesta–Kymijärvi- ja Vierumäki–Kymijärvi-linkkien matka-ajat korreloivat kohtalaisesti (korrelaatiokerroin 0,70), mutta muiden matka-aikojen välillä korrelaatio oli pienempi (korrelaatiokerroin korkeintaan 0,50). Sekä Vierumä-

en että Murhamäen LAM-pisteiden liikennemäärät korreloivat kohtalaisesti (korrelaatiokerroin 0,64–0,65) Heinola–Vierumäki-linkin matka-ajan kanssa. Muuten LAM-suureiden korrelaatiokertoimet olivat varsin pieniä.

Taulukko 8. Osalinkkien keskimatka-aikojen sekä Vierumäen (601) ja Murhamäen (628) LAM-pisteiden liikennemäärän (q), keskinopeuden (v) ja nopeuskeskihajonnan (kh) väliset korrelaatiokertoimet Heinola–Kymijärvi-suunnalla.

Etelään	He-Ky	He-Se	He-Vi	Vi-Ky	Vi-Se	Se-Ky
He-Ky	1,00	0,83	0,91	0,50	0,38	0,41
He-Se	0,83	1,00	0,97	0,34	0,49	0,16
He- Vi	0,91	0,97	1,00	0,31	0,39	0,32
Vi-Ky	0,50	0,34	0,31	1,00	0,28	0,70
Vi-Se	0,38	0,49	0,39	0,28	1,00	0,29
Se-Ky	0,41	0,16	0,32	0,70	0,29	1,00
q 601	0,32	0,37	0,64	0,27	0,32	0,14
v 601	-0,07	-0,09	-0,24	-0,09	-0,14	-0,02
kh 601	-0,02	-0,04	-0,12	-0,04	-0,05	0,00
q 628	0,35	0,43	0,65	0,26	0,30	0,08
v 628	0,06	0,07	0,12	0,03	0,06	0,01
kh 628	-0,01	-0,02	-0,05	-0,01	0,01	-0,02

	q 601	v 601	kh 601	q 628	v 628	kh 628
He-Ky	0,32	-0,07	-0,02	0,35	0,06	-0,01
He-Se	0,37	-0,09	-0,04	0,43	0,07	-0,02
He- Vi	0,64	-0,24	-0,12	0,65	0,12	-0,05
Vi-Ky	0,27	-0,09	-0,04	0,26	0,03	-0,01
Vi-Se	0,32	-0,14	-0,05	0,30	0,06	0,01
Se-Ky	0,14	-0,02	0,00	0,08	0,01	-0,02
q 601	1,00	0,02	0,07	0,67	0,21	-0,04
v 601	0,02	1,00	0,15	0,01	0,10	-0,02
kh 601	0,07	0,15	1,00	-0,02	-0,01	0,02
q 628	0,67	0,01	-0,02	1,00	0,19	0,02
v 628	0,21	0,10	-0,01	0,19	1,00	-0,08
kh 628	-0,04	-0,02	0,02	0,02	-0,08	1,00

Myös Kymijärvi–Heinola-suunnan osalinkkien matka-aikojen välinen korreloivuus selvitettiin (taulukko 9). Havaittiin, että osalinkkien matka-ajat korreloivat voimakkaasti tai vähintäänkin kohtalaisesti. Ainoan poikkeuksen suunnalla teki Vierumäki–Heinola-linkin matka-aika, joka korreloi huonosti kaikkien muiden linkkien matka-aikojen kanssa. Hollolan LAM-pisteen liikennemäärä korreloi matka-aikojen kanssa paremmin kuin Vierumäen LAM-pisteen liikennemäärä.

Taulukko 9. Osalinkkien keskimatka-aikojen sekä Vierumäen (601) ja Hollolan (424) LAM-pisteiden liikennemäärän (q), keskinopeuden (v) ja nopeuskeskihajonnan (kh) väliset korrelaatiokertoimet Kymijärvi–Heinola-suunnalla.

Pohjoiseen	Ky-He	Ky-Vi	Ky-Se	Se-He	Se-Vi	Vi-He
Ky-He	1,00	0,96	0,81	0,90	0,94	0,35
Ky-Vi	0,96	1,00	0,90	0,92	0,96	0,24
Ky-Se	0,81	0,90	1,00	0,61	0,81	0,21
Se-He	0,90	0,92	0,61	1,00	0,96	0,38
Se-Vi	0,94	0,96	0,81	0,96	1,00	0,25
Vi-He	0,35	0,24	0,21	0,38	0,25	1,00
q 601	0,33	0,31	0,31	0,28	0,31	0,13
v 601	-0,38	-0,50	-0,32	-0,35	-0,52	-0,12
kh 601	-0,09	-0,13	-0,08	-0,08	-0,13	-0,02
q 424	0,49	0,52	0,42	0,44	0,53	0,16
v 424	-0,02	-0,01	-0,01	0,00	-0,01	-0,01
kh 424	0,03	0,00	0,03	0,03	-0,01	0,03

	q 601	v 601	kh 601	q 424	v 424	kh 424
Ky-He	0,33	-0,38	-0,09	0,49	-0,02	0,03
Ky-Vi	0,31	-0,50	-0,13	0,52	-0,01	0,00
Ky-Se	0,31	-0,32	-0,08	0,42	-0,01	0,03
Se-He	0,28	-0,35	-0,08	0,44	0,00	0,03
Se-Vi	0,31	-0,52	-0,13	0,53	-0,01	-0,01
Vi-He	0,13	-0,12	-0,02	0,16	-0,01	0,03
q 601	1,00	-0,39	-0,08	0,52	0,14	-0,05
v 601	-0,39	1,00	0,09	-0,36	0,02	-0,01
kh 601	-0,08	0,09	1,00	-0,10	-0,03	0,03
q 424	0,52	-0,36	-0,10	1,00	0,07	0,02
v 424	0,14	0,02	-0,03	0,07	1,00	-0,66
kh 424	-0,05	-0,01	0,03	0,02	-0,66	1,00

Osalinkkien matka-aikojen ennustemalleja varten selvitetiin vielä, kuinka hyvin edellisen minuutin aikana kunkin linkin loppupään ohittaneiden ajoneuvojen keskimatka-aika korreloi eri osalinkeille seuraavan minuutin aikana lähtevien ajoneuvojen keskimatka-ajan kanssa (taulukot 10 ja 11). Taulukosta 10 voidaan havaita Heinola-alkuisten linkkien matka-aikojen voimakas korreloivuus. Näillä linkeillä muut osalinkit ja muilla linkeillä kaikki osalinkit korreloivat ennustettavien matka-aikojen kanssa suhteellisen heikosti (korrelaatiokertoimet alle 0,50).

Taulukko 10. Eri osalinkeille lähtevien ajoneuvojen keskimatka-aikojen sekä Vierumäen (601) ja Murhamäen (628) LAM-pisteiden liikennetietojen korrelaatio osalinkkien loppuun edeltävän minuutin aikana saapuneiden ajoneuvojen keskimatka-ajan kanssa Heinola–Kymijärvi-suunnalla.

Saapuneiden ajoneuvojen linkki	Lähtevien ajoneuvojen linkki (etelään)					
	He-Ky	He-Se	He-Vi	Vi-Ky	Vi-Se	Se-Ky
He-Ky	0,68	0,76	0,87	0,35	0,37	0,19
He-Se	0,79	0,85	0,92	0,37	0,32	0,21
He-Vi	0,91	0,93	0,95	0,34	0,37	0,32
Vi-Ky	0,33	0,33	0,32	0,24	0,23	0,22
Vi-Se	0,42	0,43	0,37	0,25	0,29	0,40
Se-Ky	0,20	0,17	0,32	0,28	0,27	0,27
L.määrä 601	0,31	0,36	0,63	0,26	0,31	0,14
K.nopeus 601	-0,06	-0,07	-0,21	-0,13	-0,10	-0,03
Nop.k.haj. 601	-0,03	-0,04	-0,11	-0,07	-0,05	0,00
L.määrä 628	0,36	0,42	0,65	0,26	0,29	0,07
K.nopeus 628	0,06	0,07	0,13	0,02	0,05	0,00
Nop.k.haj. 628	-0,02	-0,02	-0,05	0,03	0,01	-0,03

Kymijärvi–Heinola-suunnalla tilanne oli hieman parempi (taulukko 11). Ainoastaan Vierumäki–Heinola-linkille ei löytynyt voimakkaasti korreloivia osalinkkejä. Muilla linkeillä muut osalinkit korreloivat vähintäänkin kohtalaisesti (korrelaatiokerroin yli 0,60) Vierumäki–Heinola-linkkiä lukuun ottamatta.

Taulukko 11. Eri osalinkeille lähtevien ajoneuvojen keskimatka-aikojen sekä Vierumäen (601) ja Hollolan (424) LAM-pisteiden liikennetietojen korrelaatio osalinkkien loppuun edeltävän minuutin aikana saapuneiden ajoneuvojen keskimatka-ajan kanssa Kymijärvi–Heinola-suunnalla.

Saapuneiden ajoneuvojen linkki	Lähtevien ajoneuvojen linkki (pohjoiseen)					
	Ky-He	Ky-Vi	Ky-Se	Se-He	Se-Vi	Vi-He
Ky-He	0,87	0,91	0,78	0,83	0,92	0,25
Ky-Vi	0,93	0,94	0,86	0,93	0,94	0,27
Ky-Se	0,81	0,90	0,93	0,61	0,80	0,22
Se-He	0,82	0,88	0,62	0,88	0,92	0,25
Se-Vi	0,90	0,91	0,79	0,94	0,96	0,27
Vi-He	0,22	0,21	0,21	0,25	0,24	0,15
L.määrä 601	0,33	0,31	0,31	0,28	0,30	0,14
K.nopeus 601	-0,28	-0,42	-0,27	-0,29	-0,47	-0,12
Nop.k.haj. 601	-0,05	-0,10	-0,06	-0,05	-0,11	-0,01
L.määrä 424	0,43	0,48	0,39	0,40	0,50	0,15
K.nopeus 424	-0,01	-0,02	-0,02	0,00	-0,02	-0,02
Nop.k.haj. 424	0,03	0,00	0,03	0,04	-0,01	0,02

On mielenkiintoista havaita, että edellä mainitut linkit, joilla edellisen minuutin aikana kunkin linkin loppupään ohittaneiden ajoneuvojen keskimatka-aika ei korreloinut eri osalinkeille seuraavan minuutin aikana lähtevien ajoneuvojen keskimatka-ajan kanssa, ovat kaikki linkkejä, joilla ei ollut käytännöllisesti katsoen lainkaan ruuhkaa tarkastelujakson aikana (kaikki Vierumäki-alkuiset linkit sekä Seesta–Kymijärvi). Korreloimattomuus viittaa siis ilmeisesti siihen, että matka-ajat ovat vaihdelleet satunnaisesti normaalin matka-ajan ympärillä.

Eri linkeiltä määritettyjen viimeisen viiden minuutin ja 10/20 ajoneuvon matka-aikojen mediaanien sekä viimeisen minuutin ja 10/20 ajoneuvon matka-aikojen keskihajontojen korrelaatio ennustettaviin matka-aikoihin on esitetty osalinkkien osalta liitteessä A.

Syötteiksi päätettiin valita kaikki niiden linkkien keskimatka-aikojen aikasarjat, joiden korrelaatio ennustettavan suureen suhteen oli vähintään 0,20 (taulukot 10 ja 11). Matka-aikojen mediaaneista ja keskihajonnoista (taulukko 6) valittiin ne, jotka korreloivat ennustettavan matka-ajan kanssa parhaiten, olettaen, että korrelaatiokerroin oli vähintään 0,20. Matka-aikamediaaneista jätettiin kuitenkin huomiotta ne, jotka korreloivat erittäin voimakkaasti minuutin keskimatka-ajan kanssa (korrelaatiokerroin yli 0,95, taulukko 7). Jos kaksi mediaania korreloi ennustettavan suureen kanssa yhtä voimakkaasti, valittiin syötteisiin se, joka korreloi keskimatka-ajan kanssa vähemmän.

Yhteenvedot valituista syötesuureista on esitetty taulukoissa 12 ja 13. Malleja opetettaessa syötteistä piti kuitenkin jättää LAM-pisteen 628 tiedot pois, sillä ilmaisain ei ollut toiminnassa kesän 2000 aikana ja suureen mukaan ottaminen olisi karsinut käytettävissä olevaa aineistoa turhan paljon.

Taulukko 12. Heinola–Kymijärvi-suunnan eri ennustemalleille valittavat syötesuureet. A.s. tarkoittaa minuuttihavainnoista muodostettua aikasarjaa, 10/20 v. tarkoittaa 10/20 viimeistä havaintoa ja 5 min viimeistä viittä minuuttia.

Etelään		Ennustelinkki					
		He-Ky	He-Se	He-Vi	Vi-Ky	Vi-Se	Se-Ky
He-Ky	Keskiarvo	a.s.	a.s.	a.s.	a.s.	a.s.	-
	Mediaani	20 v.	10 v.	20 v.	20 v.	20 v.	20 v.
	Keskihajonta	10 v.	10 v.	20 v.	-	10 v.	-
He-Se	Keskiarvo	a.s.	a.s.	a.s.	a.s.	a.s.	a.s.
	Mediaani	10 v.	10 v.	10 v.	20 v.	20 v.	10 v.
	Keskihajonta	5 min	5 min	10 v.	-	-	-
He-Vi	Keskiarvo	a.s.	a.s.	a.s.	a.s.	a.s.	a.s.
	Mediaani	-	-	-	-	-	-
	Keskihajonta	10 v.	10 v.	10 v.	-	-	-
Vi-Ky	Keskiarvo	a.s.	a.s.	a.s.	a.s.	a.s.	a.s.
	Mediaani	10 v.	20 v.	20 v.	20 v.	20 v.	20 v.
	Keskihajonta	20 v.	20 v.	20 v.	20 v.	20 v.	20 v.
Vi-Se	Keskiarvo	a.s.	a.s.	a.s.	a.s.	a.s.	a.s.
	Mediaani	20 v.	20 v.	20 v.	20 v.	20 v.	10 v.
	Keskihajonta	20 v.	20 v.	20 v.	10 v.	10 v.	20 v.
Se-Ky	Keskiarvo	a.s.	-	a.s.	a.s.	a.s.	a.s.
	Mediaani	20 v.	20 v.	20 v.	20 v.	20 v.	20 v.
	Keskihajonta	-	20 v.	20 v.	-	10 v.	-
601	Liikennemäärä	a.s.	a.s.	a.s.	a.s.	a.s.	-
	Keskinopeus	-	-	a.s.	-	-	-
	Nopeuskeskihajonta	-	-	-	-	-	-
628	Liikennemäärä	a.s.	a.s.	a.s.	a.s.	a.s.	-
	Keskinopeus	-	-	-	-	-	-
	Nopeuskeskihajonta	-	-	-	-	-	-

Taulukko 13. Kymijärvi–Heinola-suunnan eri ennustemalleille valittavat syötesuureet. A.s. tarkoittaa minuuttihavainnoista muodostettua aikasarjaa, 10/20 v. tarkoittaa 10/20 viimeistä havaintoa ja 5 min viimeistä viittä minuuttia.

Pohjoiseen		Ennustelinkki					
		Ky-He	Ky-Vi	Ky-Se	Se-He	Se-Vi	Vi-He
Ky-He	Keskiarvo	a.s.	a.s.	a.s.	a.s.	a.s.	a.s.
	Mediaani	10 v.	10 v.	10 v.	10 v.	10 v.	20 v.
	Keskihajonta	5 min	5 min	5 min	5 min	5 min	-
Ky-Vi	Keskiarvo	a.s.	a.s.	a.s.	a.s.	a.s.	a.s.
	Mediaani	-	-	-	-	-	-
	Keskihajonta	5 min	5 min	5 min	5 min	5 min	-
Ky-Se	Keskiarvo	a.s.	a.s.	a.s.	a.s.	a.s.	a.s.
	Mediaani	-	-	-	-	-	-
	Keskihajonta	20 v.	20 v.	5 min	20 v.	20 v.	-
Se-He	Keskiarvo	a.s.	a.s.	a.s.	a.s.	a.s.	a.s.
	Mediaani	20 v.	10 v.	20 v.	10 v.	20 v.	20 v.
	Keskihajonta	5 min	5 min	-	5 min	5 min	-
Se-Vi	Keskiarvo	a.s.	a.s.	a.s.	a.s.	a.s.	a.s.
	Mediaani	-	-	-	-	-	-
	Keskihajonta	5 min	5 min	5 min	20 v.	20 v.	-
Vi-He	Keskiarvo	a.s.	a.s.	a.s.	a.s.	a.s.	-
	Mediaani	20 v.	20 v.	20 v.	20 v.	20 v.	20 v.
	Keskihajonta	20 v.	20 v.	20 v.	20 v.	20 v.	-
601	Liikennemäärä	a.s.	a.s.	a.s.	a.s.	a.s.	-
	Keskinopeus	a.s.	a.s.	a.s.	a.s.	a.s.	-
	Nopeuskeskihajonta	-	-	-	-	-	-
424	Liikennemäärä	a.s.	a.s.	a.s.	a.s.	a.s.	-
	Keskinopeus	-	-	-	-	-	-
	Nopeuskeskihajonta	-	-	-	-	-	-

6.1.2 Kehä I:n mallit

Kehä I:llä oli tarkoituksena ennustaa yksittäisille osalinkeille tai osalinkkiyhdistelmille lähdössä olevien ajoneuvojen keskimatka-aikoja viimeisten matka-aikamittausten ja LAM-tietojen perusteella aivan kuten valtatiellä 4 (ennustemallit). Lisäksi oli tarkoitus estimoida tietyn linkin keskimatka-aika muiden linkkien matka-aikojen ja LAM-pistetietojen perusteella (estimaattimallit).

Kehä I:n aineisto aggregoitiin viiden minuutin jaksoissa. Valtatiellä 4 käytetty yhden minuutin jakso olisi ollut liian pieni, koska Kehä I:llä linkin päästä päähän kulkevan liikenteen osuus on pienempi kuin valtatiellä 4 ja kohteessa on muita häiriötekijöitä, kuten liikennevaloja.

Syötemuuttujien valintaa varten aggregoidusta aineistosta selvitettiin, kuinka eri muuttujat korreloivat keskenään. Saman suunnan linkkien matka-ajat korreloivat kohtalaisesti keskenään (korrelaatiokertoimet 0,40 ja 0,60, taulukko 14). Otaniemi–Konala-linkin matka-aika korreloi molemmilla suunnilla LAM-pisteiden liikennemäärien kanssa voimakkaammin kuin Konala–Pukinmäki-linkin matka-aika. Korrelaatio LAM-pisteissä mitattujen pistenopeuksien tai pistenopeuden keskihajonnan kanssa oli erittäin pieni.

Taulukko 14. Osalinkkien keskimatka-aikojen sekä Konalan (126) ja Pukinmäen (148) LAM-pisteiden liikennemäärän (q), keskinopeuden (v) ja nopeuskeskihajonnan (kh) välinen korrelaatio.

Itään	Ot-Ko	Ko-Pu	q 126	v 126	kh 126	q 148	v 148	kh 148
Ot-Ko	1,00	0,60	0,31	-0,09	-0,06	0,31	-0,02	-0,06
Ko-Pu		1,00	0,15	-0,06	-0,03	0,14	-0,01	-0,03
q 126			1,00	0,17	-0,19	0,90	-0,16	-0,11
v 126				1,00	-0,20	-0,12	0,01	0,03
kh 126					1,00	-0,20	0,07	0,05
q 148						1,00	-0,17	-0,11
v 148							1,00	-0,34
kh 148								1,00

Länteen	Pu-Ko	Ko-Ot	q 126	V 126	kh 126	Q 148	V 148	kh 148
Pu-Ko	1,00	0,40	0,24	-0,12	-0,03	0,23	-0,02	-0,04
Ko-Ot		1,00	0,40	-0,20	-0,08	0,45	-0,02	-0,06
q 126			1,00	0,03	-0,20	0,92	-0,04	-0,18
v 126				1,00	-0,15	-0,22	-0,01	0,03
kh 126					1,00	-0,20	0,04	0,08
q 148						1,00	-0,02	-0,19
v 148							1,00	-0,46
kh 148								1,00

Kunkin linkin keskimatka-aikojen korrelaatio matka-ajan keskihajonnan ja mediaanimatka-ajan kanssa selvitettiin myös. Korrelaatio oman linkin matka-aikojen hajontasuureiden kanssa oli voimakkaampi kuin saman suunnan toisen linkin matka-ajan hajonnan kanssa (taulukko15). Keskimatka-aika korreloi hyvin voimakkaasti saman linkin mediaanimatka-ajan kanssa.

Taulukko 15. Eri linkeiltä laskettujen matka-aikasuureiden korrelaatio estimoitavien tai ennustettavien matka-aikojen kanssa.

Ennuste-/estimaattilinkki Lähtötietolinkki	Ot-Ko		Ko-Pu		Pu-Ko		Ko-Ot	
	O-K	K-P	O-K	K-P	P-K	K-O	P-K	K-O
Viimeisen viiden minuutin matka-aikojen keskihajonta	0,45	0,15	0,27	0,32	0,44	0,11	0,16	0,29
Viimeisen 10 ajoneuvon matka-aikojen keskihajonta	0,56	0,14	0,34	0,42	0,55	0,12	0,28	0,36
Viimeisen 20 ajoneuvon matka-aikojen keskihajonta	0,66	0,18	0,38	0,43	0,57	0,16	0,31	0,45
Viimeisen viiden minuutin matka-aikojen mediaani	0,99	0,59	0,60	0,97	0,99	0,40	0,40	0,99
Viimeisen 10 ajoneuvon matka-aikojen mediaani	0,97	0,62	0,59	0,87	0,94	0,42	0,38	0,96
Viimeisen 20 ajoneuvon matka-aikojen mediaani	0,96	0,64	0,58	0,83	0,88	0,41	0,37	0,94

Vastaavalla tavalla kuin valtatielle 4 myös Kehä I:n malleille valittiin syötteiksi parametrit, joiden korrelaatiokertoimet ennustettavan/estimoitavan matka-ajan kanssa olivat vähintään 0,20 ja muiden syötteiden kanssa korkeintaan 0,95. Syötteiden perustaksi valittiin jälleen saman suunnan keskimatka-ajat. Matka-aikojen mediaanit korreloivat keskimatka-aikojen kanssa niin voimakkaasti (korrelaatiokerroin yli 0,95, taulukko 15), ettei niitä otettu syötteisiin mukaan. Keskihajonnoista valittiin ne, jotka korreloivat ennustettavan matka-ajan kanssa parhaiten, olettaen, että edellä esitetyt ehdot täyttyivät.

Yhteenveto valituista syötesuureista on esitetty taulukossa 16. On huomattava, että malleissa, joissa tietyn linkin matka-aikaa yritetään estimoida muista mittasuureista, ei käytetty syötteinä mitään kyseisen linkin matka-aikasuureista. Ennustemallit sitä vastoin perustuivat kaikkiin valittuihin syötesuureisiin. Aikasarjan pituutena käytettiin kolme kertaa viittä minuuttia.

Taulukko 16. Eri linkkien ennuste- ja estimaattimalleille valittavat syötesuureet. Estimaahteissa tosin oman linkin matka-aikasuureita ei käytetä hyväksi. 20 v. tarkoittaa 20 viimeistä havaintoa.

			Ennuste-/estimaattilinkki			
			Ot-Ko	Ko-Pu	Pu-Ko	Ko-Ot
Itään	Ot-Ko	Keskiarvo Keskihajonta	aikasarja 20 v.	aikasarja 20 v.	- -	- -
	Ko-Pu	Keskiarvo Keskihajonta	aikasarja -	aikasarja 20 v.	- -	- -
Länteen	Pu-Ko	Keskiarvo Keskihajonta	- -	- -	aikasarja 20 v.	aikasarja 20 v.
	Ko-Ot	Keskiarvo Keskihajonta	- -	- -	aikasarja -	aikasarja 20 v.
126		Liikennemäärä Keskinopeus	aikasarja -	- -	aikasarja -	aikasarja aikasarja
148		Liikennemäärä	aikasarja	-	aikasarja	aikasarja

6.2 Ennusteet ja estimaatit

6.2.1 Valtatie 4

Valtatielle 4 tehtiin matka-ajan ennustemalleja. Mallit ennustivat tarkasteluhetkeä seuraavan minuutin aikana linkeille lähtevien ajoneuvojen keskimatka-aikaa tarkasteluhetkeä edeltävien minuuttien aikana linkkien loppupisteet ohittaneiden ajoneuvojen matka-aikojen ja LAM-pistetietojen perusteella. Syötteet valittiin taulukoiden 12 ja 13 mukaisesti. Murhamäen LAM-pisteen (628) tietoja ei kuitenkaan käytetty, koska liikennetiedot puuttuivat lähes kokonaan (23.6., 4.–18.7. ja 28.7.–7.8.2000) kesän 2000 aineistosta. Kokeiltavaksi jäi käytettävän aikasarjan optimaalinen pituus. Malleissa kokeiltavien aikasarjojen pituuksiksi valittiin kolme, neljä ja viisi minuuttia.

Mallien syötteet koottiin periaatteella, että muuttujan arvo pysyy samana, kunnes saadaan uusi havainto. Neuroverkon opetusjoukon vasteet kuitenkin koottiin ainoastaan mittausuloksista, eikä niissä sovellettu vastaavaa päivitysperiaatetta. Aineistosta siis karsittiin pois minuutit, joilta ei ollut yhtään havaintoa vasteen määrittämiseksi sekä ne jaksot, joissa jollain linkillä ei ollut päivitetty syötteitä yli puoleen tuntiin. Jäljelle jääneiden havaintojen määrät sekä neuroverkon eri kerrosten neuronien määrät eri malleille on esitetty taulukoissa 17 ja 18.

Taulukko 17. Heinola–Kymijärvi-suunnan eri ennustemallien tekemisessä käytettävissä olleiden havaintojen määrä (N) sekä mallin syöte-, piilo- ja vasteneuronien määrät. Piiloneuronien maksimimäärä rajoitettiin kuitenkin 20:een (*).

Aikasarjan pituus		Etelään					
		He-Ky	He-Se	He-Vi	Vi-Ky	Vi-Se	Se-Ky
5 min	N	12 713	9 983	18 580	15 682	13 856	11 527
	syöte	45	41	51	42	44	32
	piilo	14	12	18	19	16	19
	vaste	1	1	1	1	1	1
4 min	N	12 738	9 997	18 654	15 743	13 916	11 557
	syöte	38	35	43	35	37	27
	piilo	17	15	22*	23*	19	23*
	vaste	1	1	1	1	1	1
3 min	N	12 764	10 013	18 733	15 798	13 968	11 590
	syöte	31	29	35	28	30	22
	piilo	21*	18	28*	30*	24*	28*
	vaste	1	1	1	1	1	1

Taulukko 18. Kymijärvi–Heinola-suunnan eri ennustemallien tekemisessä käytettävissä olleiden havaintojen määrä (N) sekä mallin syöte-, piilo- ja vasteneuronien määrät. Piiloneuronien maksimimäärä rajoitettiin kuitenkin 20:een (*).

Aikasarjan pituus		Pohjoiseen					
		Ky-He	Ky-Vi	Ky-Se	Se-He	Se-Vi	Vi-He
5 min	N	30 323	42 541	43 695	39 748	51 691	50 106
	syöte	54	54	53	54	54	28
	piilo	29*	40*	42*	38*	49*	96*
	vaste	1	1	1	1	1	1
4 min	N	30 342	42 563	43 726	39 767	51 719	50 181
	syöte	45	45	44	45	45	23
	piilo	35*	49*	52*	46*	60*	119*
	vaste	1	1	1	1	1	1
3 min	N	30 361	42 584	43 753	39 781	51 739	50 241
	syöte	36	36	35	36	36	18
	piilo	44*	62*	66*	58*	76*	156*
	vaste	1	1	1	1	1	1

Ennusteiden hyvyttä voidaan tarkastella kahdella eri tavalla: tilastollisesti ja matka-aikatiedotuksen hyvyyden näkökulmasta. Mallien tekemien ennusteiden tilastollisia tunnuslukuja on esitetty taulukoissa 19 ja 20. Keskimääräinen virhe ja suhteellinen virhe kertovat, onko mallilla taipumus yli- tai aliarvioida ennusteita. Keskimääräinen neliövirhe sekä virheen ja suhteellisen virheen itseisarvo taas kuvaavat sitä, kuinka virheet jakautuvat. Mallien paremuusjärjestys on harvoin yksikäsitteinen, sillä se vaihtelee usein tarkasteltavan kriteerin mukaan.

Heinola–Kymijärvi-suunnalla (taulukko 19) Heinola–Kymijärvi- ja Heinola–Seesta-linkkien malleista parhaaksi osoittautui malli, jossa käytettiin neljän minuutin aikasarjaa. Heinola–Vierumäki-linkin malleissa optimaalisimmaksi aikasarjan pituudeksi osoittautui kolme minuuttia. Vierumäki–Seesta- ja Seesta–Kymijärvi-linkeillä se taas oli viisi minuuttia ja Vierumäki–Kymijärvi-linkillä neljän ja viiden minuutin aikasarjat antoivat yhtä hyvän tuloksen. Kaikilla linkeillä parhaat mallit tekivät 0,5–0,6 prosentin keskimääräisen suhteellisen virheen ja suhteellisen virheen itseisarvo oli noin kuusi prosenttia.

Taulukko 19. Eri Heinola–Kymijärvi-suunnan (etelään) ennustemallien hyvyys keskimääräisellä neliövirheellä (MSE), virheellä (ME), virheen itseisarvolla (MAE), suhteellisella virheellä (MRE), suhteellisen virheen itseisarvolla (MARE) ja sovitustasteella (R) määritettynä. Parhaat mallit on merkitty lihavoidulla reunuksella.

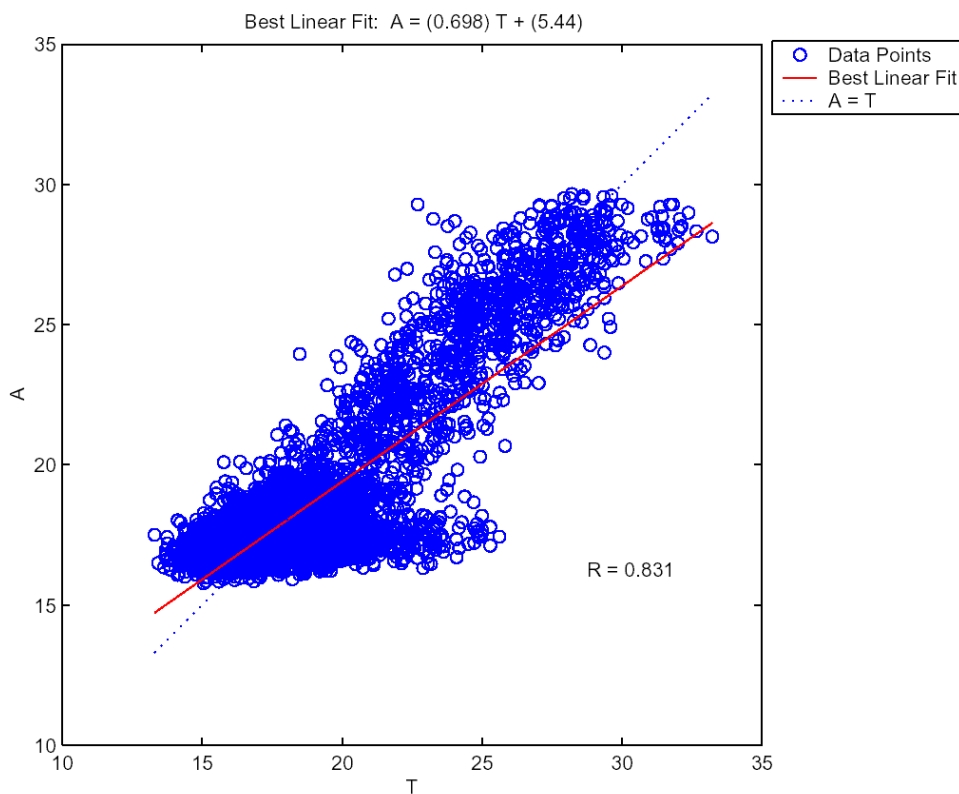
Aikasarjan pituus		He-Ky	He-Se	He-Vi	Vi-Ky	Vi-Se	Se-Ky
5 min	MSE	2,1	1,0	0,4	0,7	0,2	0,2
	ME	0,0	0,0	0,0	0,0	0,0	0,0
	MAE	1,1	0,8	0,5	0,6	0,3	0,3
	MRE	0,6 %	0,5 %	0,9 %	0,5 %	0,5 %	0,6 %
	MARE	6,1 %	6,0 %	7,1 %	5,6 %	5,8 %	5,9 %
	R	0,829	0,907	0,940	0,421	0,540	0,377
4 min	MSE	2,1	1,0	0,4	0,7	0,2	0,2
	ME	0,0	0,0	0,0	0,0	0,0	0,0
	MAE	1,1	0,8	0,5	0,6	0,3	0,3
	MRE	0,6 %	0,5 %	0,8 %	0,5 %	0,6 %	0,7 %
	MARE	6,0 %	5,9 %	7,1 %	5,6 %	5,8 %	5,9 %
	R	0,831	0,908	0,941	0,424	0,541	0,380
3 min	MSE	2,1	1,0	0,4	0,7	0,2	0,2
	ME	0,0	0,0	0,0	0,0	0,0	0,0
	MAE	1,1	0,8	0,5	0,6	0,3	0,3
	MRE	0,7 %	0,7 %	0,6 %	0,7 %	0,6 %	0,6 %
	MARE	6,1 %	6,0 %	6,9 %	5,7 %	5,8 %	6,0 %
	R	0,830	0,907	0,945	0,420	0,535	0,368

Kymijärvi–Heinola-linkin malleissa (taulukko 20) aikasarjan pituudella ei ollut vaikutusta tuloksiin. Kymijärvi–Seesta-linkillä optimaaliseksi aikasarjan pituudeksi osoittautui neljä minuuttia ja muilla Kymijärvi–Heinola-suunnan linkeillä viisi minuuttia. Tulokset olivat suunnilleen samanlaisia kuin päinvastaisella suunnalla. Kaikilla linkeillä parhaat mallit tekivät 0,5–0,7 prosentin keskimääräisen suhteellisen virheen ja suhteellisen virheen itseisarvo oli noin kuusi prosenttia.

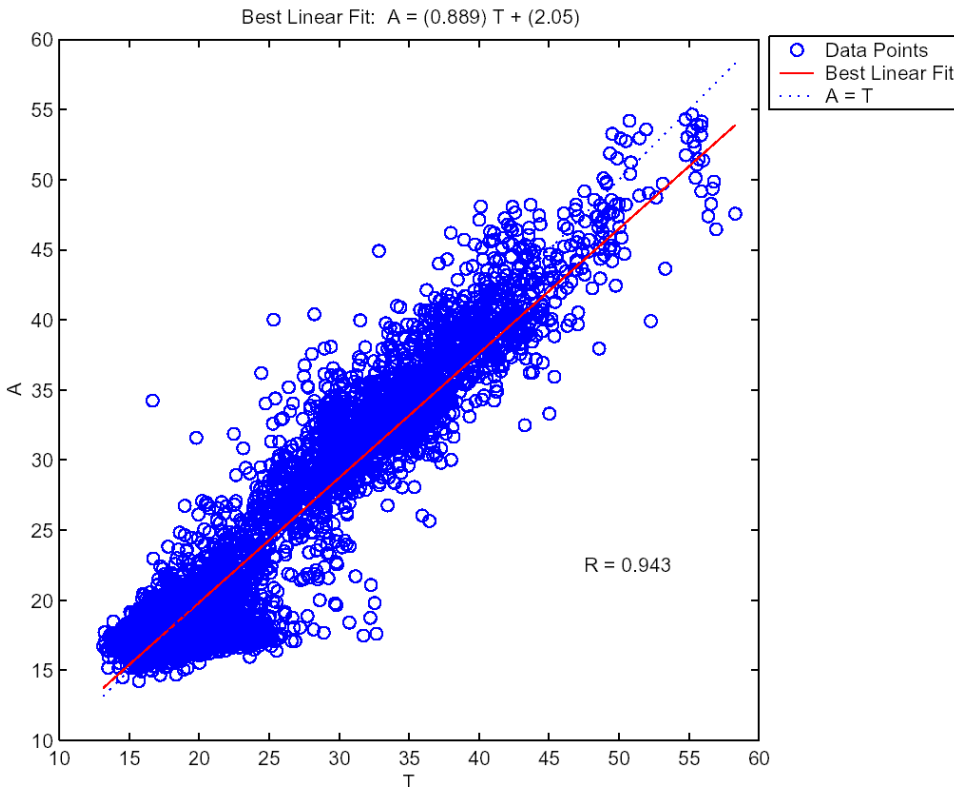
Taulukko 20. Eri Kymijärvi–Heinola-suunnan (pohjoiseen) ennustemallien hyvyys keskimääräisellä neliövirheellä (MSE), virheellä (ME), virheen itseisarvolla (MAE), suhteellisella virheellä (MRE), suhteellisen virheen itseisarvolla (MARE) ja sovitustasteella (R) määritettynä. Parhaat mallit on merkitty lihavoidulla reunuksella.

Aikasarjan pituus		Ky-He	Ky-Vi	Ky-Se	Se-He	Se-Vi	Vi-He
5 min	MSE	2,5	1,3	0,4	1,0	0,5	0,2
	ME	0,0	0,0	0,0	0,0	0,0	0,0
	MAE	1,1	0,8	0,4	0,7	0,4	0,4
	MRE	0,6 %	0,7 %	1,1 %	0,5 %	0,6 %	0,5 %
	MARE	6,0 %	5,9 %	6,7 %	5,6 %	6,1 %	5,5 %
	R	0,943	0,969	0,910	0,964	0,981	0,357
4 min	MSE	2,5	1,3	0,4	1,1	0,5	0,2
	ME	0,0	0,0	0,0	0,0	0,0	0,0
	MAE	1,1	0,8	0,4	0,7	0,4	0,4
	MRE	0,6 %	0,7 %	0,7 %	0,6 %	0,7 %	0,5 %
	MARE	6,0 %	6,0 %	6,5 %	5,6 %	6,1 %	5,6 %
	R	0,943	0,969	0,909	0,963	0,981	0,357
3 min	MSE	2,5	1,3	0,4	1,1	0,5	0,2
	ME	0,0	0,0	0,0	0,0	0,0	0,0
	MAE	1,1	0,8	0,4	0,7	0,4	0,4
	MRE	0,6 %	0,7 %	0,8 %	0,5 %	0,7 %	0,6 %
	MARE	6,0 %	6,0 %	6,5 %	5,6 %	6,2 %	5,6 %
	R	0,943	0,968	0,912	0,963	0,980	0,351

Kuvissa 14 ja 15 sekä liitteessä B on esitetty kullekin linkille sovitussasteella (taulukot 19 ja 20) mitattuna paras malli ja, kuinka sen ennustamat havainnot osuvat mitattuihin arvoihin. Osa malleista ennustaa matka-ajan erittäin hyvin (esimerkiksi Kymijärvi–Heinola-linkki kuvassa 15). Osassa malleista ennusteet osuvat pääosin hyvin, mutta ennustettujen matka-aikojen vaihteluväli on rajautunut siten, että kaikkein pienimmät ja suurimmat matka-ajat jäävät pois (esimerkiksi Heinola–Kymijärvi-linkki kuvassa 14). Näitä molempia edellä kuvattuja tilanteita voidaan pitää tyydyttävänä. Mukana on kuitenkin myös malleja, joiden opetusjoukossa ruuhkahavaintoja on ollut erittäin vähän, jos ollenkaan (esimerkiksi Vierumäki–Seesta-linkki liitteessä B). Nämä mallit eivät ole näin ollen oppineet ennustamaan ruuhkatilanteita oikein, vaan malli ennustaa matka-ajan samalla tavalla kaikissa olosuhteissa. Opetusjoukkoa varten pitäisi koota lisähavaintoja ruuhkaisista olosuhteista, jos mallin halutaan kykenevän ennustamaan myös ruuhkatilanteissa.



Kuva 14. Heinola–Kymijärvi-linkillä mitatut (vaaka-akseli) ja ennustetut (pysty-akseli) matka-ajat. Aikasarja pituutena on käytetty neljää minuuttia.



Kuva 15. Kymijärvi–Heinola-linkillä mitatut (vaaka-akseli) ja ennustetut (pysty-akseli) matka-ajat. Aikasarjan pituutena on käytetty viittä minuuttia.

Molemmassa kuvissa 14 ja 15 on havaittavissa joukko havaintoja, joissa ennustemalli ei ole kyennyt ennakoimaan ruuhkaa (ennustettu matka-aika alle 20 minuuttia, mutta mitattu yli 22 minuuttia). Heinola–Kymijärvi-linkillä kaikki nämä havainnot ovat sellaisia, joissa yksittäisen minuutin aikana on mitattu selvästi suurempi matka-aika kuin ennen sitä tai sen jälkeen. Kysymyksessä on joko erittäin paikallinen ja nopeasti selvinnyt häiriö tai tilanne, jossa raaka-aineiston suodatus on tehty liian karkeasti. Suodatettaessa aineistoon jää aina havaintoja, joista on vaikea sanoa, onko kyseessä poikkeava havainto vai ei. Tällaisia tapauksia ovat ajoneuvot, jotka on jostain syystä pysähtyneet hetkeksi välille, tai rekisterikilpien väärin tulkinnoista aiheutuvat muista havainnoista vain hieman poikkeavat virheelliset matka-ajat. Kun esimerkiksi Heinola–Kymijärvi-linkin minuuttihavainnoista 73 prosenttia (Kymijärvi–Heinola-linkillä 76 prosenttia) perustui yhteen ainoaan ajoneuvohavaintoon, tällaisen virheen vaikutus näkyy.

Myös Kymijärvi–Heinola-linkillä on runsaasti edellä mainitun kaltaisia yksittäisiä pidempiä matka-aikahavaintoja, joihin malli ei reagoi. Tämän lisäksi joukossa on myös pieniä ruuhkia tai suurempien ruuhkien alkua, joiden alusta malli ”myöhästyy”. Kyseiset ruuhkat sijoittuvat tyypillisesti Seestan ja Vierumäen välille, jolloin ruuhka alkaa selvästi ennusteen teko hetkeä myöhemmin (ensimmäiset mitattavat merkit näkyvät noin 10–20 minuuttia ennusteen teko hetken jälkeen).

Edellä esitetyistä tyyppin yksi virhe (yksittäinen piikki) ei ole vakava, mutta tyyppin kaksi virheet (ruuhkan alusta myöhästyminen) pitäisi korjata. Jos mitaustulokset alkavat antaa ensimmäisiä merkkejä jostain poikkeavasta vasta noin 10–20 minuuttia ennusteen tekohetken jälkeen, on kuitenkin vaikea keksiä keinoja, jolla tästä ongelmasta päästäisiin eroon.

Mallien antamien ennusteiden hyvyttä mitattiin myös matka-aikatiedotuksen näkökulmasta eli sillä, kuinka suuri osa ennusteista olisi mennyt oikein ja kuinka suuri osa olisi ollut liian pieniä tai suuria. Tarkastelu tehtiin erikseen kaikille havainnoille ja ruuhkatilanteille. Laskennassa oletettiin, että opastetaululla olisi näytetty matka-ajan alarajana 0,9 kertaa kyseiselle linkille ennustettua matka-aikaa ja ylärajana 1,1 kertaa ennustettua matka-aikaa (kymmenen prosentin marginaali) kuitenkin siten, etteivät rajat voisi olla pienempiä kuin nopeusrajoituksen mukainen matka-aika. Jos siis mitattu matka-aika on ollut pienempi kuin opastetaulun antama matka-ajan alaraja, ja alaraja on näyttänyt minimiä, tieto on tulkittu oikeaksi.

Kaikkien linkkien matka-aikatieto olisi ollut kesinä 2000 ja 2001 oikeassa yli 95 prosenttia ajasta (taulukot 21 ja 22) ja Heinola–Vierumäki-suunnalla yli 97 prosenttia ajasta. Aikasarjan pituudella ei ollut suurta merkitystä.

Taulukko 21. Heinola–Kymijärvi-suunnan eri ennustemallien hyvyys mitattuna oikein ennustettujen (o%), liian pieneksi arvioitujen (p%) ja liian suureksi arvioitujen (s%) matka-aikojen osuutena ajasta koko kesien 2000 ja 2001 aineistossa. Parhaat mallit on merkitty lihavoidulla reuoksella.

Aikasarjan pituus		He-Ky	He-Se	He-Vi	Vi-Ky	Vi-Se	Se-Ky
5 min	o%	98,1	97,0	96,7	99,5	99,5	100,0
	p%	1,2	1,7	1,7	0,5	0,4	0,0
	s%	0,7	1,3	1,7	0,0	0,1	0,0
4 min	o%	98,4	97,2	96,9	99,5	99,5	100,0
	p%	1,1	1,8	1,5	0,5	0,4	0,0
	s%	0,5	1,1	1,5	0,0	0,1	0,0
3 min	o%	98,3	97,1	97,1	99,5	99,5	100,0
	p%	1,2	1,7	1,5	0,5	0,4	0,0
	s%	0,6	1,2	1,5	0,0	0,2	0,0

Taulukko 22. Kymijärvi–Heinola-suunnan eri ennustemallien hyvyys mitattuna oikein ennustettujen (o%), liian pieneksi arvioitujen (p%) ja liian suureksi arvioitujen (s%) matka-aikojen osuutena ajasta koko kesien 2000 ja 2001 aineistossa. Parhaat mallit on merkitty lihavoidulla reuoksella.

Aikasarjan pituus		Ky-He	Ky-Vi	Ky-Se	Se-He	Se-Vi	Vi-He
5 min	o%	97,4	96,3	96,4	97,8	95,5	99,8
	p%	1,5	1,9	1,8	1,2	2,1	0,2
	s%	1,1	1,7	1,8	1,0	2,4	0,0
4 min	o%	97,4	96,4	96,4	97,7	95,5	99,8
	p%	1,5	1,9	1,8	1,3	2,1	0,2
	s%	1,1	1,7	1,8	1,0	2,5	0,1
3 min	o%	97,4	96,4	96,5	97,7	95,4	99,8
	p%	1,5	2,1	1,8	1,3	2,1	0,2
	s%	1,1	1,6	1,7	1,1	2,5	0,1

Tasaista ruuhkatonta liikennettä on helppo ennustaa ja tämän takia edellä kuvattu tarkastelu tehtiin myös ruuhkaliikenteelle. Ruuhkalla tarkoitetaan tässä yhteydessä mitattua matka-aikaa, joka vastaa alhaisempaa nopeutta kuin näyttötaulun matka-aikaminimin ylärajaa vastaava nopeus. Tulokset on esitetty taulukoissa 23 ja 24. Taulukoihin on tummennettu ne mallit, joita ei voitu testata kunnolla ruuhkahavaintojen vähäisyyden takia.

Oikein ennustettujen matka-aikojen osuus vaihteli voimakkaasti linkiltä toiselle (taulukot 23 ja 24). Ruuhkaennusteissa Heinola–Kymijärvi-linkin matka-aikaennusteet osoittautuivat parhaiksi (oikein ennustettujen osuus 78,8 prosenttia ajasta) ja Kymijärvi–Seesta-linkin huonoimmiksi (oikein ennustettujen osuus 39,3 prosenttia ajasta). Muista linkeistä Heinola–Seesta-, Kymijärvi–Heinola- ja Seesta–Heinola-linkeillä ruuhkassa oikein ennustettujen osuus oli yli 70 prosenttia. Vierumäelle päättyvillä linkeillä osuus oli 60–70 prosenttia.

Taulukko 23. Heinola–Kymijärvi-suunnan eri ennustemallien kyky ennustaa ruuhkaliikennettä mitattuna oikein ennustettujen (o%), liian pieneksi arvioitujen (p%) ja liian suureksi arvioitujen (s%) matka-aikojen osuutena ajasta. Tummennetuilla linkeillä ei ollut ruuhkaa. Parhaat mallit on merkitty lihavoidulla reunuksella.

Aikasarjan pituus		He-Ky	He-Se	He-Vi	Vi-Ky	Vi-Se	Se-Ky
5 min	o%	76,3	72,5	64,0			
	p%	20,4	20,3	21,9			
	s%	3,2	7,2	14,2			
4 min	o%	78,8	72,8	66,4			
	p%	18,8	20,8	20,4			
	s%	2,5	6,4	13,2			
3 min	o%	78,1	72,5	68,4			
	p%	19,2	19,6	19,5			
	s%	2,7	7,9	12,1			

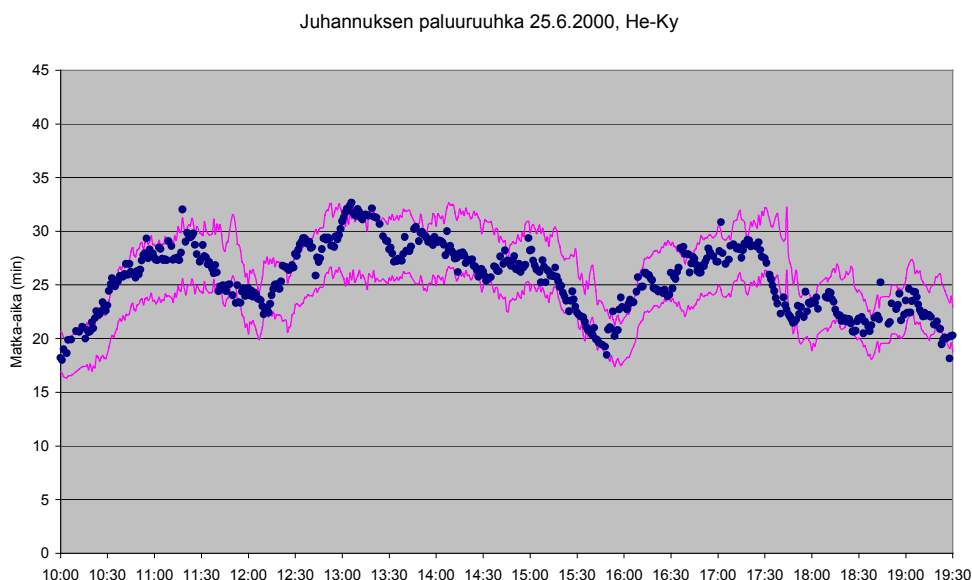
Taulukko 24. Kymijärvi–Heinola-suunnan eri ennustemallien kyky ennustaa ruuhkaliikennettä mitattuna oikein ennustettujen (o%), liian pieneksi arvioitujen (p%) ja liian suureksi arvioitujen (s%) matka-aikojen osuutena ajasta. Tummennetulla linkillä ei ollut ruuhkaa. Parhaat mallit on merkitty lihavoidulla reunuksella.

Aikasarjan pituus		Ky-He	Ky-Vi	Ky-Se	Se-He	Se-Vi	Vi-He
5 min	o%	69,5	64,1	37,7	76,5	62,8	
	p%	21,4	24,5	45,2	15,4	25,0	
	s%	9,1	11,4	17,1	8,2	12,3	
4 min	o%	68,9	64,2	36,6	74,7	62,1	
	p%	22,2	24,2	45,2	16,9	24,6	
	s%	8,9	11,6	18,2	8,3	13,2	
3 min	o%	70,8	63,0	39,3	75,3	60,5	
	p%	20,8	25,9	44,5	16,2	25,4	
	s%	8,4	11,0	16,3	8,5	14,1	

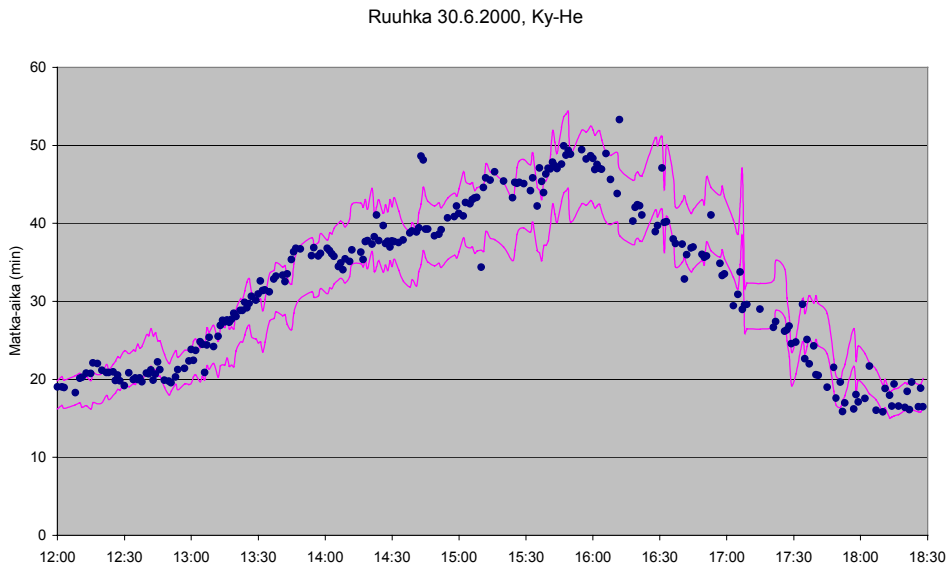
On huomattava, että liitteessä B esitetyistä malleista huonoiten ennustavat ovat täsmälleen samat mallit, joille ei ruuhkattomuuden vuoksi määrittää osuuksia taulukkoihin 23 ja 24. Ns. huonosti ennustavien mallien linkeillä ei siis esiinny varsinaista ruuhkaa (nopeus ei laske alle nopeusrajoituksen), vaan mitatuissa matka-ajoissa esiintyvä vaihtelu on satunnaisvaihtelua, jota malli ei toista.

Taulukoissa 21 ja 22 esitetyt koko aineistosta lasketut osuudet olivat samaa suuruusluokkaa kuin luvussa 5 tarkastellulla nykyjärjestelmällä. Ruuhkaliikenteessä nykyjärjestelmä antoi todellisuutta vastaavaa matka-aikatietoa Heinola–Kymijärvi-linkille 50 prosenttia ja Kymijärvi–Heinola-linkille 33 prosenttia ajasta. Ruuhkaliikenteessä neuromallit (taulukot 23 ja 24) siis olivat selvästi parempia kuin nykyjärjestelmä (Heinola–Kymijärvi 79 prosenttia ja Kymijärvi–Heinola 71 prosenttia).

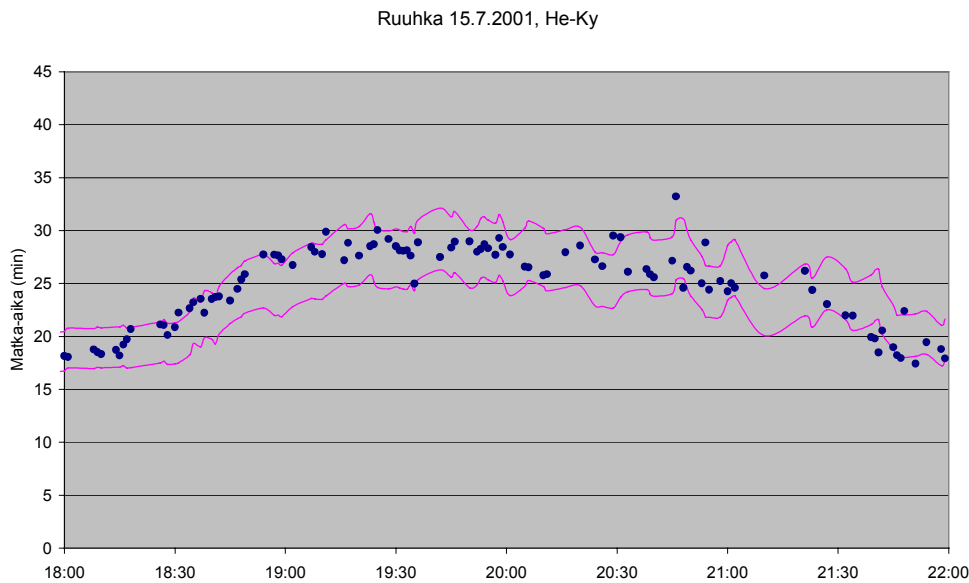
Esimerkit matka-aikaennusteista pahoin ruuhkaantuneina päivinä on esitetty kuvissa 16 ja 17. Molempien linkkien mallit kykenevät ennustamaan ruuhkan tyydyttävästi, Kymijärvi–Heinola-linkin malli hieman paremmin kuin Heinola–Kymijärvi-linkin malli. Vertailemalla kuvia 16 ja 17 kuviin 6 ja 7 näkee selvästi, että neuromalli kykenee ennakoimaan ruuhkan nykyjärjestelmää paremmin. Kuvissa 18 ja 19 on esimerkit neuromallilla ennustetuista tavallisista viikonloppuruuhkista. Kymijärvi–Heinola-linkin malli suoriutuu esimerkin lievemmästä ruuhkasta paremmin kuin juhannuksen pahasti ruuhkautuneesta tilanteesta.



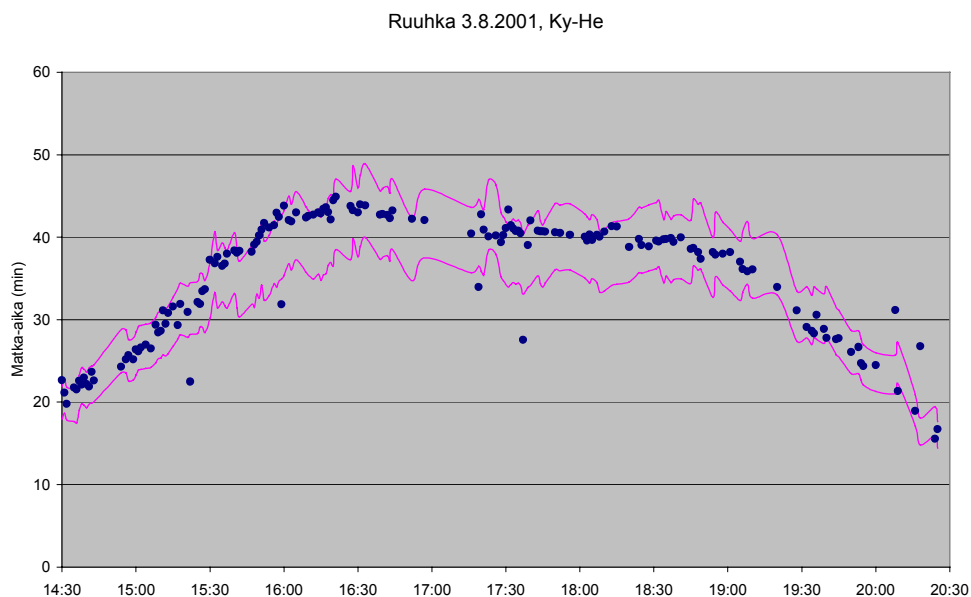
Kuva 16. Neuromallin ennustama matka-aika (ylä- ja alaraja) sekä mitatut arvot (pisteet) juhannuksen paluuruuhkassa 25.6.2000 Heinola–Kymijärvi-linkillä. Malli perustuu neljän minuutin aikasarjaan ja kaikkien linkkien matka-aikatietoihin.



Kuva 17. Neuromallin ennustama matka-aika (ylä- ja alaraja) sekä mitatut arvot (pisteet) perjantairuuhkassa 30.6.2000 Kymijärvi–Heinola-linkillä. Malli perustuu kolmen minuutin aikasarjaan ja kaikkien linkkien matka-aikatietoihin.



Kuva 18. Neuromallin ennustama matka-aika (ylä- ja alaraja) sekä mitatut arvot (pisteet) sunnuntairuuhkassa 15.7.2001 Heinola–Kymijärvi-linkillä. Malli perustuu neljän minuutin aikasarjaan ja kaikkien linkkien matka-aikatietoihin.



Kuva 19. Neuromallin ennustama matka-aika (ylä- ja alaraja) sekä mitatut arvot (pisteet) perjantairuuhkassa 3.8.2001 Kymijärvi–Heinola-linkillä. Malli perustuu kolmen minuutin aikasarjaan ja kaikkien linkkien matka-aikatietoihin.

6.2.2 Kehä I

Kehä I:lle tehtiin sekä estimaatti- että ennustemallit. Estimaattimalleilla pyrittiin estimoimaan tietyn linkin matka-aika muiden mittaustietojen perusteella tilanteessa, jossa kyseisen linkin matka-ajan mittaaminen ei onnistu. Malleissa estimoidaan tietyn linkin matka-aikaa saman ja edeltävien jaksojen muiden liikennetietojen perusteella. Ennustemallit olivat vastaavia kuin valtielle 4 tehdyt. Niissä pyrittiin ennustamaan tietylle linkille lähtevien ajoneuvojen keskimatka-aika viimeisten mitattujen suureiden perusteella. Ennustemalleissa olivat syötteinä tarkasteltavan linkin omat viimeksi mitatut matka-aikatiedot toisin kuin estimaattimalleissa, joissa nämä tiedot puuttuivat.

Neuroverkkomallien eri kerrosten neuronien määrät on esitetty taulukossa 25. Syötteiden määrä jäi malleilla pieneksi (estimaattimalleilla 4–13 ja ennustemalleilla 8–17) muuttujien huonon korrelaation takia. Havaintojoukko oli kaikilla malleilla suuri ja tämä johti suureen piiloneuronien maksimimäärään. Määrä rajoitettiin kuitenkin 20:een, jottei opetusprosessi muodostuisi liian raskaaksi.

Taulukko 25. Kehä I:n ennuste- ja estimaattimallien tekemisessä käytettävissä olleiden havaintojen määrä (N) sekä mallien syöte-, piilo- ja vastoneuronien määrät. Malleissa piiloneuronien maksimimäärä rajoitettiin 20:een (*).

		Ot-Ko	Ko-Pu	Pu-Ko	Ko-Ot
Estimaatti	N	17 367	26 297	14 022	13 534
	syöte	9	4	9	13
	piilo	78*	218*	63*	45*
	vaste	1	1	1	1
Ennuste	N	17 066	25 742	13 706	13 310
	syöte	13	8	13	17
	piilo	56*	128*	45*	34*
	vaste	1	1	1	1

Ennustemallit antoivat tilastollisesti tarkastellen selvästi parempia tuloksia kuin estimaattimallit (taulukot 26 ja 27). Keskimääräinen suhteellisen virheen itseisarvo oli estimaattimalleilla Otaniemen ja Konalan välillä noin kymmenen prosenttia ja Konalan ja Pukinmäen välillä viidestä kuuteen prosenttia. Ennustemalleilla vastaavat luvut olivat vajaat kahdeksan prosenttia sekä neljästä viiteen prosenttia.

Taulukko 26. Kehä I:n eri estimaattimallien hyvyys tilastollisilla virhetermeillä mitattuna.

	Ot-Ko	Ko-Pu	Pu-Ko	Ko-Ot
Keskimääräinen neliövirhe	3,0	0,3	0,8	1,9
Keskimääräinen virhe	0,0	0,0	0,0	0,0
Keskivirheen itseisarvo	0,8	0,3	0,4	0,7
Keskimääräinen suhteellinen virhe	2,5 %	0,7 %	1,0 %	2,2 %
Keskimääräisen suhteellisen virheen itseisarvo	10,4 %	5,1 %	6,1 %	9,9 %

Taulukko 27. Kehä I:n eri ennustemallien hyvyys tilastollisilla virhetermeillä mitattuna.

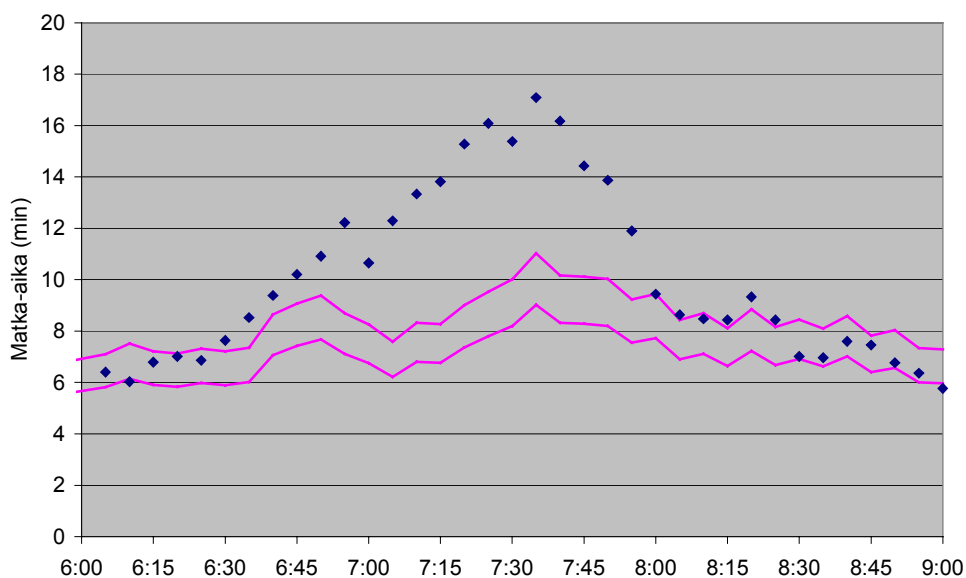
	Ot-Ko	Ko-Pu	Pu-Ko	Ko-Ot
Keskimääräinen neliövirhe	1,2	0,2	0,4	0,7
Keskimääräinen virhe	0,0	0,0	0,0	0,0
Keskivirheen itseisarvo	0,6	0,2	0,3	0,5
Keskimääräinen suhteellinen virhe	0,9 %	0,4 %	0,8 %	0,9 %
Keskimääräisen suhteellisen virheen itseisarvo	7,5 %	4,2 %	5,0 %	7,7 %

Kun mallien hyvyttä mitattiin oikein ennustettujen matka-aikojen osuutena, koko aineistoa tarkastellen Kehä I:n estimaattimallit pärjäsivät kohtalaisesti (oikein ennustettujen osuus oli noin 90 prosenttia, taulukko 28). Ruuhkavainoja tarkasteltaessa tulos on kuitenkin hyvin erilainen. Ainoastaan korkeintaan kolmannes ennusteista osui oikeaan. Tässä ruuhkarajana käytettiin Otaniemen ja Konalan välillä vähintään kahdeksan minuutin sekä Konalan ja Pukinmäen välillä vähintään kuuden minuutin matka-aikaa. Rajat arvioitiin aineistosta silmämääräisesti.

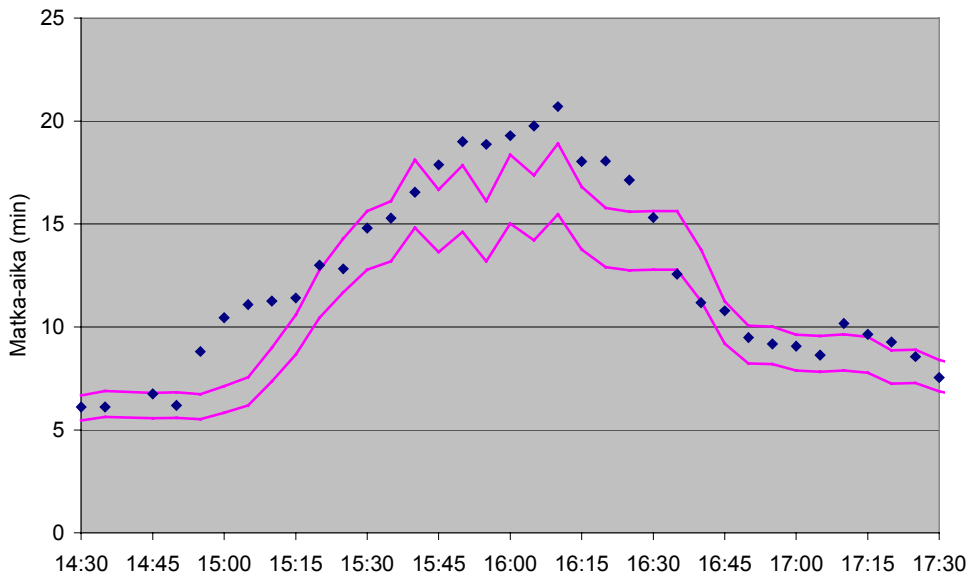
Taulukko 28. Eri estimaattimallien hyvyys mitattuna oikein arvioitujen (o%), liian pieneksi arvioitujen (p%) ja liian suureksi arvioitujen (s%) matka-aikojen osuutena ajasta koko aineistossa ja ruuhka-aineistossa.

		Ot-Ko	Ko-Pu	Pu-Ko	Ko-Ot
Kaikki	o%	88,0	94,7	91,7	90,0
	s%	5,6	2,1	4,5	3,6
	p%	6,4	3,3	3,8	6,4
Ruuhka	o%	30,5	34,3	25,8	30,1
	s%	14,3	3,0	9,4	6,4
	p%	55,3	62,7	64,8	63,5

Estimaattimallien ruuhka-ajan huonot tulokset on nähtävissä selvästi yksittäisiä ruuhkia tarkastelemalla. Osassa ruuhkista estimointimalli ei havainnut ruuhkaa lainkaan (kuva 20), välillä taas estimointimalli antoi kohtalaisen hyvän estimaatin matka-ajasta (kuva 21). Tähän on syynä se, että toisinaan ruuhka on paikallinen, eikä näy Konalan LAM-pisteellä eikä Pukinmäki-Konala-linkin matka-ajassa, joiden perusteella tarkasteltava matka-aika estimoidaan.



Kuva 20. Konala–Otaniemi-linkille 12.8.2000 aamuruuhkan aikaan mitatut (pisteet) ja estimoidut (ylä- ja alaraja) matka-ajat.



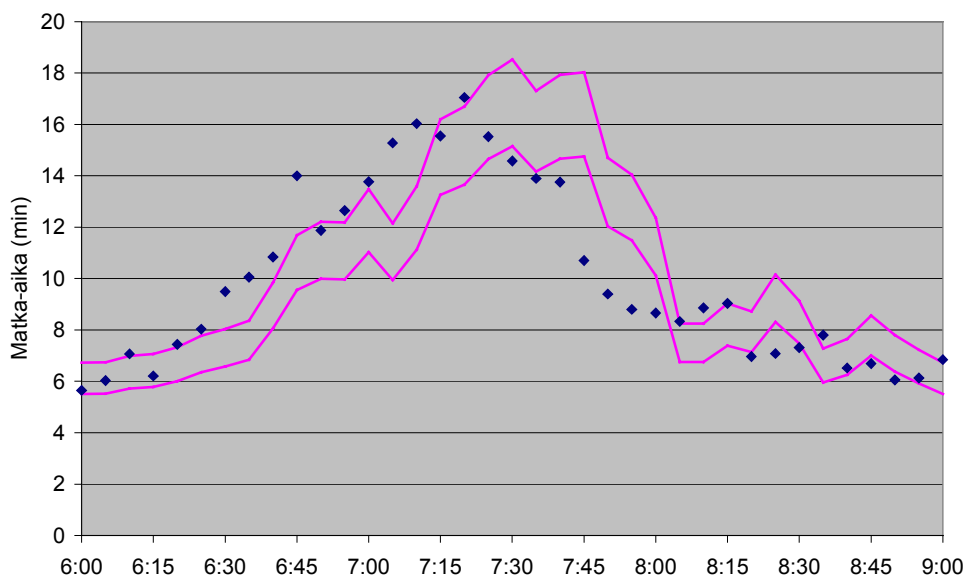
Kuva 21. Konala–Otaniemi-linkille 21.9.2000 iltaruuhkan aikaan mitatut (pisteet) ja estimoidut (ylä- ja alaraja) matka-ajat.

Ennustemallit olivat myös oikein ennustettujen osuudella mitattuna parempia kuin estimaattimallit (taulukot 28 ja 29). Kaikilla malleilla koko aineistosta määritetty oikein ennustettujen osuus oli yli 90 prosenttia. Ruuhka-aikanakin oikein ennustettujen osuus oli 40–50 prosenttia. Edellä esitetyt tulokset on nähtävissä myös liitteen B kuvaajista.

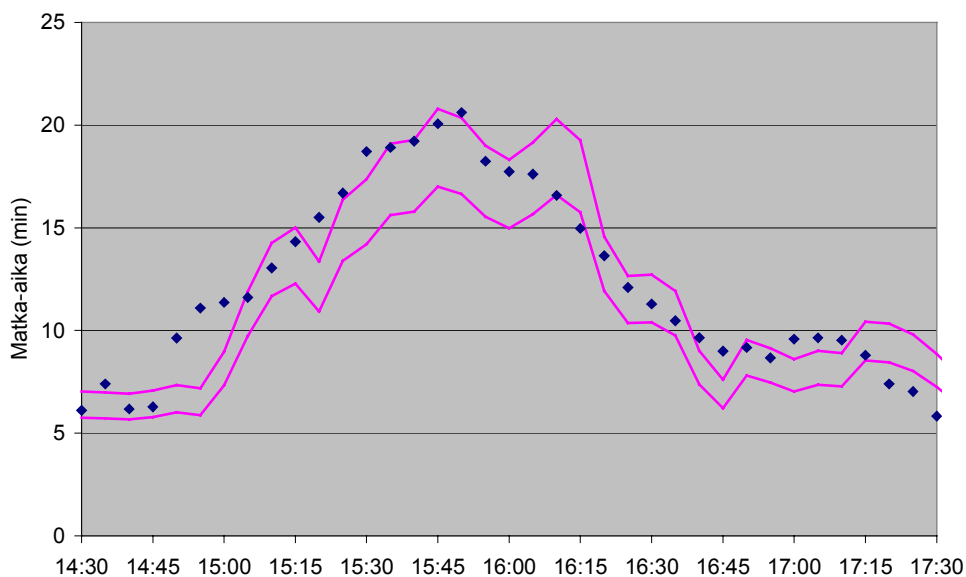
Taulukko 29. Eri ennustemallien hyvyys mitattuna oikein ennustettujen (o%), liian pieneksi arvioitujen (p%) ja liian suureksi arvioitujen (s%) matka-aikojen osuutena ajasta koko aineistossa ja ruuhka-aineistossa.

		Ot-Ko	Ko-Pu	Pu-Ko	Ko-Ot
Kaikki	o%	91,8	96,5	94,3	92,9
	s%	3,5	1,1	2,9	2,5
	p%	4,6	2,4	2,8	4,6
Ruuhka	o%	43,5	46,4	39,1	45,9
	s%	17,0	6,8	11,0	10,3
	p%	39,6	46,8	49,9	43,8

Kuvia 20 ja 21 vastaavat ruuhkat on esitetty ennustemallien osalta kuvissa 22 ja 23. Kuvista on havaittavissa viive, jolla ennuste tulee mitattujen arvojen perässä. Ruuhkan ajankohdan vaihesiirto ennuste- ja estimaattimallien välillä johtuu siitä, että estimaattimallit estimoivat tietyn viisiminuuttisen aikana linkin loppupisteen ohittaneiden ajoneuvojen keskimatka-ajan, kun taas ennustemallit ennustavat seuraavan viisiminuuttisen aikana linkille lähtevien ajoneuvojen keskimatka-ajan.



Kuva 22. Konala–Otaniemi-linkille 12.8.2000 aamuruuhkan aikaan mitatut (pisteet) ja ennustetut (ylä- ja alaraja) matka-ajat.



Kuva 23. Konala–Otaniemi-linkille 21.9.2000 iltaruuhkan aikaan mitatut (pisteet) ja ennustetut (ennustetut) matka-ajat.

Kehä I:llä ruuhka on varsin säännönmukaista, ja tästä syystä ennustemalleja yritettiin parantaa antamalla syöteinä ainoastaan suureen kulloinenkin erotus vastaavan ajankohdan keskiarvosuureesta. Tarkastelussa erotettiin perjantai, lauantai ja sunnuntai tai pyhäpäivä muista päivistä.

Tilastollisilla mittareilla määriteltynä ajankohtatietoja hyväksi käyttävän Otaniemi–Konala-mallin tulos (taulukko 30) oli selkeästi huonompi kuin mallin, jonka syötteissä ei ollut huomioitu ajankohtaa (taulukko 27). Muilla malleilla ero ei ollut yhtä suuri, mutta kuitenkin samansuuntainen.

Taulukko 30. Kehä I:n eri ajankohtatietoa hyväksi käyttävien ennustemallien hyvyys tilastollisina virhetermeillä mitattuna.

	Ot-Ko	Ko-Pu	Pu-Ko	Ko-Ot
Keskimääräinen neliövirhe	4,6	0,2	0,4	0,7
Keskimääräinen virhe	0,2	0,0	0,0	0,0
Keskivirheen itseisarvo	1,6	0,2	0,3	0,6
Keskimääräinen suhteellinen virhe	-0,2 %	0,4 %	0,3 %	0,8 %
Keskimääräisen suhteellisen virheen itseisarvo	22 %	4,4 %	5,6 %	7,8 %

Otaniemi–Konala-mallin tulokset (taulukko 31) olivat myös oikein ennustettujen osuudella mitattuna selkeästi huonompia kuin mallien, jotka eivät käyttäneet hyväkseen ajankohtatietoja (taulukko 29). Muilla malleilla tilanne oli samansuuntainen joskaan ei yhtä selvä, kun tarkastellaan koko aineistosta määritettyjä osuuksia. Ajankohtatietoja hyväksi käyttävillä malleilla ruuhkaliikenteessä määritetyt osuudet olivat kuitenkin parempia.

Taulukko 31. Kehä I:n eri ajankohtatietoja hyväksi käyttävien ennustemallien hyvyys mitattuna oikein ennustettujen (o%), liian pieneksi arvioitujen (p%) ja liian suureksi arvioitujen (s%) matka-aikojen osuutena ajasta koko aineistossa ja ruuhka-aineistossa.

		Ot-Ko	Ko-Pu	Pu-Ko	Ko-Ot
Kaikki	o%	71,9	96,1	92,3	92,9
	s%	24,5	1,7	5,4	3,4
	p%	3,6	2,3	2,4	3,8
Ruuhka	o%	31,9	52,7	42,3	54,6
	s%	31,1	3,8	15,6	9,4
	p%	33,3	43,5	42,2	36,0

Koska ruuhkaliikenteen ennusteet ovat tärkeämpiä kuin keskimääräiset ennusteet, voidaan muiden kuin Otaniemi–Konala-mallin osalta määritellä ajankohtatietoja hyväksi käyttävät mallit paremmiksi kuin ilman näitä tietoja tehtävät mallit.

Verrataan Kehä I:n mallien suoritusta valtatie 4 mallien suorituksiin (taulukot 23 ja 24). Valtatiellä 4 kaikilla linkeillä ei ollut riittävästi ruuhkaa, jotta oikein ennustettujen osuuksia olisi voitu määritellä ruuhka-aineistolle. Kehä I:n mallien kanssa vertailukelpoisia ovat valtatie 4 osalinkeistä peräkkäisten kamerapisteiden väliset mallit (Kymijärvi–Seesta, Seesta–Vierumäki ja Vierumäki–Heinola molempiin suuntiin).

Kehä I:llä mallit ennustivat 42–55 prosenttia oikein ruuhka-ajasta. Valtatiellä 4 vastaavan mittaisen linkkien malleista Kymijärvi–Seesta-malli oli hieman huonompi kuin Kehä I:n mallit: Kymijärvi–Seesta-malli oli ruuhka-aikaan oikeassa 39 prosenttia ajasta. Seesta–Vierumäki- ja Heinola–Vierumäki-mallit sitä vastoin olivat parempia kuin Kehä I:n mallit: ne olivat ruuhka-aikaan oikeassa 63–68 prosenttia ajasta. Valtatie 4 pidempien osalinkkien mallit ennustivat hieman paremmin. Ne olivat oikeassa 64–79 prosenttia ajasta.

7 SUMEAT MALLIT

7.1 Sumean ennustemallin rakentaminen

Neuroverkkomallin lisäksi muodostettiin sumeaan logiikkaan perustuva matka-ajan ennustamismalli valtatielle 4 Kymijärveltä Heinolaan. Tavoitteena oli tutustua sumean logiikan tarjoamiin mahdollisuuksiin lyhyen aikavälin matka-aikaennusteen tekemisessä sekä analysoida muodostetun mallin hyvyttä ja verrata saatuja tuloksia vastaavan neuroverkkomallin antamiin tuloksiin.

Ennuste perustui koko tieosuudelta ja sen kolmelta osalinkiltä viimeksi mitattuihin matka-aikatietoihin sekä yhdessä pisteessä mitattuihin nopeuksiin ja liikennemääriin. Mallin tehtävänä oli ennustaa linkille tarkasteluhetkeä seuraavan jakson aikana lähtevien ajoneuvojen keskimatka-aika. Aggregointijaksona käytettiin viittä minuuttia. Malli perustui kesän 2000 aikana kerättyyn aineistoon.

Liikennemäärä-, pistenopeus- ja matka-aikatiedoista muodostunut aineisto koottiin matriisiksi, jonka koko oli noin $6\,500 \times 6$ (taulukko 32). Kukin muuttuja oli omassa sarakkeessaan ja vaakariveinä olivat saman minuutin jakson aikana saatujen havaintojen keskiarvot. Koska kaikilta viiden minuutin jaksoilta ei jokaiselle muuttujalle ollut saatu arvoa, poistettiin matriisista ne rivit, joissa vähintään 3 muuttujasta oli saanut arvon 0. Sen jälkeen korvattiin jäljelle jääneet nollahavainnot niitä edeltäneen viiden minuutin jakson vastaavalla arvolla, mikäli viimeisestä todellisesta havainnosta ei ollut kulunut kauempaa kuin 30 minuuttia. Aineistojoukko pieneni näin noin 4 500 riviin. Lisäksi osalinkkien matka-ajat järjestettiin suuruusjärjestykseen, mikä vähensi tarvittavien sumeiden päättelysääntöjen lukumäärää merkittävästi. Jatkossa muuttujiin viitataan taulukon 32 mukaisesti sanoilla syötemuuttuja ja tulostuuttuja.

Taulukko 32. Esimerkki matriisiksi kootusta aineistosta.

Liikenne- määrä (ajon./5 min)	Piste- nopeus (km/h)	Syötemuuttujat			Linkiltä pois- tuneiden ajon. m.aika (min)	Tulosmuuttuja Linkille lähtevien ajoneuvojen matka-aika (min)
		Osalinkkien matka-ajat suuruusjärjestyksessä (min) pisin	2. pisin	lyhin		
73	59,7	26,1	20,6	6,6	48,3	53,4
95	62,5	27,2	21,2	6,7	48,9	54,9
69	63,3	24,5	21,5	6,7	50,3	55,7
75	60,4	23,4	21,5	6,9	51,3	55,2
85	57,0	24,1	20,6	7,1	52,0	55,2
83	59,5	24,4	18,7	6,7	52,3	56,1
79	54,0	24,7	18,4	6,9	53,1	56,6
84	55,8	25,2	20,3	7,0	50,3	56,3
86	51,7	24,3	21,0	7,1	50,2	55,7
88	56,9	23,2	22,2	6,7	50,6	52,4

Mallin rakentaminen aloitettiin tutkimalla aineistoa Microsoft Excel -taulukko-laskentaohjelmalla. Muuttujien saamat arvot jaettiin luokkiin sen perusteella, miten edelliseltä viiden minuutin jaksolta saadut mittaustulokset näyttivät vaikuttavan seuraavan viiden minuutin aikana matkaan lähteneiden ajoneuvojen matka-aikaan. Luokkia muokkaamalla saatiin jokaiselle muuttujalle sarja kolmion muotoisia jäsenyysfunktioita, ja niille annettiin sanalliset kuvaukset. Graafinen esitys luokista ja jäsenyysfunktioista on koottu liitteeseen C.

Ennusteen haluttiin osuvan luokkiin, jotka valittiin seuraavasti: alle 20 min, 20–25 min, 25–30 min, 30–40 min, 40–50 min ja yli 50 min. Normaali matka-aika tiejaksolla on noin 20 minuuttia. Pienin informaatiotaululla näytettävä matka-ajan ennuste on 20–25 minuuttia, joka vastaa nopeusrajoitusta. Luokittelua vastaavat havaintomäärät on esitetty taulukossa 33.

Taulukko 33. Ennusteiden luokittelua vastaavat havaintomäärät.

	Ennusteluokat (min)						Yhteensä
	< 20	20–25	25–30	30–40	40–50	> 50	
Havaintomäärä	4 226	200	19	43	39	14	4 541
Osuus	93,1 %	4,4 %	0,4 %	0,9 %	0,9 %	0,3 %	100,0 %

GUHA (General Unary Hypotheses Automaton) on perinteiseen kaksiarvoiseen logiikkaan perustuva menetelmä, jonka tarkoituksena on generoida empiirisestä aineistosta yleisluontoisia hypoteeseja eli löytää mahdollisia muuttujien välisiä yhteyksiä ja lakeja. (Hájek 1993.)

GUHA prosessoi aineistosta koottua matriisia, joka edustaa kaksiarvoisten (*dichotomous*) muuttujien käyttäytymistä. Mikäli muuttujalla siis on tietyssä tapauksessa (tietyllä rivillä) jokin ominaisuus, vastaava matriisin alkio saa arvon 1. Muussa tapauksessa sen arvo on 0. Käyttäjän tehtäväksi jää reaaliarvoisten muuttujien luokittelu, jonka perusteella GUHA päättää, kuuluuko muuttuja tiettyyn luokkaan (saa arvon 1) vai ei (saa arvon 0). Sen jälkeen GUHA generoi hypoteesit, jotka ovat muotoa $A \sim S$ ("A liittyy S:ään"), jossa A (*antecedent*) on jokin tiettyyn syötemuuttujaan liittyvä luokka ja S (*succedent*) on seuraus, tulostuuttujan luokka. Jokaista hypoteesia $A \sim S$ arvioidaan taulukon 34 avulla. (Hájek 1993.)

Taulukko 34. Hypoteesin arviointi (Hájek 1993).

Muuttuja	S	ei S	Yhteensä
A	a	b	a + b
ei A	c	d	c + d
Yhteensä	a + c	b + d	n

Taulukossa 34 a, b, c ja d ovat havaitut frekvenssit. Esimerkiksi b on sellaisten havaintojen lukumäärä, joissa tietty syötemuuttuja kuuluu luokkaan A, mutta joiden tulostuuttuja ei kuulu luokkaan S, ja $n = a + b + c + d$. Hypoteesi hyväksytään tai hylätään päätösfunktion $F(a, b, c, d)$ perusteella. $F(a, b, c, d)$ voi saada arvon 1 tai 0. Jos $F(a, b, c, d) = 1$, voidaan todeta, että aineistoon perustuen $A \sim S$ on tosi. (Hájek 1993.)

Muuttujat luokiteltiin yksikäsitteisiin, jäsenyysfunktioita mukaileviin luokkiin. Koska luokittelu jouduttiin perustamaan silmämääräiseen arviointiin, muodostettiin kolme luokitteluvaihtoehtoa, joita kaikkia tutkittiin GUHA:n avulla. Tulosten toivottiin auttavan sumeiden päättelysääntöjen muodostamisessa. Luokitteluvaihtoehdot ja GUHA-analyysien tulokset ovat taulukossa 35.

Taulukko 35. GUHA-analyysissä käytetty luokittelu.

	Liikennemäärä	Pistenopeus	Osalinkkien matka-ajat	Koko linkin matka-aika	Ennuste
Sarja 1	0–37 38–51 > 51	< 85 85–90,5 90,5–120	0–7 7–12 12–18 > 18	0–20 20–25 25–30 30–40 40–50 > 50	0–20 20–25 25–30 30–40 40–50 > 50
Sarja 2	0–39 40–53 > 53	< 82 82–91 91–120	0–7,5 7,5–13 13–19 > 19		
Sarja 3	0–35 36–49 > 49	< 87 87–90 90–120	0–6,5 6,5–11 11–17 > 17		

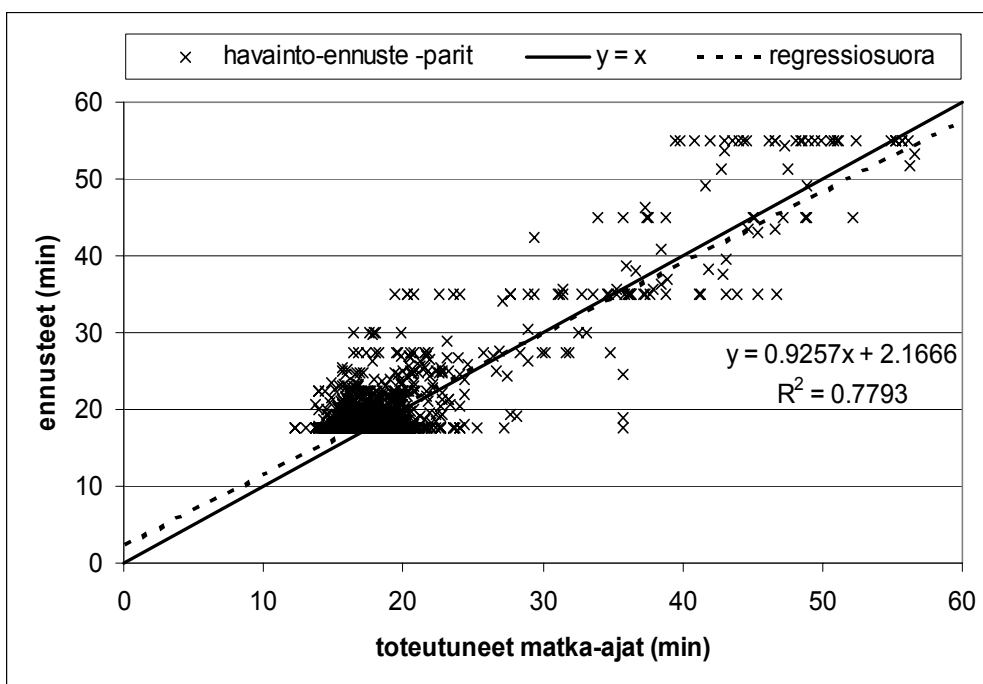
GUHA-analyysin tuloksena saadaan todennäköisyys $P(S|A)$ eli tässä tapauksessa vastaus kysymykseen: ”Jos muuttujan x (esim. liikennemäärä) arvo kuuluu luokkaan A , niin millä todennäköisyydellä seuraavan viiden minuutin aikana matkaan lähtevien ajoneuvojen matka-aika osuu luokkaan S ?”. Tulokset yhdistävät yhden tai korkeintaan kaksi syötemuuttujaa yhteen tulosmuuttujaan Ensimmäisten ajojen yhteydessä todennäköisyyksien minimiarvoksi oli määritelty 0,90, mistä johtuen saatiin vain luokkaan ”alle 20 min” liittyviä hypoteeseja, sillä valtaosa havainnoista kuului tähän luokkaan (taulukko 33). Toisella ajokerralla minimiarvoksi asetettiin 0,20, jolloin saatiin myös muihin ennusteluokkiin liittyviä hypoteeseja (taulukko 35). Kolmannen ajon yhteydessä aineistosta poistettiin havainnot, joissa ennuste osui luokkaan ”alle 20 min”, jolloin matriisin rivien lukumäärä supistui noin 300:aan. Minimiarvona tässä ajossa oli 0,6, jolloin muihin ennusteluokkiin liittyvät hypoteesit vahvistuivat. Sumeat päättelysäännöt muodostettiin GUHA-menetelmällä saatuja hypoteeseja apuna käyttäen (liite C).

7.2 Sumea ennustemalli ja sen hyvyys

Mallia testattiin käyttäen Matlab-ohjelmiston sumeaa Fuzzy Logic Toolbox -työkalua. Syötemuuttujina käytettiin kahta pisintä osalinkin matka-aikaa sekä koko tieosuuden matka-aikaa, sillä muiden muuttujien (pistenopeus, liikennemäärä, lyhin osalinkin matka-aika) arvoilla ei GUHA-analyysin perusteella juuri ollut vaikutusta lähtevien ajoneuvojen matka-aikaan (taulukko 35). Tämän voidaan katsoa johtuvan siitä, että yhdessä pisteessä mitatut liikennemäärä- ja pistenopeustiedot eivät anna oikeaa kuvaa koko tiejakson liikennetilanteesta. Lisäksi viimeisen osalinkin (Vierumäki–Heinola E) matka-aika pysyi normaalilla tasolla koko mittausjakson ajan muiden osalinkkien liikennetilanteesta riippumatta, joten se ei vaikuttanut matka-ajan ennustamiseen mitenkään.

Vertailun vuoksi tehtiin kaksi ajoa, joista ensimmäisessä malli toimi Gödelin algebran mukaisesti. Toisessa mallin toiminnan perustana oli sumea kokonaissimilaarisuus. Täsmällistäminen tehtiin laskemalla Gödel-mallissa päätelyn tuloksena saadun sumean pinta-alan painopiste (COG) ja similaarisessa mallissa maksimien keskiarvo (MOM). Käytännössä täsmällistämismenetelmien erilaisuus ei vaikuta tulosten vertailukelpoisuuteen, sillä jäsenyysfunktioiden symmetrisyyden (tasakylkisiä kolmioita) takia similaarisuuden perustuvan päätelyn tuloksena saadun pinta-alan painopiste sijoittuu aina samaan kohtaan kuin maksimien keskiarvo.

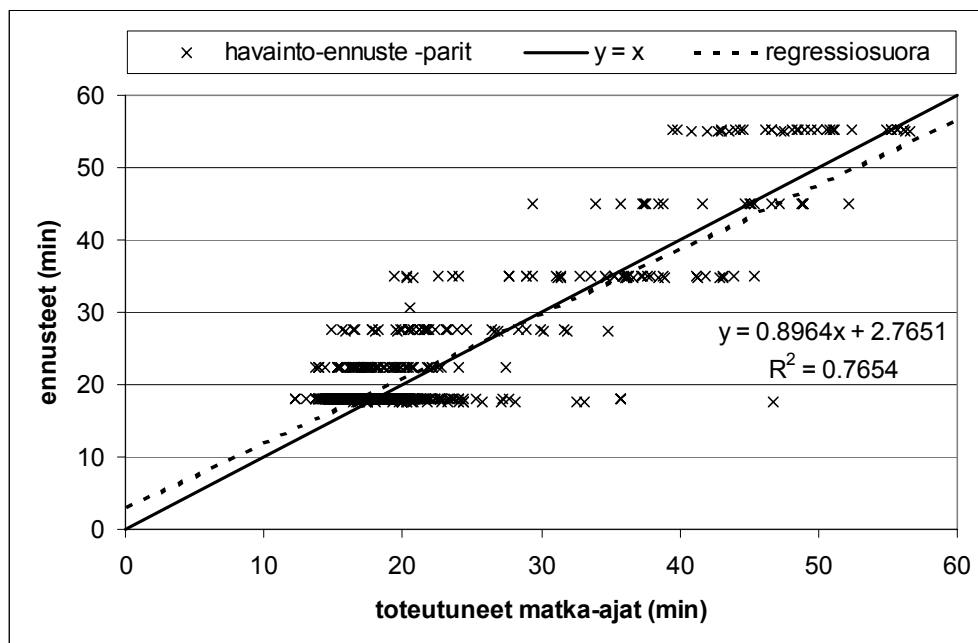
Ajojen tulokset näkyvät kuvissa 24 ja 25 sekä taulukoissa 36 ja 37. Kuvista ja taulukoista nähdään, että eri menetelmillä saadut tulokset poikkeavat toisistaan vain vähän. Gödel-mallin antamat tulokset sijoittuvat hieman paremmin suoralle $y = x$ kuin similaarisen mallin. Myös korrelaatiokerroin on Gödel-mallissa aavistuksen suurempi kuin similaarisessa mallissa. Kuvassa 26 ja taulukossa 38 ovat kesällä 2000 käytössä olleen (ns. nykyjärjestelmä), osalinkkien matka-aikojen kymmenen minuutin mediaanien summaan perustuvan estimoinnin antamat matka-aikatiedot ja niiden paikkansa pitävyys. Kuvasta 26 nähdään, että nykyjärjestelmän estimaatit osuvat sumeita ennusteita huonommin suoralle $y = x$ ja korrelaatiokerroin on pienempi kuin sumeilla ennusteilla.



Kuva 24. Toteutuneet ja ennustetut matka-ajat sumean ennustemallin toimiesä Gödelin algebran periaatteiden mukaisesti.

Taulukko 36. Tulokset sumean matka-ajan ennustemallin toimiessa Gödelin algebran periaatteiden mukaisesti.

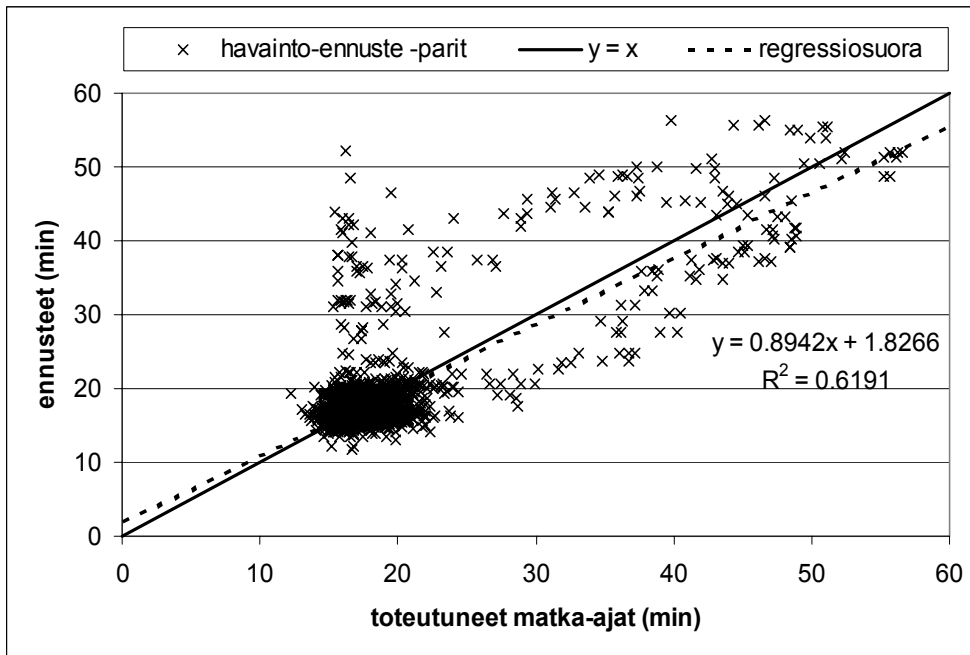
Mitattu matka-aika	Ennuste (min)						Rivin summa
	< 20	20–25	25–30	30–40	40–50	> 50	
< 20 min	4 045	163	12	6	-	-	4 226
20–25 min	132	40	22	6	-	-	200
25–30 min	4	1	7	6	1	-	19
30–40 min	2	1	4	27	7	2	43
40–50 min	-	-	-	9	10	20	39
> 50 min	-	-	-	-	1	13	14
Sarakkeen summa	4 183	205	45	54	19	35	4 541



Kuva 25. Toteutuneet ja ennustetut matka-ajat sumean ennustemallin toimies-
sa sumean similaarisuuden periaatteiden mukaisesti.

Taulukko 37. Tulokset sumean matka-ajan ennustemallin toimies-
sa sumean similaarisuuden periaatteiden mukaisesti.

Mitattu matka-aika	Ennuste (min)						Rivin summa
	< 20	20–25	25–30	30–40	40–50	> 50	
< 20 min	4 159	54	12	1	-	-	4 226
20–25 min	159	14	20	7	-	-	200
25–30 min	6	1	6	5	1	-	19
30–40 min	5	-	4	25	7	2	43
40–50 min	1	-	-	8	10	20	39
> 50 min	-	-	-	-	1	13	14
Sarakkeen summa	4 330	69	42	46	19	35	4 541



Kuva 26. Nykyjärjestelmän mukainen, osalinkkien matka-aikojen summaan perustuva matka-aikatieto ja toteutuneet matka-ajat.

Taulukko 38. Nykyjärjestelmän mukainen, osalinkkien matka-aikojen summaan perustuva matka-aikatieto.

Mitattu matka-aika	Järjestelmän antama (min)						Rivin summa
	< 20	20–25	25–30	30–40	40–50	> 50	
< 20 min	3 569	81	8	31	10	1	3 700
20–25 min	141	38	1	9	2	-	191
25–30 min	4	6	-	3	5	-	18
30–40 min	-	10	5	9	17	3	44
40–50 min	-	-	1	19	24	8	52
> 50 min	-	-	-	-	2	11	13
Sarakkeen summa	3 714	135	15	71	60	23	4 018

Similaarisen mallin toiminnan ja käytetyn täsmällistämismenetelmän takia ennusteet asettuvat luokkien keskikohtiin. Similaarinen malli käyttää ennusteen muodostamiseen vain yhtä sääntöä eli yhtä tulosmuuttujaa. Koska tulosmuuttujien jäsenyysfunktiot ovat symmetrisiä, sijoittuu maksimien keskiarvo jäsenyysasteesta riippumatta aina samaan (jäsenyysastetta 1 vastaavaan) kohtaan kyseisen tulosmuuttujan tullessa valituksi.

Taulukoissa 39 ja 40 on esitetty mallien antamien ennusteiden oikeellisuus prosentteina. Yleisesti ennusteet näyttävät toteutuvan parhaiten luokissa ”alle 20 min” ja ”yli 50 min”. Molemmassa luokissa yli 90 prosenttia ennusteista osuu oikeaan. Myös matka-ajoista, jotka kuuluvat luokkaan ”30–40 min” ennustetaan 60 prosenttia oikein. Luokissa ”20–25 min” ja ”25–30 min” ennusteet näyttävät hajoavan eniten. Taulukossa 41 ovat nykyjärjestelmän antamat estimaatit.

Taulukko 39. Ennusteiden toteutuminen prosentteina Gödelin algebran mukaisessa mallissa.

Mitattu matka-aika	Ennuste (min)					
	< 20	20–25	25–30	30–40	40–50	> 50
< 20 min	95,7 %	3,9 %	0,3 %	0,1 %	-	-
20–25 min	66,0 %	20,0 %	11,0 %	3,0 %	-	-
25–30 min	21,1 %	5,3 %	36,8 %	31,6 %	5,3 %	-
30–40 min	4,7 %	2,3 %	9,3 %	62,8 %	16,3 %	4,7 %
40–50 min	-	-	-	23,1 %	25,6 %	51,3 %
> 50 min	-	-	-	-	7,1 %	92,9 %
Osuus kaikista ennusteista	92,1 %	4,5 %	1,0 %	1,2 %	0,4 %	0,8 %

Taulukko 40. Ennusteiden toteutuminen prosentteina similaarisessa mallissa.

Mitattu matka-aika	Ennuste (min)					
	< 20	20–25	25–30	30–40	40–50	> 50
< 20 min	98,4 %	1,3 %	0,3 %	0,0 %	-	-
20–25 min	79,5 %	7,0 %	10,0 %	3,5 %	-	-
25–30 min	31,6 %	5,3 %	31,6 %	26,3 %	5,3 %	-
30–40 min	11,6 %	-	9,3 %	58,1 %	16,3 %	4,7 %
40–50 min	2,6 %	-	-	20,5 %	25,6 %	51,3 %
> 50 min	-	-	-	-	7,1 %	92,9 %
Osuus kaikista ennusteista	95,4 %	1,5 %	0,9 %	1,0 %	0,4 %	0,8 %

Taulukko 41. Nykyjärjestelmän mukaisen matka-aikatiedon toteutuminen.

Mitattu matka-aika	Ennuste (min)					
	< 20	20–25	25–30	30–40	40–50	> 50
< 20 min	96,5 %	2,2 %	0,2 %	0,8 %	0,3 %	0,0 %
20–25 min	73,8 %	19,9 %	0,5 %	4,7 %	1,0 %	-
25–30 min	22,2 %	33,3 %	-	16,7 %	27,8 %	-
30–40 min	-	22,7 %	11,4 %	20,5 %	38,6 %	6,8 %
40–50 min	-	-	1,9 %	36,5 %	46,2 %	15,4 %
> 50 min	-	-	-	-	15,4 %	84,6 %
Osuus kaikista ennusteista	92,4 %	3,4 %	0,4 %	1,8 %	1,5 %	0,6 %

Similaarinen malli näyttää ennustavan paremmin luokkaa ”alle 20 min” kuin Gödel-malli, joka puolestaan ennustaa paremmin luokkia ”20–25 min”, ”25–30 min” ja ”30–40 min”. Luokat ”40–50 min” ja ”yli 50 min” molemmat mallit ennustavat yhtä hyvin. Nykyjärjestelmä estimoii sumeita malleja paremmin luokkaan ”40–50 min” kuuluvia matka-aikoja. Muissa luokissa nykyjärjestelmän antama estimaatti toteutuu sumeiden mallien ennusteita huonommin tai lähes yhtä hyvin.

Neuromallin ja sumean mallin hyvyttä vertaillaan taulukossa 42 tilastollisten tunnuslukujen avulla. Keskimääräinen ja suhteellinen virhe kertovat, onko mallilla taipumus yli- tai aliarvioida ennusteita. Keskimääräinen neliövirhe sekä virheen ja suhteellisen virheen itseisarvot kuvaavat virheiden jakautumista. Taulukosta voidaan nähdä, että kaikki sumean mallin tuloksille laskeut virheet ovat neuromallin vastaavia virheitä suurempia. Tässä tarkastelussa Gödel-malli osoittautui jonkin verran similaarista mallia paremmaksi.

Taulukko 42. Sumean mallin ja neuromallin hyvyys keskimääräisellä neliövirheellä (MSE), virheellä (ME), virheen itseisarvolla (MAE), suhteellisella virheellä (MRE) ja suhteellisen virheen itseisarvolla (MARE) määritettyinä.

	MSE	ME	MAE	MRE	MARE
Sumea (Gödel)	4,8	0,8	1,5	5,3 %	8,4 %
Sumea (simil.)	5,1	0,9	1,6	5,9 %	8,8 %
Neuromalli	2,6	0,0	1,2	0,6 %	6,4 %

Mallien hyvyttä mitattiin myös matka-aikatiedotuksen näkökulmasta. Ennusteille määritettiin kymmenen prosentin marginaali ja tutkittiin, montako prosenttia vastaavista havainnoista osui tähän marginaaliin. Laskennassa otettiin kuitenkin huomioon se, että pienimmän tienkäyttäjille tiedotustaululla näytettävän, nopeusrajoitusta vastaavan ennusteen alaraja oli 20 minuuttia. Jos siis mitattu matka-aika ja ennuste olivat alle 20 minuuttia, tulkittiin informaatio oikeaksi. Koska tasaista ruuhkatonta liikennettä on suhteellisen helppo ennustaa, tehtiin sama tarkastelu erikseen myös ruuhkaliikenteelle. Ruuhkaliikenteeseen katsottiin kuuluvan miniminäytön ylärajaa (25 min) suuremmat havainnot. Analyysin tulokset näkyvät taulukossa 43.

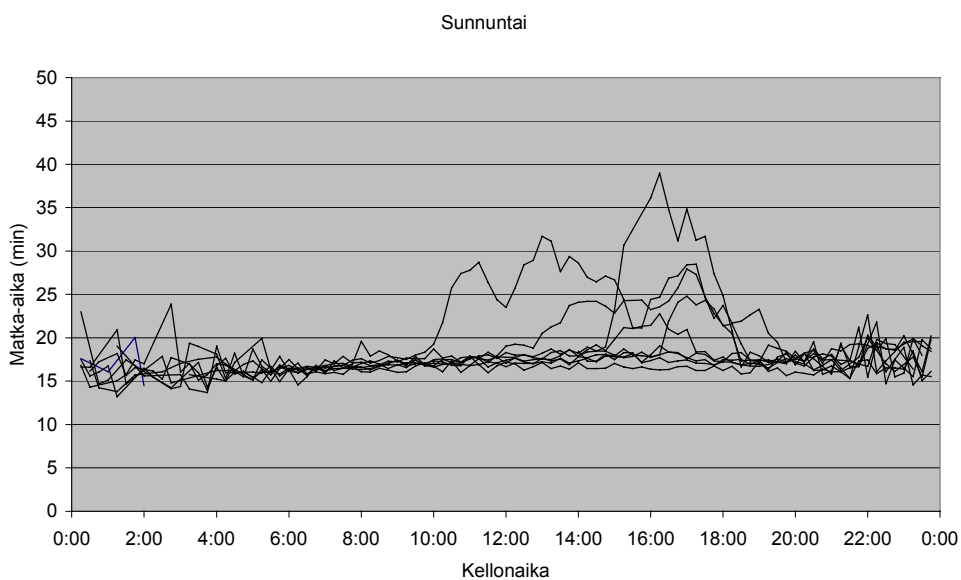
Taulukko 43. Mallien hyvyys matka-aikatiedotuksen näkökulmasta. Oikein ennustettujen (o%), liian suuriksi arvioitujen (s%) ja liian pieniksi arvioitujen (p%) matka-aikojen prosenttiosuudet.

	Kaikki havainnot			Ruuhkat		
	o%	s%	p%	o%	s%	p%
Sumea (Gödel)	93,6	3,2	3,2	49,6	32,2	18,3
Sumea (simil.)	93,0	2,9	4,1	42,6	32,2	25,2
Neuromalli	98,4	0,5	1,2	65,6	5,4	29,0
Nykyjärjestelmä	95,4	2,4	2,2	32,9	49,3	29,4

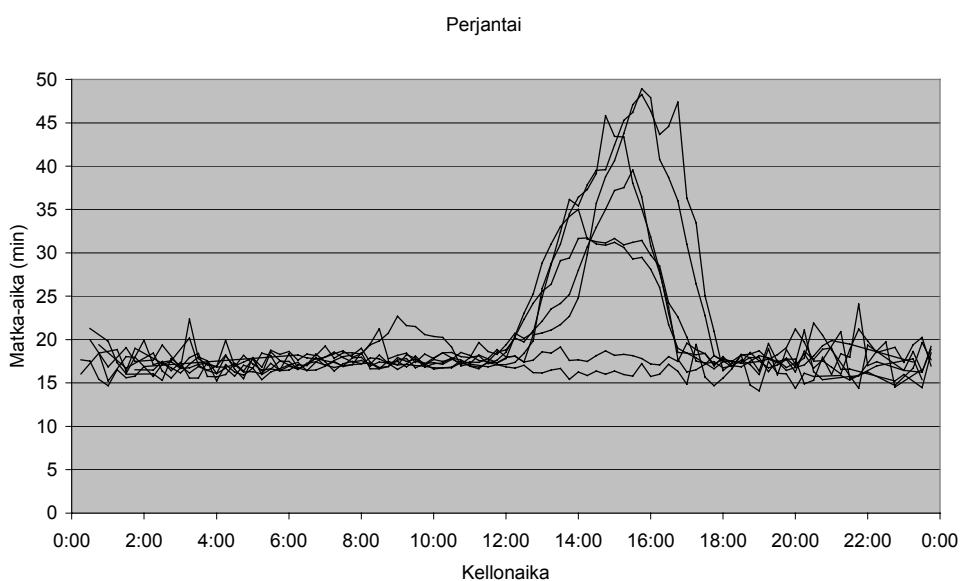
Taulukon 43 perusteella voidaan todeta, että neuromalli toimi sumeaa mallia paremmin myös matka-aikatiedotuksen näkökulmasta tarkasteltuna. Sumeita malleja keskenään vertailtaessa nähdään, että Gödelin algebran mukainen malli on myös matka-aikatiedotuksen näkökulmasta hiukan similaarista parempi. Nykyjärjestelmä estimoii ruuhkaliikennettä sumeita malleja huonommin. Neuromalli osoittautuu tässäkin tarkastelussa muita järjestelmiä paremmaksi.

8 RUUHKAPÄIVÄMALLIT

Valtatien 4 molempien suuntien matka-aikoja tarkasteltiin viikonpäivien funktiona kesän 2000 aineistosta. Heinola–Kymijärvi-suunnalle muutamaa satunnaisista ruuhkautumista lukuun ottamatta ainoa päivä, jolloin mitattiin pitkiä matka-aikoja, oli sunnuntai (kuva 27). Kymijärvi–Heinola-suunnalla ruuhkat olivat pahempia ja ruuhkautumista tapahtui enemmän, mutta sielläkin ainoastaan perjantaisin ruuhkautuminen oli säännöllistä (kuva 28).



Kuva 27. *Heinola–Kymijärvi-linkin keskimääräiset matka-ajat sunnuntaisin kesällä 2000.*



Kuva 28. *Kymijärvi–Heinola-linkin keskimääräiset matka-ajat perjantaisin kesällä 2000.*

Heinola–Kymijärvi- ja Kymijärvi–Heinola-linkeille tehtiin ruuhkamallit kokoomalla opetusjoukko ainoastaan niiden päivien aineistoista, jolloin matka-aika oli noussut päiväsaikaan jossain vaiheessa selvästi yli 20 minuuttiin. Opetusjoukkoon koottiin klo 9:00–24:00 väliset havainnot. Mikäli matka-aika kuitenkin oli kohonnut jo ennen klo 9:00, mukaan otettiin havainnot ruuhkan alusta lähtien. Tällaisia päiviä oli muutama.

Heinola–Kymijärvi-linkillä ruuhkapäiviä oli yhteensä kahdeksan. Kymijärvi–Heinola-linkillä ruuhkapäiviä oli enemmän eli 23 kappaletta. Havaintojen määrät ja näiden avulla määritetyt neuronimäärät on esitetty taulukossa 44. Kaikissa malleissa piiloneuronien määrä jäi selvästi pienemmäksi kuin kaikkien päivien malleissa (taulukot 17 ja 18).

Taulukko 44. Heinola–Kymijärvi- ja Kymijärvi–Heinola-linkkien ruuhkamallien neuronien ja havaintojen määrät.

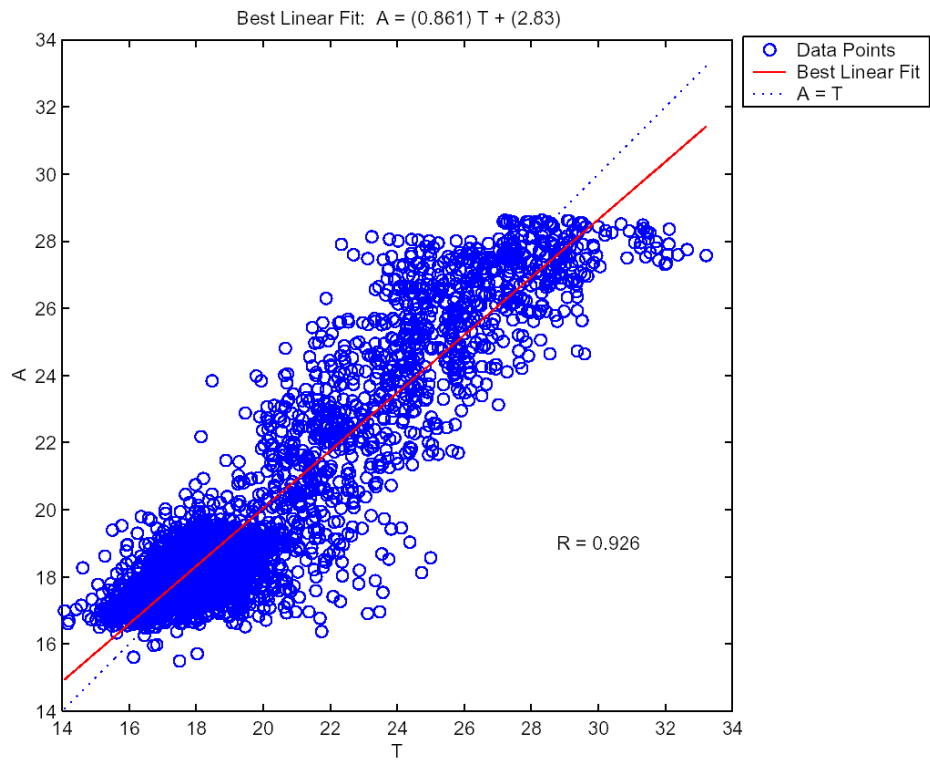
	Heinola–Kymijärvi			Kymijärvi–Heinola		
	5 min	4 min	3 min	5 min	4 min	3 min
Aikasarja	5 min	4 min	3 min	5 min	4 min	3 min
Havaintoja	3 064	3 070	3 074	7 042	7 051	7 053
Syötöneuroneja	45	38	31	54	45	36
Piiloneuroneja	3	4	5	6	8	10
Vasteneuroneja	1	1	1	1	1	1

Taulukkoon 45 on koottu molempien suuntien mallien tulokset ruuhka-ajan liikenteelle määritettynä. Tässä ruuhka-ajalla tarkoitetaan havaintoja, joissa mitattu matka-aika oli suurempi kuin muuttuvan opasteen näyttämän matkajan minimiylärajaa vastaavaa matkanopeutta vastaava matka-aika.

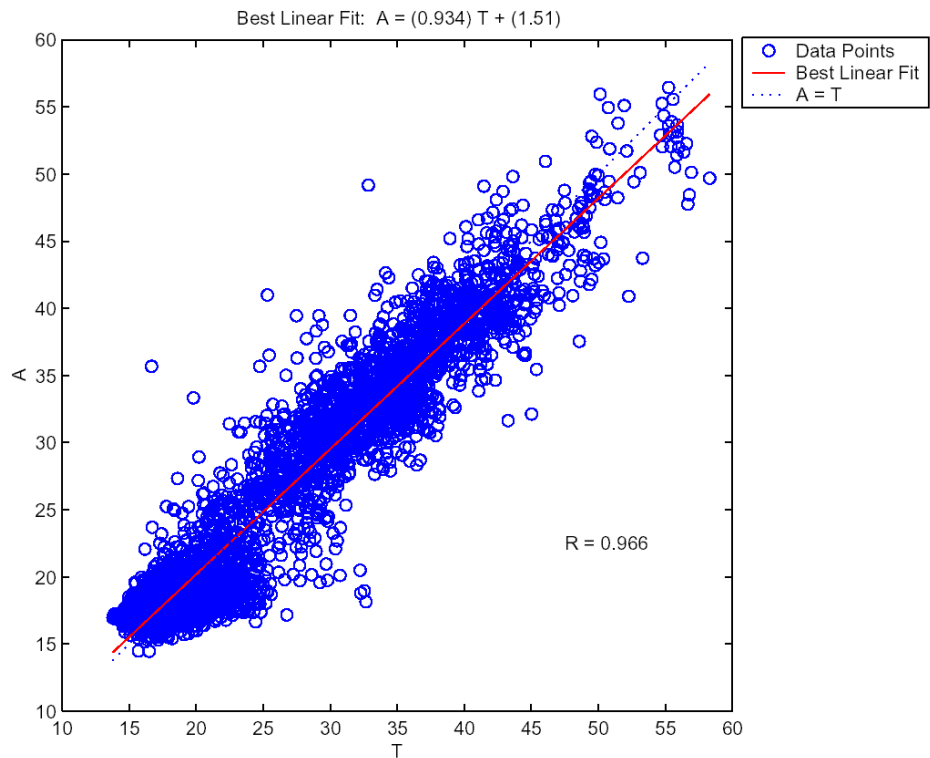
Taulukko 45. Osuudet ajasta, jotka Heinola–Kymijärvi- ja Kymijärvi–Heinola-linkkien ruuhkamallit olisivat ennustaneet ruuhka-ajan matka-ajan oikein, liian pieneksi tai liian suureksi. Parhaat mallit on merkitty.

	Heinola–Kymijärvi			Kymijärvi–Heinola		
	5 min	4 min	3 min	5 min	4 min	3 min
Aikasarjan pituus	5 min	4 min	3 min	5 min	4 min	3 min
Oikein	80,7 %	83,8 %	81,4 %	70,9 %	73,4 %	72,6 %
Liian suuri	3,2 %	3,6 %	3,7 %	9,6 %	9,4 %	9,5 %
Liian pieni	16,1 %	12,6 %	14,8 %	19,5 %	17,1 %	17,9 %

Heinola–Kymijärvi-linkillä parhaaksi osoittautui malli, joka perustui neljän minuutin aikasarjaan. Tällä mallilla 83,8 prosenttia ruuhka-ajan ennusteista olisi osoittautunut oikeiksi. Myös Kymijärvi–Heinola-linkillä paras tulos saatiin neljän minuutin aikasarjalla. Tällöin 73,4 prosenttia ruuhka-ajan ennusteista olisi osoittautunut oikeiksi. Molemmilla linkeillä päästiin siis parempaan tulokseen kuin parhaalla koko päivän aineistolla opetetulla mallilla (taulukot 23 ja 24). Ruuhkamallien avulla ennustetut matka-ajat on esitetty mitatun matka-ajan funktiona kuvissa 29 ja 30.



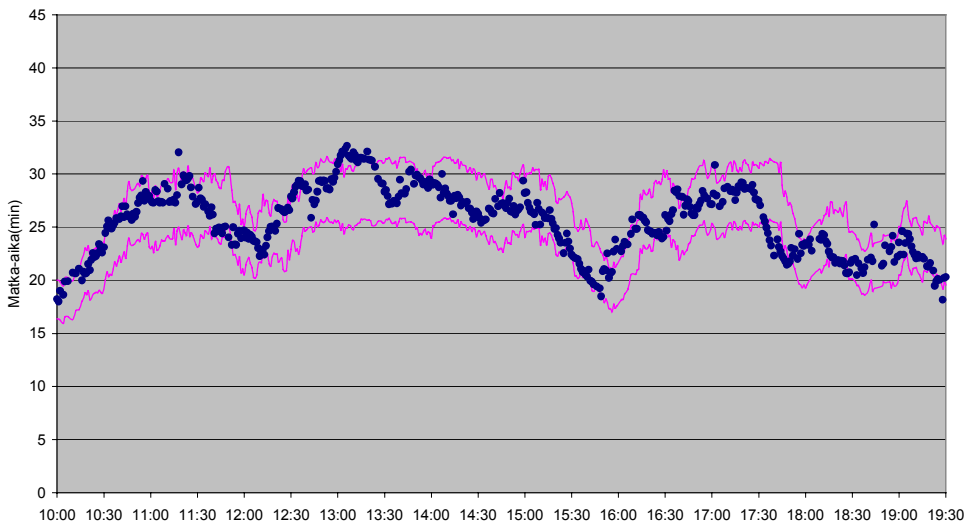
Kuva 29. Heinola–Kymijärvi-linkillä ruuhkamallin ennustamat ja mitatut matkajat. Mallissa aikasarjan pituus oli kolme minuuttia.



Kuva 30. Kymijärvi–Heinola-linkillä ruuhkamallin ennustamat ja mitatut matkajat. Mallissa aikasarjan pituus oli kolme minuuttia.

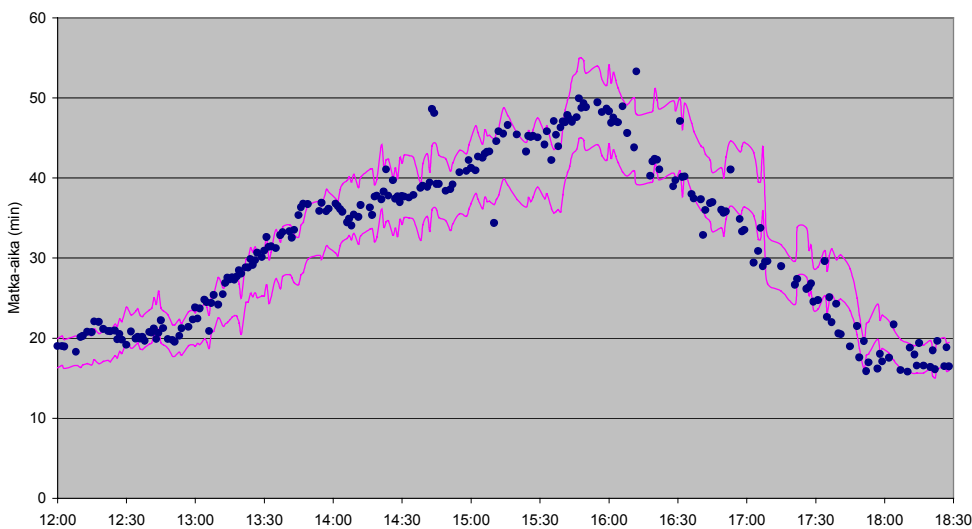
Ruuhkamalleille piirrettiin vastaavat ruuhkakuvat (kuvat 31–34) kuin perusmalleille (kuvat 16–19). Tietyissä tilanteissa ruuhkamallin ennusteet ovat vaakaampia kuin perusmallien. Suuria eroja ei kuitenkaan ollut. Heinola–Kymijärvi-linkin juhannusruuhkan ennustamisessa perusmallilla oli joitain ongelmia (kuva 16). Ruuhkamallin ennuste (kuva 31) oli jonkin verran parempi kuin perusmallin.

Juhannuksen paluuruuhka 25.6.2000, He-Ky

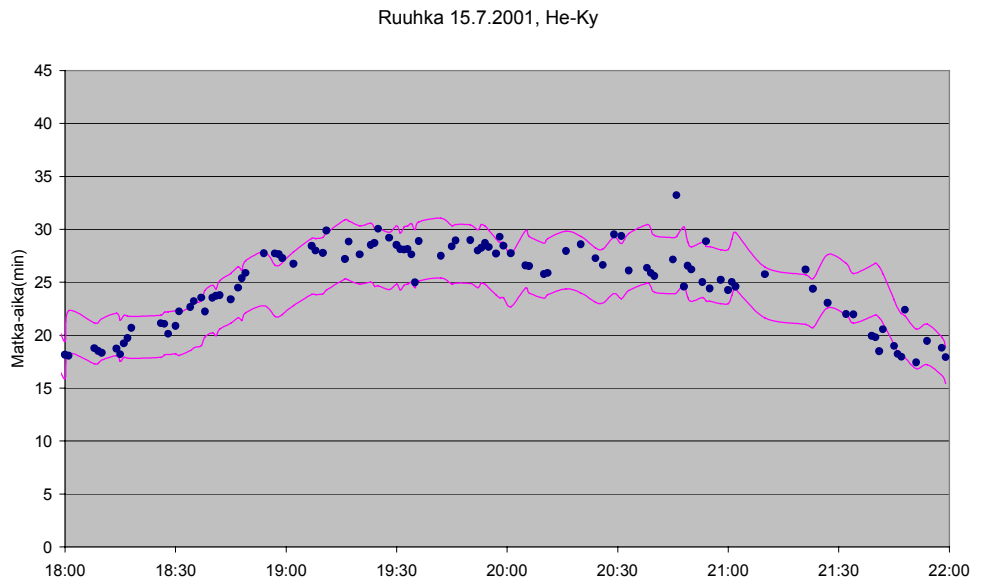


Kuva 31. Ruuhkamallin ennustama matka-aika (ylä- ja alaraja) sekä mitatut arvot (pisteet) juhannuksen paluuruuhkassa 25.6.2000 Heinola–Kymijärvi-linkillä. Malli perustuu neljän minuutin aikasarjaan ja kaikkien linkkien matka-aikatietoihin.

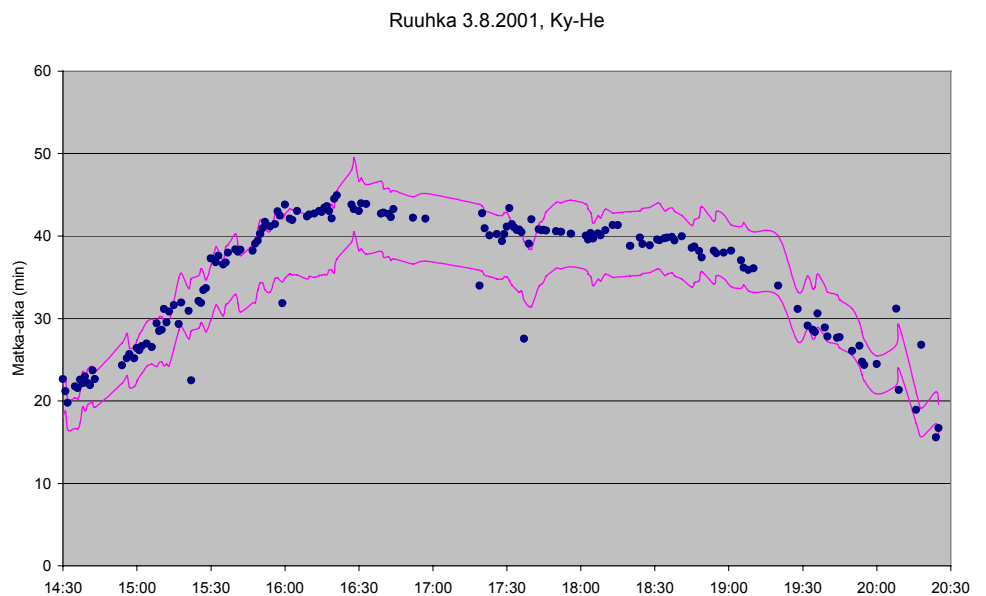
Ruuhka 30.6.2000, Ky-He



Kuva 32. Ruuhkamallin ennustama matka-aika (ylä- ja alaraja) sekä mitatut arvot (pisteet) perjantairuuhkassa 30.6.2000 Kymijärvi–Heinola-linkillä. Malli perustuu neljän minuutin aikasarjaan ja kaikkien linkkien matka-aikatietoihin.



Kuva 33. Ruuhkamallin ennustama matka-aika (ylä- ja alaraja) sekä mitatut arvot (pisteet) sunnuntairuuhkassa 15.7.2001 Heinola–Kymijärvi-linkillä. Malli perustuu neljän minuutin aikasarjaan ja kaikkien linkkien matka-aikatietoihin.



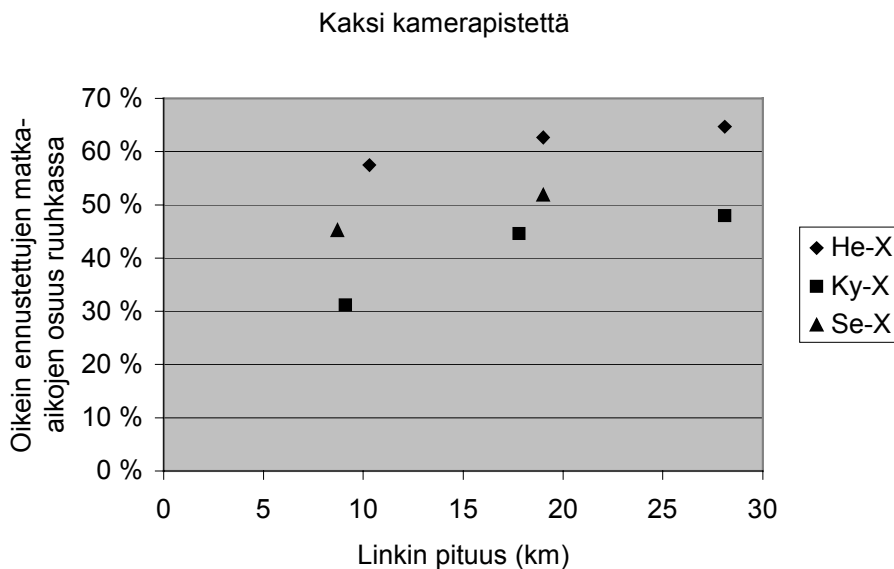
Kuva 34. Ruuhkamallin ennustama matka-aika (ylä- ja alaraja) sekä mitatut arvot (pisteet) perjantairuuhkassa 3.8.2001 Kymijärvi–Heinola-linkillä. Malli perustuu neljän minuutin aikasarjaan ja kaikkien linkkien matka-aikatietoihin.

9 MITTAUSJÄRJESTELMÄN RAKENNE

Tutkimuksessa haluttiin selvittää, kuinka mittausjärjestelmän rakenne vaikuttaa ennusteen laatuun. Selvitettäviä asioita olivat, kuinka kamerapisteidien etäisyys eli linkin pituus, kamerapisteidien määrä ja LAM-pistetietojen käyttö vaikuttavat ennusteen laatuun. Tuloksista on hyötyä myös sen arvioimisessa, kuinka herkkiä mallit ovat yksittäisten mittapisteiden vioittumiselle.

Käytettävissä oli kolme 8,7–10,3 kilometrin mittaista linkkiä sekä linkkiyhdistelmät. Pisin linkki (Heinola–Kymijärvi ja Kymijärvi–Heinola) oli siis 28,1 kilometriä pitkä ja keskimittaiset 19,0 sekä 17,8 kilometriä pitkiä. LAM-pisteet sijaitsivat Vierumäellä sekä Lahden eteläpuolella. Murhamäen LAM-piste Heinolan pohjoispuolella jätettiin tarkastelun ulkopuolelle, koska se oli epäkunnossa kesän 2000 aineiston keruujakson aikana. Mallien hyvyys mitattiin ruuhkassa oikein ennustettujen matka-aikojen osuudella.

Ensin selvitettiin matka-aikaennusteen hyvyttä tilanteessa, jossa matka-aika perustuisi ainoastaan kahden kamerapisteen tietoihin eli matka-aikatietoja saataisiin ainoastaan ennustelinkiltä. Tulokseksi saatiin, että matka-aikaennuste oli sitä parempi, mitä pidempi linkki oli (kuva 35).



Kuva 35. Ruuhkassa oikein ennustettujen matka-aikojen osuus eri linkeillä, kun ennuste perustuu kahden kamera-aseman tietoihin. He-X tarkoittaa Heinolasta alkavia linkejä.

Toiseksi selvitettiin, kuinka ylimääräinen kamerapiste vaikuttaa ennusteseen ja mihin tämä kamerapiste tulisi sijoittaa. Ylimääräisen tai ylimääräisten kamerapisteidien antamat lisäyötiedot paransivat tuloksia ja ennustemalli osoittautui pääosin sitä paremmaksi, mitä useampi kamerapiste sillä oli käytettävissä (taulukko 46). Poikkeuksina olivat Heinola–Kymijärvi- ja Heinola–Vierumäki-linkkien mallit. Ensimmäisellä toinen lisäkamera antoi hieinan huonomman tuloksen kuin parhaiten sijoitettu yksittäinen lisäkamera. Jälkimmäisessä mallissa toinen lisäkamera ei parantanut mallin suoritusta.

Taulukko 46. Ruuhkassa oikein ennustettujen matka-aikojen osuudet eri linkkien malleilla, jotka perustuvat 2–4 kamerapisteeseen. Lisäkameroiden sijainti ja se, ovatko ne linkillä, ennen sitä vaiko sen jälkeen on annettu suluissa.

Linkki	Pituus	2 kameraa		1 lisäkamera		2 lisäkameraa
Ky-He	28,1 km	48 %	60 % (Se, linkillä)	60 % (Vi, linkillä)		69 %
Ky-Vi	17,8 km	45 %	54 % (Se, linkillä)	52 % (He, jälkeen)		59 %
Ky-Se	9,1 km	31 %	33 % (Vi, jälkeen)	34 % (He, jälkeen)		40 %
Se-He	19,0 km	52 %	61 % (Ky, ennen)	64 % (Vi, linkillä)		73 %
Se-Vi	8,7 km	45 %	54 % (Ky, ennen)	50 % (He, jälkeen)		59 %
He-Ky	28,1 km	65 %	79 % (Vi, linkillä)	69 % (Se, linkillä)		78 %
He-Se	19,0 km	63 %	72 % (Vi, linkillä)	65 % (Ky, jälkeen)		74 %
He-Vi	10,3 km	57 %	61 % (Se, jälkeen)	61 % (Ky, jälkeen)		61 %

Näytti siltä, että linkillä oleva ylimääräinen kamera paransi tuloksia enemmän kuin sen ulkopuolella oleva mittauspiste (taulukko 46). Vaikutti myös siltä, että lisäkamera, joka sijaitti lähempänä linkin alku- kuin loppupistettä, johti parempiin tuloksiin kuin lisäkamera, joka sijaitti lähempänä linkin loppu- kuin alkupistettä.

Kuinka LAM-aseman liikennetiedot sitten vaikuttavat tuloksiin? Mallit perustuivat LAM-tietojen lisäksi ainoastaan kahden kamerapisteen antamiin tietoihin. Mallilla oli siis syötteinä LAM-tietojen lisäksi ainoastaan tarkasteltavan linkin matka-aikatiedot.

LAM-tiedot eivät parantaneet ennustetta yhtä paljon kuin lisäkamera (taulukot 46 ja 47). Pistekohtainen liikennetieto kuitenkin paransi mallien suoritusta lukuun ottamatta linkin Kymijärvi–Seesta malleja, joita LAM-pistetieto huononsi hieman (taulukko 47). Ennen Lahtea sijainneen LAM-aseman liikennetiedot vaikuttivat parantavan tulosta enemmän kuin Vierumäen LAM-aseman tiedot. Tämä lienee seurausta ruuhkan sijainnista (ennen Vierumäkeä). Kymijärvi–Seesta-linkkiä lukuun ottamatta kahden LAM-aseman tiedot johtivat parempiin tuloksiin kuin yksittäisen LAM-aseman tiedot.

Taulukko 47. Ruuhkassa oikein ennustettujen matka-aikojen osuudet eri linkkien malleilla, jotka perustuvat kahteen kamerapisteeseen ja nollasta kahteen LAM-pisteeseen. LAM-piste 601 sijaitsee Vierumäellä ja 424 Lahden eteläpuolella.

Linkki	Pituus	Ei LAM	601	424	601 ja 424
Ky-He	28,1 km	48 %	51 %	51 %	55 %
Ky-Vi	17,8 km	45 %	48 %	51 %	54 %
Ky-Se	9,1 km	31 %	30 %	30 %	29 %
Se-He	19,0 km	52 %	55 %	60 %	62 %
Se-Vi	8,7 km	45 %	48 %	49 %	52 %
He-Ky	28,1 km	65 %	65 %	-	-
He-Se	19,0 km	63 %	64 %	-	-
He-Vi	10,3 km	57 %	59 %	-	-

10 JOHTOPÄÄTÖKSIÄ JA SUOSITUKSIA

10.1 Neuroverkkomallit

Sekä perusmallien että muutaman kameran tietoihin perustuvien mallien tuloksia tarkasteltaessa havaittiin, että mitä pidempi linkki on, sitä paremmin ennusteet onnistuivat. Tämä saattaa kuulostaa ristiriitaiselta, mutta tulos on järkevä. Matka-ajan varianssi ei kasva suhteessa linkin pituuteen, vaan varianssin osuus pienenee, kun linkki pitenee. Pitkän linkin matka-aika on siis vakaampi (= helpompi ennustaa) kuin lyhyen linkin matka-aika. Oma osuutensa tulokseen on myös sillä, että pitkät linkit on jaettu kamerapisteillä osalinkkeihin ja lyhyitä ei. Pitkiltä linkeiltä saadaan siis yksityiskohtaisempaa tietoa kuin lyhyiltä linkeiltä.

Valtatielle 4 tehdyt mallit toimivat suurelta osin tyydyttävästi. Etenkin pisimpien linkkien mallien toimintaan voi olla suhteellisen tyytyväinen (Heinola–Kymijärvi-malli oli oikeassa ruuhka-aikaan 84 prosenttia ja Kymijärvi–Heinola-malli 73 prosenttia ajasta). Tavoitetta, että ennuste olisi oikeassa 90 prosenttia ruuhka-ajasta, ei tosin aivan saavutettu. On kuitenkin huomattava, että osa näistä ”väärin” ennustetuista ruuhkahavainnoista oli yksittäisiä minuutteja tai hyvin lyhyitä aikajaksoja, joiden ajaksi matka-aika oli kohonnut. Tällaisia virheitä ei voi pitää erityisen vakavina, koska kyseessä saattaa myös olla aineiston suodatuksessa jäljelle jäänyt ”poikkeava” havainto, jolloin tiellä ei tosi asiassa ollut ruuhkaa, mutta mittaus tulos perustui muutamaa hieman poikkeavaan havaintoon. Ongelma poistuisi, jos mitattujen matka-aikojen otoskoko olisi suurempi, koska yksittäisen (poikkeavan) havainnon osuus keskimatka-ajasta pienenis.

Valtatien 4 malleissa onnistuttiin siis kohtalaisen hyvin. Kehä I:n malleja ei sen sijaan voi pitää kuin korkeintaan välttävinä. Vaikka kamerapisteiden väli oli Kehä I:llä suunnilleen sama kuin valtatiellä 4, ero oli huomattava. Kehä I:llä linkkien päästä päähän ajavien osuus on pienempi kuin valtatiellä 4 ja väylällä on liikennevaloja, jotka sekoittavat liikennevirtaa. Ongelman huomasi jo korrelaatioista estimoitavan tai ennustettavan matka-ajan ja muiden mittasuureiden välillä. Valtatiellä 4 löydettiin useita muuttujia, jotka korreloivat ennustettavan matka-ajan kanssa erittäin voimakkaasti (korrelaatiokerroin yli 0,9), Kehä I:llä korrelaatio oli selvästi vähäisempää ja syötemuuttujien joukko jäi pieneksi.

Vastaus kysymykseen, voidaanko Kehä I:llä estimoida tietyn linkin matka-aika muiden matka-aika- ja LAM-tietojen perusteella, on näiden mallien perusteella kielteinen. Osalinkit ovat tämäntyyppiseen estimointiin liian pitkiä eikä päätepisteissä sijaitsevien LAM-asemien liikennetietojen voida mitenkään kuvitella edustavan koko linkin liikennetilannetta.

10.2 Sumeat ennustemallit

Epävarmuustekijöiden suuri määrä vaikeutti sumean mallin rakentamista, kun sekä jäsenyysfunktioita että sumeita päättelysääntöjä yritettiin generoida suuresta aineistojoukosta. Mikäli näistä lähtökohdista olisi haluttu rakentaa täydellinen säätokanta, sumeiden päättelysääntöjen lukumäärä olisi noussut noin 38 000:een. Käytännössä kaikki yhdistelmät muuttujien välillä eivät

kuitenkaan ole mahdollisia, sillä muuttajat eivät ole täysin riippumattomia toisistaan ja mahdollisten yhdistelmien määrä olisi paljon pienempi. Valmiissa mallissa sääntöjä oli yhdeksän. Mallin parantamiseksi sääntömäärää olisi pitänyt lisätä, mikä olisi johtanut mallin monimutkaistumiseen. Sumea logiikka soveltuukin yleensä paremmin sellaisten tilanteiden mallintamiseen, joissa päättelysäännöt ovat valmiiksi olemassa tai ainakin helposti ja yksikäsitteisesti muodostettavissa.

Tutkittaessa muuttujien välisiä yhteyksiä GUHA-menetelmällä muuttajat oli luokiteltava yksikäsitteisiin luokkiin. Luokittelu osoittautui sikäli ongelmalliseksi, että se luonnollisesti vaikutti merkittävästi analyysien tuloksiin. Luokkien rajat eivät olleet selkeästi aineistosta nähtävissä, joten ne pääteltiin silmämääräisesti, mikä vei jossain määrin pohjaa yritykseltä tutkia aineistoa systemaattisesti.

Sumean järjestelmän toiminta on monivaiheinen. Koska jokainen vaihe voidaan toteuttaa usealla eri tavalla, valittu toteutustapa vaikuttaa järjestelmän antamiin tuloksiin. Tässä työssä rakennettua mallia testattiin sekä sumeaan kokonaissimilaarisuuteen että Gödelin algebran periaatteisiin perustuvilla sumeilla päättelyjärjestelmillä. Sumeaan similaarisuuteen perustuva menetelmä poikkeaa muista toteutustavoista siinä, että sen takana on hyvin perusteltu matematiikka, mikä tosin ei kuitenkaan takaa tulosten paremmuutta muihin menetelmiin verrattuna. Gödel-järjestelmä osoittautuikin hieman similaarista paremmaksi tätä mallia testattaessa. Myös mallissa käytettyjen jäsenyysfunktioiden muoto vaikuttaa lopputulokseen. Esimerkiksi kellokäyrän tai puolisuunnikkaan muotoiset jäsenyysfunktiot olisivat saattaneet johtaa parempiin tuloksiin.

Sumean mallin rakentamisessa käytetty aineisto sisälsi vähän ruuhkatilanteita, mikä vaikeutti pitkien matka-aikojen ennustamista, sillä niitä koskevien sääntöjen ja jäsenyysfunktioiden muodostaminen oli pienen otoskoon vuoksi vaikeaa. Malli soveltuu vain aineistojoukossa esiintyneiden liikennetilanteiden käsittelemiseen eikä ole yleistettävissä muuntyyppisiin liikennetilanteisiin. Malli perustuukin täysin sen rakentamisessa käytettyyn aineistoon. Se ei kykene käsittelemään mitään erikoistilanteita eikä se varaudu siihen, että esimerkiksi jokin hahmontunnistuskameroista rikkoutuu. Sitä, millaisia tuloksia saataisiin, jos mallin käsiteltäväksi syötettäisiin jonakin muuna ajankohdana kerättyä aineistoa, ei ole toistaiseksi testattu. Jotta mallia voitaisiin käytännössä käyttää matka-ajan ennustamiseen valtatiellä 4 Lahden ja Heinolan välillä, olisi sitä parannettava huomattavasti. Joka tapauksessa sumea logiikka osoittautui jonkin verran paremmaksi ennustejärjestelmän perustaksi kuin nykyisin käytössä oleva viimeisimpien mittaustietojen summaan perustuva menetelmä.

10.3 Suosituksia

Selkeä ongelma, joka ennusteita tarkasteltaessa on havaittavissa, on, ettei malli voi ennakoita matka-ajan kasvua. Malli kykenee reagoimaan vasta, kun ensimmäiset merkit häiriöistä on mitattu. Satunnaisten liikenteen häiriöiden osalta tämä on varmastikin vain hyväksyttävä normaalina ilmiönä. Ylikysynnästä aiheutuva ruuhka sitä vastoin pitäisi kyetä ennakoimaan paremmin. Tällä hetkellä on ongelmana, että esimerkiksi valtatie 4 malleilla on tiedot ainoastaan Vierumäen LAM-pisteestä sekä Lahden ja Heinolan takaa.

Vierumäki ei ole kumpaankaan suuntaan erityinen ongelmapistee, joten se indikoi Kymijärvi–Heinola-välin ongelmia varsin huonosti. Linkin ulkopuolella olevien LAM-asemien tiedot eivät myöskään kerro ongelmista suoraan, sillä Lahteen ja Heinolaan menevät ajoneuvot ovat tiedoissa mukana eikä näistä kaupungeista tulevista ajoneuvoista ole mitään tietoa.

Ihannetilanteessa syötteenä olisi matka-aikatietojen lisäksi liikennemäärätiedot kaikista (vähintäänkin merkittävistä) sisään tulevista ja ulos menevistä sekä ainakin mahdollisten ongelmakohtien liikennevirroista. Näin saataisiin jo hieman etukäteen käsitys linkille pyrkivistä virroista, jotka ovat osaltaan vaikuttamassa matka-aikaan.

Tällainen mittausjärjestelmä mahdollistaisi paremmin myös satunnaisten häiriöiden aiheuttaman (tarkoittaa tässä muusta kuin ylikysynnästä aiheutuvan) matka-aikojen kasvun havaitsemisen nopeammin. Käytännössä tämä näkyisi sisään menevien ja ulos tulevien liikennemäärien erotuksen kasvuna (vertaa häiriöiden havaitsemisalgoritmit).

Toinen tärkeä asia, jossa liikennemäärätieto voisi auttaa, on sen arvioiminen, ovatko mitatut matka-ajat järkeviä. Jos liikennemäärästä ei ole mitään tietoa, on vaikea sanoa, johtuvatko pitkät matka-ajat mahdollisesti ylikysynnästä, satunnaisesta häiriöstä vai kellojen kalibroinnin pettämisestä.

Ihanteellista olisi, jos liikennemäärätiedot saataisiin suoraan matka-aikojen mittausjärjestelmästä. Näin välttyttäisiin päällekkäisiltä järjestelmiltä ja säästettäisiin kustannuksissa. Tämä tosin vaatisi järjestelmän parantamista siten, että suurempi osa ajoneuvoista tulisi havaituksi ja että kaikkia kaistoja seurattaisiin. Nykyinen matka-aikajärjestelmä kykenee hyvissä olosuhteissa lukemaan yksittäisessä kamerapisteessä noin 60 prosenttia seuratun kaistan rekisterikilvistä. Käytännössä tämä osuus on noin 40 prosenttia. Todellinen otoskoko on kuitenkin merkittävästi pienempi, koska olosuhteet eivät aina ole hyvät eikä valtatiellä 4 seurata kaikkien kaistojen liikennettä.

Pieni otoskoko ei ole ongelma, jos halutaan estimoida keskimääräinen matka-aika. Liikennemääräestimaattiin otoskoko ei kuitenkaan riitä. Samoin matka-ajan hajonnan estimointiin otoskoko on liian pieni. Suuri hajonta voisi olla merkinä häiriöistä tai muutoksista liikennetilanteessa.

Valtatiellä 4 kameraväli oli suhteellisen sopiva. Tieto ongelmista tuli kohtuullisella viiveellä ja kamerapisteeet jakoivat linkin osiin varsin onnistuneesti. Tosin se, ettei ohituskaistoja seurata, pienentää otoskokoja ja aiheuttaa otokseen harhaa, koska nopeimmat ajoneuvot jäävät herkästi havaitsematta. Olisikin tärkeää, että kamerat asennettaisiin kaikille kaistoille, jos halutaan tarkka kuva matka-ajasta muulloinkin kuin ruuhkassa, jolloin ohituskaistan merkitys pienenee.

Kehä I:llä seurattiin kaikkia kaistoja, mikä on hyvä asia. Jos Kehä I:n tyyppisessä kohteessa halutaan ennustaa matka-aikoja, pitäisi kamerapisteitä kuitenkin olla enemmän. Kehä I:llä esimerkiksi linkki Otaniemen vesitornin luota Konalaan pitää sitä vastoin sisällään liikaa tapahtumia. Välille mahtuu useita suuria liittymiä (suurimpina Turun väylä ja Vanha Turuntie) ja useat liikennevalot. Hyvä nyrkkisääntö voisi olla esimerkiksi, että maksimi kameravälinä voidaan pitää kahden merkittävän liittymän väliä. Jos liittymiä kuitenkin on harvassa, kamerat pitää asentaa sen verran tiheään, ettei havainnoinnin viive kasva liian suureksi.

Kamerapisteiden lisäksi Kehä I:ltä kaivattaisiin tarkempaa tietoa liikennemääristä. Esimerkiksi Otaniemi–Konala-linkin liikennetilanteen hahmottamiseen ei riitä linkin matka-aikatiedon lisäksi tieto liikennemääristä Konalassa. Kokonaiskuvan hahmottamiseksi tieto kaikista merkittävistä liikennevirroista olisi oleellinen. Erityisen oleellista tämä on silloin, jos matka-aikaa kuitenkin haluttaisiin estimoida raportissa kuvatulla tavalla.

Havaituista puutteista huolimatta, etenkin valtatie 4 kohteessa, tulokset ovat sen verran lupaavia, että mallista kannattaisi tehdä kokeiltavaksi tuotantoversio. Tuotantoversiossa joudutaan käymään läpi koko prosessin ajantasaisen, jatkuva-aikaisen toiminnan asettamat vaatimukset ja reunaehdot. Lisäksi tiedon saannin viive aiheuttaa käytetylle aggregointijaksolle muutostarpeita.

11 YHTEENVETO

Tärkein yksittäinen objektiivinen liikenteen sujuvuuden mittari on matka-aika. Liikenteen sujumisella on tärkeä merkitys teollisuuden ja elinkeinoelämän kuljetuksille, joille sujuvuus ilmenee kuljetusten varmuutena eli matka-aika-budjetin pitävyytenä. Ennustetun matka-ajan merkitys on liikenneoperaattorin näkökulmasta katsottuna erittäin merkittävä, sillä juuri se vaikuttaa päätöksentekoon tienkäyttäjille tarjottavasta tiedosta. Tiedotuksella pyritään vaikuttamaan tienkäyttäjien odotuksiin matka-ajan suhteen ja siten tavoitettavien aikoihin. Täten pyritään parantamaan koettua sujuvuutta.

Tämän tutkimuksen tarkoituksena oli selvittää lyhyen aikavälin matka-aikaennusteiden tekemistä sekä sitä, kuinka hyvin matka-aika saadaan estimoitua pistemittausten avulla. Tutkimuskohteina olivat valtatie 4 Lahden ja Heinolan välillä sekä Kehä I Otaniemen ja Pukinmäen välillä. Molemmat koejakso on varustettu matkanopeuden seurantajärjestelmillä, joiden avulla matkaajat mitataan automaattisesti rekisteritunnusten perusteella.

Ennustemalleja tehtiin kahdella eri menetelmällä: neuroverkkoihin ja sumeaan logiikkaan perustuen. Mallien hyvyttä verrattiin olemassa olevan järjestelmän antaman tiedon tarkkuuteen. Tällä hetkellä tiedotusjärjestelmä kertoo tienkäyttäjille arvion Lahden ja Heinolan välisestä matka-ajasta. Matka-aikaestimaatti perustuu viimeisimpiin mittauksiin.

Neuroverkkoihin perustuvien matka-ajan ennustemallien osalta tutkimuksen tarkoituksena oli selvittää, mitkä muuttujat korreloivat parhaiten kulloinkin tarkasteltavan linkin matka-ajan kanssa ja minkälaiset syötteet johtavat parhaaseen tulokseen. Tutkimuksen tarkoituksena oli myös selvittää, mitkä tekijät korreloivat matka-ajan kanssa parhaiten silloin, kun matka-aika estimoidaan neuroverkolla muiden linkkien matka-aikojen ja pistekohtaisten mittausten avulla.

Sumeiden mallien osalta oli tarkoitus selvittää syy-seuraussuhteita matka-ajan kehityksessä ja tutustua sumean logiikan antamiin mahdollisuuksiin lyhyen aikavälin ennusteiden tekemisessä. Lopuksi oli tarkoitus selvittää, riittäisikö ennusteiden tarkkuus sovelluksen käyttöön.

Ennustemalleina käytettiin monikerrosperseptronineuroverkkoja niiden aikaisemmissa tutkimuksissa saamien hyvien tulosten perusteella. Sumeaan logiikkaan perustuvat mallit tehtiin Gödelin algebran sekä Lukasiewiczin algebran ja sumean similaarisuuden periaatteiden mukaisesti.

Neuroverkkomallien syötemuuttujat valittiin sen perusteella, kuinka hyvin ne korreloivat ennustettavan matka-ajan ja muiden syötemuuttujaehdokkaiden kanssa. Syötteiksi valittiin muuttujat, joiden korrelaatio ennustettavaan matka-aikaan oli voimakas, mutta jotka eivät korreloineet liian voimakkaasti muiden syötteiden kanssa.

Valtatielle 4 tehtiin matka-ajan ennustemalleja. Mallit ennustivat tarkasteluhetkeä seuraavan minuutin aikana linkeille lähtevien ajoneuvojen keskimatka-aikaa tarkasteluhetkeä edeltävien minuuttien aikana linkkien loppupisteet ohittaneiden ajoneuvojen matka-aikojen perusteella. Lisäksi syötteissä oli pistekohtaisia liikennetietoja Vierumäeltä ja Lahden eteläpuolelta.

Ennusteiden hyvyttä tarkasteltiin kahdella eri tavalla: tilastollisesti ja matka-aikatiedotuksen hyvyyden näkökulmasta. Tilastolliset suureet kertoivat, onko mallilla taipumus yli- tai aliarvioida ennusteita ja kuinka virheet jakautuvat mitattujen arvojen ympärille. Mallien tekemien ennusteiden hyvyttä mitattiin myös matka-aikatiedotuksen näkökulmasta eli sillä, kuinka suuri osa ennusteista olisi mennyt oikein ja kuinka suuri osa poikennut varmuusvälin ylä- tai alapuolelle. Tasaista ruuhkatonta liikennettä on helppo ennustaa, ja tämän takia edellä kuvattu tarkastelu tehtiin koko aineiston lisäksi myös pelkälle ruuhkaliikenteelle.

Oikein ennustettujen matka-aikojen osuus vaihteli voimakkaasti linkiltä toiselle Heinola–Kymijärvi-linkin ruuhka-ajan matka-aikaennusteet osoittautuivat parhaiksi (oikein ennustettujen osuus 78,8 prosenttia ajasta) ja Kymijärvi–Seesta-linkin huonoimmiksi (oikein ennustettujen osuus 39,3 prosenttia ajasta). Muista linkeistä Heinola–Seesta-, Kymijärvi–Heinola- ja Heinola–Seesta-linkeillä ruuhkassa oikein ennustettujen osuus oli yli 70 prosenttia. Vierumäelle päättyvillä linkeillä osuus oli 60–70 prosenttia.

Kehä I:lle tehtiin sekä estimaatti- että ennustemallit. Estimaattimalleilla pyrittiin estimoimaan tietyn linkin matka-aika muiden mittaustietojen perusteella tilanteessa, jossa kyseisen linkin matka-ajan mittaaminen ei onnistu. Malleissa siis estimoidaan tietyn linkin matka-aikaa saman ja edeltävien jaksojen muiden liikennetietojen perusteella. Ennustemallit olivat vastaavia kuin valtatielle 4 tehdyt. Tuloksena oli, että ennustemallit antoivat selvästi parempia tuloksia kuin estimaattimallit.

Neuroverkkomallin lisäksi muodostettiin sumeaa logiikkaan perustuva matka-ajan ennustamismalli valtatie 4 linkille, joka ulottui Kymijärveltä Heino-laan. Tavoitteena oli tarkastella syy-seuraussuhteita matka-ajan kehityksessä, tutustua sumeaa logiikkaan tarjoamiin mahdollisuuksiin lyhyen aikavälin matka-aikaennusteen tekemisessä sekä analysoida muodostetun mallin hyvyttä ja verrata saatuja tuloksia vastaavan neuroverkkomallin antamiin tuloksiin.

Kaikki sumeaa mallin tuloksille lasketut tilastolliset virheet ovat neuromallin vastaavia virheitä suurempia. Neuromalli toimi sumeaa mallia paremmin myös matka-aikatiedotuksen näkökulmasta tarkasteltuna. Sumeita malleja keskenään vertailtaessa nähdään, että Gödelin algebran mukainen malli oli hiukan similaarista mallia parempi. Nykyjärjestelmä ennusti ruuhkaliikennettä sumeita malleja huonommin.

Heinola–Kymijärvi- ja Kymijärvi–Heinola-linkeille tehtiin neuroverkkoihin perustuvat ruuhkamallit kokoamalla opetusjoukko ainoastaan niiden päivien aineistoista, jolloin matka-aika oli noussut päiväsaikaan jossain vaiheessa selvästi yli 20 minuuttiin. Molemmilla linkeillä päästiin näiden mallien avulla parempaan tulokseen kuin parhaalla koko päivän aineistolla opetetulla mallilla.

Tutkimuksessa haluttiin lopuksi selvittää, kuinka mittausjärjestelmän rakenne vaikuttaa ennusteen laatuun. Asiaa tutkittiin valtatiellä 4. Ensin oletettiin, että matka-aikatietoja saataisiin ainoastaan ennustelinkiltä. Tällaisessa tilanteessa havaittiin, että matka-aikaennuste oli sitä parempi, mitä pidempi linkki oli. Toiseksi selvitettiin, kuinka ylimääräinen kamerapiste vaikuttaa ennusteseen ja mihin tämä kamerapiste tulisi sijoittaa. Havaittiin, että ylimääräisen tai ylimääräisten kamerapisteiden antamat lisäsyötiedot paransivat tuloksia

ja ennustemalli osoittautui pääosin sitä paremmaksi, mitä useampi ylimääräinen kamerapiste sillä oli käytettävissä. Vaikutti siltä, että linkillä oleva ylimääräinen kamera paransi tuloksia enemmän kuin sen ulkopuolella oleva mittauspiste ja että lisäkamera, joka sijaitsi lähempänä linkin alku- kuin loppupistettä, johti parempiin tuloksiin kuin lisäkamera, joka sijaitsi lähempänä linkin loppu- kuin alkupistettä.

Kuinka LAM-aseman liikennetiedot sitten vaikuttavat tuloksiin? Mallit perustuivat LAM-tietojen lisäksi ainoastaan kahden kamerapisteen antamiin tietoihin. LAM-tiedot eivät parantaneet ennustetta yhtä paljon kuin lisäkamera, mutta jonkin verran yhtä poikkeusta lukuun ottamatta. Ennen Lahtea sijainneen LAM-aseman liikennetiedot vaikuttivat parantavan tulosta pohjoiseen menevässä liikenteessä enemmän kuin Vierumäen LAM-aseman tiedot. Kahden LAM-aseman tiedot johtivat parempiin tuloksiin kuin yksittäisen LAM-aseman tiedot.

Valtatiellä 4 saavutetut tulokset olivat suhteellisen tyydyttäviä. Kehä I:n malleja sitä vastoin voi pitää korkeintaan välttävinä. Kehä I:llä matka-aikojen seurantajärjestelmän yksityiskohtaisuus ei selvästikään riitä ennusteiden tekoon ja kamerapisteitä pitäisi olla enemmän. Molempien kohteiden ennusteet paranisivat varmasti, mikäli liikennetilannekuvaa täydennettäisiin matkakatietojen lisäksi liikennemäärätiedoilla.

Ihannetilanteessa syötteenä olisi matka-aikatietojen lisäksi liikennemäärätiedot kaikista (vähintäänkin merkittävistä) sisään tulevista ja ulos menevistä sekä ainakin mahdollisten ongelmakohtien liikennevirroista. Näin saataisiin jo hieman etukäteen käsitys linkille pyrkivistä virroista, jotka ovat osaltaan vaikuttamassa matka-aikaan. Tällainen mittausjärjestelmä mahdollistaisi paremmin myös satunnaisten häiriöiden aiheuttaman matka-aikojen kasvun havaitsemisen nopeammin. Toinen tärkeä asia, jossa liikennemäärätieto voisi auttaa, on sen arvioiminen, ovatko mitatut matka-ajat järkeviä.

Havaituista puutteista huolimatta, etenkin valtatie 4 kohteessa, tulokset ovat sen verran lupaavia, että mallista kannattaisi tehdä kokeiltavaksi tuotantoversio. Tuotantoversiossa joudutaan käymään läpi koko prosessin ajantasaisen, jatkuva-aikaisen toiminnan asettamat vaatimukset ja reunaehdot. Lisäksi tiedon saannin viive aiheuttaa käytetylle aggregointijaksolle muutostarpeita.

LÄHDELUETTELO

CASSIDY M, WINDOVER J (1995). *A Method for Assessing the Dynamics of Freeway Traffic Flow*. Transportation research record. No. 1484 (July 1995). s. 73-79.

DEMUTH H, BEALE M (1998). *Neural Networks Toolbox for Use with Matlab*. User's Guide, Version 3. The Math Works Inc. s. 5-1 – 5-58.

ELORANTA T (1999). *Rekisterikilpien tunnistukseen perustuva liikenteen automaattinen matkanopeuden seuranta*. Tielaitoksen selvityksiä 46/1999. Uudenmaan tiepiiri, Tielaitos, Helsinki. 149 s.

HÁJEK P, SOCHOROVÁ A, ZVÁROVÁ J (1993). *GUHA for personal computers*. Reprinted from Computational Statistics & Data Analysis. UIVT Praha, Czech Republic. 3 s.

HAUGEN T (1996). *Section Data. Possibilities and experiences*. SINTEF Civil and Environmental Engineering, Transport Engineering, Norway. 16 s.

HAUKILEHTO K (2001). *Sumea similaarisuus*. Sovelletun matematiikan erikoistyö, Tampereen teknillinen korkeakoulu, Matematiikan laitos. 7 s.

HÄMEEN TIEPIIRI (2000). *Vt 4 Lahti-Heinola matka-ajan seuranta- ja informaatiojärjestelmän toiminnan arviointi*. Tielaitoksen selvityksiä 58/2000. Hämeen tiepiiri, Tielaitos, Tampere. 46 + 61 s.

INNAMAA S (1999). *Automaattiset liikenteenohjaus- ja liikenneinformaatiojärjestelmät*. Tielaitoksen selvityksiä 28/1999, Tielaitos, Helsinki. 136 s.

INNAMAA S, PURSULA M (2000). *Liikennemäärän ja nopeuden lyhyen aikavälin ennustaminen*. Tielaitoksen selvityksiä 54/2000. Tielaitos, Helsinki. 101 + 3 s.

KILJUNEN M, SUMMALA H (1996). *Ruuhkaisuuden kokeminen ja liikennetiedottaminen*. Tielaitoksen selvityksiä 25/1996, Tielaitos, Keskushallinto, Helsinki. 77 s.

KLIR G, ST. CLAIR U, YUAN B (1997). *Fuzzy Set Theory – Foundations and Applications*. Prentice Hall PTR, Upper Saddle River, NJ. 245 s.

LEE S, KIM D, KIM J, CHO B (1998). *Comparison of Models for Predicting Short-Term Travel Speeds*. Conference CD-ROM, 5th World Congress on Intelligent Transport Systems, 12 – 16 October 1998, Seoul, Korea. 9 s.

LIIKENNEMINISTERIÖ (1997). *Liikennetelematiikan kehittämisen suuntaaminen, Osaraportti 2, Liikennetelematiikan kehittäminen vuoteen 2005*. Liikenneministeriön julkaisu B:37/97, Liikenneministeriö, Helsinki. 110 s.

LUOMA S (1998). *Tieliikenteen sujuvuus ja sen mittaaminen*. Diplomityö, Teknillinen korkeakoulu, Liikennetekniikka. 122 s.

OJA E (1998). *Neuraalilaskennan perusteet*. Luentomoniste, Teknillinen korkeakoulu, Informaatiotekniikan laboratorio. 199 s.

PFANNERSTILL E (1991). *Measuring Section-Related Traffic Data by Correlation Methods - a New Approach to Traffic Monitoring and Control*. Applications of Advanced Technologies in Transportation Engineering: Proceedings of the Second International Conference, New York: American Society of Civil Engineers. s. 1 - 5.

PURSULA M (1995). *Sumeat joukot ja sumeat menetelmät liikennetekniikassa*. Teknillinen korkeakoulu, Liikennetekniikka, Opetusmoniste 14. Otaniemi. 192 s.

SMITH B, DEMETSKY M (1994). *Short-Term Traffic Flow Prediction: Neural Network Approach*. Transportation Research Record 1453. S. 98 - 104.

SMITH B, DEMETSKY M (1997). *Traffic Flow Forecasting: Comparison of Modeling Approaches*. Journal of Transportation Engineering, Vol. 123, No. 4, July / August 1997. S. 261 - 266.

TIEHALLINTO (2001). *Tiehallinnon liikenteen hallinnan toimintalinjat – taustaraportti*. Tiehallinto, Liikenteen palvelut. 53 + 36 s.

TIELAITOS (1999). *Liikenteen automaattinen mittaus 1998*. Tielaitoksen sisäisiä julkaisuja 52/1999, Tielaitos, Helsinki. 240 s.

TURUNEN E (2001). *Asiantuntijan ajattelun mallintaminen sumean logiikan avulla*. Opetusmoniste. Liikennetekniikan vaihtuva opintojakso, syksy 2001.

WESTERMAN M (1995). *Real-Time Traffic Data Collection for Transportation Telematics*. Delft University of Technology, Delft. 266 + 43 s.

YAGER R, FILEV D (1994). *Essentials of Fuzzy Modeling and Control*. John Wiley, New York. 388 s.

LIITTEET

- Liite A: Korrelaatiot ennustettavan matka-ajan kanssa
- Liite B: Mitatut ja neuromalleilla ennustetut matka-ajat
- Liite C: Sumeiden ennustemallien jäsennyysfunktiot ja sumeat päätte-
lysäännöt

KORRELAATIOT ENNUSTETTAVAN MATKA-AJAN KANSSA

Taulukoissa on esitetty valtatie 4 eri linkeiltä laskettujen suureiden korrelaatio ennustettavan linkin matka-ajan kanssa. Suureissa on viiveenä yksi minuutti. Taulukon yläpuolelle on merkitty ennustelinkki.

Heinola–Seesta

	He-Ky	He-Se	He-Vi	Vi-Ky	Vi-Se	Se-Ky
Viimeisen viiden minuutin matka-aikojen keskihajonta	0,00	0,24	0,09	0,29	-0,12	-0,03
Viimeisten 10 ajoneuvon matka-aikojen keskihajonta	-0,21	-0,01	-0,35	-0,50	-0,42	-0,22
Viimeisten 20 ajoneuvon matka-aikojen keskihajonta	-0,17	0,21	-0,07	-0,59	-0,49	-0,23
Viimeisen minuutin matka-aikojen mediaani	0,91	0,97	0,97	0,43	0,49	0,21
Viimeisten 10 ajoneuvon matka-aikojen mediaani	0,94	0,97	0,98	0,57	0,58	0,28
Viimeisten 20 ajoneuvon matka-aikojen mediaani	0,93	0,96	0,98	0,61	0,64	0,32

Heinola–Vierumäki

	He-Ky	He-Se	He-Vi	Vi-Ky	Vi-Se	Se-Ky
Viimeisen viiden minuutin matka-aikojen keskihajonta	-0,20	0,16	0,10	-0,25	-0,15	-0,08
Viimeisten 10 ajoneuvon matka-aikojen keskihajonta	-0,48	-0,16	-0,36	-0,51	-0,39	-0,45
Viimeisten 20 ajoneuvon matka-aikojen keskihajonta	-0,52	0,02	-0,07	-0,60	-0,47	-0,49
Viimeisen minuutin matka-aikojen mediaani	0,89	0,91	0,95	0,42	0,46	0,42
Viimeisten 10 ajoneuvon matka-aikojen mediaani	0,89	0,90	0,94	0,56	0,57	0,52
Viimeisten 20 ajoneuvon matka-aikojen mediaani	0,89	0,89	0,93	0,62	0,65	0,59

Vierumäki–Kymijärvi

	He-Ky	He-Se	He-Vi	Vi-Ky	Vi-Se	Se-Ky
Viimeisen viiden minuutin matka-aikojen keskihajonta	-0,03	0,16	0,03	-0,13	-0,12	-0,02
Viimeisten 10 ajoneuvon matka-aikojen keskihajonta	-0,15	-0,14	-0,09	-0,22	-0,20	-0,18
Viimeisten 20 ajoneuvon matka-aikojen keskihajonta	-0,19	-0,04	0,04	-0,24	-0,19	-0,14
Viimeisen minuutin matka-aikojen mediaani	0,39	0,42	0,32	0,27	0,33	0,42
Viimeisten 10 ajoneuvon matka-aikojen mediaani	0,39	0,42	0,31	0,41	0,50	0,49
Viimeisten 20 ajoneuvon matka-aikojen mediaani	0,39	0,42	0,31	0,44	0,54	0,51

Vierumäki–Seesta

	He-Ky	He-Se	He-Vi	Vi-Ky	Vi-Se	Se-Ky
Viimeisen viiden minuutin matka-aikojen keskihajonta	-0,03	0,12	0,05	-0,20	-0,18	-0,05
Viimeisten 10 ajoneuvon matka-aikojen keskihajonta	-0,22	-0,07	-0,15	-0,25	-0,26	-0,21
Viimeisten 20 ajoneuvon matka-aikojen keskihajonta	-0,21	0,02	0,03	-0,27	-0,25	-0,21
Viimeisen minuutin matka- aikojen mediaani	0,38	0,22	0,44	0,27	0,40	0,29
Viimeisten 10 ajoneuvon matka-aikojen mediaani	0,38	0,21	0,44	0,39	0,49	0,39
Viimeisten 20 ajoneuvon matka-aikojen mediaani	0,38	0,22	0,43	0,41	0,52	0,40

Seesta–Kymijärvi

	He-Ky	He-Se	He-Vi	Vi-Ky	Vi-Se	Se-Ky
Viimeisen viiden minuutin matka-aikojen keskihajonta	-0,01	0,14	-0,02	-0,13	-0,09	-0,01
Viimeisten 10 ajoneuvon matka-aikojen keskihajonta	-0,08	0,00	-0,17	-0,28	-0,20	0,06
Viimeisten 20 ajoneuvon matka-aikojen keskihajonta	-0,07	0,02	-0,04	-0,29	-0,25	0,10
Viimeisen minuutin matka- aikojen mediaani	0,22	0,24	0,33	0,23	0,51	0,32
Viimeisten 10 ajoneuvon matka-aikojen mediaani	0,23	0,22	0,33	0,40	0,56	0,41
Viimeisten 20 ajoneuvon matka-aikojen mediaani	0,24	0,21	0,33	0,42	0,55	0,43

Kymijärvi–Vierumäki

	Ky-He	Ky-Vi	Ky-Se	Se-He	Se-Vi	Vi-He
Viimeisen viiden minuutin matka-aikojen keskihajonta	0,45	0,46	0,55	0,47	0,51	-0,07
Viimeisten 10 ajoneuvon matka-aikojen keskihajonta	-0,03	0,19	0,46	0,11	0,38	-0,24
Viimeisten 20 ajoneuvon matka-aikojen keskihajonta	0,04	0,34	0,60	0,22	0,45	-0,31
Viimeisen minuutin matka- aikojen mediaani	0,93	0,96	0,92	0,90	0,93	0,28
Viimeisten 10 ajoneuvon matka-aikojen mediaani	0,93	0,95	0,92	0,90	0,93	0,42
Viimeisten 20 ajoneuvon matka-aikojen mediaani	0,92	0,94	0,92	0,89	0,92	0,50

Kymijärvi–Seesta

	Ky-He	Ky-Vi	Ky-Se	Se-He	Se-Vi	Vi-He
Viimeisen viiden minuutin matka-aikojen keskihajonta	0,27	0,37	0,49	0,19	0,46	-0,05
Viimeisten 10 ajoneuvon matka-aikojen keskihajonta	0,03	0,17	0,28	0,02	0,33	-0,23
Viimeisten 20 ajoneuvon matka-aikojen keskihajonta	0,13	0,34	0,43	0,13	0,41	-0,28
Viimeisen minuutin matka- aikojen mediaani	0,82	0,88	0,94	0,61	0,79	0,28
Viimeisten 10 ajoneuvon matka-aikojen mediaani	0,82	0,87	0,94	0,61	0,79	0,40
Viimeisten 20 ajoneuvon matka-aikojen mediaani	0,81	0,85	0,93	0,61	0,79	0,44

Seesta–Heinola

	Ky-He	Ky-Vi	Ky-Se	Se-He	Se-Vi	Vi-He
Viimeisen viiden minuutin matka-aikojen keskihajonta	0,46	0,47	0,34	0,35	0,44	-0,10
Viimeisten 10 ajoneuvon matka-aikojen keskihajonta	0,07	0,29	0,29	0,09	0,36	-0,27
Viimeisten 20 ajoneuvon matka-aikojen keskihajonta	0,17	0,45	0,50	0,22	0,52	-0,34
Viimeisen minuutin matka- aikojen mediaani	0,85	0,94	0,60	0,91	0,96	0,33
Viimeisten 10 ajoneuvon matka-aikojen mediaani	0,85	0,94	0,60	0,92	0,95	0,47
Viimeisten 20 ajoneuvon matka-aikojen mediaani	0,84	0,93	0,60	0,91	0,95	0,53

Seesta–Vierumäki

	Ky-He	Ky-Vi	Ky-Se	Se-He	Se-Vi	Vi-He
Viimeisen viiden minuutin matka-aikojen keskihajonta	0,53	0,52	0,48	0,49	0,48	-0,09
Viimeisten 10 ajoneuvon matka-aikojen keskihajonta	0,03	0,31	0,43	0,11	0,34	-0,25
Viimeisten 20 ajoneuvon matka-aikojen keskihajonta	0,13	0,47	0,64	0,27	0,48	-0,32
Viimeisen minuutin matka- aikojen mediaani	0,94	0,95	0,83	0,94	0,97	0,32
Viimeisten 10 ajoneuvon matka-aikojen mediaani	0,94	0,95	0,83	0,94	0,97	0,46
Viimeisten 20 ajoneuvon matka-aikojen mediaani	0,93	0,94	0,83	0,94	0,96	0,51

Vierumäki–Heinola

	Ky-He	Ky-Vi	Ky-Se	Se-He	Se-Vi	Vi-He
Viimeisen viiden minuutin matka-aikojen keskihajonta	0,13	0,11	0,10	0,11	0,11	-0,11
Viimeisten 10 ajoneuvon matka-aikojen keskihajonta	-0,01	0,07	0,11	-0,01	0,12	-0,10
Viimeisten 20 ajoneuvon matka-aikojen keskihajonta	0,00	0,11	0,17	0,05	0,15	-0,12
Viimeisen minuutin matka- aikojen mediaani	0,29	0,31	0,29	0,30	0,30	0,23
Viimeisten 10 ajoneuvon matka-aikojen mediaani	0,29	0,30	0,29	0,29	0,29	0,29
Viimeisten 20 ajoneuvon matka-aikojen mediaani	0,29	0,30	0,30	0,29	0,29	0,31

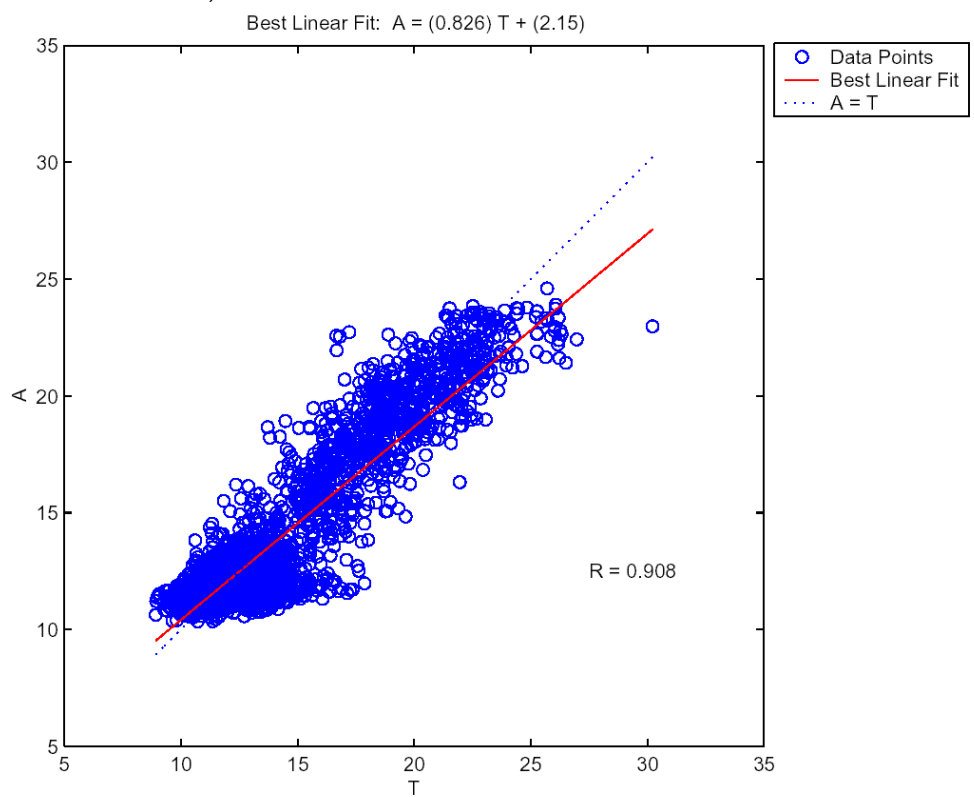
MITATUT JA NEUROMALLEILLA ENNUSTETUT MATKA-AJAT

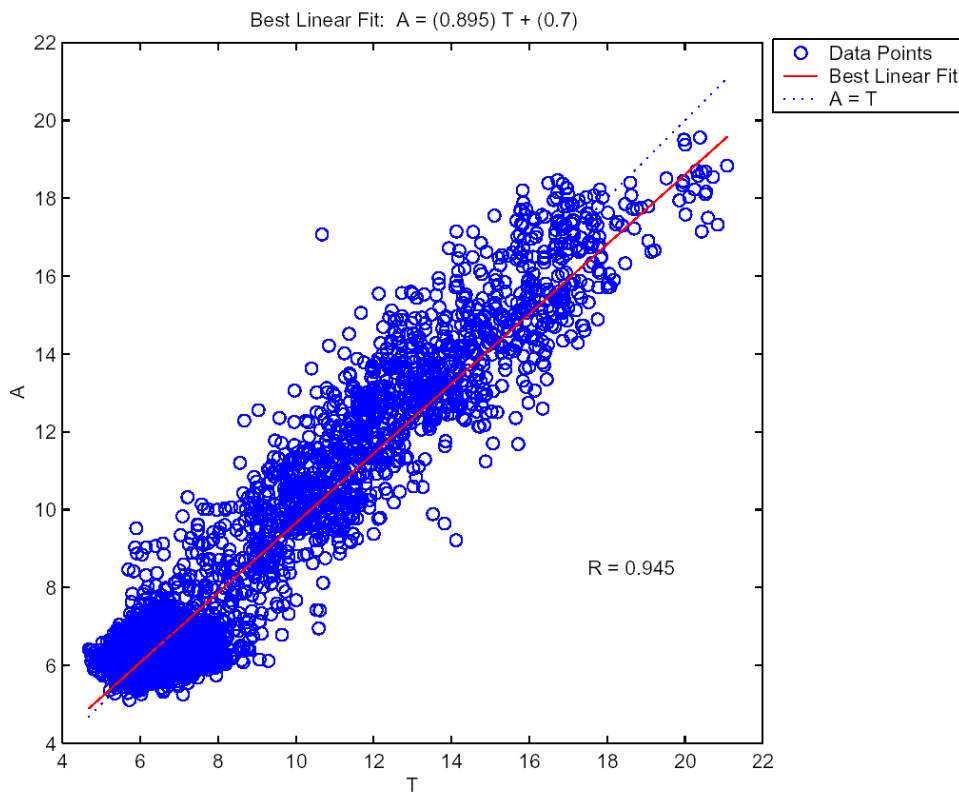
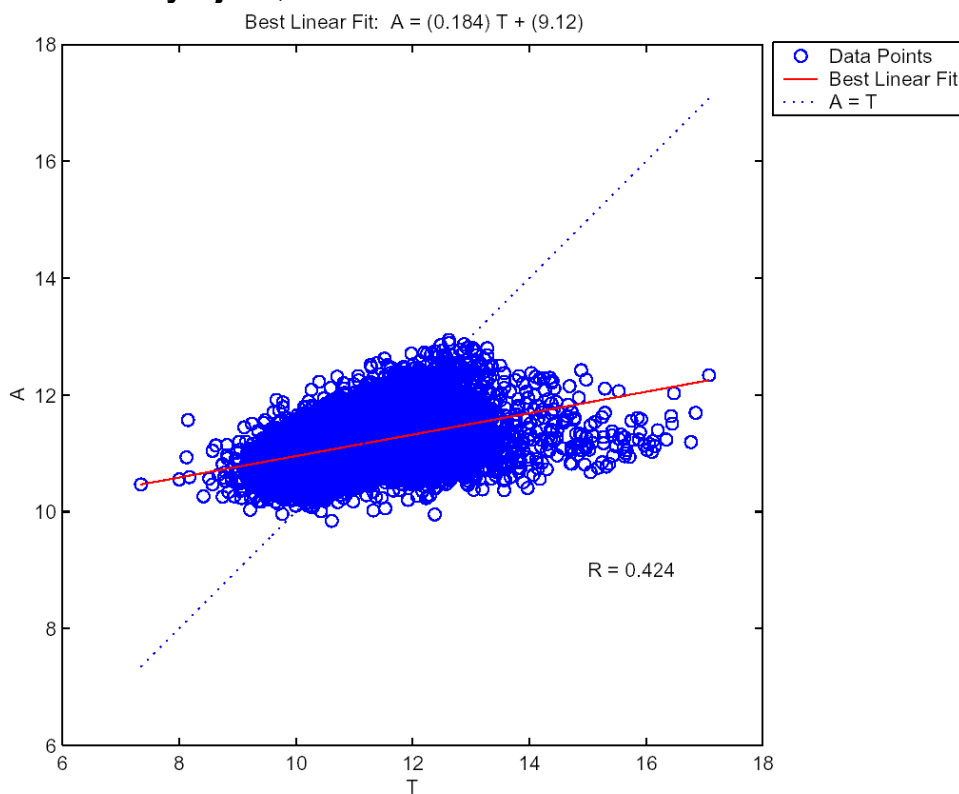
Valtatie 4

Kuvissa on esitetty eri linkeiltä mitatut (vaaka-akseli) ja ennustetut (pystyakseli) matka-ajat. Kuvan yläpuolelle on merkitty ennustelinkki ja käytetty aika-sarjan pituus.

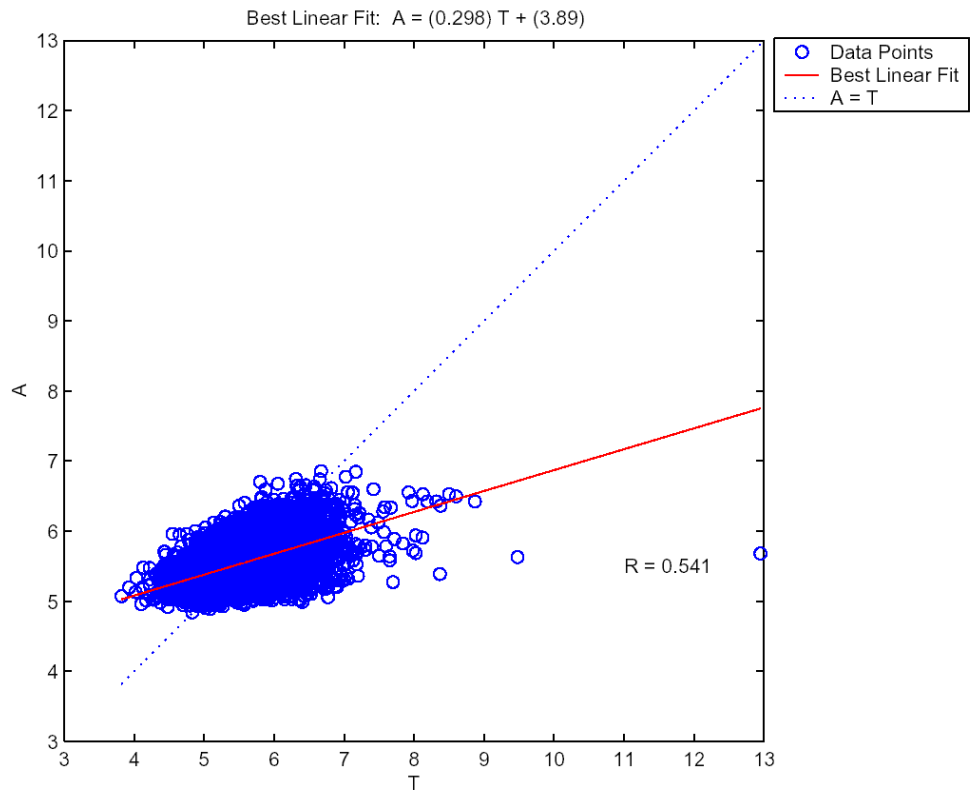
Etelään

Heinola–Seesta, 4 minuuttia

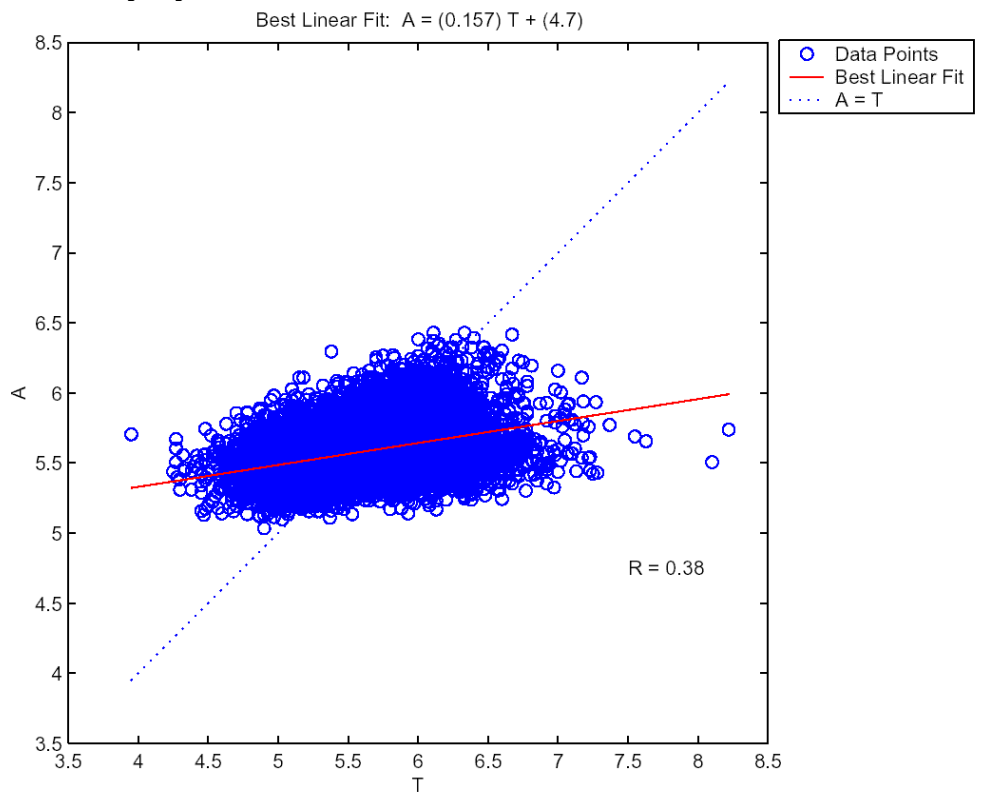


Heinola–Vierumäki, 3 minuuttia**Vierumäki–Kymijärvi, 4 minuuttia**

Vierumäki–Seesta, 4 minuuttia



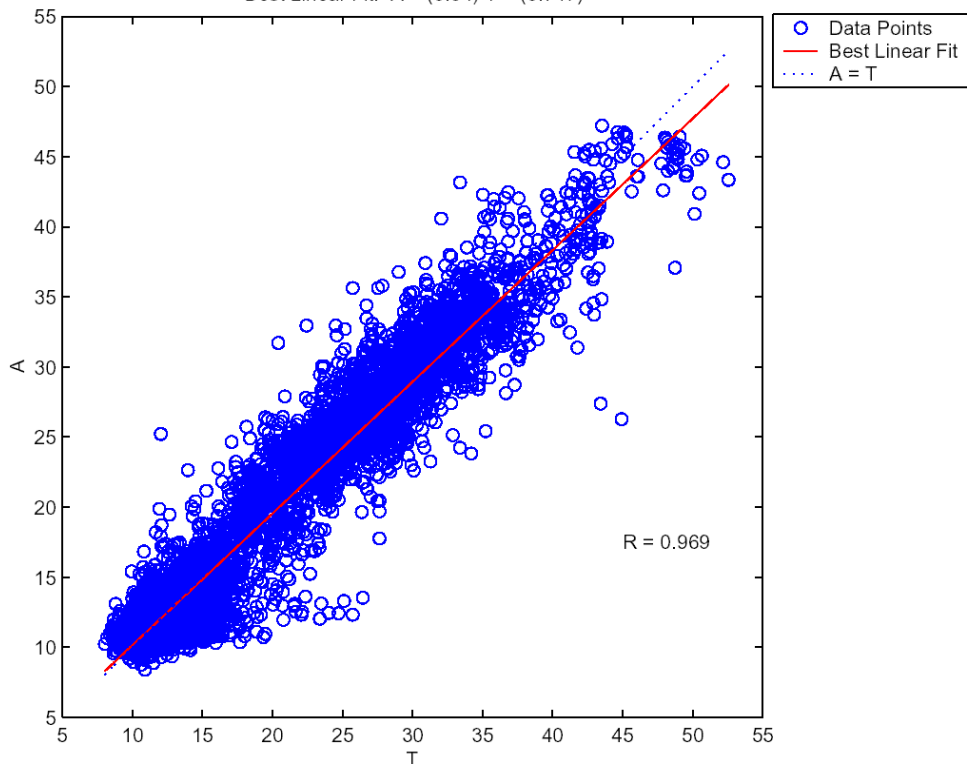
Seesta–Kymijärvi, 4 minuuttia



Pohjoiseen

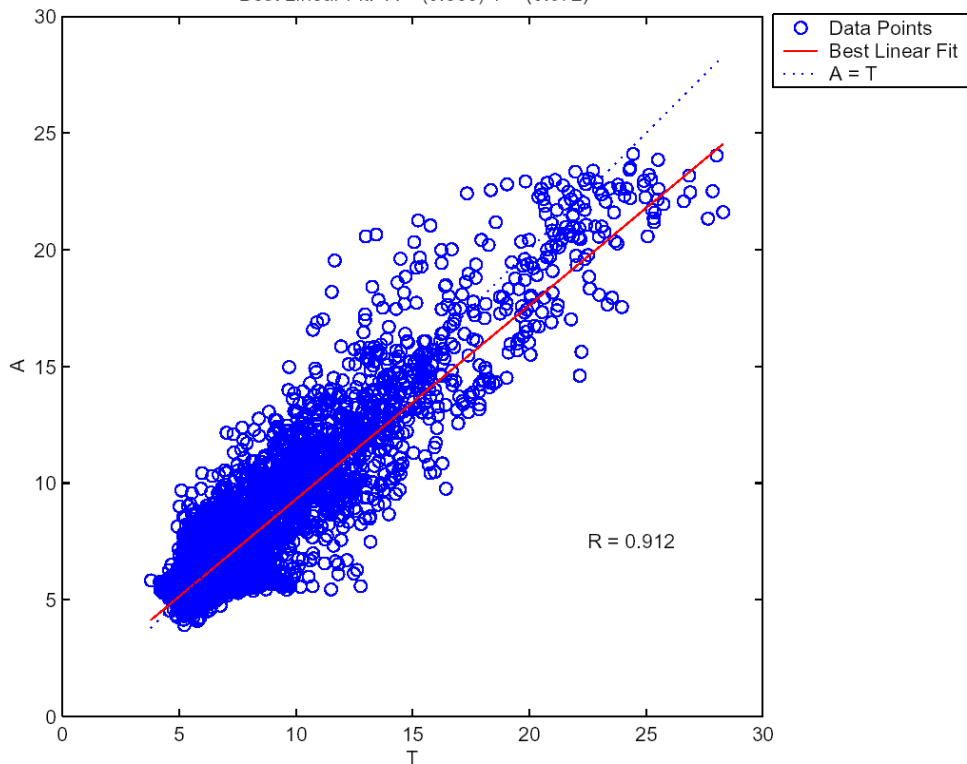
Kymijärvi–Vierumäki, 5 minuuttia

Best Linear Fit: $A = (0.94) T + (0.747)$

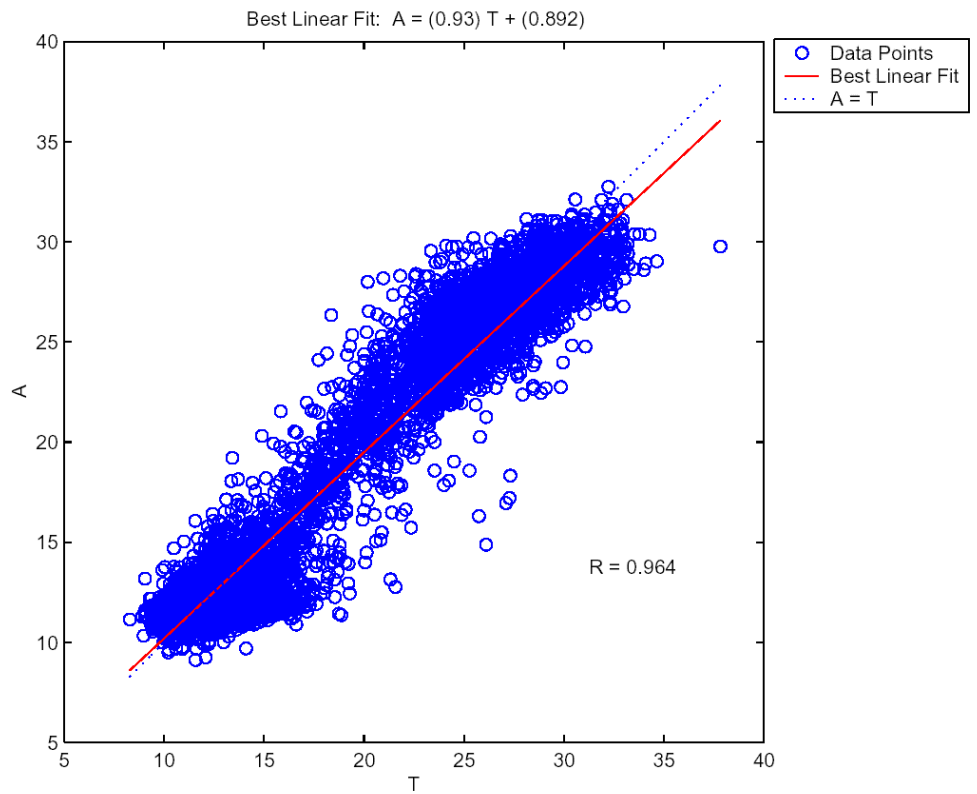


Kymijärvi–Seesta, 3 minuuttia

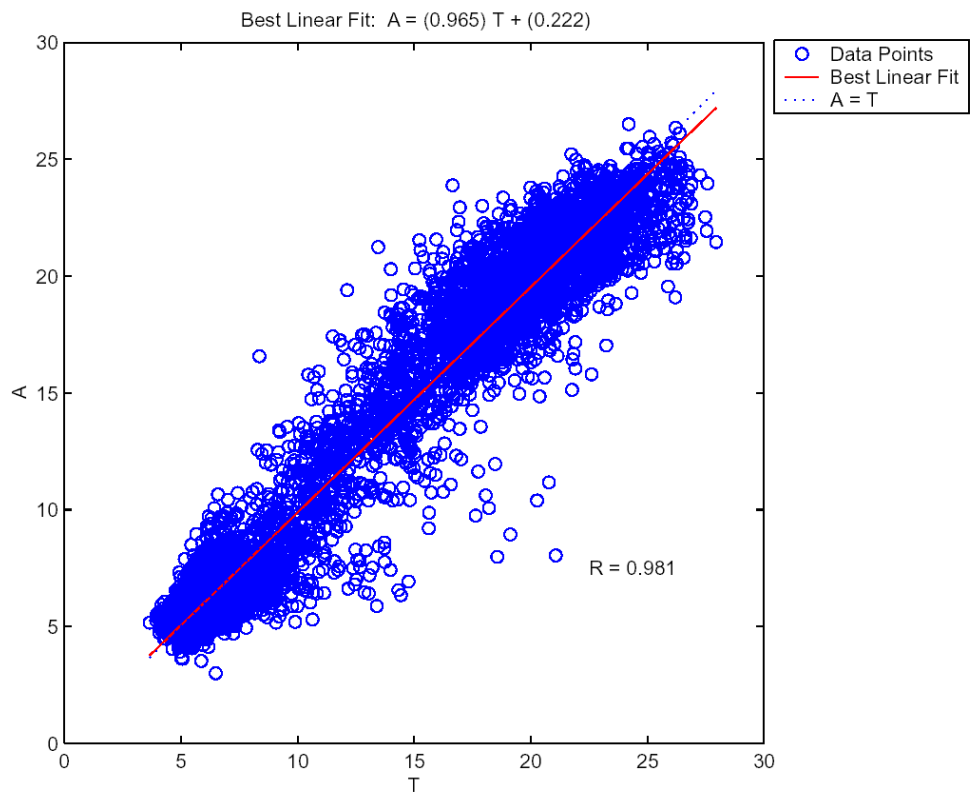
Best Linear Fit: $A = (0.833) T + (0.972)$



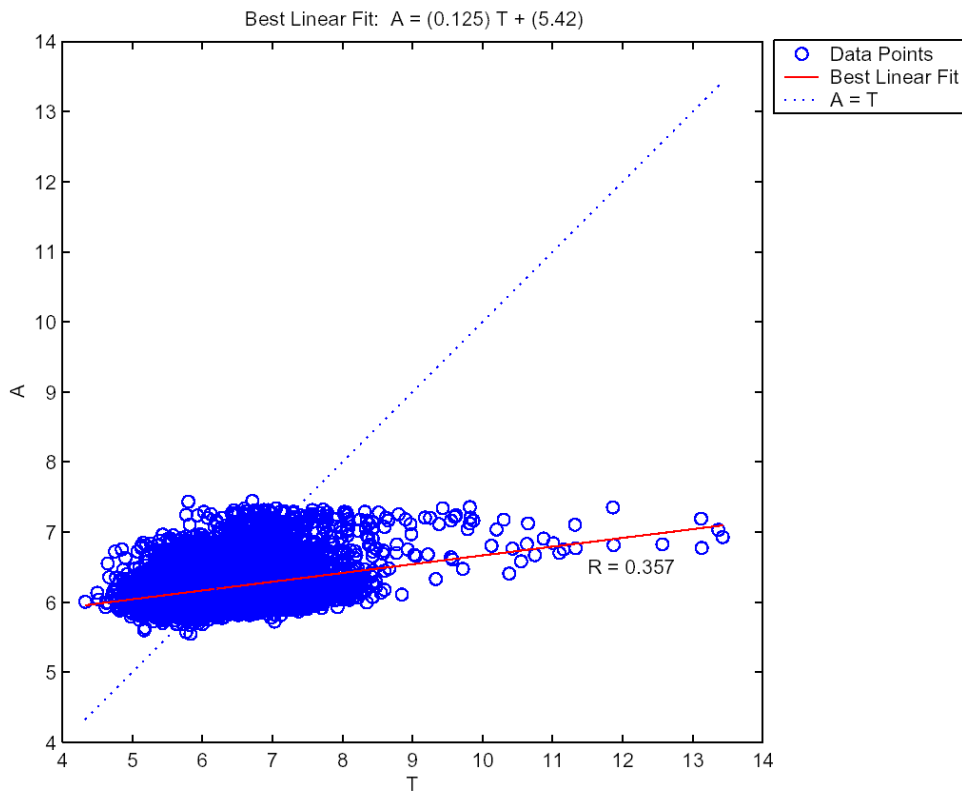
Seesta–Heinola, 5 minuuttia



Seesta–Vierumäki, 5 minuuttia



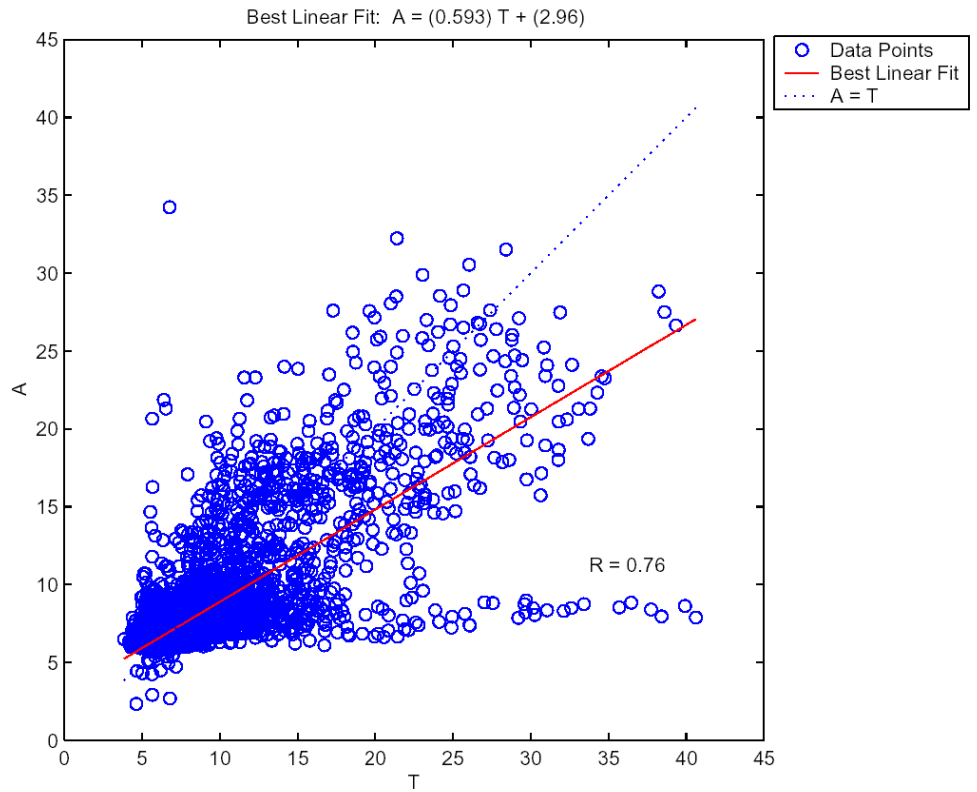
Vierumäki–Heinola, 5 minuuttia



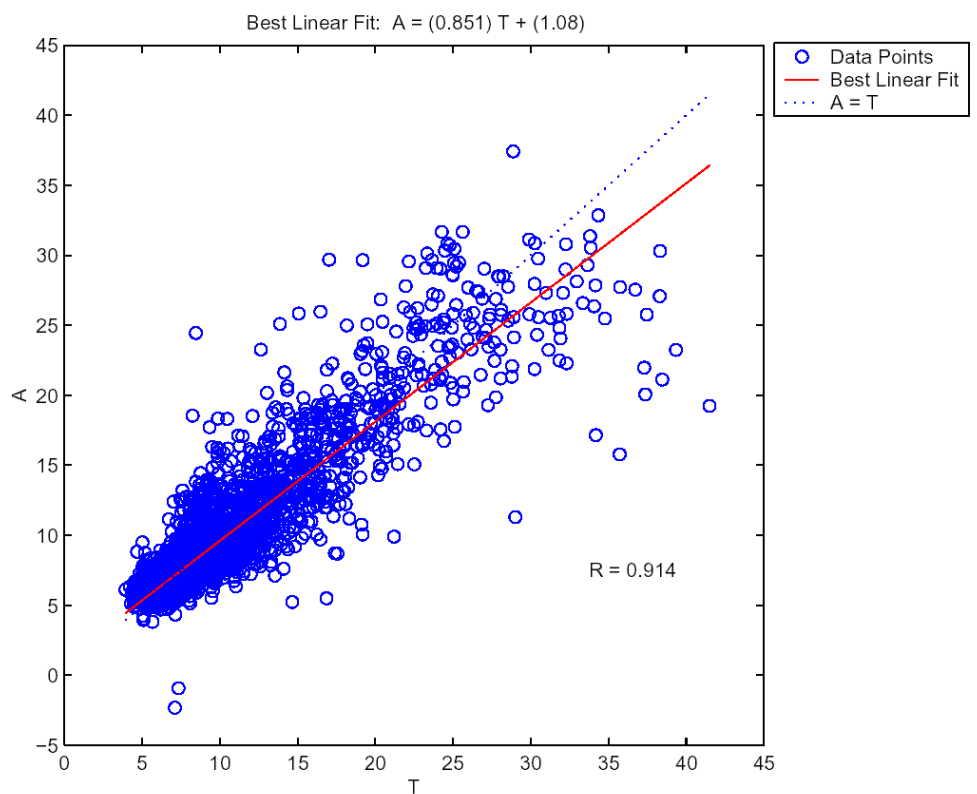
Kehä I

Itään

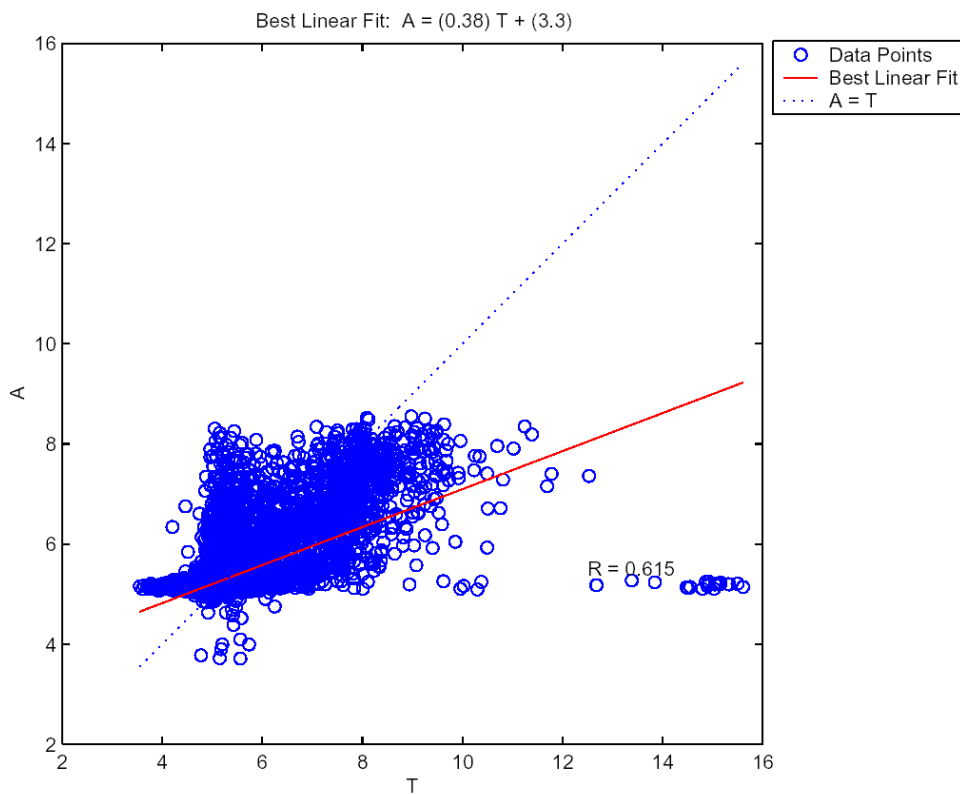
Otaniemi-Konala, estimaattimalli



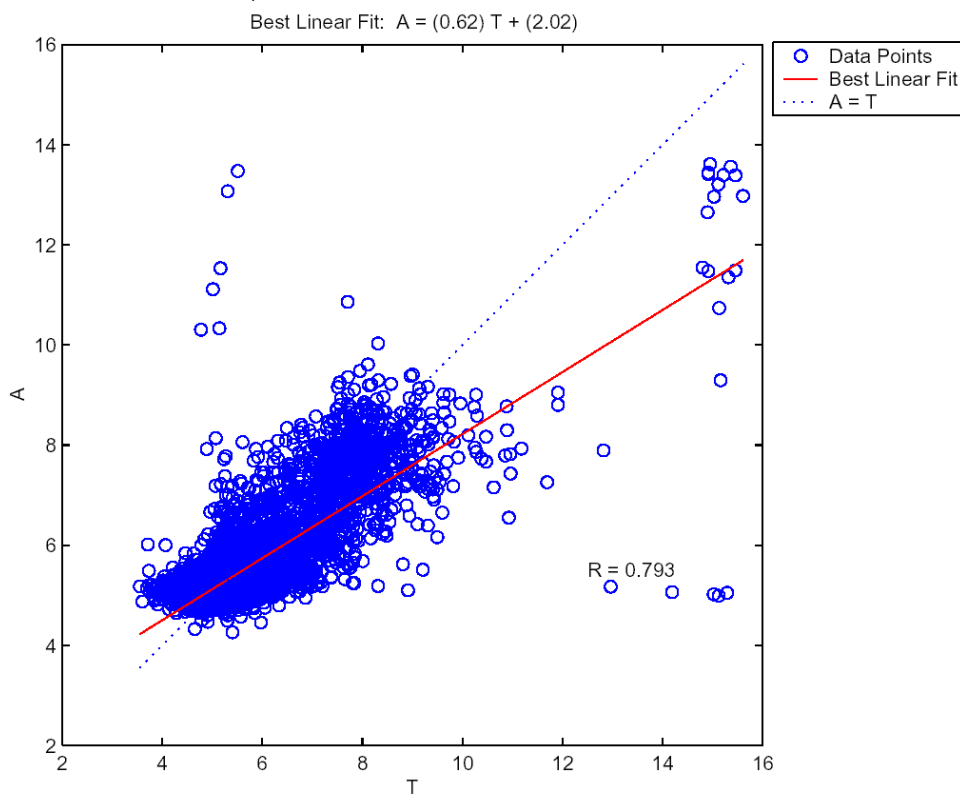
Otaniemi-Konala, ennustemalli



Konala-Pukinmäki, estimaattimalli

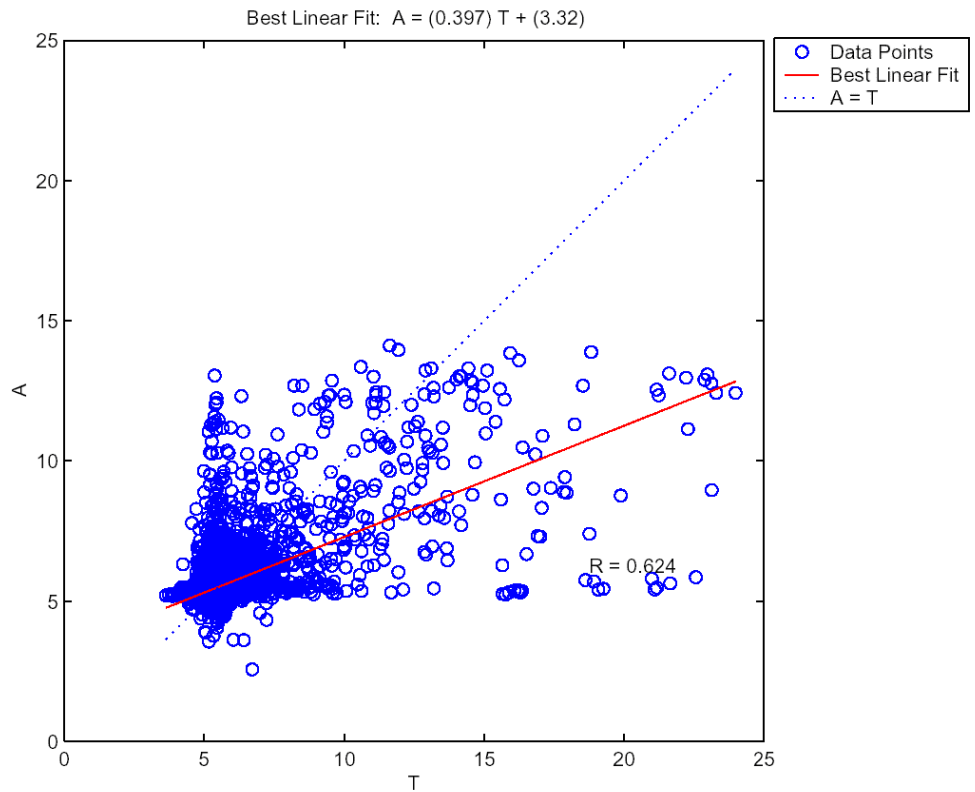


Konala-Pukinmäki, ennustemalli

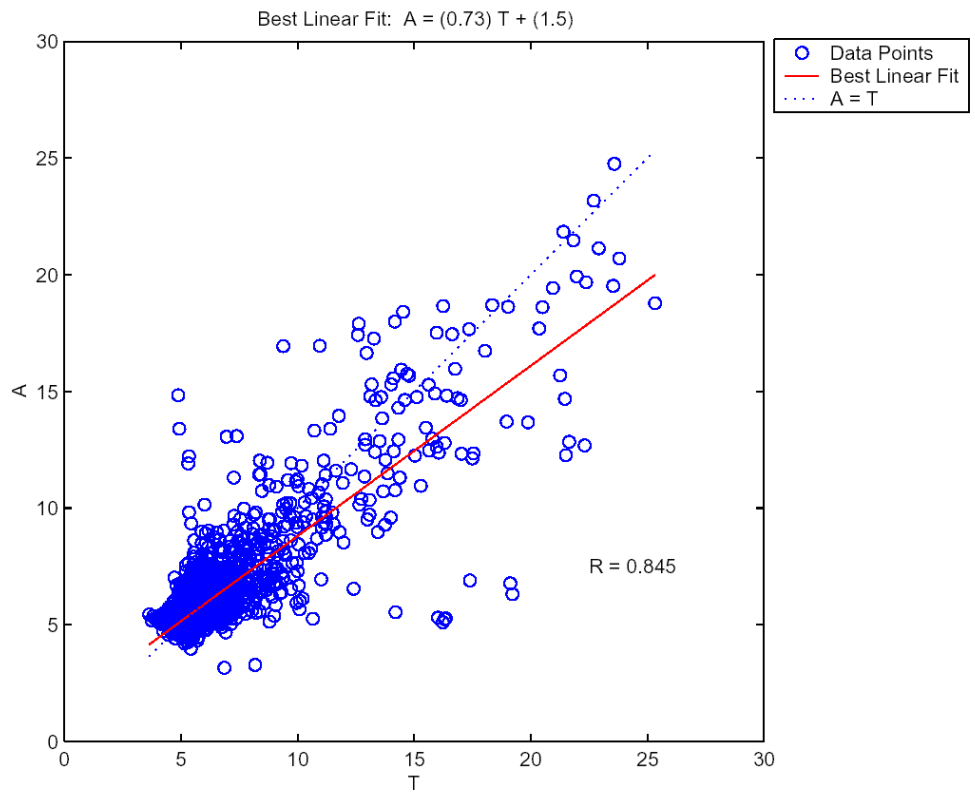


Länteen

Pukinmäki- Konala, estimaattimalli

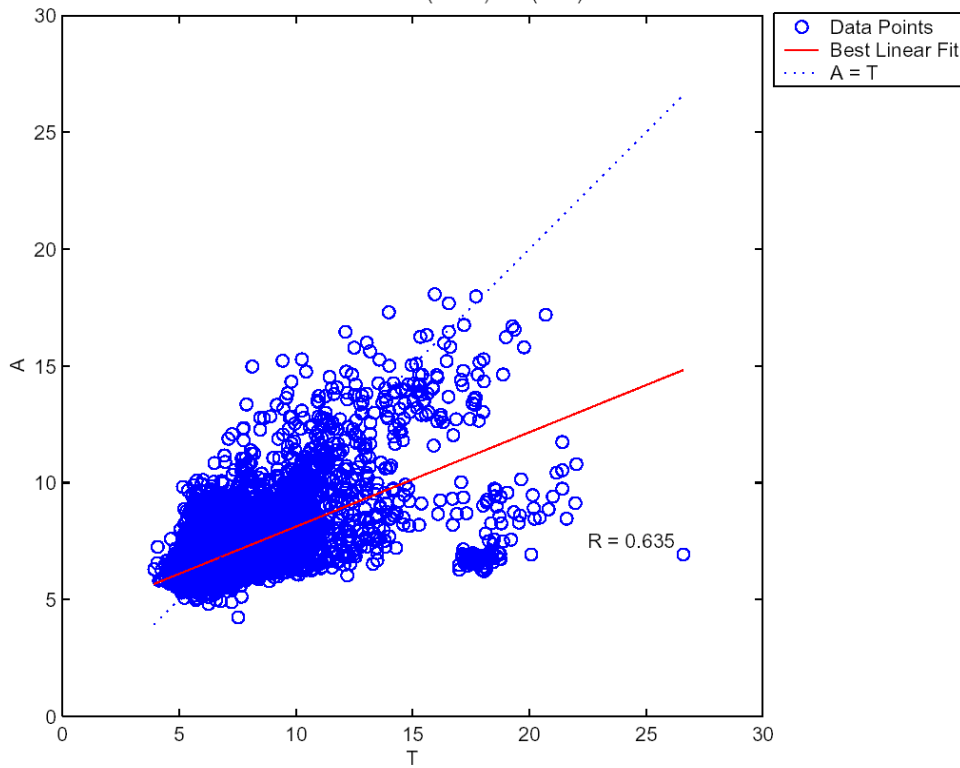


Pukinmäki- Konala, ennustemalli



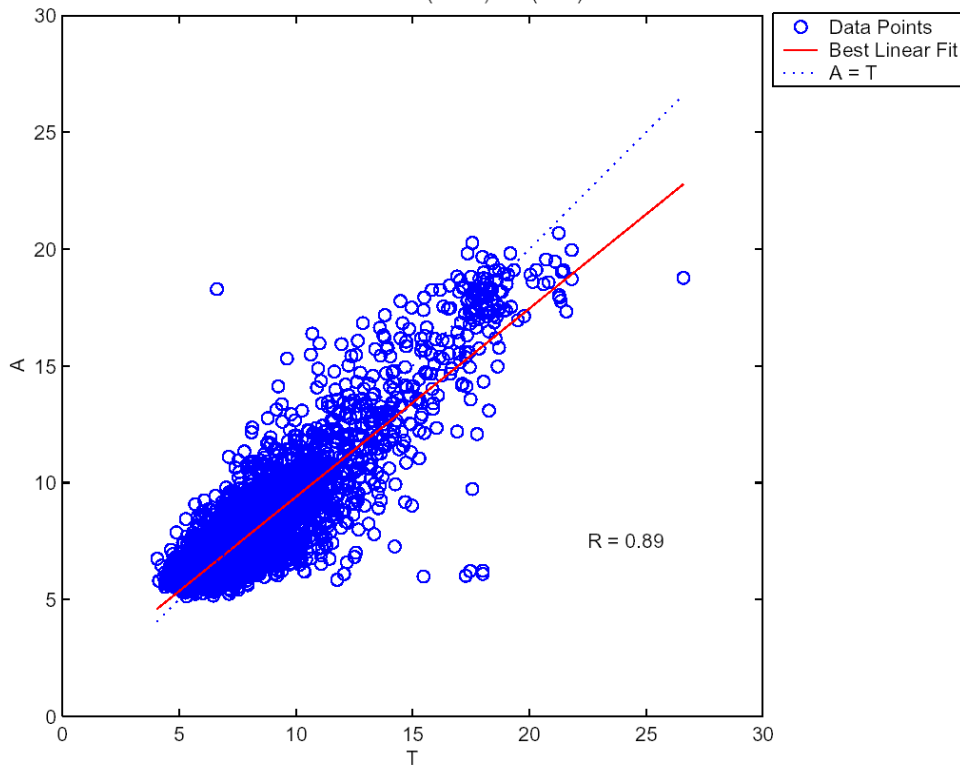
Konala–Otaniemi, estimaattimalli

Best Linear Fit: $A = (0.404) T + (4.08)$



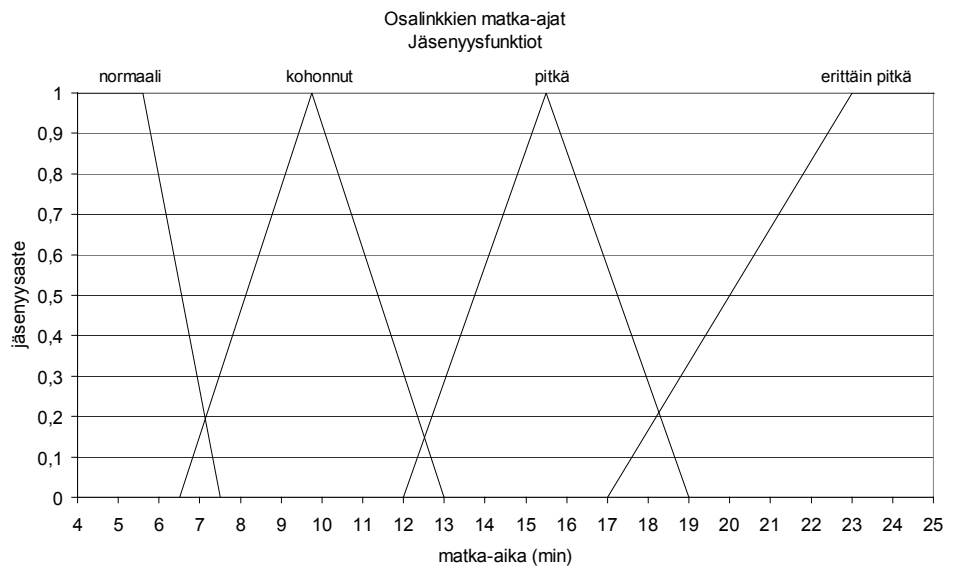
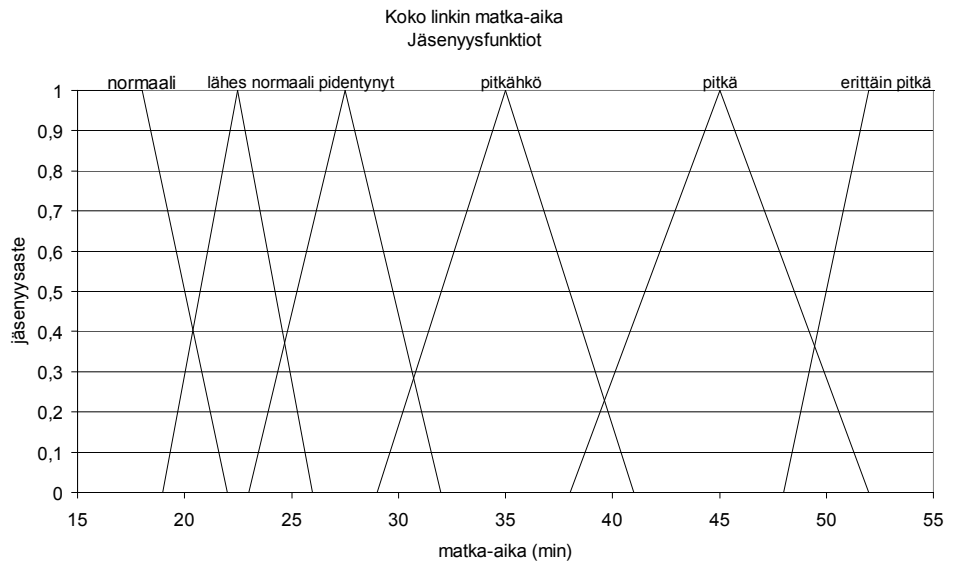
Konala–Otaniemi, ennustemalli

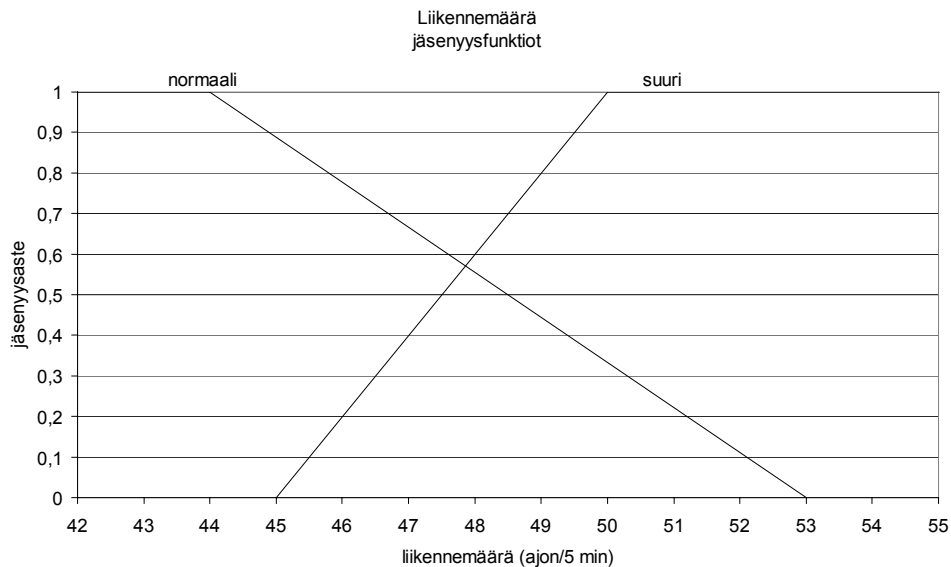
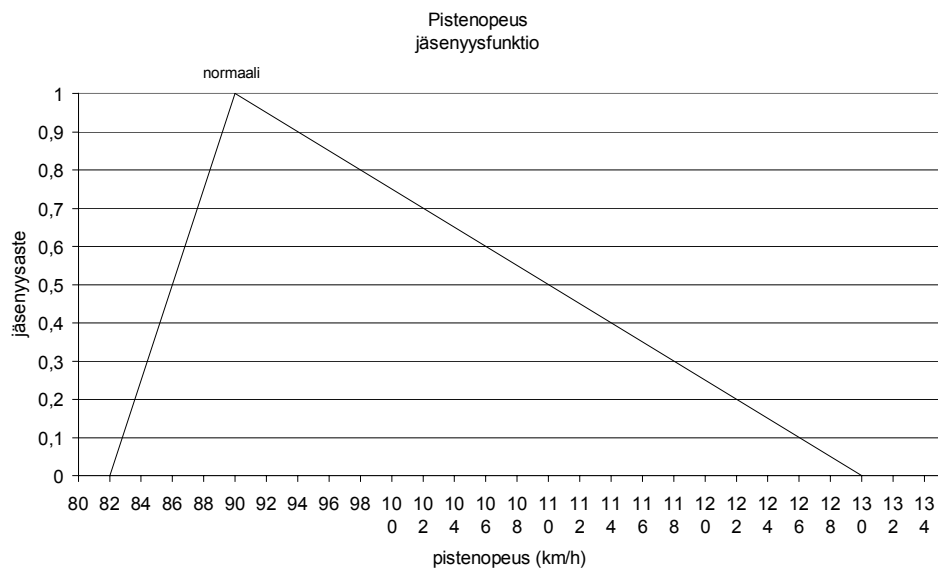
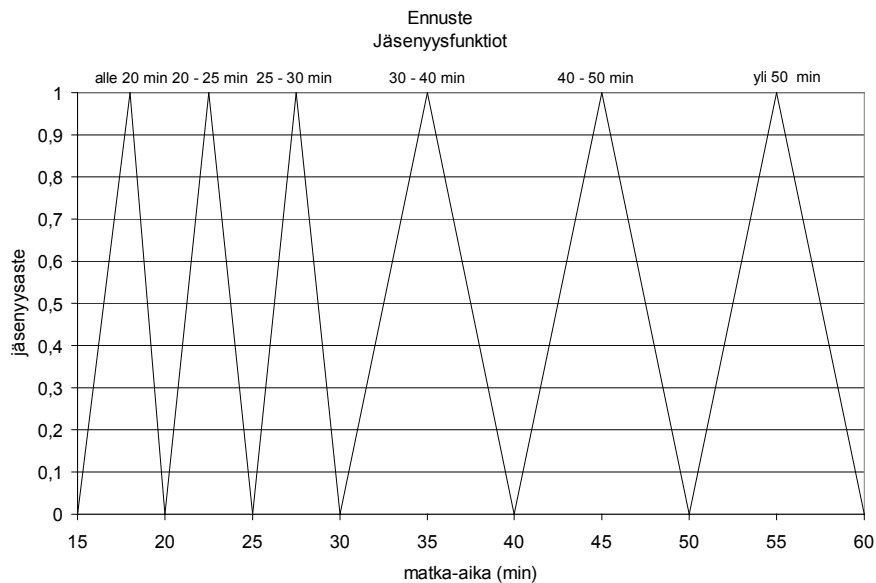
Best Linear Fit: $A = (0.807) T + (1.32)$



SUMEIDEN ENNUSTEMALLIEN JÄSENYYSFUNKTIOT JA SUMEAT PÄÄTTELYSÄÄNNÖT

Jäsenyyssfunktiot





Sumeat päättelysäännöt

- 1) Jos osalinkkien matka-ajat ovat normaalit, niin ennuste < 20 min.
- 2) Jos kahden osalinkin matka-ajat ovat erittäin pitkät, niin ennuste on > 50 min.
- 3) Jos yhden osalinkin matka-aika on erittäin pitkä ja yhden osalinkin matka-aika on kohonnut, niin ennuste on 30–40 min.
- 4) Jos yhden osalinkin matka-aika on erittäin pitkä ja yhden osalinkin matka-aika on normaali ja koko linkin matka-aika on pitkä, niin ennuste 30–40 min.
- 5) Jos yhden osalinkin matka-aika on kohonnut ja yhden osalinkin matka-aika on normaali ja koko linkin matka-aika on normaali, niin ennuste on 20–25 min.
- 6) Jos yhden osalinkin matka-aika on kohonnut ja yhden osalinkin matka-aika on normaali ja koko linkin matka-aika on lähes normaali, niin ennuste on 25–30 min.
- 7) Jos yhden osalinkin matka-aika on erittäin pitkä ja yhden osalinkin matka-aika on pitkä, niin ennuste on 40–50 min.
- 8) Jos yhden osalinkin matka-aika on pitkä ja yhden osalinkin matka-aika on normaali ja koko linkin matka-aika on pitkä, niin ennuste on 25–30 min.
- 9) Jos yhden osalinkin matka-aika on pitkä ja yhden osalinkin matka-aika on kohonnut, niin ennuste on 30–40 min.



ISSN 1457-9871
ISBN 951-726-863-7
TIEH 3200739