

# **Teiden kunnan ennustemallit kirjallisuudessa sekä riskianalyysiin soveltuvan ennustemallin kehittäminen Suomen tieverkostolle**

## **Loppuraportti**

**8.5.2009**

Toimeksiantajat: Pöyry Infra Oy  
Tiehallinto

Projektityöryhmä 3: Lars Baarman (projektipäällikkö)

Olli Eskola

Noora Hyttinen

Janne Laitonen

Juha Martikainen

Martti Paatela

Tommi Rantanen

## Sisällysluettelo

1	Johdanto.....	4
Osa I: Tieverkon kunnan stokastinen ennustemalli ja sen soveltaminen riskianalyyysiin.....		
2	Mallin kehittäminen .....	6
2.1	Datan tarkastelu.....	6
2.2	Mallin lähtökohdat.....	7
2.3	Mallin muodostaminen .....	9
2.4	Todennäköisyyksien estimointi.....	10
2.5	Osaverkkojen luokittelutekijöiden valinta .....	13
2.6	Tulokset.....	14
2.7	Uuden mallin tuloksien vertailu Tiehallinnon nykyisiin ennusteisiin....	16
2.8	Pohdintaa ja johtopäätökset .....	17
3	Ennustemallin käyttö riskianalyyysissä .....	21
3.1	Riskianalyysi .....	21
3.2	Tulokset.....	28
3.2.1	Kuntoluokitusjakauman kehittyminen tasaisuuden perusteella.....	29
3.2.2	Kuntoluokitusjakauman kehittyminen urasyvyyden perusteella ...	33
Osa II: Tieverkostojen ennustemallit kirjallisuudessa .....		
4	Teiden kunnossa pitojärjestelmä, PMS.....	40
5	Bayes-verkot .....	41
6	Markov-mallit .....	43
6.1	Yleistä.....	43
6.2	Markov-mallit teiden kunnan ennustamisessa .....	43
6.3	Tilasiirtotodennäköisyysmatriisin määrittelemineen .....	44
6.4	Tuloksia.....	47
7	Neuroverkot .....	48
7.1	Neuroverkon mallintamisen historia .....	48
7.2	Keinotekoisien neuroverkon malli .....	48
7.2.1	Neuroverkkomallin muodostaminen.....	51

7.2.2	Neuroverkkomallin heikkoudet.....	52
7.2.3	Neuroverkkomallin vahvuudet .....	53
7.3	Neuroverkkomallin koulutus.....	54
7.3.1	Vaihtoehto backpropagation-koulutusalgoritmille.....	56
7.4	Kokemuksia ja oppeja aikaisemmista tutkimuksista .....	57
7.5	Tutkimusasetelma neuroverkkokatsaukselle kirjallisuuskatsauksen perusteella .....	58
8	Alustava tarkastelu neuroverkkomallin käytöstä Suomen tieverkon ennustemallina .....	59
9	Yhteenveto .....	62
10	Viitteet .....	64
	LIITE 1 – Putoamistodennäköisyyksien kuvaajat.....	67
	LIITE 2 – Java koodi havaintojen keräämiseen .....	75

## 1 Johdanto

Tieinfrastruktuuria voidaan pitää yhtenä nyky-yhteiskunnan toimivuuden perusedellytyksenä. Tämän vuoksi on tärkeää, että tieverkkoa pidetään hyvässä kunnossa ja sen kunnan muuttumista tarkkaillaan säännöllisesti. Tieverkoston kuntoon vaikuttaa monia ulkoisia tekijöitä, jotka ovat toisaalta luonteeltaan hyvin stokastisia. Tämä tekee tieverkoston tarkan tutkimisen monimutkaiseksi ja osaltaan myös lisää tieverkoston valvontakäytäntöjen kehittämisen tärkeyttä. Tämä Operaatiotutkimuksen projektityöseminaarille tehty harjoitustyö pyrkii osaltaan tarjoamaan oman panostuksensa Suomen tieverkoston tilan kehitysmallin laatimiselle ja käyttöönnotolle. Tämän harjoitustyön toimeksiantajana toimii Pöyry Infra Oy ja Tiehallinto.

Tieverkoston kunnan valvonta on vuosien kuluessa tullut yhä systemaattisempaa ja keinojen kirjo on kasvanut muun tekniikan ja tieteen kehityksen mukana. Nykyään käytetään lähinnä malleja, jotka ottavat huomioon erilaisten ulkoisten tekijöiden stokastisen luonteen ja pyrkivät tekemään ennusteita teiden kunnan kehitykselle tulevaisuudessa. Käytettyjä menetelmiä ovat muun muassa Bayes-verkot, Markov-ketjut ja neuroverkot. Erityisesti tietokoneiden laskentatehon kasvaessa on saavutettu suurempia tarkkuuksia tieverkoston kunnan ennustemalleille ja kyetty ottamaan useimpia muuttuvia tekijöitä huomioon ennusteissa. Tämän harjoitustyön tarkoituksena on ensinnäkin kartoittaa, miten tieverkoston kuntoa on mallinnettu ja miten sen muuttumista ennustettu aikaisemmin perehtyen asianmukaiseen kirjallisuuteen ja lisäksi muodostaa malli, joka kuvaa Suomen tieverkoston kehitystä sekä kuntojakautumia.

Suuresta datamäärästä johtuen tieverkoston tarkka tutkiminen ja ennusteiden laatiminen vaatii paljon aikaa. Runsaasta datasta tulee myös hankaluuksia monimutkaisempien mallinnustekniikoiden soveltamisessa, esimerkiksi neuroverkkomallien käyttö juuri tämäntyyppisen mallintamiseen voi olla ja onkin varsin haasteellista todella suurten laskentatehovaatimusten vuoksi. Tämän takia muiden vaihtoehtoisten menetelmien, joiden laskentatarkkuus on verrattain hyvä, kehittäminen on todella tärkeää. Mallien on oltava käyttökelpoisia, mutta niiden ennusteominaisuudet eivät saa olla huonompia kuin monimutkaisempien menetelmien.

Tässä työssä muodostamme Suomen tieverkostolle eräänlaisen Markov-ketjun sovellutuksen. Markov-mallissamme on kaksi ulottuvuutta, teiden kuntoluokka sekä ikä. Mallin tarkoituksena on esittää teiden kuntoluokkien muuttumista ajan suhteen. Tarkoituksenamme on määrittää todennäköisyydet, joilla tietyn ikäinen tie, joka on tietyssä kuntoluokassa tällä hetkellä, siirtyy alempaan kuntoluokkaan. Täten saadaan malli, joka kuvaa teiden kunnan muuttumista ajan kuluessa ja joka antaa suuntaviivoja tulevaisuuden tienkunnostus

toimenpiteille. Työn toisessa osassa teemme kokeiluin neuroverkkojen avulla. Kokeilemme neuroverkkomallin soveltuvuutta kuvaamaan Suomen tieverkoston muuttumista ajan kuluessa ja vertaamme mallin toimivuutta ja hyvyttä varsinaiseen malliimme.

Harjoitustyö jakaantuu kahteen osaan: Ensimmäisessä osassa muodostamme ensisijaisen mallimme ja tutkimme sen hyvyttä. Toisessa osassa käsittelemme laajemmin aiheeseen liittyvää kirjallisuutta ja tutkimme, miten kirjallisuudessa tieverkostojen kuntoa on pyritty mallintamaan. Lisäksi teemme kokeilun Suomen tieverkoston datalla neuroverkkomallin avulla. Lopuksi esittelemme tärkeimmät johtopäätökset ja implikaatiot muun muassa tulevaisuuden teiden kunnan ennustamiselle ja tieverkoston yleisen tilan kartoittamiselle.

# Osa I: Tieverkon kunnan stokastinen ennustemalli ja sen soveltaminen riskianalyysiin

## 2 Mallin kehittäminen

Päällystetyn tien rappeutuminen ilmenee pituus- (IRI) ja poikkisuuntaisena (tien urat) epätasaisuutena sekä pinnan rikkoutumisena. IRI on lyhenne sanoista international roughness index ja sen avulla saadaan määrällistä tietoa teiden karkeudesta. Rappeutumista aiheuttavat yleisesti liikenteen kuormitus, ilmasto ja pohjamaan rakennekerrosten sekä päällysteen ominaisuudet. (*Tieomaisuuden yhtenäinen kuntoluokitus, Tiehallinnon selvityksiä 57/2005*)

Tässä osassa (osa I) esitetään projektin aikana tehty tiestön rappeutumista kuvaava matemaattinen malli ja sen osalta keskeiset tulokset. Lopuksi tarkastelemme myös ennustemallin soveltamista riskianalyysiin ja teemme ennusteiden pohjalta riskianalyysin tulevaisuutta varten. Aluksi tarkastelemme projektityöryhmälle toimitettua dataa, jonka jälkeen esitellään mallin rakenne ja laskentamenetelmät. Luokittelutekijöiden valintaa ja jaotteluprosessia perustellaan alaluvussa "Osaverkkojen luokittelutekijöiden valinta" ja lopulliset valinnat esitellään luvussa "Tulokset". Mallin ja osatulosten pohdintojen jälkeen sovellamme mallin tuloksia riskianalyysiin.

### 2.1 Datan tarkastelu

Suomen tieverkko on toimitetussa csv-muotoisessa tekstitiedostossa jaettu sadan metrin tiepätkiin, joista ilmoitetaan erinäisiä teknisiä tietoja, muun muassa tiepiiri, nopeusrajoitus, keskimääräinen vuorokausiliikenne (KVL), päällysteluokka, ikä sekä uran ja tasaisuuden (IRI) kuntoluokat yhdeltä mittauskerralta. Teiden kunnan mittauksesta saadaan millimetritulokset tiepätkän urien syvyydelle ja vastaavasti myös tasaisuudelle. Näiden tulosten perusteella annetaan tiepätkille kuntoluokat, riippuen minkä rajan sisällä mittaus on. Kuntoluokka viisi on korkein kuntoluokka. Alhaisimmat kuntoluokat 1 ja 2 merkitsevät jo, että tie on huoltotoimenpiteen tarpeessa.

Tiedosto, kuten myös tämä projekti, rajoittuu käsittelemään vain päällystettyjä asfalttibetoni- (AB) sekä pehmeitä asfalttibetoni- (PAB) autoteitä. Kukin tiepätkä on omalla rivillään, joten tiedosto on kohtuullisen suuri sisältäen reilut puoli miljoonaa riviä. Joidenkin tiepätkien osalta oli käytössä vain osa tiedoista, joten kyseiset rivit jouduttiin hylkäämään. Puutteet olivat kuitenkin niin pieniä, että ne eivät aiheuttaneet suurempia ongelmia.

Alkuperäinen idea oli mallintaa tiestön rappeutumista Bayes-verkkojen avulla ilman aitoa historiadataa. Kantavana ajatuksena tässä lähestymistavassa oli, että samanlaiset eri-ikäiset tiepätkät samaistetaan yhdeksi ”virtuaalitieksi”, jonka rappeutumishistoria tiedettäisiin eri-ikäisten tiepätkien ansiosta. Dataan tutustumisen myötä kuitenkin huomattiin, että kuntoluokitusten muutos ei ollut odotettua, vaan vaihteli hyvin voimakkaasti. Tästä syystä otettiin käyttöön myös aikaisempi mittaustulos eli kunkin tiepätkän osalta käytettiin kahta mittaustulosta, joiden väli on muutamia vuosia. Uuden datan myötä mallinnuksessa voitiin käyttää oikeaa, vaikkakin lyhyttä, aikasarjaa. Samalla myös mallinnustapa muuttui, mutta samanlaiset tiepätkät samaistetaan ja luokitellaan edelleen samaan osaverkkoon.

## 2.2 Mallin lähtökohdat

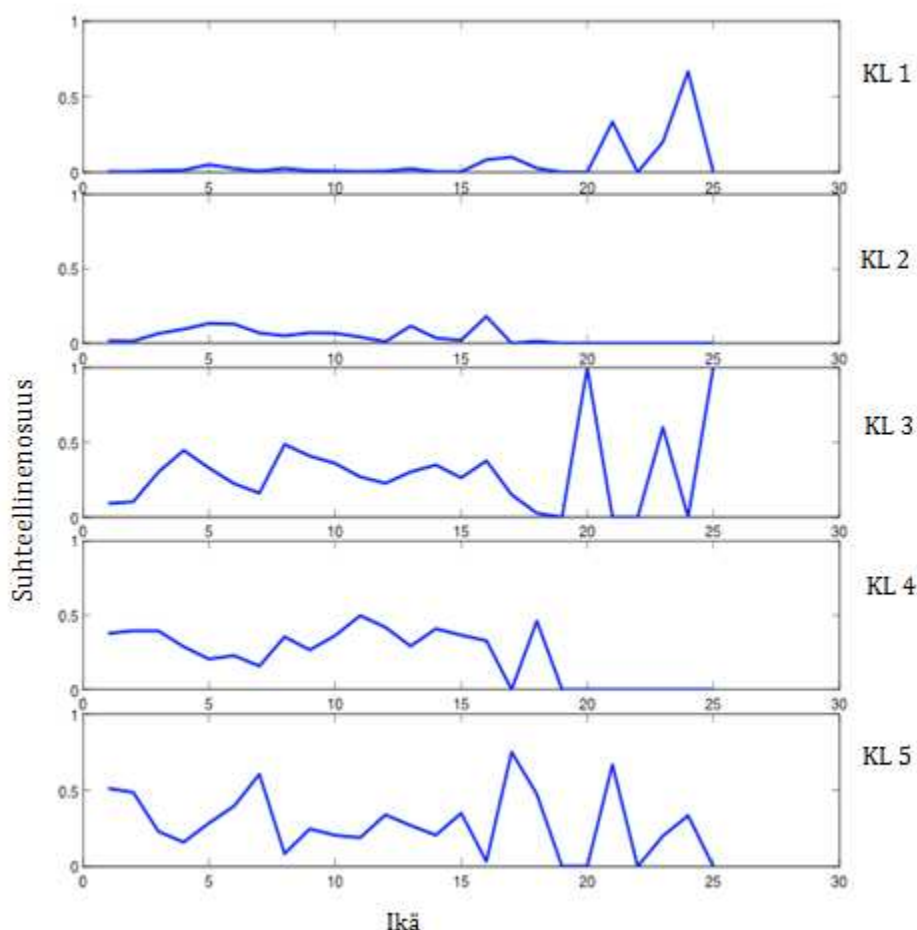
Tässä luvussa tarkastellaan kuntoennustemallin rakentamiseen valitun lähestymistavan järkevyyttä ja matemaattisia perusteita. Lisäksi käsitellään alkuperäiseen lähestymistapaan perustuvan mallin matemaattinen validointi, jonka perusteella ensimmäinen suunniteltu malli hylättiin.

Työn yhtenä matemaattisesti haastavana tarkoituksena oli rakentaa malli ja riskiarvio tieverkon kunnan ennustamiseen. Tässä alkuperäinen ajatus oli käyttää mallinnukseen yhden mittauskerran dataa. Tällöin ajateltaisiin, että eri-ikäisiä, mutta muuten samanlaisia teitä yhteen kokoamalla voitaisiin luoda näennäinen historiatieto tietyn tietyypin kunnan kehityksestä.

Aiottu lähestymistapa loi vastaavasti merkittäviä odotuksia ja vaatimuksia yhden mittauskerran datalle; tapauksia eri osaverkoista ja kaikista tarkasteltavista ikäluokista vähintäänkin tulisi olla riittävästi. Alkuperäisen datan validoinnissa ja datan viitoittamisessa johtopäätöksissä oli kuitenkin selviä ristiriitaisuuksia tiestön odotetun käyttäytymisen ja toimeksiantajan asiantuntemuksen perusteella tiedettyjen lainalaisuuksien kanssa. Toimeksiantajan ja kurssihenkilökunnan kanssa käytyjen keskustelujen pohjalta oli alkuperäisen lähestymistavan perustana oletus, että kaikki samaan osaverkkoluokkaan kuuluvat, mutta eri-ikäiset teiden havainnot yhtenä ajanhetkenä antaisivat luotettavaa tietoa osaverkon teiden kunnan kehityksestä. Tästä voidaan edelleen johtaa tarkistettavia lisäoletuksia. Esimerkiksi kyseinen päättelytapa indikoi, että ikäjakauman ensimmäisen vuoden kohdalla havaittaisiin vain kuntoluokkia neljä ja viisi.

Validoitaessa dataa edellä mainitun oletuksen mukaan, huomattiin seuraavat oletuksen kyseenalaistavat tai kumoavat seikat: kuntoluokan viisi osuus kasvaa välillä merkittävästi, vaikka sen oli odotettu laskevan. Vaikka dataa on 30 vuoden ikäisiin teihin asti, havaitaan jo 15 vuoden kohdalla kuntoluokkien suhteellisten hallitsevuuden vaihtelevan niin suuresti, että havaintoja ei ole mallinrakennukseen riittävästi. Lisäksi toimeksiantajaa nimenomaisesti

kiinnostavista kuntoluokista 1 ja 2 on ensinnäkin muutamia havaintoja heti ikäjakauman alussa, ja toisaalta kuntoluokasta kokonaisuudessaankin on varsin vähän havaintoja luotettavan prosenttiosuuden määrittämiseksi. Nämä validointiprosessin löydökset on havainnollistettu myös Kuvassa 1. Mainittujen havaintojen perusteella päädyimme hylkäämään ajatuksen eri-ikäisten teiden käytöstä historiadataan korvaamiseksi.



**Kuva 1** Tiestön urasyvyyden kuntoluokituksen kehitys eräässä osaverkossa. Vaaka-akselilla ikä, pystyakselilla todennäköisyys, että kuntoluokka heikkenee vuoden aikana.

Koska selvästikään yksi mittaus tiestä ei riitä kunnan aikakehityksen mallintamiseen, tarvittiin historiadataa. Uutena lähestymistapana mallinnukseen päätettiin laskea historiatiedon perusteella todennäköiset polut, joita pitkin kunnan putoamista voidaan ennustaa ja arvioida näiden perusteella siirtymätodennäköisyydet kuntoluokkien välillä. Uuden lähestymistavan rakenne on havainnollistettu Kuvassa 2.

Tällä uudella mallilla on monia positiivisia puolia: samankaltainen polkumalli on ollut aikaisemminkin yleisessä käytössä, ei pelkästään tien kunnan ennustamisessa, vaan myös monissa muissa vastaavissa sovelluksissa. Tämä antaa vankan kuvan mallin monipuolisuudesta ja toisaalta myös sen



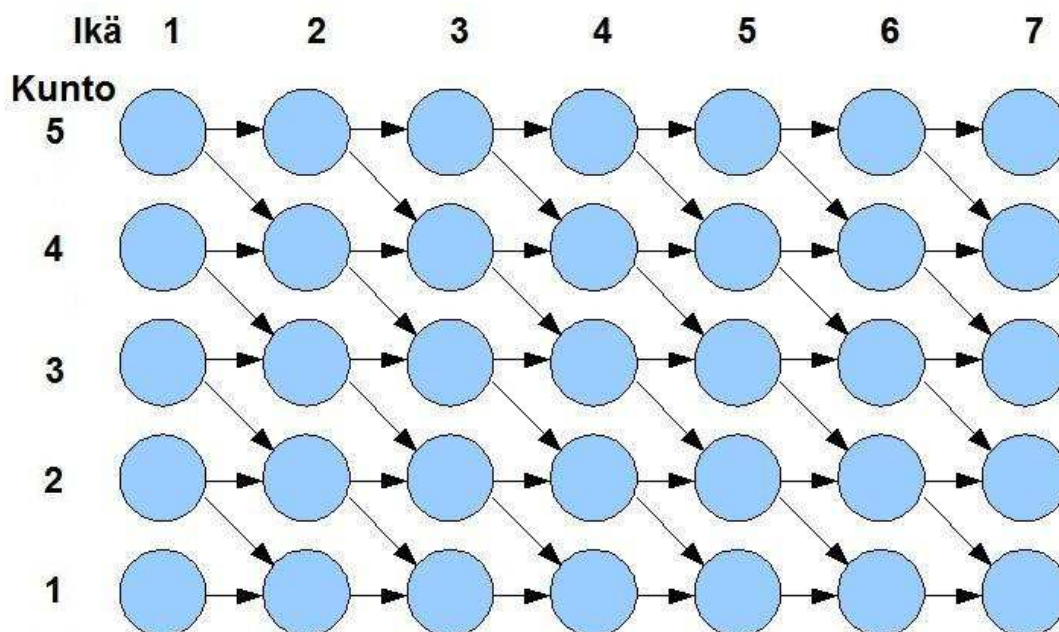
luotettavuudesta. Lisäksi kahden mittauskerran mukaan ottaminen kuvaa eksplisiittisemmin tiestön todellista muuttumista yli ajan. Ensimmäisessä mallivaihtoehdossa jouduimme luottamaan vain yhteen mittauskertaan, mikä pakotti meidät muodostamaan kyseenalaisen tavan kuvata tiestön kunnan muutosta todennäköisyysjakaumien avulla. Uudemmassa mallissa saadaan kahden eri mittauskerran avulla todennäköisyysjakauma muodostettua paljon loogisemmin ja suoraviivaisemmin, mikä tekee uudesta mallista paljon luotettavamman.

### 2.3 Mallin muodostaminen

Uuden datan kanssa päätimme yhdessä toimeksiantajan sekä kurssihenkilökunnan kanssa mallintaa kunnan huononemista eräänlaisella Markov-mallilla. Tiepätkien tasaisuudelle sekä urautuneisuudelle teimme rakenteeltaan samanlaiset mallit. Mallien rakenteiden voidaan ajatella koostuvan kahdesta ulottuvuudesta: ajasta edellisestä toimenpiteestä tiepätkälle sekä tiepätkän uran tai tasaisuuden kuntoluokka. Aika lasketaan kokonaisina vuosina ja kuntoluokat on määritelty kokonaisluvuiksi.

Jokainen aika-kuntoluokka-pari toimii Markov-mallissa omana tilana. Kuntoluokkia on sekä urille että tasaisuudelle viisi kappaletta. Luokat kuvataan numeroina 1-5. Mitä korkeampi numero, sitä paremmassa kunnossa tiepätkä on. Kuntoluokka 2 tarkoittaa, että tietä kuuluisi korjata. Kuntoluokka 1 tarkoittaa, että tiepätkä on jo aliarvoisessa kunnossa. Aikaulottuvuudessa ei teoriassa tarvitsisi olla ylärajaa. Datasta kuitenkin selviää, että vain yksittäiset tiepätkät ovat odottaneet yli 30 vuotta ilman toimenpiteitä. Täten rajaamme mallimme suoraan 40 vuoteen.

Siirtymismahdollisuuksia jokaisesta tilasta on yksi tai kaksi siten, että aikaulottuvuudessa siirrytään yksi vuosi eteenpäin ja kuntoluokka joko pysyy samana tai putoaa tasan yhden kuntoluokan verran. Tämä on asiantuntijamme ehdottama oletus, joka helpottaa todennäköisyyksien estimointia. Emme siis pidä mahdollisena, että kuntoluokka voi parantua itsestään. Tilat, joista on vain yksi siirtymätodennäköisyys, ovat ne, joissa kuntoluokka on jo huonoin mahdollinen, eli 1. Mallin rakenne on visualisoitu Kuvassa 2.



Kuva 2 Markov-mallin rakenne. Kuvassa näkyy vain mallin alkuosa.

Saamassamme historiadatassa on eri tiepätkien kuntomittaukset vuosilta 2003 – 2008 sekä tietoa milloin edelliset toimenpiteet on tehty. Sillä tiedolla voimmekin arvioida miten isot ovat todennäköisyydet kuntoluokan heikkenemiselle yhtenä tai useampana vuotena. Todennäköisyydet arvioidaan erikseen jokaiselle kuntoluokka-ikä-tilalle, eli  $(KL,t)$ -tilalle.

Kun tiedetään kuinka monta tienpätkää kussakin tilassa on nykyhetkenä, voidaan estimoida siirtymätodennäköisyydet ja edelleen simuloida tulevaisuuden kuntoennusteet.

## 2.4 Todennäköisyyksien estimointi

Lähtökohtana on kerätä saadusta datasta havaintoja kuinka tiet ovat kehittyneet kyseisessä tilassa. Koska dataa on paljon, voimme olettaa, että voimme jakaa sen eri luokkiin, joissa on eri tienpätkät eri luokittelutekijöiden mukaan jaoteltuna. Näitä luokkia kutsutaan osaverkoiksi. Pyrimme estimoimaan kunkin osaverkon kaikille tiloille omat kunnan heikkenemistodennäköisyydet.

Estimoimme todennäköisyyksiä keräämällä havaintoja siitä, miten monta tienpätkää datassa on pysynyt tietyssä kuntoluokassa tietyn ikäisenä – ja vastaavasti montako tienpätkää oli sen ikäisenä pudonnut alempaan kuntoluokkaan. Laskemalla näiden suhteen saamme estimoitua putoamismistodennäköisyyden. Datassa oli vain muutama tienpätkä, joka oli odottanut seuraavaa toimenpidettä yli 20 vuotta. Tämän takia on odotettavaa, että emme saa luotettavia tilansiirtotodennäköisyyksiä kyseisen ajankohdan jälkeen. Toisaalta koska teitä näissä tiloissa on vähemmän, todennäköisyydet eivät ole erityisen tärkeitä simuloitaessa Suomen päällystetyn tieverkon kehitystä. Päätimme toimeksiantajan kanssa, että kun datamäärä alkaa

silmämääräisesti pienentyä liikaa ja sen myötä tulokset heitellä, voimme katkaista todennäköisyysfunktion ja pitää putoamistodennäköisyyttä vanhemmille teille vakiona.

Datatiedoston lukemista varten kehitimme Java-ohjelmointikielellä ohjelman, joka löytyy kommentoituna Liitteestä 2. Ohjelma käsittelee datatiedostosta yksittäistä tiepätkeä vastaavaa riviä yksi kerrallaan. Mikäli kyse on käyvästä tiepätkestä, niin se otetaan mukaan todennäköisyyksien estimointiin havaintona. Ohjelmasta voi valita onko tutkittavana urien vai tasaisuuden kehitys.

Käyppyyt määrätty monesta eri vaatimuksesta. Ensinnäkin kyseiselle tiepätkelle on oltava kaksi kuntomittausta kyseisen historiadatan aikana, eli vuosina 2003–2008. Tiepätke ei kuitenkaan ole käypä, mikäli sille on tehty muita toimenpiteitä mittausten välisinä vuosina. Tiepätke ei myöskään ole käypä, mikäli se on rappeutunut kunnonkehitysoletuksiemme vastaisesti. Havainnoksi hyväksyttävän tiepätken kuntoluokka on siis saanut laskea korkeintaan mittausten välisten vuosien verran, eikä se ole saanut parantua. Hyväksytyjen sekä kehitysoletuksien takia hylättyjen tiepätkien määrät näkyvät Taulukossa 1. Asiantuntija-arvion mukaan syy suhteelliseen suureen osuuteen nouseviin kuntoluokkiin on se, että tiepätkelle tehtyä toimenpidettä ei ole rekisteröity.

**Taulukko 1 Hyväksytyjen sekä kehitysoletuksien takia hylättyjen tiepätkien määrät**

	Urat	Tasaisuus
Rivien kokonaismäärä	512136	512136
Väh. 2 mittausta, eikä toimenpiteitä näiden välillä	333303	330088
Hyväksytyt	324547	312662
Nouseva kuntoluokka	8708	17340
Liiallisesti putoava	48	86

Mittauksia ei ole suoritettu vuosittain vaan yleensä kyseisellä aikajaksolla on kaksi mittaustulosta. Ensimmäisestä ja viimeisestä mittauksesta selvitämme, mistä kuntoluokasta tiepätke on pudonnut mihin kuntoluokkaan. Ottamalla huomioon mittausvuodet sekä edellisen toimenpiteen vuoden, saamme tietää ajan edellisestä toimenpiteestä.

Esimerkkitapaus: tiepätken edellinen toimenpide on tehty vuonna 2001 ja uran kuntoluokan mittauksia on tehty vuosina 2004 ja 2007, joista kummallakin kerralla tulos oli kuntoluokka 4. Tällöin tallennetaan havaintoihin pysymiset tiloihin (4, 3), (4, 4), (4, 5) sekä (4, 6).

Huomaamme, että kaikissa tapauksissa, joissa mittausten välinen aika  $\Delta t$  on suurempi kuin yksi vuosi, ja missä kuntoluokan muutos  $\Delta KL$  on välillä  $1 \leq \Delta KL <$

$\Delta t$ , emme voi yksikäsitteisesti sanoa mitä polkua tiepätkän kuntoluokka on kehittynyt.

Esimerkiksi tapauksessa, jossa  $\Delta KL=1$  ja  $\Delta t=3$  on mahdollista, että tien kuntoluokka on pudonnut ensimmäisenä, toisena tai kolmantena vuonna. Tässä tapauksessa polkuja on siis kolme.

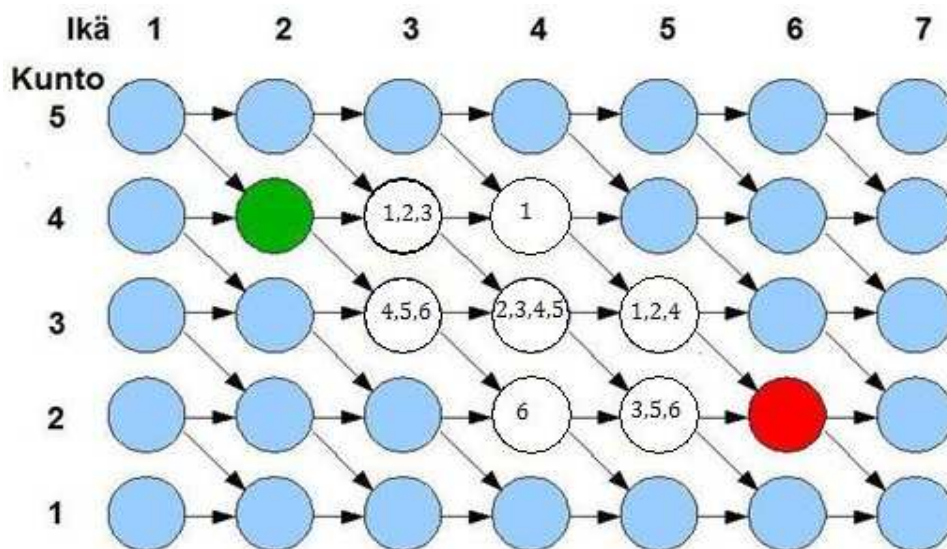
Tätä ajatusta voidaan yleistää, jolloin saadaan kombinatoriikan avulla selvitettyä, moneenko eri järjestykseen voidaan järjestää  $\Delta KL$  määrä putoamisia ja  $\Delta t - \Delta KL$  määrä pysymisiä. Tuloksena on luonnollisesti seuraava kaava:

$$\text{Mahdollisten polkujen lukumäärä} = \frac{\Delta t!}{(\Delta t - \Delta KL)! \Delta KL!} \quad (1)$$

Putoamistodennäköisyyksiä estimoitaessa annetaan kunkin tiepätkän jokaiselle polulle sama painoarvo siten, että yhden tiepätkän polkujen yhteenlaskettu painoarvo on aina 1.

Kuvassa 3 on esitetty tapaus, jossa ensimmäisessä mittauksessa on kaksi vuotta vanha kuntoluokan 4 tie ja toisessa mittauksessa neljä vuotta myöhemmin se on kuntoluokassa 2. Eli se siirtyy tilasta (4, 2) tilaan (2, 6) kuutta mahdollista polkua pitkin. Mahdolliset kehityspolut ovat:

1. (4, 2) → (4, 3) → (4, 4) → (3, 5) → (2, 6)
2. (4, 2) → (4, 3) → (3, 4) → (3, 5) → (2, 6)
3. (4, 2) → (4, 3) → (3, 4) → (2, 5) → (2, 6)
4. (4, 2) → (3, 3) → (3, 4) → (3, 5) → (2, 6)
5. (4, 2) → (3, 3) → (3, 4) → (2, 5) → (2, 6)
6. (4, 2) → (3, 3) → (2, 4) → (2, 5) → (2, 6)



Kuva 3 Esimerkkihavainto (4, 2) → (2, 6), jossa polut merkitty numeroin 1-6

Vastaavasti kunnan pysymis- ja putoamishavaintoja tulee Taulukkojen 2 ja 3 mukaisesti. Laskettua yhteen jokaisen käyvän tiepätkän pysymis- ja putoamishavainnot, voimme vertailla havaintojen osuuksia keskenään. Suhteellinen osuus putoamishavaintoja paljastaa todennäköisyyden vastaavanlaisen tiepätkän kuntoluokituksen putoamiselle yhden vuoden aikana. Esimerkissämme saisimme tämän yhden havainnon perusteella, että esimerkiksi tilassa (4,3) sijaitsevan tiepätkän putoamistodennäköisyys olisi  $0.3333/(0.1667+0.3333)=2/3$ , eli 66,7%. Tilassa (4,4) se olisi 100%, ja tilassa (4,5) se ei olisi määritelty. Putoamistodennäköisyys lasketaan kuitenkin vasta kun kaikista tienpätkistä yhteenlasketut havainnot on saatu, jolloin havaintoja on jopa tuhansia.

**Taulukko 2 Esimerkkitiepätkästä kerätyt pysymishavainnot**

KL \ t	1	2	3	4	5	6
5	0	0	0	0	0	0
4	0	3/6	1/6	0	0	0
3	0	0	2/6	2/6	0	0
2	0	0	0	1/6	3/6	0
1	0	0	0	0	0	0

**Taulukko 3 Esimerkkitiepätkästä kerätyt putoamishavainnot**

KL \ t	1	2	3	4	5	6
5	0	0	0	0	0	0
4	0	3/6	2/6	1/6	0	0
3	0	0	1/6	2/6	3/6	0
2	0	0	0	0	0	0
1	0	0	0	0	0	0

## 2.5 Osaverkkojen luokittelutekijöiden valinta

Tiettyjen luokittelutekijöiden avulla tiepätkät samaistetaan eli luokitellaan samaan osaverkkoon, jolloin saadaan aikaiseksi suurempi havaintomäärä ja pidempi aikasarja tiettyntyyppisestä tiestöstä.

Luokittelutekijöiden valinta on subjektiivinen päätös, mutta joitain tekijöitä voi pitää lähtökohtaisesti päätettyinä. Näitä ovat liikenteen kuormitus ja päällysteen ominaisuudet (katso mainintaa luvun alussa). Tätä luokittelujakoa noudattaa toimeksiantajalta saatu osaverkkoluokittelu, jossa on kaksi tekijää: päällystetyyppi (AB tai PAB) ja keskimääräinen vuorokausiliikenne (AB:lle luokittelurajat ovat 1000–3000, 3000–6000, 6000–12000 ja yli 12000 sekä vastaavasti PAB:lle alle 350, 350–1000, 1000–3000). Lisätarkasteluihin valittiin nopeusrajoitus ja maantieteellinen sijainti (Pohjois-Suomi, Etelä- ja Keski-Suomi sekä rannikko) sekä hoitoluokka, viimeisin toimenpide ja raskaan liikenteen

KVL. Myös mainittujen tekijöiden yhdistelmiä kokeiltiin ja kaikkiaan erilaisia osaverkkoja muodostettiin noin viisikymmentä erikseen uralle ja tasaisuudelle.

Luokittelutekijöiden alustava valinta perustui osittain toimeksiantajalta saatuihin kommentteihin ja ehdotuksiin sekä Tiehallinnon selvitykseen 57/2005, jossa mainittiin tien rappeutumista aiheuttavia tekijöitä. Liikenteen kuormitusta tulkittiin lisäävän KVL, raskas KVL ja nopeusrajoitus. Päälysteen ominaisuuksiksi luokiteltiin päälystetyyppi ja viimeisin toimenpide sekä osittain myös hoitoluokka. Ilmaston ja pohjamaan ominaisuuksien tutkimiseen käytettiin maantieteellistä sijaintia.

Päätösprosessi luokittelutekijöiden lopullisesta valinnasta perustui muun muassa havaintojen määrään, putoamistodennäköisyyksien eroihin osaverkkojen ja keskimääräisten arvojen välillä sekä putoamistodennäköisyyksien kehityksen selkeyteen. Tietysti kriteerinä oli myös kombinaatioiden mielekkyys ja maalaisjärki. Tietyissä osaverkoissa havaintojen määrä oli niin vähäinen, että todennäköisyyksien mielekkyys oli kyseenalainen, jolloin tällaiset osaverkkojaottelut päätettiin hylätä. Putoamistodennäköisyyksien osalta tutkittiin, onko eri osaverkkojen sekä koko datasta laskettujen todennäköisyyksien välillä merkittävää eroa. Päätösten tukena käytettiin Excel-makroa, jonka avulla todennäköisyyksien erot voitiin ilmaista värein lukujen sijasta. Positiivisen ja negatiivisen eron ilmaisuun käytettiin eri värejä ja eron voimakkuus korreloi värin voimakkuuden kanssa. Näin voitiin hyödyntää ihmisen hahmontunnistuskykyä ja näköaistin nopeutta, jotta suuresta määrästä osaverkkoja voitaisiin löytää ne osaverkot ja luokittelutekijät, jotka aiheuttivat eroa ja jakoperusteita. Putoamistodennäköisyyksistä piirrettiin myös aikasarjoihin verrattavia kuvaajia ja tavoitteena oli saada näistä mahdollisimman sileitä hyppelehtivän ja kohinaisen kuvaajan sijaan. Lopullinen ratkaisu perustui luokittelutekijöiden mielekkyyteen ja määrään.

## 2.6 Tulokset

Osaverkkojen tarkastelun yhteydessä havaittiin, että viimeisin toimenpide tai raskas KVL eivät olleet hyviä luokittelutekijöitä. Myös tien hoitoluokka osoittautui huonoksi vaihtoehdoksi, varsinkin kun lisäselvityksissä kävi ilmi, että se määräytyy suoraan keskimääräisestä vuorokausiliikenteestä. Myös maantieteellisellä jaolla oli yllättävän pieni vaikutus, mutta kahtiajako Oulun ja Lapin tiepiirit yhdistettynä verrattuna muuhun Suomeen vaikutti sen verran selvältä, että se hyväksyttiin tiettyihin luokittelukombinaatioihin.

Nopeusrajoituksen vaikutus havaittiin tietyissä osaverkoissa. Varsinkin nopeusrajoitusten ääripäiden havaittiin vaikuttavan kohtuullisen suurilla KVL:n arvoilla. Luokittelurajaksi valittiin 85 km/h, koska se erottelee suuret nopeudet (100 ja 120 km/h) alhaisemmista nopeuksista. Rajoja olisi voinut määrätä

useampiakin mutta osaverkkojen määrä haluttiin pitää kohtuullisena ja tarkempi jako olisi tehnyt osaverkoista havaintojen mielessä liian pieniä.

Selvät luokittelutekijät KVL ja päällystetyyppi säilytettiin sekä urien että tasaisuuden luokittelussa, mutta KVL:n luokkarajoihin tehtiin pieniä muutoksia ja esimerkiksi päällystetyypin PAB osalta luokkia oli pakko yhdistellä tarpeeksi suuren havaintomäärän aikaansaamiseksi. Luokittelu vaihtelee myös hieman urien ja tasaisuuden välillä, sillä tarkasteluissa havaittiin urien kehityksen vaativan tarkempaa jakoa, kun taas tasaisuus käyttäytyi yleisesti selkeämmin.

Lopullinen jaottelu on esitetty Taulukoissa 4 ja 5. Liitteen 1 Kuvissa 38–52 on esitetty kuntoluokkien 3, 4 ja 5 putoamistodennäköisyydet yhtä luokkaa alemmas tien iän funktiona. Kuvia tulkittaessa on muistettava, että kuvaajat eivät ole puhtaita aikasarjoja. Putoamistodennäköisyyden ei siis välttämättä tarvitse olla nouseva käyrä. Asian voi tulkita siten, että tietyssä luokassa kauan pysynyt tie pysyy siinä hyvin todennäköisesti jatkossakin. Kuvista on havaittavissa esimerkiksi urien kohdalla selvä ero päällysteen suhteen: Asfalttibetonilla päällystetyt kuntoluokan 5 tiet putoavat luokan alaspäin todennäköisemmin kuin kuntoluokan 4 tai 3. Pehmeä asfalttibetonilla päällystettyjen teiden kohdalla tilanne on täysin päinvastainen. Kuvaajista tai osaverkkojaottelusta ei kuitenkaan tehdä tämän syvällisempää analyysia, vaan jatkotarkastelu jätetään asiantuntijoille.



Taulukko 4 Lopullinen osaverkkojaottelu uraisuuden osalta.

Osaverkon nimi	Luokittelutekijät
Ura AB alle 1500	Päällyste: AB, KVL < 1500
Ura AB 1000-3000 alle 85 km/h	Päällyste: AB, 1000 < KVL < 3000, nopeus < 85 km/h
Ura AB 1000-3000 yli 85 km/h	Päällyste: AB, 1000 < KVL < 3000, nopeus > 85 km/h
Ura AB 3000-6000 alle 85 km/h	Päällyste: AB, 3000 < KVL < 6000, nopeus < 85 km/h
Ura AB 3000-6000 yli 85 km/h	Päällyste: AB, 3000 < KVL < 6000, nopeus > 85 km/h
Ura AB yli 6000	Päällyste: AB, KVL > 6000
Ura PAB Oulu ja Lappi	Päällyste: PAB, Tiepiirit: Oulu ja Lappi
Ura PAB muu Suomi	Päällyste: PAB, Tiepiirit: muu Suomi

Taulukko 5 Lopullinen osaverkkojaottelu tasaisuuden osalta.

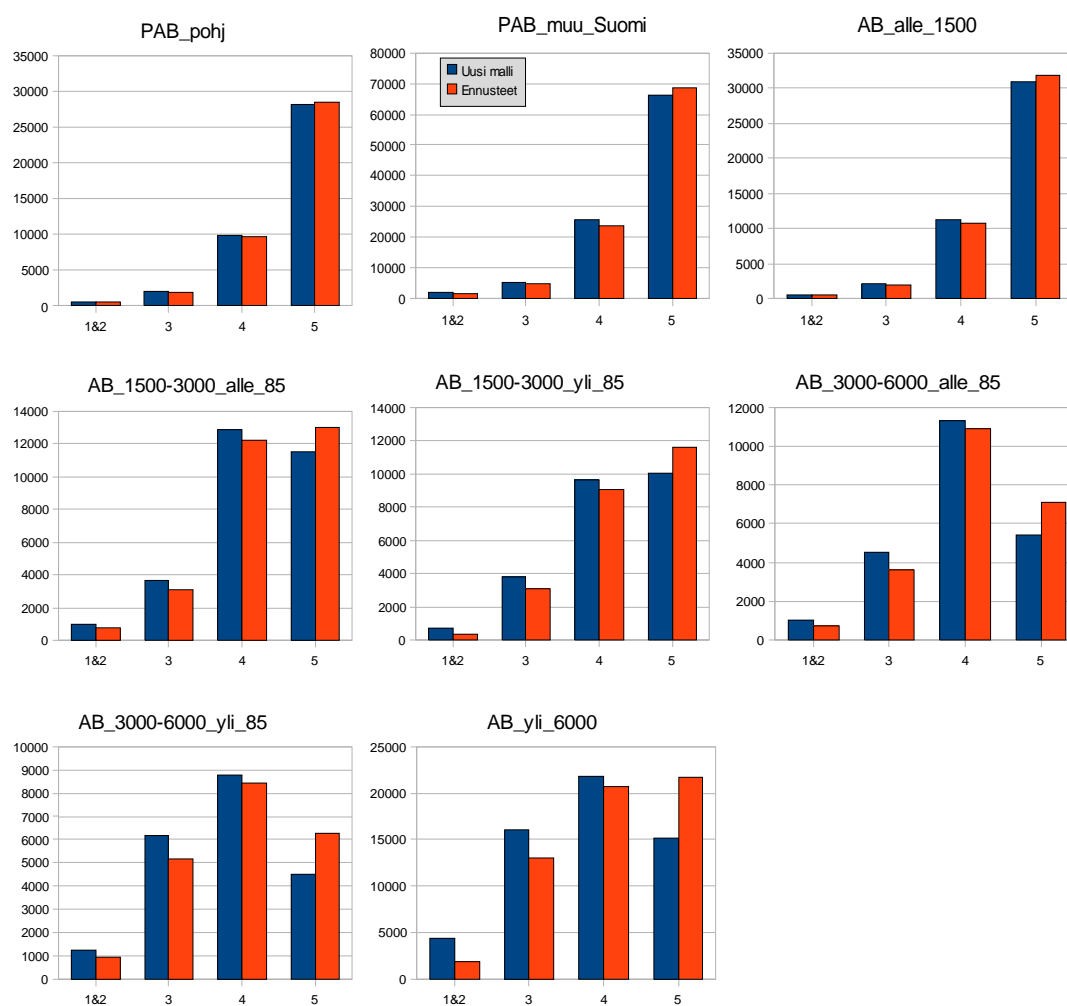
Osaverkon nimi	Luokittelutekijät
Tasaisuus AB alle 1500	Päällyste: AB, KVL < 1500
Tasaisuus AB 1000-3000	Päällyste: AB, 1000 < KVL < 3000
Tasaisuus AB yli 3000	Päällyste: AB, KVL > 3000
Tasaisuus PAB alle 350 Oulu ja Lappi	Päällyste: PAB, KVL < 350, Tiepiirit: Oulu ja Lappi
Tasaisuus PAB alle 350 muu Suomi	Päällyste: PAB, KVL < 350, Tiepiirit: muu Suomi
Tasaisuus PAB yli 350 Oulu ja Lappi	Päällyste: PAB, KVL > 350, Tiepiirit: Oulu ja Lappi
Tasaisuus PAB yli 350 muu Suomi	Päällyste: PAB, KVL > 350, Tiepiirit: muu Suomi

## 2.7 Uuden mallin tuloksien vertailu Tiehallinnon nykyisiin ennusteisiin

Saadussa datassa oli mittausdatan lisäksi Tiehallinnon aikaisemmillä ennustemenetelmillä tehtyjä ennusteita nykyhetkestä. Ennusteita oli tosin vain osalle teistä. Simuloimalla jokaisen ennustetun tien edellisen mittauksen nykyhetkeen uudella menetelmällä voidaan verrata miten näiden menetelmien tulokset suhtautuvat toisiinsa. Huomaamme kuvan 4 pylväsdigrammeista, miten eri osaverkkojen vanhemman menetelmän ja uuden menetelmän ennusteet poikkeavat tietyllä säännönmukaisuudella. On varsin selkeää, miten osaverkot poikkeavat toisistaan enemmän ja enemmän sen mukaan, miten paljon niillä ajetaan ja kuinka suurella nopeudella. Uuden menetelmän tiepätkien määrä kuntoluokassa viisi on matalampi kuin aikaisemmalla menetelmällä



ennustettu määrä. Vastaavasti muissa kuntoluokissa olevien tiepätkien määrä on yhteensä saman verran enemmän. Tämä saattaa johtua esimerkiksi siitä, ettei aikaisemmassa millimetri-ekstrapolaatiossa oteta huomioon, paljonko teillä ajetaan, tai muita vastaavia ominaisuuksia. Toisaalta voi myös olla, että ennusteissa on otettu huomioon tulevia korjauksia. Huolimatta siitä, missä syy piilee, on syytä tarkastella malleja kriittisesti, jotta eroavaisuuteen johtava syy selviää.



Kuva 4 Eri menetelmillä ennustetut osaverkkojen kuntoluokkajakaumat uralle. Vaaka-akselilla kuntoluokat ja pystyakselilla tienpätkien määrät

## 2.8 Pohdintaa ja johtopäätökset

Mallin rakentamisen lähtökohtana oli tehdä kuntoluokitukseen pohjautuva ennustemalli ison tieverkoston kuntojakauman kehitykselle. Mallimme on varsin yksinkertainen ja selkeä ja soveltuu varsin hyvin ison tieverkoston kuntojakauman kehityksen ennustamiseen. Kun on tietoa millaisia teitä Suomessa on ja missä määrin, niin on varsin helppoa laskea kuntojakauman ennusteen odotusarvo. Koska tilasiirtotodennäköisyydet ovat helposti käsiteltäviä ja ymmärrettäviä, on tätä mallia myös mahdollista käyttää

analysoitaessa riskejä ja erilaisia worst-case -skenaarioita, mikä ei joka mallilla ole niin helppoa.

On tärkeätä ymmärtää, että kuntoluokkiin perustuva ennustemalli ei kuitenkaan sovellu niin hyvin yksittäisten teiden kunnan heikkenemisen ennustamiseen. Jos ollaan kiinnostuneita yhden tiepätkän tarkemmasta ennusteesta, niin on lähdettävä liikkeelle millimetri-luvuista, jotka ovat kuntoluokituksen taustalla. Esimerkiksi uran millimetri-ekstrapolaatio paljastaa toki paremmin, milloin tietty tiepätkä on lähestymässä kahden kuntoluokan välistä rajaa. Yksinkertaisimmillaan millimetri-ekstrapolaatio ei kuitenkaan ota millään tavalla huomioon miten eri ominaisuuksia omaavat tiet saattavat kehittyä eri tavalla.

Mallimme todennäköisyyksiä estimoitaessa teemme piste-estimaatin, jonka tarkkuutta ei kuitenkaan tiedetä tarkasti. Tämä johtuu osittain siitä, että mallimme rakenne on itse kehitetty eikä tämän työn puitteissa voitu tutkia esimerkiksi luottamusvälejä matemaattisesti. Toisaalta voi huomauttaa, että syyn voidaan katsoa johtuvan itse asiassa ongelman luonteesta ja datan keräämiskäytännöistä. Tiet korjataan, kun ne ovat tietyssä kunnossa, mutta tähän vaikuttaa kuitenkin moni muukin asia; pidemmän tieosuuden ollessa suurelta osin huonossa kunnossa, korjataan usein myös niitä tiepätkiä, jotka ovat hyväksyttävässä kunnossa. Selvää on, että korjaukset sekä mittausten ajankohdat eivät ole sattumanvaraisia, jolloin itse asiassa jäljellä olevien tiepätkien perustuva ennustemalli ei ole harhaton. Harhattomuus on tietenkin matemaattisesti katsottuna tärkeä oletus ennustusmallia tehtäessä.

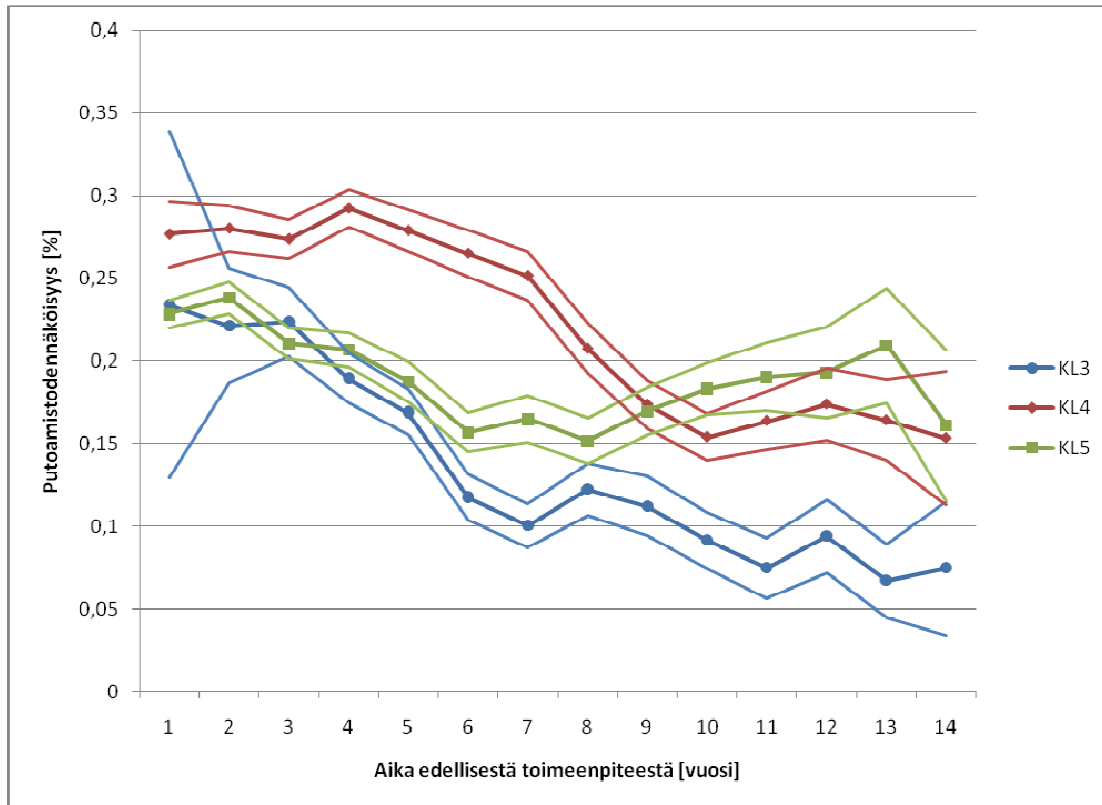
Jos kuitenkin oletamme, että kerätty data ja siihen pohjautuva malli ovat harhattomia, olemme lähempänä todennäköisyyksien tarkkuuden selvittämistä. Tällöin mittausten välinen aika aiheuttaa sen, että luottamusvälin laskeminen ei ole mahdollista tavallisimpia menetelmiä käyttäen. Mikäli mittauksia tehtäisiin vuosittain, voisi olettaa tietyn osaverkon tietyn (KL, ikä)-pisteen havainnot Bernoulli-jakauman realisaationa. Tällöin voisimme laskea todennäköisyyden luottamusvälin rajoja luottamustasolla  $(1-a)$  kaavalla (Mellin, Ilkka, Tilastolliset mallit ja kaavat 2008):

$$p \pm z_{a/2} \sqrt{\frac{p(1-p)}{n}} \quad (2)$$

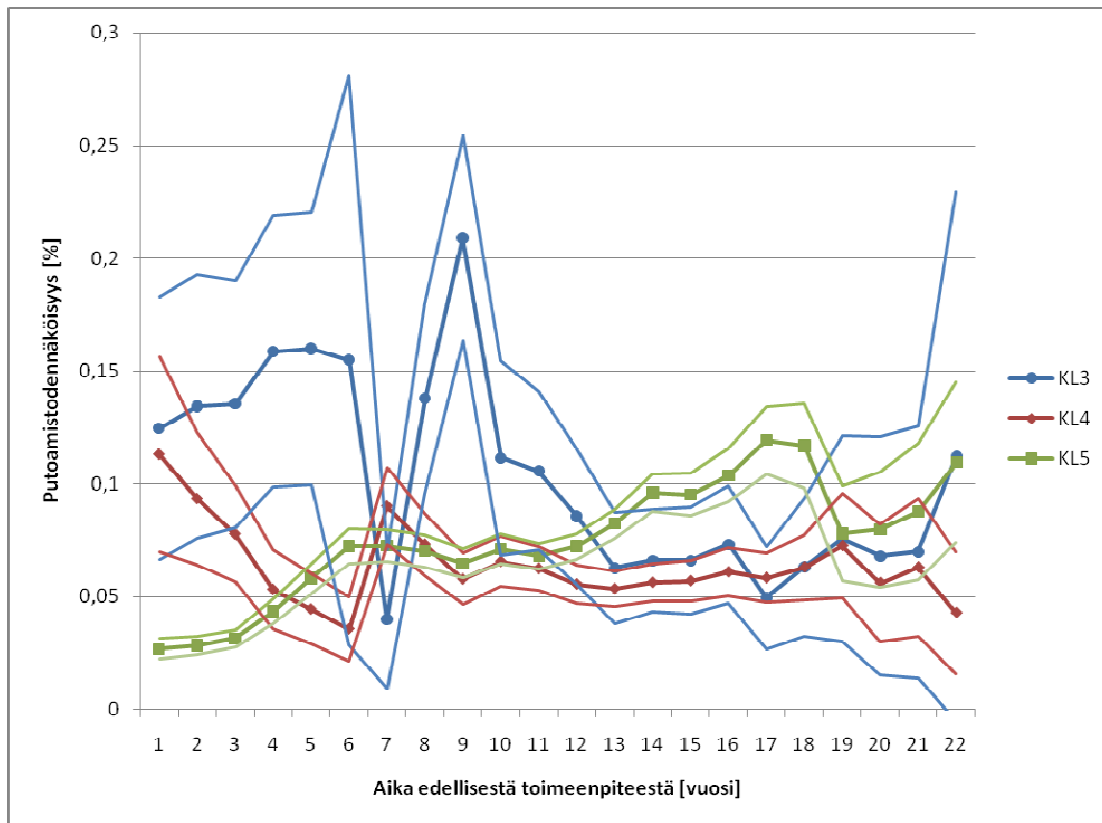
Tässä  $p$  on odotusarvon estimaattori ja  $n$  on havaintojen lukumäärä. Muuttuja  $z_{a/2}$  on luottamustasoon  $(1-a)$  liittyvä luottamuskerroin standardoidusta normaalijakaumasta  $N(0,1)$ .

Vaikka on täysin selvää, että tapamme laskea havaintoja ei sovi aukottomasti yhteen tämän kaavan kanssa, tarkastamme minkälaisia tuloksia se antaisi. Nämä

tulokset saattavat hyvinkin olla suuntaa antavia todennäköisyyksien tarkkuudesta. Tuloksen näkee alla olevista kuvaajista (kuvat 5 ja 6). Erityisen huomattavaa on miten aivan alkupään ja myös loppupään luottamusväli laajenee suuresti, havaintojen pienen määrän takia. Parhaimmillaan luottamusvälin tarkkuus on pienempi kuin  $\pm 0,004$ . Huomaamme myös miten suuri merkitys osaverkkojaottelulla on havaintomäärien tärkeyden takia.



Kuva 5 Yllä olevasta kuvaajasta näkee suuntaa antavan 95 % luottamusvälin uran heikkenemisen todennäköisyydelle osaverkoissa AB&KVL>6000.



Kuva 6 Yllä olevasta kuvaajasta näkee suuntaa antavan 95 % luottamusvälin uran heikkenemisen todennäköisyydelle osaverkoissa AB&KVL<1500.

Eri malleja on varsin helppoa laskea erilaisille osaverkoille, joiden kautta voi huomioida muutamaa kuntoon vaikuttavaa tekijää, kuten liikenteen määrää, nopeusrajoituksia, sijaintia ja niin edelleen. Jakamalla data liian moneksi osaverkoksi, ennustemenetelmän käyttö on kuitenkin työläämpää ja lisäksi havaintomäärän vähäisyys voi johtaa huonoon tarkkuuteen. Jotta eri vaikuttajien tarkka syy-seuraus suhde tulisi ilmi, pitää kuitenkin käyttää malleja, jotka pohjautuvat millimetri-mittauksiin eikä karkeisiin kuntoluokkiin. Toisaalta voi kysyä miten suuri tarve on ottaa kovin monta vaikuttajaa huomioon mallissa, jolla haluamme estimoida vain ison tieverkoston kuntojakaumaa, eikä sinänsä yksittäisen tien kuntokehitystä. Esimerkiksi neuroverkko-malli voisi ottaa muita tekijöitä selkeästi paremmin huomioon. Toisaalta huonona puolena olisi kuitenkin, että tieto ei tule mallista ulos käytännöllisellä tavalla. Silloin eri muuttujien vaikutukset selviävät ainoastaan keinotekoisella matemaattisella sovituksella dataan.

Pidämme mallia varsin hyvin käyttötarkoitukseensa sopivana. Erityisesti voimme pitää mallia riskianalyysiä helposti mahdollistavana. Siten yksi sen parhaita vahvuuksia onkin käyttää sitä korjauspäätöksien tekemisen tukena.

## 3 Ennustemallin käyttö riskianalyyssissä

### 3.1 Riskianalyysi

Riskianalyysin tavoitteena on arvioida kunnan putoamisia kuvaavien piste-estimaattien hajontaa ja antaa selvä kuva siitä, kuinka suuri joukko tiestöstä voisi huonossa tapauksessa pudota. Lähestymistavaksi on valittu tieverkon kuntokehityksen simuloiminen useamman vuoden päähän sekä worst-case -skenaarioiden tarkastelu simulaatioiden lopputulemista. Riskin mittarina tarkastellaan kuntoluokkakohtaista lukuarvoa joka kertoo montako tietä 95 %:n todennäköisyydellä enintään putoaa kunkin aikahorisontin kuluessa. Riskianalyysi ei käsittele projektisuunnitelmassa esitettyjä, itse projektiin liittyviä riskejä joiden toteutumista arvioitiin väliraportissa. Työssä lähtökohtana olevaan kuntoennustemalliin liittyviä riskejä puolestaan on arvioitu Mallin lähtökohdat – kappaleessa.

**Tiehallinnon tavoitteena** on saada teiden kunnan putoamisesta ennusteet usean, jopa kymmenen vuoden päähän. Vastaavasti halutaan riskiarvio pudonneiden teiden määrästä samalta ajalta. Riskin mittarista ja sen parametreina olevasta luottamustasosta ja aikahorisontista on sovittu tilaajan kanssa.

**Datan** riskitarkastelun pohjana on sama kuin ennustemallia rakennettaessa. Kuntoennustemallilla on laskettu kunkin kunto- ja ikäluokan tyypilliselle tienpätkälle todennäköisyys pudota alempaan kuntoluokkaan ja lisäksi tiedetään kussakin luokassa mallinnushetkellä olevien teiden määrä. Nämä määrät ja todennäköisyydet muodostavat kunnan putoamisille teoreettiset binomitodennäköisyysjakaumat, joiden tarkasteluun tämä, kuten mikä tahansa muukin riskianalyysin lähestymistapa tyypillisesti perustuu.

**Menetelmänä** riskin mittaamiseen on siis kuntoputoamisten määrä joka 95 % todennäköisyydellä enintään havaitaan valitulla aikahorisontilla. Muita yleisiä mutta kenties vähemmän informatiivisia vaihtoehtoja olisivat yksinkertaiset keskihajonnan tai luottamusvälien tarkastelu. Vaikka riskiluku perustuu luottamustason ja aikahorisontin subjektiiviseen valintaan, antaa vastauksena saatu luku varsin hyvän kuvan siitä minkälaisia lopputulemia huonossa tapauksessa on odotettavissa. Lähestymistapa on yleismaailmallisesti tunnettu (muun muassa erilaisten porfolioiden arvonmuutokset ja riskit) ja sen käyttöä tukee laaja akateeminen kirjallisuus (mm. Jorion, 1997; Stulz, 2001).

Hieman matemaattisemmin esitettynä riskianalyysin perusta on seuraava: Työssä aiemmin estimoidulla ennustemallilla on saatu kunto- ja ikäluokkakohtaiset putoamistodennäköisyydet  $p$ . Kustakin kuntoluokasta tiedetään lisäksi tienpätkien määrä  $n$  ja alempaan kuntoluokkaan putoavien

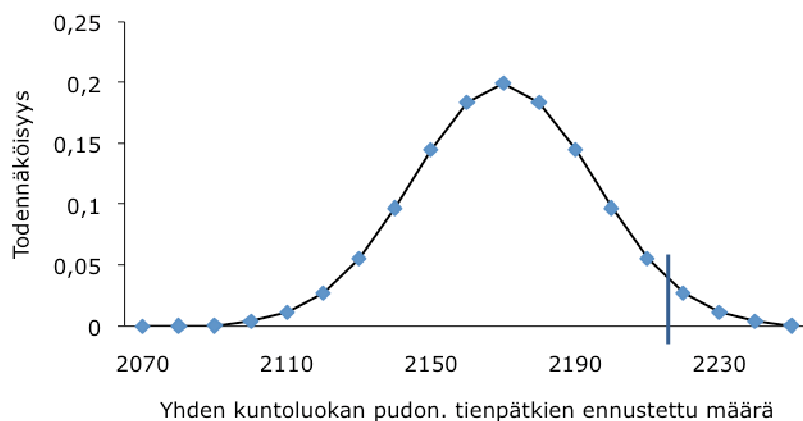
teiden määrän  $X$  oletetaan noudattavan binomitodennäköisyysjakaumaa  $X \sim \text{Bin}(n, p)$  jonka edellytykset ovat:

- i) Vain kaksi mahdollista lopputulosta (kunto putoaa tai säilyy) kullekin satunnaiskokeelle
- ii) Kunnan putoamisen todennäköisyys  $p$  sama kullakin satunnaiskokeella tai simulaatiolla
- iii) Tehdään yhteensä ennalta määrätty  $K$  toistoa; tässä simulointia
- iv) Nämä  $K$  toistoa ovat riippumattomat

Tällöin voidaan teiden kunnan putoamisten mahdolliset määrät (eli vaihtoehdot että tienpätkän kunto putoaa tai säilyy) simuloida poimintoina binomijakaumasta jonka parametrit  $n$  ja  $p$  tunnetaan datan (tienpätkiä  $n$  kpl) ja mallin estimoinnin (siirtymistodennäköisyydet  $p$ ) perusteella. Tällöin

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad 0 \leq k \leq n \quad (3).$$

Binojakauman (3) antamia eri tienpätkien putoamismäärien todennäköisyyksiä on havainnollistettu kuvassa 7. Kuvasta päästään myös hyvin käsiksi riskin mittariin, eli pudonneiden tienpätkien odotettuun enimmäismäärään 95% todennäköisyydellä, jota kuvastaa sininen pystyviiva jakauman oikeassa hännässä.



**Kuva 7** Kuntoluokasta viisi kuntoluokkaan neljä putoavaksi ennustettavien tienpätkien määrä eräässä osaverkossa. Pudonneiden tienpätkien määrä noudattaa binomijakaumaa  $\sim \text{Bin}(n, p)$ , joka suurilla  $n$  arvoilla lähestyy normaalijakaumaa kuten kuvassa. Jakauman hännästä voidaan lukea montako tienpätkää putoaa 0,05 todennäköisyydellä, tai kääntäen montako tienpätkää enintään putoaa 95% todennäköisyydellä, mitä kuvaa sininen pystyviiva.

Kuvan 7 tyyppisestä approksimatiivisesta normaalijakaumasta, tai tarkemmin jakauman parametreista, voitaisiin myös analyttisesti laskea 95% todennäköisyyttä vastaava kuntoputoamisten enimmäismäärä. Kun havaintojen määrä on riittävän suuri on saadaan normaalijakaumaa lähestyvän jakauman

keskihajonnaksi  $\sigma = \sqrt{np(1-p)}$ , ja kuntoputoamisten odotusarvo puolestaan on  $np$ . Lisäksi tiedetään että jakaumassa (kuva 7) 95% todennäköisyysalueen rajalla tapahtuva lopputulema on 1.65 keskihajonnan päässä odotusarvosta. Tällöin riskiluku saataisiin kertomalla pudonneiden tienpätkien odotettu lukumäärä putoamisten keskihajonnalla ja 1.65 (joka vastaa normaalijakauman 95% todennäköisyysalueen jälkeen jäävän hännän havaintoja) ts. riski:  $1,65 \times \sigma \times \text{pudonneiden tienpätkien odotettu lkm}$ .

Ongelma analyttisessä lähestymistavassa on kuitenkin mm. oletukset havaintojen normalisuudesta ja keskihajonnasta, joiden vuoksi simulaatio on matemaattisesti validimpi lähestymistapa (vaikkakin tässä tapauksessa antaa varsin samankaltaisen vastauksen). Riskisimuloinnissa vältytään ottamasta kantaa teiden kuntokehityksen todelliseen (teoreettiseen) jakaumaan, kuten siihen noudattavatko kuntoputoamiset normaalijakaumaa.

Simulaatiossa riskilukuun päästään kiinni laittamalla esim. tuhannen simulaatiokierroksen antamat kuntoputoamismäärät järjestykseen pienimmästä suurimpaan putoamisskenaarioon (vrt. kuvan 7 x-akseli), ja poimimalla näistä 950:nes skenaario eli  $0.95 \cdot 1000$ :s tapaus. Tämä vastaa 95% todennäköisyydellä enintään putoavien teiden määrää. Vastaavalla tavalla voitaisiin tarvittaessa lukea myös 90 % tai 99 % luottamustasoja vastaavat riskiluvut järjestettyjen skenaarioiden joukosta.

**Useiden tienpätkien ja ikäluokkien simulointi yhdessä:** Tässä tarkastelussa ei lasketa riskilukua yksittäiselle tienpätkälle, vaan kokonaisten kuntoluokkien muutokselle yli ajan. Tosin mallin muuttujista on myös saatavissa komponenttikohtaiset riskiluvut, jos halutaan selvittää esimerkiksi minkä ikäiset tiepätät vaikuttavat tulokseen eniten. Ennen simulaation kulun tarkempaa selitystä, pohditaan kuitenkin vielä yhtä riskilaskennan kannalta tärkeää huomiota.

Kuten minkä tahansa portfolion riskiä laskettaessa, hajauttaminen tai tekijöiden määrän lisääminen vähentää kokonaisriskiä, kunhan tekijöiden korrelaatio on alle 1. Tässä työssä ei tienpätkien kuntomuutosten korrelaatiota kuitenkaan tiedetä, eikä toisaalta saadakaan selville, kun ennustemallissa yhden vuoden havaintopisteeseen on koottu useina eri aikoina mitattuja tiepätä. Riskiarvioinneissa komponenttien välinen korrelaatio on kuitenkin merkittävä tekijä ja ratkaisuna tähän on työssä päädytty tarkastelemaan kolmen eri skenaarion vertailua: ensimmäisessä oletetaan vahva korrelaatio, toisessa ei lainkaan korrelaatiota, ja kolmannessa eli jatkuvassa menetelmässä arvotaan tuhat riippumatonta kymmenen vuoden polkua (takempi kuvaus jäljempänä).

Jos oletetaan esimerkiksi että sattuu jokin teitä rappeuttava tapahtuma, paha routakausi tai muuta vastaavaa, joka vaikuttaa kaikkiin alueen teihin samalla tavalla, niin eri tienpätkille simuloidun putoamisen tulisi korreloida laskennassakin. Vastaavasti, jos simuloidaan kaikkien tienpätkien huononemispolut erikseen, eli oletetaan korrelaatio nollassa, ja sitten summataan lopputulemat yli polkujoukon, niin vähiten huonontuneet polut vaimentavat eniten huonontuneiden (worst-case) vaikutusta. Tällöin lopputuloksena saatu riskiarvio on muita menetelmiä pienempi, mahdollisesti todellisuutta vähättelevä. Tarkastellaan seuraavaksi eri lähestymistapoja tarkemmin ja valitaan niistä nähdäksemme sopivin jäljempänä seuraavien tulosten laskemiseen.

### Heikon korrelaation simulaatio

Oletetaan aluksi että kunkin ikäluokan tienpätkien putoaminen kuntoluokasta toiseen ei riipu muista ikäluokista. Toisin sanoen: se että vuoden ikäisiä tienpätkiä putoaa jossain simuloidussa skenaariossa (eli yhdellä kierroksella) paljon, ei ennakoisi että kahden vuoden ikäisiä putoaisi samalla kierroksella paljon. Puhumme tällöin heikon korrelaation simulaatiosta, ja sen kulku selviää parhaiten kahdesta alla olevasta taulukosta selityksineen. Näissä tarkastellaan yksinkertaisuuden vuoksi yhtä osaverkkoa, yhtä aika-askelta (i.e. 1 vuotta simulaatiossa eteenpäin) ja yhden kuntoluokan simuloituja ennusteita pudonneiden tienpätkien (100m) määrälle tänä aikana. Lähtötietoina tiedetään tienpätkien määrä vuoden alussa ja kunnan putoamistodennäköisyys kunkin ikäluokan tienpätkille.

**Taulukko 6a.** Ensimmäisessä sarakkeessa on simulointikierroksen  $n=1$  tulokset siitä, monenko tienpätkän (per ikäluokka) ennustetaan vuoden aikana putoavan alempaan kuntoluokkaan. Sarakkeen luvut on arvottu binomitodennäköisyysjakaumasta kun putoamistodennäköisyys ja tienpätkien määrä vuoden alussa tunnetaan. Kun toistetaan arvonta 1000 kertaa, saadaan eri simuloinneilla muiden esimerkkisarakeiden mukaisia tuloksia. Kokonaisriskin kannalta kiinnostavaa on kuntoluokasta pudonneiden tienpätkien määrä yhteensä, eli summat alimmalla rivillä. Näiden summien joukko lähestyy normaalijakaumaa, kun simulaatiokierroksia on tarpeeksi.

Alkuperäiset, järjestämättömät simulaatituloset kun simuloitiin pudonneiden tienpätkien määrää KL5:stä KL4:ään eräissä osaverkossa							
Tienpätkien ikä	Simulaatiokierros $n=1,2,\dots,N$						
	$n=1$	$n=2$	$n=3$	...	$n=159$	...	$n=1000$
1	106	95	120	...	123	...	117
2	144	130	155	...	159	...	158
3	181	199	183	...	185	...	163
4	129	129	135	...	139	...	116
5	103	113	98	...	101	...	93
...							
25	17	15	21	...	23	...	19
<b>Pudonneita yhteensä</b>	<b>2180</b>	<b>2181</b>	<b>2212</b>	...	<b>2230</b>	...	<b>2166</b>



Taulukko 6b. Kun kaikki simulointikierrokset Taulukon 6a tapaan on laskettu, järjestetään sarakkeet 'Pudonneita yhteensä' summan perusteella pienimmästä suurimpaan. 1000:sta simulaatiosta suurin pudonneiden määrä on tällä kertaa sattunut simulaatiossa n=702 (ei kuvassa) ja järjestyksessä 950:ksi suurin on sattunut kierroksella n=159. Tämä järjestämällä löydetty tulos (2230 pudonnutta) on tulokseksi saatu pudonneiden tienpätkien määrä 95% riskitasolla.

Järjestetään ylemmän kuvan sarakkeet 'pudonneita yhteensä' summan perusteella pienimmästä suurimpaan.							
Tienpätk. ikä	n=1000	n=1	n=2	...	n=3	...	n=159
1	117	106	95		120		123
2	158	144	130		155		159
3	163	181	199		183		185
4	116	129	129		135		139
5	93	103	113		98		101
...							
25	19	17	15		21		23
<b>Pudonneita yhteensä</b>	2166	2180	2181	...	2212	...	<b>2230</b>

Kun edellä kuvattu yhden aika-askeleen laskenta on valmis, saadaan riskiennuste 10 vuoden yli samalla menettelyllä. Ainoastaan vuosittaisten simulaatioiden lähtökohtana oleva kuntoluokan tienpätkien määrä muuttuu (kuntoluokan 5 tapauksessa aina vähentyen). Eli seuraavana vuonna pudotetaan simulaation taulukossa 6b osoittamat määrät alempaan luokkaan, tehdään simulaatiot uudestaan, ja toistetaan menettely edelleen seuraaville 8 vuodelle.

Heikon korrelaation simulaatiossa kokonaisriski pienenee, kun eri-ikäiset tienpätkät putoavat toisistaan (tai esim. yhtä aikaa vaikuttavista sääoloista tms.) riippumatta. Näin ollen kun samalla simulaatiokierroksella jossain ikäluokassa putoaa poikkeuksellisen paljon ja toisessa taas erityisen vähän tienpätkiä, putoamiset kumoavat toistensa vaikutuksia pienentäen tarkasteltavia summia. Erityisesti riskiluvun kannalta tämä stokastisen vaihtelun vaimentava vaikutus on tietenkin haitallista, ja saattaa arvioida riskejä alakanttiin. Parempi arvio riskien mittaamiseksi saadaan seuraavalla muutoksella lähestymistapaan.

### Vahvan korrelaation oletus:

Riski- ja simulaatiometodologian kirjallisuudessa (Jorion, 1997) muuttujien (tai tässä tapauksessa tienpätkien putoamiset eri ikäluokissa) välinen korrelaatio on otettu huomioon kertomalla sopivalla korrelaatiomatriisilla arvottuja satunnaislukuja. Aiemmin kuvatun mukaisesti tämä lähestymistapa ei kuitenkaan ollut toteutettavissa tässä.

Tässä arvotaan joka simulaatiokierroksella kustakin ikäluokasta putoamaan muiden ikäluokkien putoamisista riippumaton määrä teitä, jolloin vuoden ikäisiä

teitä voi pudota paljon ja kaksivuotiaita vähän (kompensoiden ko. simulaatiokierroksen koko kuntoluokan kaikenikäisten putoamisten summaa). Vahvan simulaation lähestymistavassa unohdetaan kuitenkin tämä summa hetkeksi, ja pureudutaan pienempiin komponentteihin eli tienpätkien putoamisiin eri ikäluokittain. Kun pienempiä komponentteja järjestetään suuruusjärjestykseen riskilukua varten, on myöskin onnekkaiden ja onnettomampien kuntokehitysten kompensoiva vaikutus pienempi. Allaolevat taulukot 6 c. ja 6 d. selventävät lähestymistavan.

Taulukko 6c. Esimerkki simulaatioiden järjestämisestä vahvan korrelaation luomiseksi. Simuloidaan ensin kunkin ikäluokan kuntoputoamisten määrää 1000 eri simulaatiokierrosta binomijakaumasta Bin(n,p), aivan kuten taulukossa 6a. Vahvempaan korrelaatioon päästään seuraavaksi järjestämällä simulaatiotulokset taulukon 6d. kuvaamalla tavalla.

Alkuperäiset, järjestämättömät simulaatiotulokset kun simuloitiin pudonneiden tienpätkien määrää KL5:stä KL4:ään eräässä osaverkossa							
Tienpätkien ikä	Simulaatiokierros n=1,2,...,N						
	n=1	n=2	n=3	...	n=159	...	n=1000
1	106	95	120	...	123	...	117
2	144	130	155	...	159	...	158
3	181	199	183	...	185	...	163
4	129	129	135	...	139	...	116
5	103	113	98	...	101	...	93
...							
25	17	15	21	...	23	...	19
<b>Pudonneita yhteensä</b>	<b>2180</b>	<b>2181</b>	<b>2212</b>	...	<b>2230</b>	...	<b>2166</b>

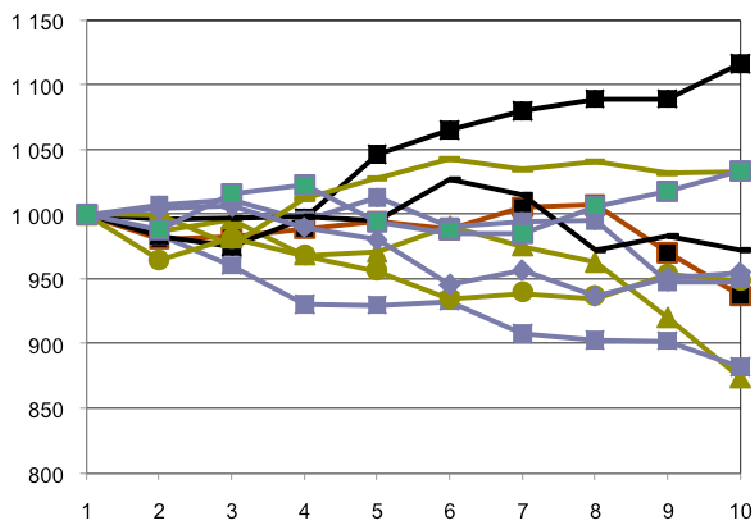
Taulukko 6d. Olkoot simulaatiokierrokset missä tahansa järjestyksessä (järjestetyinä tai ei) sarakkeissaan. Järjestetään nyt kukin rivi kerrallaan kussakin ikäluokassa pudonneiden tienpätkien määrät pienimmästä pudonneesta määrästä suurimpaan simuloituun putoamisskenaarioon. Kun tämä on tehty kaikille riveille 1-25 erikseen, tulee 950:een sarakkeeseen kaikista simuloinneista 950:ksi huonoimmat skenaariot, implisiittisellä oletuksella että tässä riskiskenaariossa tiestö rappeutuu enemmänkin "kaikkialla samaan aikaan". 950:n sarakkeen summa vastaa nyt 95% todennäköisyydellä enintään alempana kuntoluokkaan putoavien tienpätkien määrää.

Järjestetään tällä kertaa ylemmän kuvan <b>kukin rivi</b> pienimmästä suurimpaan simulaatiotulemaan. Sarakkeita ei tarvitse enää järjestää, poimitaan vain 950:n sarakkeen summa joka sisältää nyt kunkin ikäluokan 950:ksi suurimmat tulemat.						
Tienpätkien ikä						
1	<b>95</b>	<b>106</b>	<b>117</b>	...	<b>120</b>	<b>123</b>
2	130	144	155	...	158	159
3	163	181	183	...	185	199
4	116	129	129	...	135	139
5	93	98	101	...	103	113
...						
25	15	17	19	...	21	23
<b>Pudonneita yhteensä</b>	<b>2112</b>	<b>2175</b>	<b>2204</b>	...	<b>2222</b>	<b>2256</b>

Kymmenen vuoden aikahorisontin yli kunakin vuonna 95% tn:llä enintään pudonneiden tienpätkien määrät (kaikenikäikäiset tiet yhteensä per kuntoluokka) saadaan jälleen kuten aiemminkin. Summataa siis järjestyksessä 950:ksi eniten pudonneet, eri ikäluokkien tienpätkät, kussakin kuntoluokassa. Vähennetään nämä pudonneet tiet kustakin kuntoluokasta (ja lisätään alempaan luokkaan), ja tehdään samat simulaatiot, tulosten järjestäminen sekä teiden pudotus alempaan luokkaan uudestaan jokaiselle vuodelle. Simulaatiossa 95 % riskitasoa vastaava toteutuma (jäljellä oleva tienpätkien määrä) on aina lähtökohta seuraavan vuoden simuloinnille.

Tämä lähestymistapa valittiin lopulta tulosten esittämiseen (vaikka tulokset laskettiin liitteenä toimitettuihin tiedostoihin kaikilla esitetyillä lähestymistavoilla) jotta riskejä ei aliarvioitaisi. Alla vielä yksi intuitiivinen polkulähestymistapa kuntoputoamisten kehitysten simulointiin 10 vuoden ajalle.

**Jatkuvan simulaation** menetelmässä arvotaan jokaisella simulaatiokierroksella  $K=1000$  kymmenen vuoden polku teiden kunnan kehitykselle kussakin ikä- ja kuntoluokassa. Joka vuosi pudotetaan simulaatio esittämät määrät teitä kussakin kuntoluokassa kullakin polulla. Laskennan lopuksi järjestetään polkujen loppupisteet suuruusjärjestykseen ja otetaan  $0,95 \cdot N$ :s huonoin kehityspolku. Aiempien vuosien riskit ovat tällä kehityspolulla tapahtuneet kunnan putoamiset kunakin vuonna. Havainnollistus on Kuvassa 8.



Kuva 8. Jatkuvassa simulaatiossa arvotaan työssä mallinnetusta teiden kunnan putoamisten todennäköisyysjakaumasta 1000 eri mahdollista polkua, jossa kukin polun piste vastaa kuntoluokassa vuosittain pudonneiden teiden määrää kyseisen polun tyypin kehityksen sattuessa. Teitä voi siis simulaatiossa pudota odotusarvoa vähemmän (alemmat polut) tai enemmän (mikä riskiarvion kannalta on olennaista). X-akselilla on vuodet, ja y-akselilla polkua simuloitaessa saatujen pudonneiden tienpätkien määrä kunakin vuonna.

Ylläkuvatut kolme eri oletuksista lähtevää lähestymistapaa on kukin toteutettu kullekin osaverkolle ja 10 vuoden aikahorisontille nykyhetkestä tulevaisuuteen. Aiemmin kappaleessa 'Mallin lähtökohdat' todettiin kuitenkin eri-ikäisten teiden sopivan huonosti tai olevan sopimattomia historiadataan edustamiseen jos todellisia mittauskertoja on vain yksi (tai kaksi). Tämän vuoksi myös riskiarvion laskentamenetelmään tulee suhtautua tietyin varauksin. Kahden mittauskerran historiadatasta voi olla vaarallista ennustaa tulevaisuutta 10 vuotta eteenpäin.

Vaikka teiden kunnan kehitystä voidaan approksimoida sopivilla todennäköisyysjakaumilla, riskimallinnuksessa huomio kiinnittyy kulloisenkin todennäköisyysjakauman muutamien havaintojen häntään. Tämän hännän muoto riippuu suuresti siitä, kuinka edustavia worst-case -skenaario tapauksia estimointiin käytettävästä datasta löytyy. Lyhyt data ei välttämättä sisällä lainkaan niitä huonoimpia tapauksia, joita valitulla luottamuvälillä hyvinkin voisi olettaa kerran sattuvan pitkällä aikahorisontilla. Erityisesti vuosittaisen tai muun sellaisen historiadataan puute heikentää riskiarvioita, mikäli ajatellaan esimerkiksi tietyn vuoden tai sääolosuhteiden ja muiden tekijöiden edesauttavan teiden kunnan huononemista.

Lisäksi on huomioitava, että riskimittari ei anna mitään kuvaa siitä kuinka monen tienpätkän kunto putoaa erittäin huonossa tapauksessa, jonka sattuminen jää todennäköisyysjakauman 95 % rajan takaiselle alueelle. Toisaalta, riskien arviointi esimerkiksi 99 % luottamustasolla ei tällä datamäärällä liene edes järkevää koska vähäinen historiadata sisältää huonosti edustavan joukon tarkasteltavista ääritapauksista validimpaan riskiestimointiin. Tähän riskiarvioinnissa tyypilliseen häntäjakauman ongelmaan palataan myöhemmin.

### 3.2 Tulokset

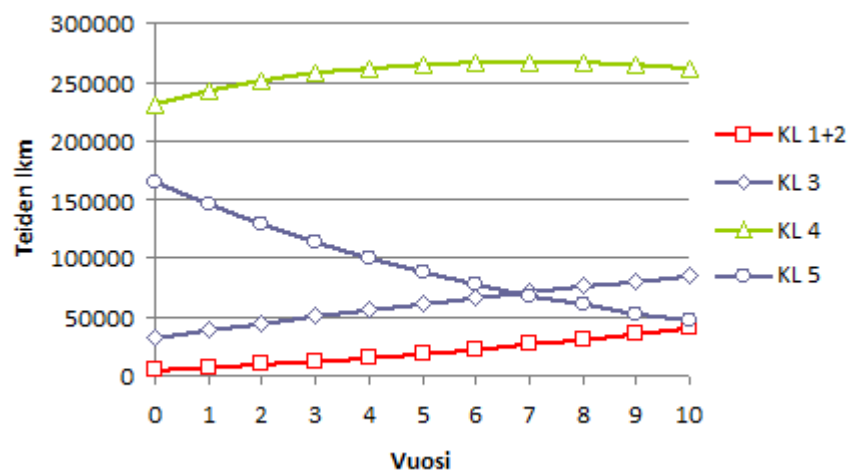
Kuntomallin riskejä simuloitiin 10 vuoden aikahorisontilla tulevaisuuteen, ja alla esitetyissä tuloksissa nähdään, kuinka monta tietä kussakin kuntoluokassa vähintään voidaan odottaa olevan jäljellä kunakin vuonna, 95 % luottamustasolla. Kuvattu riskimallinnus tehtiin erikseen kaikille osaverkoille, ja analyysissä on samalla nostettu esiin havaintoja teiden kunnan muuttumisesta urasyvyyden ja tien tasaisuuden sekä kummassakin kategoriassa vielä mm. liikennemäärien (KVL) ja päällystetyypin (PAB, AB) perusteella.

Kappaleessa esitetyissä tuloksissa keskitytään kolmesta ylläkuvatusta lähestymistavasta vahvan korrelaation antamiin tuloksiin, joka tutkitusta mittausdatasta estimoiduilla todennäköisyysjakaumilla on worst-case -skenaario. Tähän on perusteena se että halutaan välttyä esittämästä riskejä alakanttiin. Toisaalta huomattiin että ero kahden muun tavan ja worst-case lähestymistavan välillä on yleensä alle 15 %, joten näin pystymme perustellusti esittämään tiivistetyn riskianalyysin. Kuvissa 11–16, 20, 21, 26–29 %KL 1+2

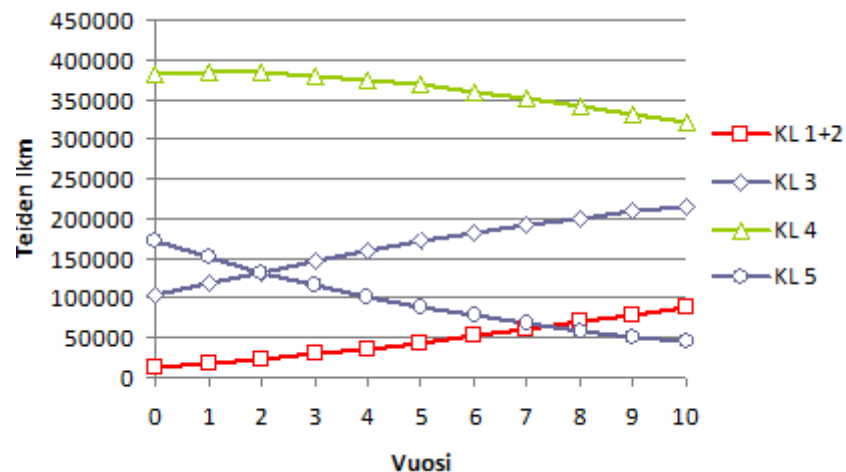
tarkoittaa simuloinnissa saatua huonojen kuntoluokkien suhteellisen osuuden kehittymistä ja %Mean 1+2 vastaavien luokkien odotusarvon kehittymistä laskettuna putoamistodennäköisyys \* teiden määrä. Muissa riskianalysiosuuden kuvissa *teiden lkm* pystyakselilla kuvastaa *teiden määrää*, joka 95 % todennäköisyydellä kuntoluokissa vähintään on jäljellä kunakin vuonna.

### 3.2.1 Kuntoluokitusjakauman kehittyminen tasaisuuden perusteella

Tasaisuuden perusteella simuloitu kuntoluokkien kehitys noudattaa kahta erilaista kehityskaarta, joskin näidenkään eroavaisuudet eivät ole erityisen vahvoja. Kehityskaaret olivat jakautuneet selkeästi tienpäällysteen mukaan, pehmeiden asfalttibetonien kehityskaari on esitetty Kuvassa 9 ja kovien asfalttibetonien kehityskaari Kuvassa 10.



Kuva 9 Tiestön määrän muutos kuntoluokissa 1-5 pehmeiden asfalttibetonien osaverkossa PAB



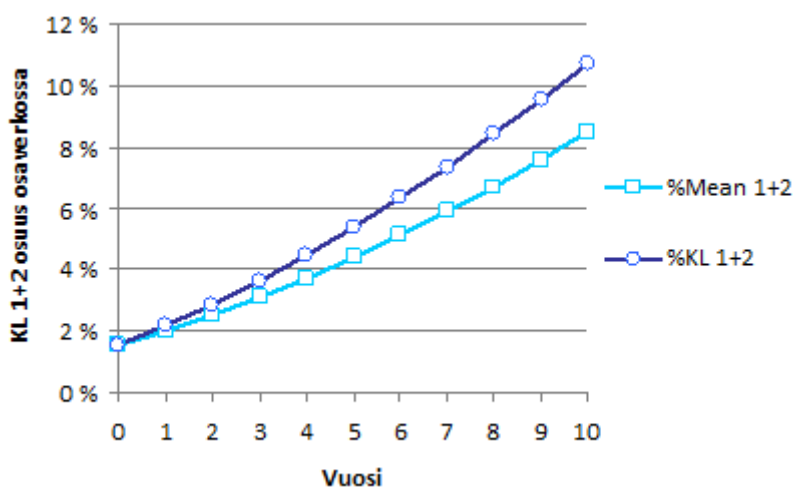
Kuva 10 Tiestön määrän muutos kuntoluokissa 1-5 kovien asfalttibetonien osaverkossa AB

Kummassakin tapauksessa tiemäärät viidennessä kuntoluokassa putoavat keskimäärin samalla tahdilla. Erityisesti, jos otetaan huomioon historiallisten datapisteiden vähyys huomioon, ei selkeää eroa voida päätellä pehmeiden ja kovien viidennessä kuntoluokkien välillä. Pikaisen tarkastelun perusteella voisi

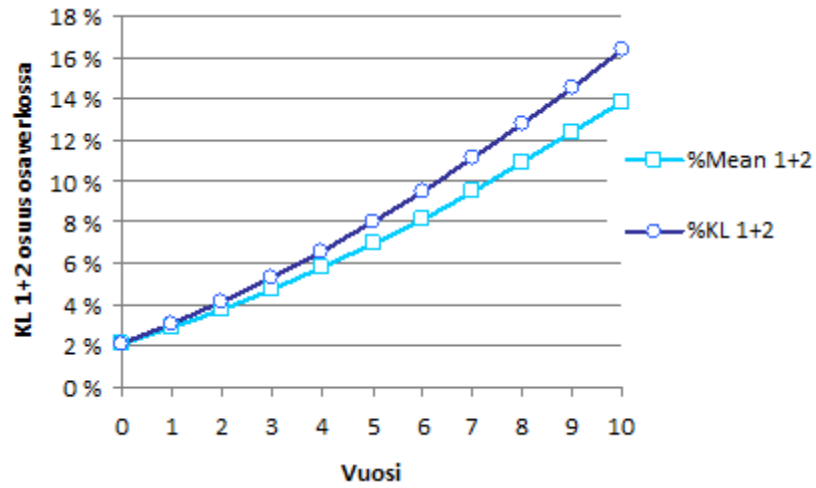
olettaa, että kuntoluokka neljän tiemäärät kehittyvät eri tavalla pehmeiden ja kovien päällysteiden välillä. Tarkemmin syy- ja seuraussuhteita jäsentämällä voidaan kuitenkin todeta että näin ei ole. Kovalla päällysteellä neljännen kuntoluokan alussa havaittu selkeä kasvu johtuu viidennen kuntoluokan suhteellisesti suuremmasta osuudesta verrattuna pehmeän päällysteen luokkaan viisi. Tämä ilmiö vaimenee kuntoluokalle kolme, jossa kehitys kahden eri päällysteen välillä ei ole selkeästi eroavainen.

Koska kovan päällysteen neljännen kuntoluokan alun kasvutrendi ei siirry kolmannelle kuntoluokalle, tunnistetaan neljännen kuntoluokan tuovan suurimman eroavaisuuden näiden kahden päällysteen välille. Kovalla päällysteellä kuntoluokasta neljä putoaa kuntoluokkaan kolme suhteellisesti selvästi vähemmän teitä kuin pehmeällä päällysteellä. Huonot kuntoluokat, yksi ja kaksi, kasvavat pehmeällä päällysteellä nopeammin, mutta tämä ilmiö selittyy kolmannen kuntoluokan suhteellisella suuremmuudella. Koska pehmeillä päällysteillä kolmannessa kuntoluokassa on suhteellisesti enemmän teitä, sieltä myös putoaa enemmän huonoihin kuntoluokkiin.

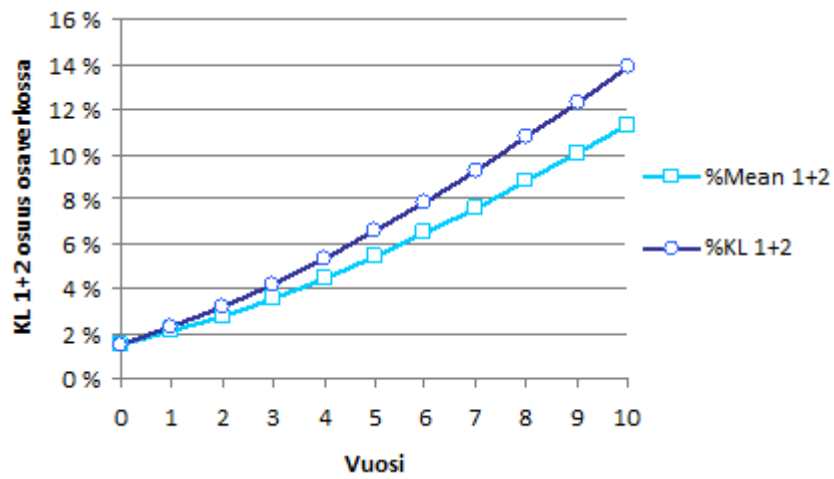
Verrattaessa simulaatiotuloksia teoreettisesti laskettuun kehityskulun odotusarvoon huomataan erikoinen ilmiö. Vahvasti korreloivan simulaation mukaan Oulussa ja Lapissa teitä siirtyy huonoihin kuntoluokkiin odotusarvoon nähden suhteellisesti enemmän kuin muualla Suomessa (Kuvat 11–14). Alle ja yli 350:n keskimääräisellä vuorokausiliikenteellä Oulun ja Lappi verrattuna muuhun Suomeen saavuttaa kymmenessä vuodessa noin viisi prosenttiyksikköä enemmän huonokuntoisia teitä suhteessa odotusarvoon. Koska liikennemäärän kahtiajako ei kerro vielä mitään liikenteen keskiarvosta, voidaan tehdä kaksi vaihtoehtoista päätelmää. Joko Lapin ja Oulun alueella tieverkostoa käytetään keskiarvoisesti enemmän kuin muualla Suomessa, tai pohjoisten seutujen sääoloilla on selkeä vaikutus teiden tasaisuuden heikkenemiseen.



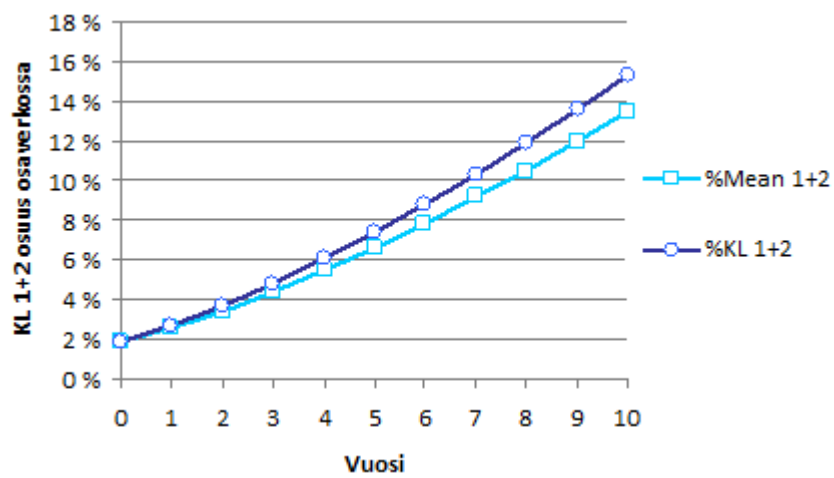
Kuva 11 PAB, Oulu ja Lappi missä KVL ovat alle 350



Kuva 12 PAB, muu Suomi missä KVL on alle 350



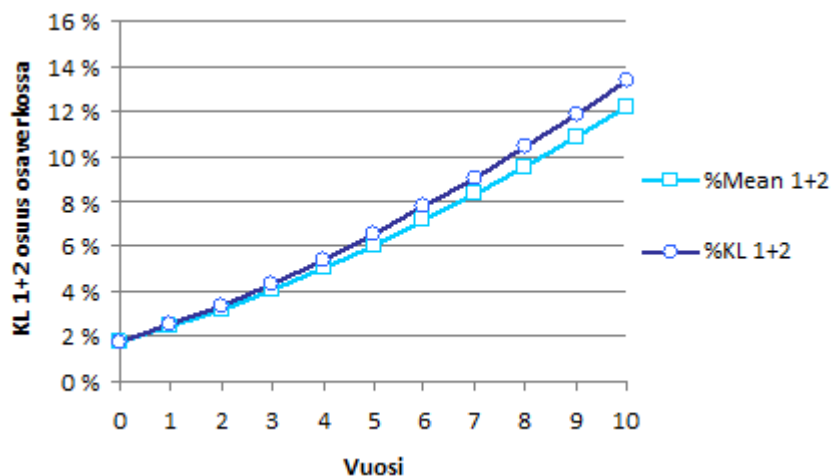
Kuva 13 PAB, Oulu ja Lappi missä KVL ovat yli 350



Kuva 14 PAB muu Suomi missä KVL on yli 350

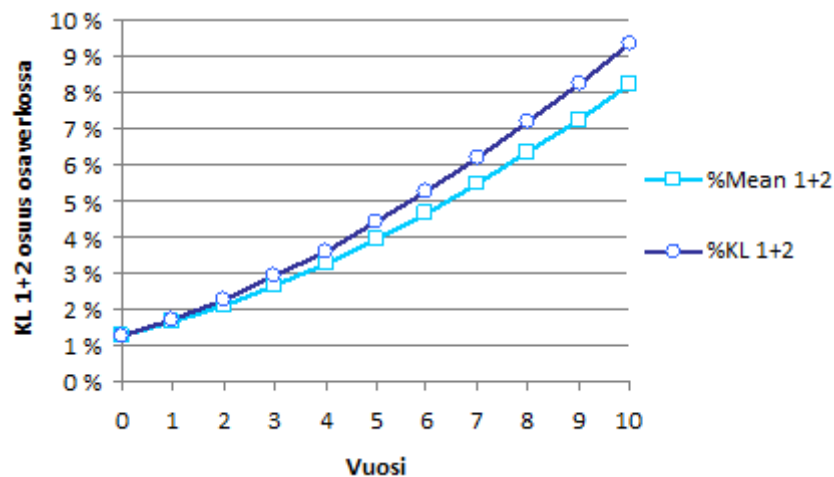
Myöskään keskimääräisen liikennemäärän suhde tieverkon huononemiseen ei ollut alkuperäisten odotusten mukainen. Teiden, joiden keskimääräinen liikennemäärä oli alle 350, huonot kuntoluokat saavuttivat kymmenessä vuodessa noin viisi prosenttiyksikköä suuremman odotusarvoon suhteutetun arvon kuin tiet yli 350 päivittäisellä liikennemäärällä. Liikennemäärä erottavana tekijänä ei ole edes täysin selitettävissä teiden alkuperäisten jakaumien eroavaisuudella, sillä ne ovat yli ja alle 350 luokissa kohtuullisen samanlaiset. Hypoteesina voisi olettaa esimerkiksi että suuremmalle liikennemäärälle suunniteltu tie on rakenteeltaan kestävämpi kuin kevyemmälle rasiukselle suunniteltu.

Kovan päällysteen kehityksessä liikennemäärällä vaikuttaa olevan selkeästi heikompi vaikutus tien kulumiseen tasaisuuden kannalta. Päiväkohtaisen liikennemäärän kasvaessa alle 1500:sta yli 3000:een huonojen teiden odotusarvoon suhteutettu määrä kasvaa jopa yhdeksän prosenttiyksikköä, mutta yli 3000 liikennemäärän aliverkossa kuntoluokkajakauma on hyvin erilainen: luokat kolme ja neljä ovat suhteellisesti selvästi pienempiä. Jos lähtötilanne olisi kaikilla rasiuksiluokilla sama, yli 3000 aliverkko huononisikin kovista päällysteistä todennäköisesti nopeiten. Eri aliverkkojen luokkakohtaisista aliverkkokuvaajista voidaankin päätellä, ettei pehmeiden ja kovien päällysteiden simuloitussa kehityksessä ei ole selkeää eroa keskenään (Kuvat 15 ja 16). Pehmeiden teiden huonojen kuntoluokkien jyrkempi kehittyminen suhteessa odotusarvoon johtuu aloitushetkenä huomattavasti suuremmasta kolmosluokan tiemäärän osuudesta koko aliverkossa.



Kuva 15 KL 1 ja 2 osuuden kasvu PAB tiestössä



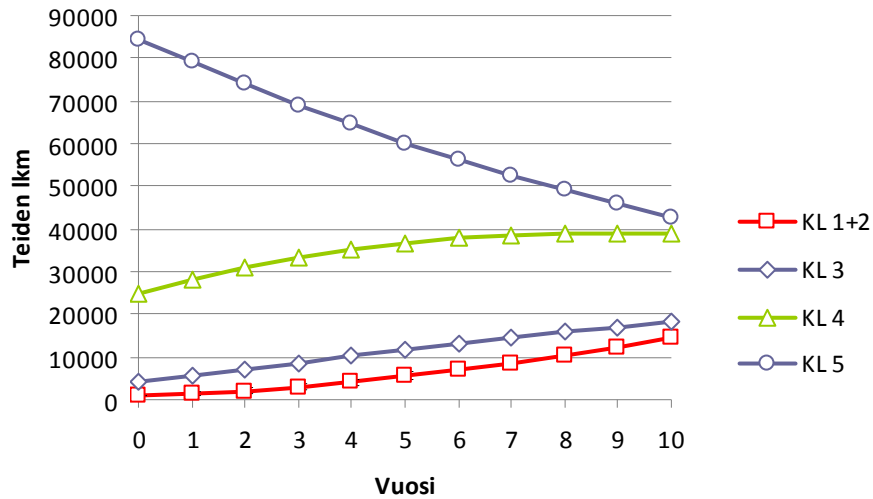


Kuva 16 KL 1 ja 2 osuuden kasvu AB tiestössä

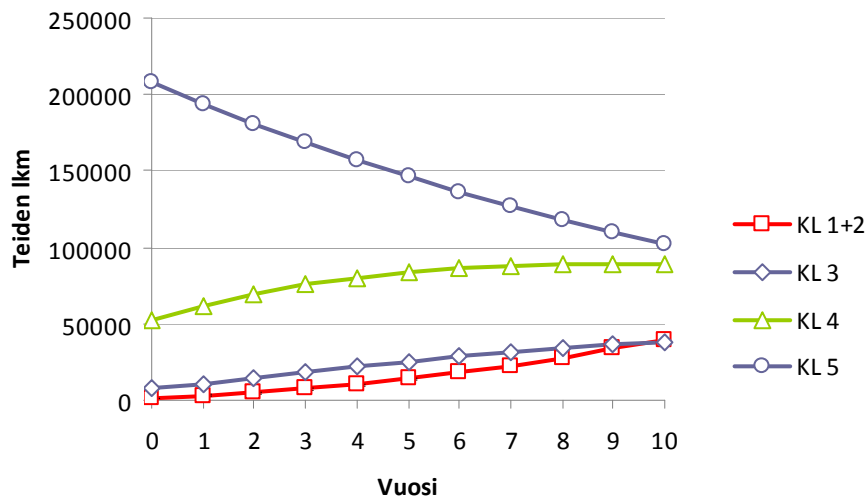
Tasaisuuden perusteella tehty kuntoluokitus ei synnytä todella selkeitä eroja kuntoluokkajakauman kehittymisessä. Yhteenvedona voidaan todeta liikennemäärän vaikuttavan teiden kulumiseen nopeammin pehmeillä päällysteillä kuin kovilla. Erityisesti jatkotarkasteluiden kannalta olisi merkittävää tarkastella onko tien pohjarakenteella ja ympäröivällä ilmastolla mallinnettava vahva vaikutus tasaisuuden rappeutumiseen. Ainakin simuloinnin perusteella näillä voisi olla mahdollisuuksia kasvattaa tien rappeutumisennusteiden selitysastetta. Lisäksi, erikoishavaintona huomattiin tasaisuuden tarkastelussa teiden pysyvän sitkeämmin kuntoluokassa neljä, jos kyseessä on kova päällyste.

### 3.2.2 Kuntoluokitusjakauman kehittyminen urasyvyyden perusteella

Käytettäessä urasyvyyttä luokkajakauman kehityskriteerinä, aliverkoissa huomataan toista erottelukriteeriämme, eli tien tasaisuutta, enemmän eroavaisuuksia eri tekijöiden vaikuttaessa. Silmiinpistävämpänä erona kovien teiden simulaatiokehityskaarista huomataan heti viidennen luokan erittäin nopea putoaminen seuraavaan kuntoluokkaan. Pehmeillä teillä lasku ei ole läheskään yhtä jyrkkä, mutta tämä voi toisaalta johtua liikennemäärä- ja nopeusrajoituseroista kovien ja pehmeiden teiden välillä.

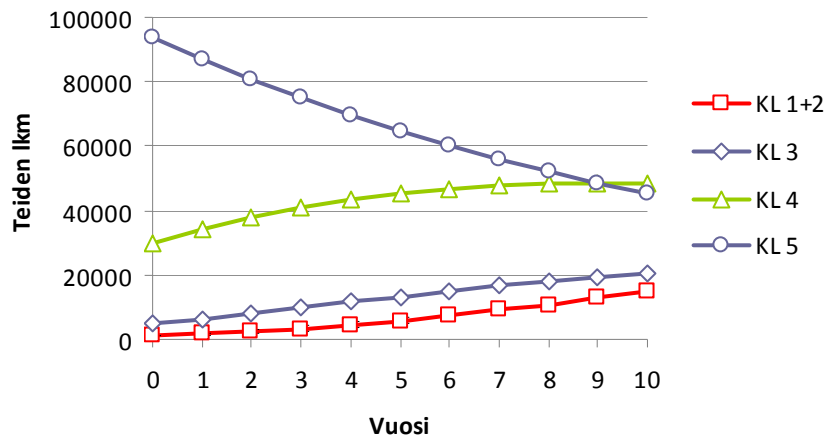


Kuva 17 PAB Oulu ja Lappi



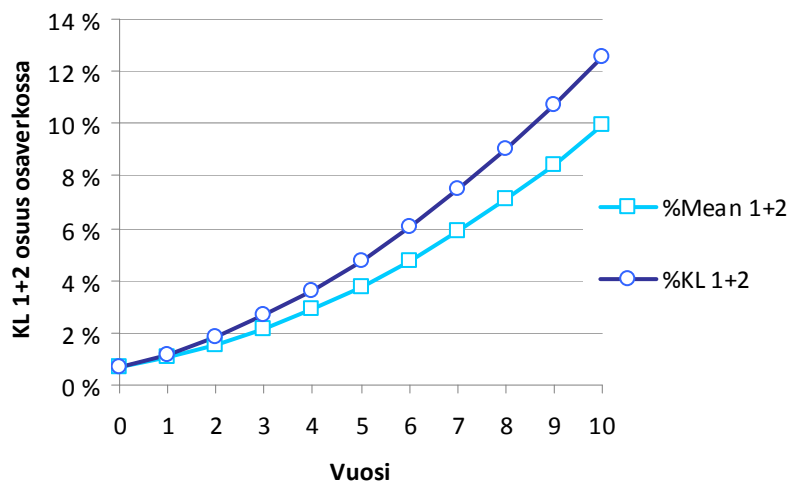
Kuva 18 PAB muu Suomi

Kuvista 17 ja 18 huomataan selkeästi kuinka tulokset urasyvyyden osaverkon kohdalla ovat ainakin maantieteellisesti kohtuullisen tasaiset. Itse asiassa pehmeiden teiden jakauma ja sen kehitys on hyvin samanlainen kovien teiden alle 1500 keskimääräisen liikenteen kanssa (Kuva 19).

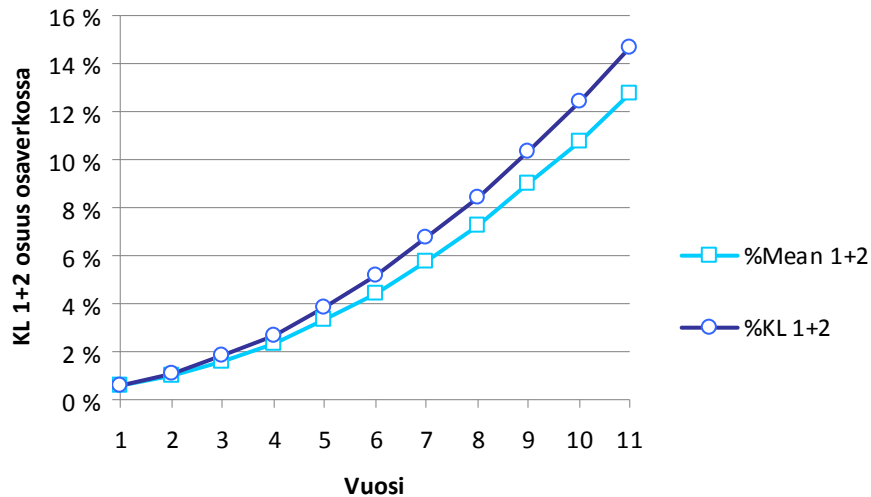


Kuva 19 AB, keskimääräinen vuorokausiliikenne alle 1500

Kuntoluokan 1 ja 2 teiden osuus koko osaverkon tiestöstä taas kehittyy pitkälti samaa tahtia kuin aiemmin tasaisuuden perusteella tehdyissä tarkasteluissa (Kuvat 20 ja 21).

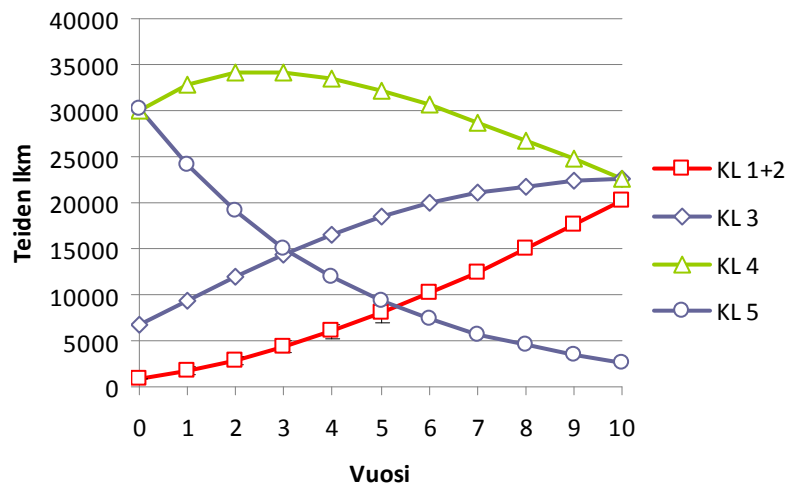


Kuva 20 KL 1 ja 2 osuuden kehitys, PAB Oulu ja Lappi

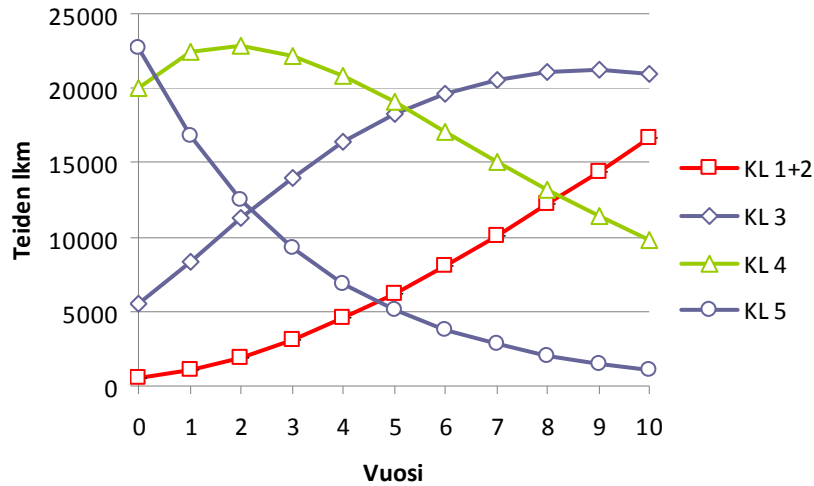


Kuva 21 KL 1 ja 2 osuus, PAB muu Suomi

Selvää eroa aikaisempaan kehitykseen aletaan kuitenkin nähdä kun liikennemäärät kasvavat. Tämä liikennemäärän kasvun vaikutus urasyvyyden kehitykseen alkaa merkitä kun vuorokausiliikenne on luokkaa 1500-3000 autoa. Nopeus taas ei ole yhtä selvästi kulumiseen ja kuntoluokan putoamisriskiin (eli 95% todennäköisyydellä kuntoluokissa jäljellä olevien teiden määrään) vaikuttava tekijä, kuten Kuvien 22 ja 23 muodosta voidaan nähdä.



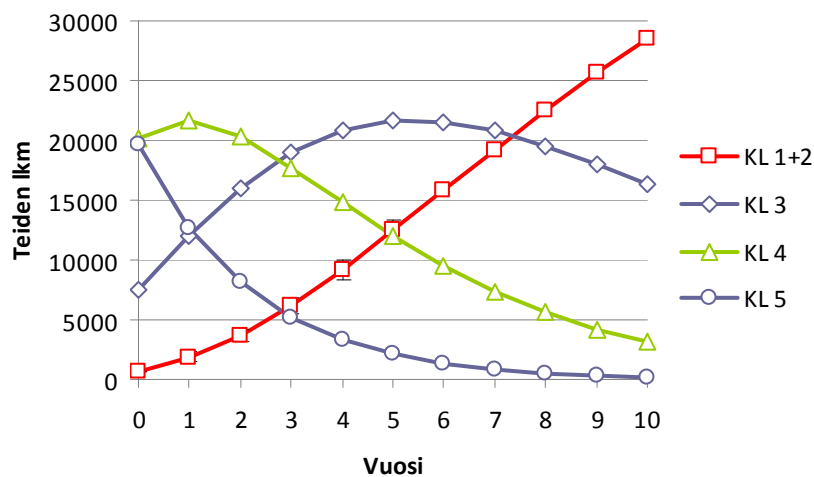
Kuva 22 AB, KVL 1500-3000, nopeudet alle 85 km/h



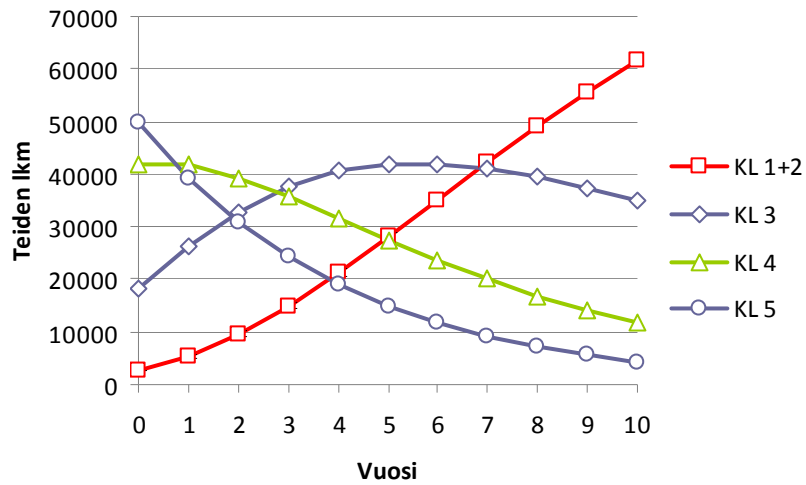
Kuva 23 AB, 1500–3000, nopeudet yli 85 km/h

Liikennemäärien edelleen kasvaessa vaikutus urasyvyyden kehitykseen ei kuitenkaan jatku lineaarisena. Ero KVL 1500-3000 ja sitä suurempien vuorokausiliikennemäärä jaottelujen välillä on vähemmän huomattava kuin ero vähiten liikennöityyn tiestöön (ts. ero jaottelujen ‘alle 1500’ ja ‘1500-3000’ välillä). Jälleen kerran nopeus on vähemmän merkitsevä tekijä; ainoastaan kuntoluokka neljän putoaminen ja kolmosten määrän lisääntyminen näyttää tuloksissa hivenero kasvavan kun nopeus on yli 85km/h osaverkossa AB, KVL 3000-6000 verrattuna vastaavaan verkkoon alle 85km/h nopeuksilla.

Kuvissa 24 ja 25 näkyy vielä kahden suurimman liikennemääräjaottelun vaikutus urien kasvun riskiin pudottaa tie alempaan kuntoluokkaan.

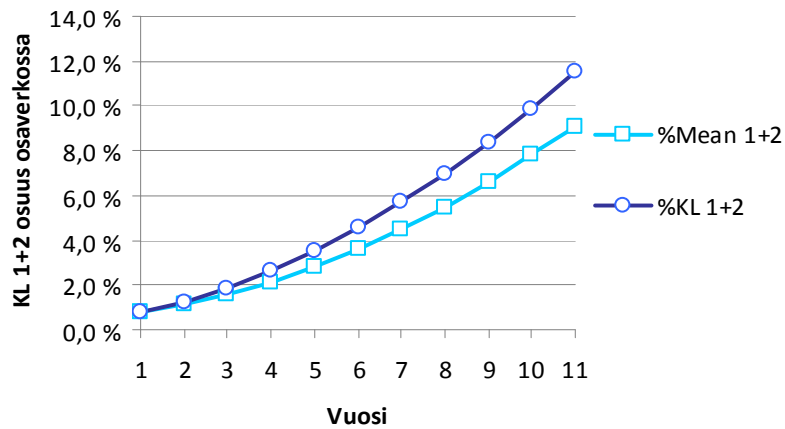


Kuva 24 AB, 3000–6000, nopeudet yli 85km/h

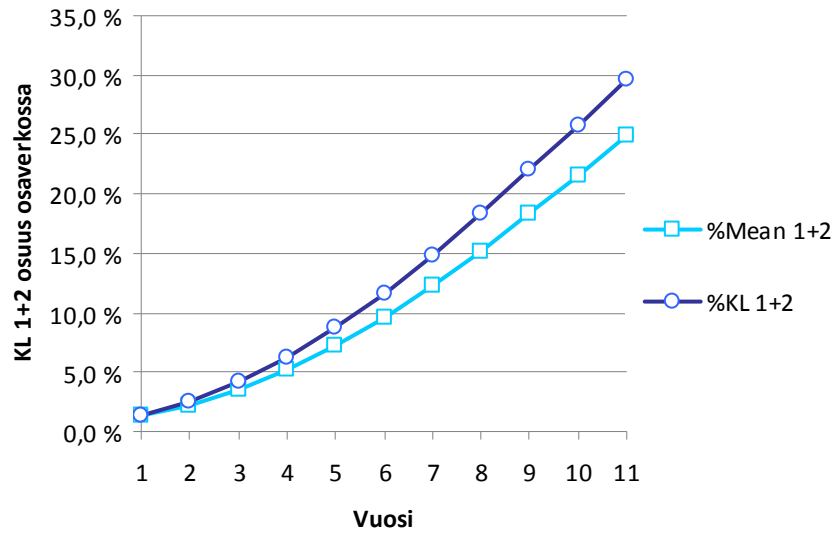


Kuva 25 AB missä KVL yli 6000

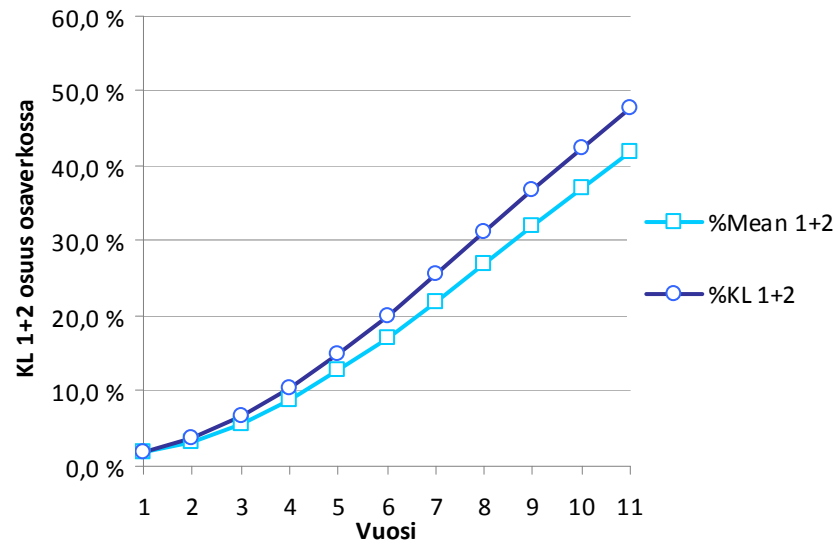
Kuten kappaleen alussa todettiin, liikennemäärät vaikuttavat selvästi kunnan putoamisriskiin urasyvyyden perusteella tehdyssä luokkajaottelussa. Tämä käy selvästi ilmi seuraavien 1 ja 2 kuntoluokan tiestöosuuskehitystä osoittavista kuvista (Kuvat 26-29), missä huomataan merkittävästi eri asteikko pystyakselilla.



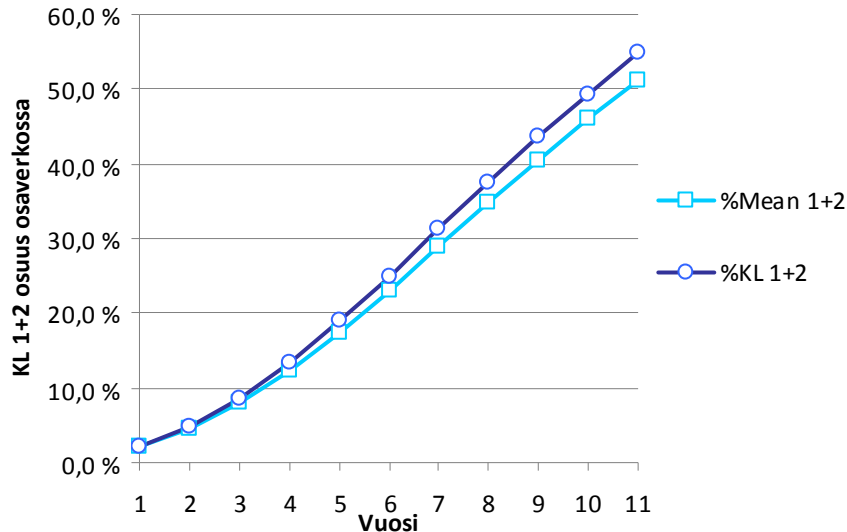
Kuva 26 AB alle 1500



Kuva 27 AB 1500–3000 alle 85km/h



Kuva 28 AB 3000–6000 alle 85km/h



Kuva 29 AB yli 6000

Jälleen kerran nopeus ei vaikuta merkittävästi, poikkeuksena kuitenkin luokka jossa vuorokausiliikennemäärä on välillä 3000-6000. Tällöin yli 85km/h nopeusluokkaa olevilla teillä kuluminen on lähes 10 prosenttiyksikköä suurempaa kymmenen vuoden aikana, kuin Kuvassa 28 esitetty kulumisen riski.

Vaikka liikennemäärä onkin selvä vaikuttaja tässä tarkasteluluokassa, selittyy osa kuvaajien eroista myös kuntoluokkajakaumien suhteen erilaisilla lähtötilanteilla. Kun kuntoluokan viisi teitä on paljon, nosta niiden kunnan putoaminen ensimmäisinä vuosina neljännen kuntoluokan teiden määrää tuloksissa.

## Osa II: Tieverkostojen ennustemallit kirjallisuudessa

### 4 Teiden kunnossa pitojärjestelmä, PMS

Teiden kunto on satunnaismuuttuja, johon vaikuttaa lukuisia ilmiöitä ja tekijöitä. Muun muassa vallitsevat sääolosuhteet voivat vahingoittaa ajoteitä hyvin voimakkaasti ja äkilliset lisäykset tien käytössä voivat rasittaa tien kestävyyttä. Teistä vastaaville tahoille lisää vastuuta ja huolenaihetta lisää myös se, että hyväkuntoiset tiet ovat perusedellytys kehittyneelle yhteiskunnalle. Ihmiset tarvitsevat hyväkuntoisia teitä matkustaessaan muun muassa töihin, lomalle tai jonkin muun tarpeen vuoksi. Lisäksi maan sisäinen logistiikka vaatii hyväkuntoisia teitä, jotta kaupankäynti yritysten ja kuntien välillä ei tyrehtyisi. Kaiken tämän vuoksi teiden kunnossapitoa on pidettävä arvossaan ja siihen on suhtauduttava vakavasti. Hyväkuntoiset tiet mahdollistavat yhteiskunnan monipuoliset aktiviteetit ja jatkuvan kehittymisen.



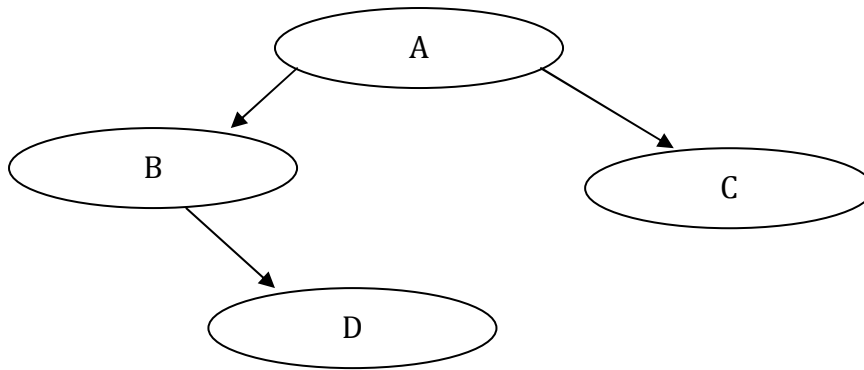
Tarve teiden kunnan ylläpidolle on herännyt jo verrattain aikaisin 1900-luvulla. Erityisesti Yhdysvalloissa on ryhdytty kaikista aikaisimmin laatimaan erilaisia suunnitelmia teiden rakentamiselle ja kunnostukselle. Eräänlaisen kaavamaisen teiden kunnossapitojärjestelmän laatiminen on ollut hyvinkin suuren mielenkiinnon alla erityisesti Yhdysvaltojen säätilojen perusteella äärimmäisissä paikoissa. Esimerkiksi Arizonassa tiet kohtaavat lukuisia erilaisia sääilmiöitä ja -olosuhteita. Toisaalta siellä on äärimmäisen helteinen ilma kun taas toisaalla siellä vallitsee hyvinkin kylmät ja ankarat olosuhteet. Tämänkaltaisissa tilanteissa teiden kunnosta vastaavat tahot joutuvat todellisten haasteiden eteen arvioidessaan, milloin teitä kannattaisi ryhtyä uusimaan ja milloin rakentamaan kokonaan uusia teitä. Tällaisia ongelmia varten on haluttu kehittää eräänlaisia päätöksentekoa tukevia systeemejä, jotka antavat vastauksia muun muassa käytettävien kunnostustekniikoiden valintaan ja kunnostustaajuuteen.

Teiden kunnossapitojärjestelmien tarkoituksena on toimia eräänlaisena päätöksentekovälineenä, joka antaa optimaalisimman toimintavaihtoehdon teiden kunnostuksen suhteen. Yleensä teiden kunnosta ja uudelleenrakentamista on rajoittamassa lukuisia tekijöitä kuten käytettävissä oleva budjetti ja haluttu hyöty tienkunnostamisesta. Tämän systeemin tarkoituksena on tuoda esille teiden tärkeyttä yhteiskunnallisena voimavarana ja lisäksi kanavoida yhä enemmän huomiota teiden kuntoon ja myös antaa vaihtoehtoja erilaisille kunnossapitoratkaisuille. Järjestelmien rakentamisessa on käytetty erityisen paljon hyödyksi Markov-mallinnusta, mutta myös erinäiset Bayes-verkko kuvaukset ja neuroverkot ovat yleistyneet todella merkittävästi. Järjestelmän idea on ottaa huomioon teiden kunnan satunnainen heikkeneminen ja pyrkiä ennustamaan mahdollisten korjaustoimenpiteiden laajuutta ja taajuutta.

Teiden kunnossapitojärjestelmien hyötynä voidaan pitää päätöksenteon systemaattisuutta ja eräänlaisen epävarmuuden hallitsemista. Ilman kaavamaista teiden kunnossapitojärjestelmää eivät teistä vastaavat tahot kykenisi tehokkaasti vastaamaan teiden kunnan heikkenemiseen ja teiden käyttäjien yhä koveneviin vaatimuksiin. Nämä vaatimukset voivat liittyä esimerkiksi teiden ajomukavuuteen ja nopeuteen. Systemaattisen mallinnuksen avulla voidaan ennakoitavasti allokoida tietty määrä budjetista teiden jatkuvaan kunnossapitoon ja loput teiden radikaalimpaan uudistamiseen. (Golabi ym., 1982; Finn, 1998.)

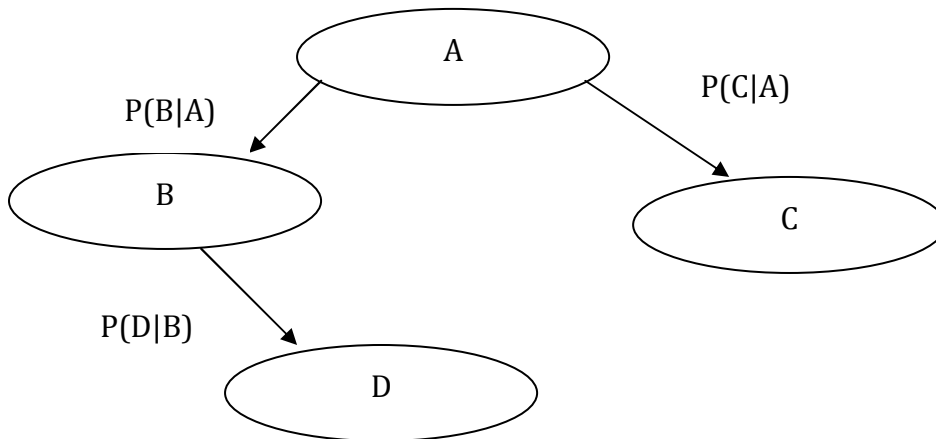
## 5 Bayes-verkot

Eri ilmiöiden todennäköisyyksien systemaattinen kuvaaminen on nostattanut tarpeen bayesiläiseen todennäköisyysmallinnukseen. Bayes-verkon idea on esittää eri ilmiöt solmuina, joita yhdistetään toisiinsa ilmiöihin ja joiden välisiä ehdollisia todennäköisyyksiä pyritään määrittämään. Alla on yksinkertainen esimerkki Bayes-verkosta:



Kuva 30 Bayes-verkon periaate. Solmuista A-D näkyy näiden väliset vaikutukset nuolina

Kuvassa 30 siis A, B, C ja D ovat satunnaismuuttujia, joilla on tietyt todennäköisyydet  $P(A)$ ,  $P(B)$ ,  $P(C)$  ja  $P(D)$ . Nuolet edustavat satunnaismuuttujien välisiä ehdollisia todennäköisyyksiä. Täydennetty Bayes-verkko on esitetty alla (Kuva 31):



Kuva 31 Bayes-verkko ja ehdolliset todennäköisyydet

Bayes-verkon avulla voidaan laskea eri ehdolliset todennäköisyydet, kunhan tiedämme ilmiöiden juuritodennäköisyydet. Kun saadaan konstruoitua verkon rakenne ja laskettua ehdolliset todennäköisyydet, määräytyy muuttujien yhteisjakauma tätä kautta helposti. Muuttujien yhteisjakauma on siis muotoa

$$P(x_1, \dots, x_n) = P(x_n | x_{n-1}, \dots, x_1) P(x_{n-1}, \dots, x_1) \quad (4)$$

Loppujen lopuksi purkamalla yllä olevaa lauseketta saadaan yhteisjakauman lausekkeeksi seuraavaa:

$$P(x_1, \dots, x_n) = \prod_{i=1, \dots, n} P(x_i | x_{i-1}, \dots, x_1) \quad (5)$$

Bayes-verkoissa tämä laskukaava yhteisjakaumalle pätee mille tahansa satunnaismuuttujien joukolle. Tämä johtaa siihen, että solmujen on oltava ehdollisesti riippumattomia toisistaan. (Charniak, 2002.)

## 6 Markov-mallit

### 6.1 Yleistä

Matematiikassa Markov-ketjulla tarkoitetaan stokastista prosessia, jolla on Markov-ominaisuus. Toisin sanoen, mikäli tiedetään nykyinen tila, tulevat tilat riippuvat vain nykyisestä tilasta ja ne ovat riippumattomia aiemmista tiloista (Durrett, 1999). Nykyinen tila pitää siis sisällään kaiken informaation, joka vaikuttaa prosessin tulevien tilojen kehitykseen. Matemaattisesti asia ilmaistaan seuraavasti. Markov-ketju on joukko satunnaismuuttujia  $X_1, X_2, X_3, \dots$ , joille pätee ehdollinen todennäköisyys

$$\Pr(X_{n+1} = x \mid X_n = x_n, \dots, X_1 = x_1) = \Pr(X_{n+1} = x \mid X_n = x_n). \quad (6)$$

Markov-ketju on nimetty venäläisen matemaatikon Andrey Markovin (1856 – 1922) mukaan. Ketjun idea on, että nykyisestä tilasta voidaan siirtyä toisiin tiloihin tietyn todennäköisyysjakauman mukaan tai vaihtoehtoisesti voidaan pysyä nykyisessä tilassa. Todennäköisyyksiä siirtyä tiloista toisiin kutsutaan tilasiirtotodennäköisyyksiksi. Markov-ketjuilla on erittäin paljon sovellusalueita, esimerkiksi fysiikassa, taloustieteissä ja jopa biologiassa. Eräs tunnetuimmista esimerkeistä on niin sanottu satunnaiskävely (*engl. random walk*). Siinä liikutaan yhtä suurilla todennäköisyyksillä nykyisestä tilasta uusiin tiloihin ja esimerkiksi osakkeiden hinnan heilahteluja sekä molekyylin kulkemaa rataa kaasussa tai nesteessä voidaan mallintaa satunnaiskävelyllä.

### 6.2 Markov-mallit teiden kunnan ennustamisessa

AASHO: n (1962) tietestistä asti tieverkon kunnan ennustamista on tutkittu erittäin intensiivisesti. Ensimmäisen kerran Markov-prosessia käytti Woodward-Clyde-Consultants- niminen yritys. Heidän edistyksellisessä työssään kehiteltiin verkko-optimointisysteemi (*engl. network optimization system, NOS*). He sovelsivat Markov-ketjuja yhdessä lineaarisen ohjelmoinnin kanssa ja se oli ensimmäinen kerta, kun menestyksekkäästi onnistuttiin luomaan suuren datamäärän optimointiin ja stokastisen ennustamisen tekniikkaan perustuva matemaattinen malli tieverkon kunnan hallintaan (Wang ym., 1994). Arizonan osavaltio Yhdysvalloissa otti NOS: n käyttöön ja ensimmäisen implementointivuoden aikana osavaltio säästi sen ansiosta 14 miljoonaa dollaria (Golabi ym., 1982), joten kustannusten kannalta systeemi oli erittäin merkittävä.

Useat tutkijat, (mm. Way ym., 1982; Butt ym., 1987; Li ym., 1996; Abaza & Ashur, 1999; Hong & Wang, 2003; Abaza, 2006), ovat käyttäneet Markov-malleja teiden kunnossapidon ennustamiseen AASHO: n sekä muiden tutkimusten innoittamana. Tosin viime vuosina muitakin malleja on alettu soveltaa yhä enenevässä määrin. Näistä mainittakoon esimerkiksi edellä mainittu bayesiläinen mallintaminen ja myöhemmin esiteltävä neuroverkot. Teiden kunnan ennustamista käsittelevät Markov-malleihin perustuvat artikkelit ovat

luonteeltaan usein case-tutkimuksia, joissa tutkitaan mallin soveltuvuutta johonkin tieverkkodataan. Tutkimukset ovat pääpiirteissään muiden kuin suomalaisten tekemiä ja tieverkkodatana on myös käytetty ulkomaista dataa. Sirviön ja Holménin (2008) tutkimus on ainoa löytämämme suomalaisten tutkimus, jonka aihepiiri on lähellä omaamme.

Markov-malli teiden kunnan ennustamiseen on stokastinen prosessi, jolla on kolme rajoitusta (Ortiz-Garcia ym., 2006):

- Teiden kunnan huonontuminen on ajan suhteen jatkuva. Kuitenkin se saadaan muutetuksi diskreetiksi, koska tieverkon kuntoa analysoidaan vain tietyin väliajoin, esimerkiksi kerran vuodessa.
- Tila-avaruus, eli mahdollisten kuntoluokitusten määrä on ääretön. Kuitenkin todellisuudessa tila-avaruutena käytetään äärellistä määrää kuntoluokkia.
- Markov-ominaisuuden oletetaan olevan voimassa.

Ensimmäinen rajoitus tulee luonnollisesti siitä, että teiden kuntoa tarkistetaan vain tietyin väliajoin. Ei ole järkeä tutkia teiden kuntoa päivittäin tai viikoittain, koska kunnan huonontuminen niin lyhyessä ajassa on todennäköisesti melko minimaalista. On kuitenkin tärkeää tehdä tutkimukset tasaisin väliajoin. Tietenkin mitä lyhyemmin väliajoin tehdään kokeita sitä sileämmäksi ja jatkuvammaksi muodostuu kunnan huonontumista ajan suhteen kuvaava käyrä. Tila-avaruudeksi saadaan periaatteessa kuinka suuri tahansa sen mukaan mitä muuttujia otetaan mukaan kunnan huonontumisen selittäjiksi. Yleensä selittäjiksi valitaan kuitenkin useista mahdollisista vaihtoehdoista vain muutama, joilla oletetaan olevan eniten merkitystä ja niiden perusteella lasketaan kuntoluokat. Kuntoluokat ovat toki karkeita, sillä esimerkiksi tässä työssä on viisi kuntoluokkaa ja varmasti voidaan spekuloida kuuluisiko joku kuntoluokka erilaisella määrittelytavalla toiseen luokitukseen. Mutta, koska datamäärä on suuri, pitää tärkeimmät tekijät seuloa joukosta ja luokitella tiet niiden perusteella. Artikkeleissa ei kyseenalaisteta Markov-ominaisuuden olemassa oloa.

### 6.3 Tilasiirtotodennäköisyysmatriisin määrittelyminen

Lähtökohtana teiden kunnan ennustamisessa on tilasiirtotodennäköisyysmatriisi. Teiden kunto jaetaan viiteen kuntoluokkaan 1, 2, 3, 4 ja 5. Vastaavat kvalitatiiviset luokitukset ovat yleensä kirjallisuudessa erittäin hyvä, hyvä, kohtalainen, huono ja erittäin huono. Tässä työssä ryhmämme käyttää kuitenkin luokitusta, joka on juuri päinvastainen, eli kuntoluokkaa 1 vastaa kvalitatiivinen luokitus erittäin huono ja kuntoluokkaa 5 vastaa luokitus erittäin hyvä. Käytämme siis tällaista luokitusta jatkossa.

Markov-ketjujen tapauksessa teiden kunto voi yhdellä aika-askeleella joko pysyä samana tai pudota yhden luokituksen huonommaksi. Esimerkiksi jos ollaan

tilassa 4, niin seuraavan askeleen jälkeen ollaan joko tilassa 4 tai 3. Siirtymisille tiloista toisiin estimoidaan todennäköisyydet. Koska mahdollisia tiloja on viisi kappaletta, niin tilasiirtotodennäköisyysmatriisin koko on  $5 \times 5$ . Seuraavassa on esitetty tilasiirtotodennäköisyysmatriisi  $P_1$ .

$$P_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ P_{21} & P_{22} & 0 & 0 & 0 \\ 0 & P_{32} & P_{33} & 0 & 0 \\ 0 & 0 & P_{43} & P_{44} & 0 \\ 0 & 0 & 0 & P_{54} & P_{55} \end{bmatrix}$$

Matriisin ylin vasen alkio on 1, koska erittäin huonolta tilalta ei päästä muihin tiloihin ilman korjaustoimenpiteitä. Esimerkiksi alkio  $P_{32}$  kuvaa todennäköisyyttä, että kuntoluokassa 3 oleva tieosuus on yhden aika-askeleen kuluttua kuntoluokassa 2. Vastaavasti alkio  $P_{33}$  kuvaa todennäköisyyttä, että yhdellä aika-askeleella pysytään yhä kuntoluokassa 3, mikäli siinä oltiin alussa. Matriisin kaikilla riveillä todennäköisyyksien summan tulee olla 1.

Markov-malleissa tilasiirtotodennäköisyysmatriisin avulla voidaan ennustaa tulevien vuosien teiden kuntoa. Ensin pitää määrittää alkujakauma teiden kunnolle, eli mitkä ovat suhteelliset osuudet alussa eri kuntoluokille. Olkoon alkujakauma  $a_0 = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)$ , missä  $\alpha_i, i = 1, 2, 3, 4, 5$  kuvaa kuntoluokkien 1 - 5 suhteellisia osuuksia alussa. Suhteelliset osuudet hetkellä  $t$  saadaan kaavalla

$$a_t = a_0 P_1^t. \tag{7}$$

Tilasiirtotodennäköisyys korotetaan siis potenssiin  $t$  ja kerrotaan alkujakaumalla. Otettaessa korjaustoimenpiteet mukaan, saadaan tilatodennäköisyysmatriisi  $P_2$ . Alkiot  $X_{ij}$ , kuvaavat ylläpito- ja korjausmuuttujia. Ylläpito- ja korjausmuuttujien ollessa mukana mallissa, voidaan jokainen päällyste nostaa tiloilta 1, 2, 3 ja 4 niitä korkeammille tiloille. Kyseessä on ongelma, jossa on ratkaistava milloin ja mitä teitä ylläpidetään ja korjataan. Useassa teiden kunnan ennustamista käsittelevässä artikkelissa käsitellään pelkästään todennäköisyyksiä  $P_{ij}$ , eli todennäköisyyksiä pysyä yhden aika-askeleen jälkeen samassa kuntoluokassa kun ennen aika-askelta ja todennäköisyyksiä, että kuntoluokitus huononee yhdellä. On kuitenkin olemassa myös artikkeleita, esimerkiksi Abaza ym. (2004) ja Abaza (2006), joissa tutkitaan teiden ylläpidon ja korjausten vaikutusta. Tällöin käytetään matriisia  $P_2$ .

$$P_2 = \begin{bmatrix} P_{11} & X_{12} & X_{13} & X_{14} & X_{15} \\ P_{21} & P_{22} & X_{23} & X_{24} & X_{25} \\ 0 & P_{32} & P_{33} & X_{34} & X_{35} \\ 0 & 0 & P_{43} & P_{44} & X_{45} \\ 0 & 0 & 0 & P_{54} & P_{55} \end{bmatrix}$$

Todennäköisyyksien  $P_{ij}$  estimointiin on olemassa useita tapoja. Mikäli historiadataa on tarpeeksi, voidaan todennäköisyydet määrittää sen avulla. Usein ei kuitenkaan ole saatavilla niin paljon dataa, että voitaisiin luotettavasti ennustaa kyseisiä todennäköisyyksiä. Tällöin voidaan joko tukeutua asiantuntija-arvioihin tai käyttää edistyneempiä menetelmiä. Asiantuntijoina voivat olla esimerkiksi teiden ylläpitoon erikoistuneet insinöörit. Asiantuntija-arviot ovat erityisen hyödyllisiä yritysille, jotka eivät omaa yksityiskohtaisia teiden kunnan tutkimusmetodeja. Koska asiantuntija-arviot ovat riippuvaisia yksittäisten insinöörien mielipiteistä, sopivia matemaattisia tekniikoita täytyy käyttää, jotta todennäköisyydet eivät ole pelkästään subjektiivisia (Bandara ym., 2001).

Abaza ym. (2004) esittelevät kolme menetelmää tilasiirtotodennäköisyyksien määrittämiseksi. Ensimmäinen menetelmä on hyvin yksinkertainen. Jos aluksi  $N_0$  päällysteyksikköä on tilalla  $i$  ja jos yhden aika-askelen jälkeen  $N_f$  päällysteyksikköä on vastaavalla tilalla, niin tilasiirtotodennäköisyydet voidaan estimoida kaavoilla

$$P_{i+1,i} = \frac{N_0 - N_f}{N_0} \quad (8)$$

ja

$$P_{i,i} = 1 - P_{i+1,i} = \frac{N_f}{N_0}. \quad (9)$$

Yllä oleva menetelmä antaa realistiset todennäköisyydet, jos tarkastelun kohteena on suuri määrä päällysteyksiköitä. Toinen menetelmä perustuu siihen, että estimoidaan periodin pituutta  $D_i$ , jonka päällysteyksikkö pysyy tilalla  $i$ . Olkoon  $t$  tapahtuneiden tilasiirtojen aikavälin pituus, jolloin seuraavalla yhtälöllä voidaan estimoida tilasiirtotodennäköisyydet

$$P_{i+1,i} = \frac{t}{D_i} \leq 1, \quad (10)$$

missä

$$\sum D_i = T$$

missä  $T$  on yleensä estimoitu aika, joka saadaan data-aineiston perusteella. Kolmas menetelmä perustuu malliin, jonka otti ensimmäisenä käyttöön Yhdysvaltain armeija (Butt ym., 1987). Siinä minimoidaan virhettä, joka saadaan päällysteen kunnan indeksin (*engl. pavement condition index, PCI*) havaitun arvon sekä sen ennusteen erotuksena.

#### 6.4 Tuloksia

Tarkasteltaessa optimaalista ylläpito- ja korjauspolitiikkaa, on useita vaihtoehtoja, joita voidaan soveltaa. Abaza ym. (2004) esittävät artikkelissaan menetelmän, jossa optimoidaan painofunktiota

$$F = \sum_{i=1}^5 w_i \times Q_i^{(n)}, \quad (11)$$

missä  $w_i$  ovat painoja ja  $Q_i^{(n)}$  ovat tilan  $i$  suhteellisia osuuksia. Rajoitteena on käytettävissä oleva budjetti. Artikkelissa käytetään kuutta funktiota, joista toisissa maksimoidaan erittäin hyvien tai hyvien teiden suhteellisia osuuksia ja toisissa minimoidaan erittäin huonojen tai huonojen teiden suhteellisia osuuksia. Myös kohtalaisten eli kuntoluokitusta 3 vastaavien teiden suhteellisen osuuden maksimointia sekä painotetun keskiarvon menetelmää testataan. Optimaalinen tulos saavutetaan maksimoimalla hyvien ja erittäin hyvien teiden suhteellisia osuuksia tai minimoimalla erittäin huonojen teiden suhteellisia osuuksia. Mikäli kustannusnäkökulma on tärkeä, on huomattava, että kustannukset vaihtelevat melko paljon eri vaihtoehdoissa.

Dussadee ja Pannapa (2007) vertailevat jatkuva- ja diskreettiaikaista mallia muun muassa niiden ennustetarkkuuden perusteella. Diskreettiaikaisena mallina käytetään Markov-mallia ja datana hyödynnetään Thaimaan hallinnolta saatua tiedataa. jatkuva-aikainen malli on parempi, mikäli tutkitaan tieverkkoja, joilla on samanlaisia ominaisuuksia keskenään, kuten että ne sijaitsevat samalla seudulla. Kuitenkin vahvan ennustettavuuden sekä muuttujien vaikutusten vertailukykyyn takia, diskreettiaikainen Markov-malli on parempi Thaimaan teiden kunnan ennustamiseen. Lisäksi malliin sisällytettäviin muuttujiin liittyvä riski ja epävarmuus pystytään ottamaan mukaan tulevissa tutkimuksissa.

Sirviön ja Holménin, jotka työskentelevät teknillisessä korkeakoulussa, tutkimuksessa vertaillaan Markov-malleja ja keinotekoisia neuroverkkoja. He tutkivat ennen kaikkea miten teiden urat sekä karkeus muuttuvat ajassa ja mittarina heillä on maailmanlaajuinen standardoitu kuntomuuttuja IRI (*engl. international roughness index*), joka mittaa ajomukavuutta. Artikkelin mukaan Markov-mallit ovat selvästi nopeampia ja yksinkertaisempia ja ne antavat parempia ennusteita kuin neuroverkot. Eräs syy tähän voi olla, että neuroverkkoja opetetaan liiaksi ja ne eivät pysty mukautumaan uuteen dataan. Markov-mallien ongelmana on huonojen teiden vähäinen osuus datan joukossa

ja tästä syystä tulokset eivät ole harhattomia ja pitkän ajan ennustaminen on hankalaa. Tätä ongelmaa voitaisiin helpottaa ottamalla maantieteellisesti suurempi otos. Tutkimus kannustaa Markov-mallien käyttöön tulevaisuudessa ja motivoi niiden tutkimiseen jatkossa.

Markov-malleja voidaan käyttää myös muiden menetelmien kanssa ennustettaessa teiden kuntoa. Näistä mainittakoon esimerkiksi sumeat joukot, joita tutkivat Bandara ja Gunaratne artikkelissaan. Myös Wang ym. tutkimus päällysteen kunnosta todennäköisyyslaskentaa apuna käyttäen on syytä mainita. Heidän artikkelissaan kehitellään niin sanottu päällysteen käyttäytymisen todennäköisyyskäyrä (*engl. probabilistic behavior curve*) ja he miettivät ideaa, että teiden palvelutasolle määritetään minimitaso, jota ei saa alittaa. Kunnostukset pitää tehdä siten, että minimitasoa ei aliteta koskaan. Artikkelin ottaa siis myös huomioon teiden laadun käyttäjien näkökulmasta.

## 7 Neuroverkot

### 7.1 Neuroverkon mallintamisen historia

Neuroverkon konsepti kehitettiin jo 1800-luvun lopulla kuvaamaan ihmismielen toimintaa. 1950-luvulla ideaa jalostettiin eteenpäin spontaanin järjestyksen käsitteellä, joka kuvaa aivojen hajautunutta neuroverkkorakennetta, jossa ei ole erityistä prosessointi keskittymää vaan kaikki tieto käsitellään kohtuullisen hajautuneesti. Lopulta vuonna 1975 kehitettiin ensimmäinen multikerroksinen neuroverkko, Cognitron. Tämä malli ei puutteistaan johtuen kuitenkaan saavuttanut suurta menestystä tiedeyhteisössä.

Vihdoin vuonna 1986 neuroverkkomallinnus saavutti selkeästi suurempaa suosiota, mitä edesauttoi tehokkaasta backpropagation-koulutusalgoritmista samaan aikaan kasvanut tietoisuus. Tähän päivään mennessä neuroverkoilla on ollut jo menestyksekkäitä sovelluksia muun muassa puheentunnistamisessa, kuvan analysoinnissa ja itsenäisesti toimivissa roboteissa.

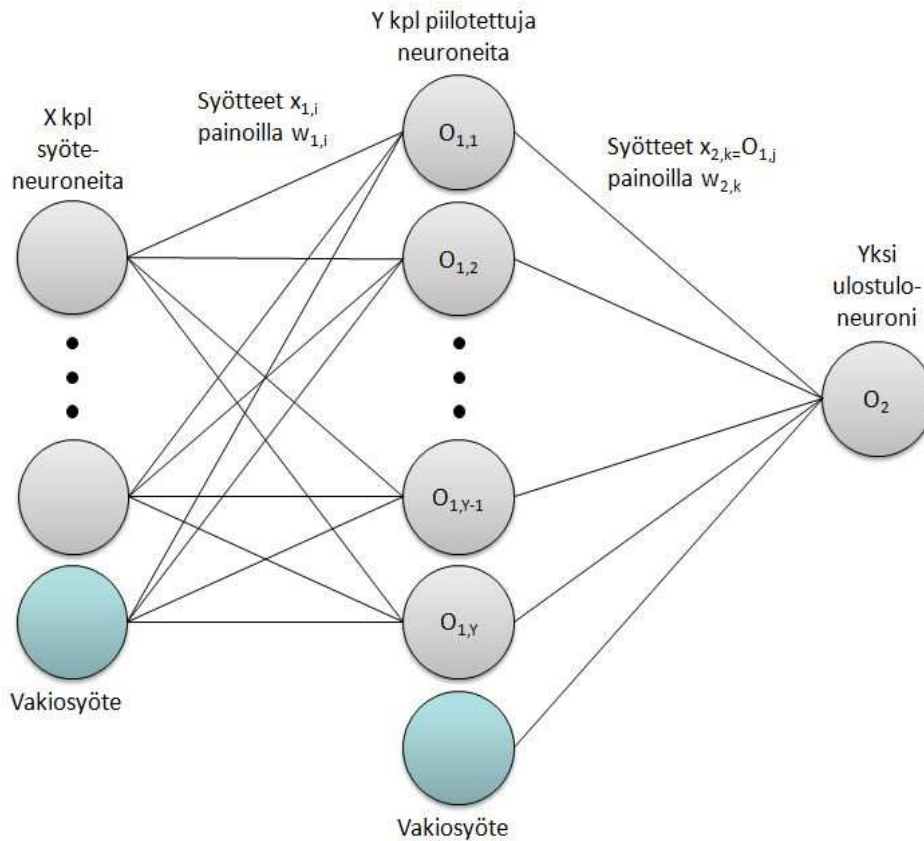
### 7.2 Keinotekoisien neuroverkon malli

Keinotekoinen neuroverkko mallintaa matemaattisesti biologisten verkkojen ongelmanratkaisumenetelmää, joskin mallin vertailtavuutta biologiseen todellisuuteen ei ole vielä todistettavasti vahvistettu. Neuroverkko onkin käytännössä yleistys hermoston ja aivojen ratkaisuprosessista. Alkuperäisenä mallina toiminut biologinen neuroverkko koostuu toisiinsa synapseilla kytketyistä neuroneista, joiden kokonaislukumäärä yhdellä ihmisellä on erittäin suuri. Lisäksi näiden kytkentöjen erilainen kirjo on hämmästyttävänkin laaja. (Basheer & Hajmeer, 2000.) Jatkossa tässä tutkimuksessa neuroverkolla tarkoitetaan keinotekoisia neuroverkkoja.



Keinotekoinen neuroverkko koostuu kolmesta pääosasta: syötöneuronikerros, piilotetut neuronikerrokset ja ulostuloneuronikerros. Neuronit ovat kytketty toisiinsa siirtofunktiolla, jotka siirtävät syötteen eteenpäin seuraaville neuroneille. Jokaiselle yksittäiselle siirtofunktiolle asetetaan oma painokerroin, jolloin vaikutukset seuraaviin neuroneihin ovat yksilölliset. Jokainen syötöneuroni on kytketty jokaiseen ensimmäisen piilotetun kerroksen neuroniin ja viimeisen piilotetun kerroksen neuronit ovat kytketty vastaavasti ulostulokerroksen neuroneihin. Piilotetun kerroksien sisäiset neuronien kytkökset voidaan määrittää erikseen, mutta tässä työssä sille ei ole tarvetta.

Haykin (2008) suosittelee että neuroverkoissa käytettäisiin kahta piilotettua kerrosta, jolloin ensimmäinen kerros tunnistaisi määritettävän mallin lokaalit ominaisuudet ja toinen kerros globaalit. Kuitenkin aikaisempien tutkimusten perusteella, voidaan vakuuttua yhden piilotetun kerroksen riittävydestä (mm. Attoh-Okine, 1999) , joten kytkentöjä piilotettujen kerroksien neuronien välille ei tarvitse suunnitella. Neuroverkon graafinen kuvaus on esitetty Kuvassa 32. Yleisen neuroverkon lisäksi mallissa on useimmiten myös mukana vakiosyötöneuronit sekä syötekerroksessa että jokaisessa piilotetussa kerroksessa. Tämä neuroni lähettää normitetulla asteikolla aina arvoa yksi ja näin siihen ei edellisten neuronien lähettämät syötteet vaikuta lainkaan. Onnistuneen koulutuksen tuloksena näiden neuronien siirtofunktion painokerroin tulisi olla nolla. Kuvaa tarkasteltaessa on kuitenkin huomioitava, että tässä työssä piilotettuja kerroksia tarvitaan vain yksi kappale.



Kuva 32 Tieverkon ennustamisessa käytettävä neuroverkkomalli

Painotettujen syötteiden summa on

$$a_{l,j} = \sum_{i=1}^n w_{l,i} x_{l,i} \quad (12)$$

muutetaan siirtofunktioina toimivien Sigmoid funktioiden avulla välille  $[0,1]$

$$O_{l,j} = f(a_{l,j}) = \frac{1}{1 + e^{-g a_{l,j}}} \quad (13)$$

Lasketun kerroksen neuronien arvot  $O$  siirretään edelleen seuraaville neuroneille painotettuina. Tämä algoritmi toistetaan kunnes saavutetaan lopulliset ulostulokerroksen neuronit. Ulostuloja voi olla useita, mutta tämän työ tarpeisiin riittää vain yksi ulostulo, tien kuntokerroin. Edellä esitetyissä funktioissa  $O$  on neuronin lopullinen arvo,  $a$  on syötteiden painotettu summa,  $x$  on syöte neuroneille ja  $g$  on neuronin Sigmoid funktion kasvu. (Yang ym., 2003.)

Verkkoa muodostettaessa on tehtävä päätöksiä koskien verkon rakennetta ja käytettäviä algoritmeja. Koska mallissa ei pyritä luomaan tieteellisesti perusteltuja riippuvuussuhteita (black box -malli) (Basheer & Hajmeer, 2000), monet päätökset joudutaan tekemään tieteellisesti tuettuina yritys ja erehdys -prosesseina. Seuraavaksi käsitelläänkin mallin erityisiä vahvuuksia ja heikkouksia, joista monet liittyvät vahvasti mallin "black box" -ominaisuuteen.

### 7.2.1 Neuroverkkomallin muodostaminen

Ensin on päätettävä kuinka montaa eri arvoa mallin halutaan ennustavan. Useimmissa infrastruktuuriin liittyvissä rappeutumisen malleissa ennustetaan vain yhtä rappeutumisastetta kuvaavaa luokkaa kerrallaan. Esimerkiksi jos tien kunto määritellään huonoimman kertoimen (uraisuus, tasaisuus, halkeamat, jne.) mukaan, jokainen kerroin ennustetaan yksitellen. Tämän jälkeen tulee tunnistaa tarpeellinen syötedata, jolla uskotaan olevan merkitystä mallin ennustustuloksiin (<http://www.mathworks.com/products/neuralnet>). Lopputuloksia varten on myös tärkeää tunnistaa ne muuttujat, joiden arvoa ei ole saatavilla, mutta joiden uskotaan olevan erittäin merkitseviä ennusteen rakentamisessa. (Roberts & Attoh-Okine, 1998.)

Data tulee myös valmistella analysointia varten. Jotta oppiminen on mahdollisemman tehokasta tietokannasta, tulee ensin tunnistaa puuttuvat arvot. Puuttuvat arvot voidaan korvata esimerkiksi lineaarisella interpolaatiolla tai muulla käyttäytymistä kuvaavalla entuudestaan tunnetulla menetelmällä. Jos dataa ei voida täydentää, tulee puuttuvia tietoja sisältävä datasarja poistaa tutkimusjoukosta. (mm. Looney, 1996.) Yleisen datan kuntotarkistuksen ja korjauksen jälkeen data tulee lisäksi optimoida mahdollisemman tehokkaaksi: saman syötedatan tekijän aikariippuvaiset datapisteet tulee tasoittaa esimerkiksi kolmen mittauspisteen liikkuvalla keskiarvolla, jolloin virheellisten mittausarvojen vaikutus heikkenee (Yang ym., 2003). Lisäksi datasta tulee poistaa selkeästi ylimääräiset tekijät, jotka eivät kasvata mallin selitysasetta: tämä voidaan tehdä esimerkiksi pääkomponenttianalyysin avulla (<http://www.mathworks.com/products/neuralnet>).

Luokkaperusteiset syötteet ja ulostulot tulisi esittää binääriluvuilla; esimerkiksi *1000, 0100, 0010* ja *0001*. Neljän numeron binääriesitysmuoto kasvattaa syötteen ulottuvuuksia, jolloin tuloksista saadaan helpommin tulkittavat. (Basheer & Hajmeer, 2000.) Jotta koulutus tapahtuu tehokkaasti, syötedata tulee myös normalisoida välille  $[0, 1]$ . Tällöin painokertoimien keskinäiset erot eivät kasva niin suuriksi kuin alkuperäisillä arvoilla. Jos painokertoimet olisivat hyvin erisuuruisia, koulutus olisi huomattavasti epävarmempi ja hitaampi prosessi. Lisäksi tilanne aiheuttaisi helposti mallin liian aikaisen saturaation, vaikka se ei kykenisikään vielä ennustamaan halutulla tasolla. Normalisointi suositellaan tehtävän kaavan (14) mukaan, jolloin pienin havaittu arvo on  $0$  ja suurin  $1$ . (Lou ym., 2001.)

$$x_i = y_1 + (y_2 - y_1) \left( \frac{z_i - z_i^{\min}}{z_i^{\max} - z_i^{\min}} \right) \quad (14)$$

$$\text{Kun } [y_1, y_2] = [0, 1], \quad x_i = \frac{z_i - z_i^{\min}}{z_i^{\max} - z_i^{\min}} \quad (15)$$

Esitettyssä normalisoinnin kaavassa  $x_i$  on normalisoitu arvo,  $[y_1, y_2]$  on normalisointiväli (tässä  $[0, 1]$ ),  $z_i$  on alkuperäinen syöte,  $z_i^{max}$  on suurin havaittu syötteen arvo ja  $z_i^{min}$  on pienin havaittu syötteen arvo. Mallia muodostettaessa tulee myös huomioida sen tuleva käyttötarkoitus ja sen mukaan päättää käytetäänkö keinotekoista dataa vai todellista mittausdataa. Keinotekoisien datan heikkoutena saattaa olla koulutetun mallin virheellisyys, kun taas oikean datan häiriöt heikentävät mallin suorituskykyä (Yang ym., 2003).

Seuraavaksi mallin rakentamisessa on tunnistettava piilotetun kerroksen neuronien määrä. Nyrkkisääntönä Nelson ja Illingworth (1990) esittää neljää neuronia jokaista syöteneuronian kohden. Valitettavasti ei voida esittää mitään yleispätevää sääntöä neuronien määrään liittyen, sillä se riippuu vahvasti ongelman monimutkaisuudesta. Piilotettujen neuronien määrä joudutaan etsimään kokeilun ja erehdyksen kautta (Nelson ja Illingworth, 1990), ja siksi mallin rakennusvaihe voi olla hyvinkin aikaa vievä. Koulutusprosessissa neuronien määrää tulee minimoida, sillä jos neuroneita on liikaa, neuroverkko pystyy opettelemaan koulutusdatan "ulkoa" ja ei näin ollen pysty ennustamaan uusien erilaisten tapausten käyttäytymistä. Tällöin koulutusprosessi hidastuu myös huomattavasti. (mm. Alsugair & Al-Quadrah, 1998.)

Jos taas neuroneita on liian vähän, neuroverkko ei pysty muodostamaan ennustusmallia, joka kuvaisi parhaiten muuttujien käyttäytymistä. Siksi tuleekin siis valita pienin neuronien määrä, jolla pystytään saavuttamaan parhaan tason ennustusaste. Rakentamista voidaan lähestyä minimaalisesta mallista, jota kasvatetaan hiljalleen kunnes virhe saadaan tarpeeksi pieneksi tai isosta mallista vähennetään vähitellen hyödyttömiä osia, kunnes optimaalinen rakenne saavutetaan. (Attoh-Okine, 1999.) Koulutuksen lähtökohdaksi painokertoimet ja mieltymykset neuroverkon siirtofunktioille kannattaa valita satunnaisluvuista, jotta ne eivät aiheuta vääristymiä tuloksissa. Tutkijat ovat yhtä mieltä että satunnaislukujen keskiarvo tulee olla nolla, mutta hankalampaa onkin valita näiden satunnaislukujen skaala. Suuri skaala nopeuttaa oppimista, mutta lisää aikaisen saturaation riskiä. Pieni skaala hidastaa oppimista, mutta ei toisaalta heikennä mallia. Tässä työssä lähdetään ASCE:n (2000) suosituksesta käyttää skaalaa  $[-0,3; 0,3]$ , jota käytämme myös katsauksessamme painokertoimien lähtökohtaisena skaalana.

### 7.2.2 Neuroverkkomallin heikkoudet

Neuroverkkomallin yhtenä suurena heikkoutena on pidetty sen kyvyttömyyttä hyödyntää jo olemassa olevaa tieteellistä tietoa – esimerkiksi eri muuttujien riippuvuussuhteista. Tästä syystä esimerkiksi neuroverkkojen hyödyntäminen robotiikassa vaatii huomattavan määrän kokemuksia, eli erilaisia datasarjoja. Datatarpeen kasvaessa aika datan luomiseksi tai keräämiseksi saattaa myös kasvaa liian pitkäksi. Tietenkin myös datan prosessointi ja verkon kouluttaminen pitkittyvät ajallisesti.

Myös oikeanmallisen neuroverkon määrittäminen on koettu ongelmalliseksi, sillä se ei perustu vahvaan tieteelliseen perusteluun. Esimerkiksi Haykinin (2008) suosittama kahden piilotetun kerroksen neuroverkkomalli on huomattu tämän kirjallisuuskatsauksen tutkimuksissa joko liikaa aikaa kuluttaviksi tai tehottomiksi. Toisaalta epäjatkuvia koulutusalgoritmeja käytettäessä kaksi kerroksisuus saattaa puolestaan olla erittäin hyödyllistä (Basheer & Hajmeer, 2000). Myöskään koulutusalgoritmin valinnalla ei ole aina tieteellistä perustelua tukena, joskin backpropagation-koulutusalgoritmi on saavuttanut selkeästi eniten menestystä. Valittu algoritmi on kyseisessä tapauksessa saattanut vain tuottaa parhaimmat tulokset eli pienimmät virheet.

Neuroverkkojen hyödyllisyydestä huolimatta ne eivät ole ylivertaisia menetelmiä, mutta asiaan perehtymättömän on helppo saada mallista liian hyvä vaikutelma. Tuleekin muistaa seuraavat huomiot:

Neuroverkkojen ominaisuudet verrattuna polynomiapproksimaatioihin (Attoh-Okine, 1999).

- Eivät ole ylivertaisia
- Tarvitsevat vähintään yhtä paljon koulutussarjoja
- Ovat yhtä herkkiä koulutusdatan vääristymille
- Niiden saavuttaminen ei ole aina yhtä helppoa

### 7.2.3 Neuroverkkomallin vahvuudet

Neuroverkon yhtenä suurimpana vahvuutena on pidetty sen kykyä kuvata epälineaarisia funktioita. Siksi verkko pystyykin sovittumaan mitä monimutkaisimpiin ongelmiin. Myös tienpäällysteen kulumisen ennustamisessa tämä ominaisuus tulee tarpeeseen, sillä kulumisen ei koostu ainoastaan lineaarisista riippuvuussuhteista. (Roberts & Attoh-Okine, 1998.) Malli saavuttaa kohtuullisen varmuustason jo melko pienellä määrällä dataa sekä suorituskyky on kohtuullisen vakaa vaikka syötettä sekoittaisikin suurella syötteeseen lisätyllä kohinalla (mm. Fwa & Chan, 1993).

Mallin toimintavarmuus riippumatta ilmiön asiantuntijatiedon saatavuudesta on myös koettu useassa tapauksessa tarpeettomaksi. Ennusteet pystytään tällöin luomaan ilman ilmiön dynamiikan tuntemista. Usein dynamiikan ymmärtäminen ja matemaattinen mallintaminen perinteisin menetelmin voi olla liian haastavaa, joten joudutaan tekemään yksinkertaistuksia, jotka heikentävät mallin tarkkuutta. (Fwa & Chan, 1993.) Tämän ansiosta verkko on helppo mukauttaa uusiin tilanteisiin ja täysin uudenlaiseen tai -tyyppiseen dataan. Verkkoahan voidaan hyödyntää ennusteiden luomisessa tietämättä eri tekijöiden välisiä matemaattisia riippuvuuksia. Tällöin myös uuden datan hyödyntäminen ennusteissa ei vaadi vaikutussuhteiden asiantuntija-arviointia.

### 7.3 Neuroverkkomallin koulutus

Koulutuksen tarkoituksena on määrittää oikeat painot,  $w_i$ , eri siirtofunktioille. Neuroverkkojen koulutusalgoritmit voidaan jakaa kolmeen pääluokkaan: valvottu, valvomaton ja vahvistettu. Verkkojen koulutuksessa käytetään hyväksi algoritmista riippuvaista kustannusfunktiota  $C$  ja optimaalista ratkaisua  $f^*$

$$C: F \rightarrow R$$

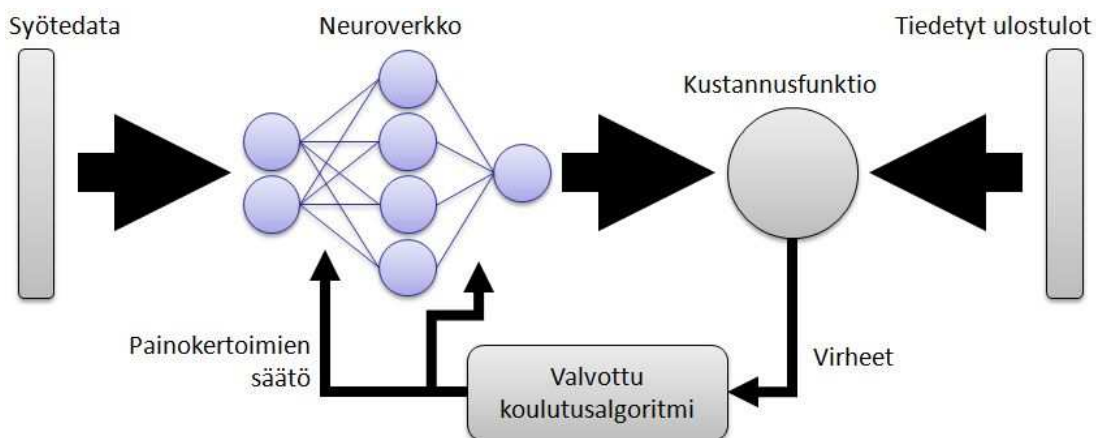
$$f^* \in F,$$

niin että optimaalinen ratkaisu minimoi kustannusfunktion

$$C(f^*) \leq C(f). \quad (16)$$

Kustannusfunktion tarkka muoto riippuu käytetystä koulutusalgoritmista. Toisin sanoen koulutusalgoritmi käy läpi kaikkien ratkaisujen joukkoa, kunnes löytää ratkaisun, joka aiheuttaa pienimmän mahdollisen kustannuksen eli poikkeaman toivotusta arvosta.

Valvotussa koulutuksessa kustannusfunktion tulee ilmentää verkon ulostulojen ja ennalta tiedettyjen tulosten epäsopivuutta. Algoritmin toimintaperiaate on esitetty Kuvassa 33. Koulutusdatan on oltava erillinen joukko lopullisesta testausdatasta, sillä testausdatalla varmistetaan verkon toimivuus ennalta kokemattomissa tilanteissa. Tätä katsausta varten tarkastelluissa aikaisemmissa tutkimuksissa koulutusdatan osuus koko datasta on ollut noin puolet. Joukon tulee sisältää tarpeeksi erilaisia ryhmiä, mutta samaan aikaan jokaisesta ryhmästä olisi oltava useampi joukko, joka edustaa samanlaista käyttäytymistä. Datasta tulisi lisäksi erottaa vielä kolmas joukko, jolla voidaan lopullisesti vahvistaa mallin hyvyys. Looney (1996) ehdottaa 65 % koulutukseen, 25 % testaukseen ja 10 % vahvistamiseen.



Kuva 33 Valvotun kouluttamisen toimintaperiaate

Koulutusta varten on määriteltävä myös sopiva momenttitermi ja oppimismopeus. Momenttitermin käyttöä suositellaan, jos datan uskotaan sisältävän virheitä; virheettömällä datalla momenttitermi onkin usein jätetty pois. Suuri momenttitermi ehkäisee mallin pysähtymistä paikallisiin minimeihin, mutta saattaa estää ratkaisun löytymistä. Termin suuruudeksi on ehdotettu kaikkea väliltä 0-1. Oppimismopeuden arvo määrittelee askeleen suuruuden, jolla painokertoimia korjataan oikeaan suuntaan, ja myös nopeudeksi on suositeltu kaikkea 0,1 ja 10 väliltä. Nopeuden tulee olla suuri kun ollaan kaukana ratkaisusta ja pieni kun lähestytään ratkaisua. Liian suuri arvo voi saada verkon värähtelemään ratkaisun eri puolilla ja näin ei ikinä saavuta ratkaisua. Iso tutkimuksen kohde onkin kuinka selvittää ennusteen etäisyys ratkaisusta. (mm. Attoh-Okine, 1999.)

Näitäkin termejä määriteltäessä suositun lähestymistapa on ollut yritys ja erehdys. (mm. Attoh-Okine, 1999.) Zupan and Gasteiger (1993) suosittelevat että oppimismopeuden ja momenttitermin summa olisi yksi. Tarkasteltaessa eri tekijöiden vaikutusta teiden kunnan ennustamisverkon kouluttamisessa, huomattiin hyviksi oppimismopeuden arvoiksi 0,2-0,5 ja momenttitekijän arvoksi 0,4-0,5. (Attoh-Okine, 1999.)

Valvomattomassa oppimisessä kustannusfunktio  $C$  perustuu ennako-oletuksiin siitä kuinka mallin tulisi käyttäytyä ja sisältää tällöin ulostulofunktion  $f$  ja sisääntulon  $x$ . Tällöin tulee määrittää millä kustannusfunktiolla malli saa pienimmät virheet ennusteen ja todellisten arvojen välillä. Vahvistetussa oppimisessä koulutuksen rangaistukset ja palkinnot luodaan yhteistoiminnassa ympäristön kanssa. Pyritään siis tunnistamaan millä rangaistuksilla ja palkinnoilla pystytään minimoimaan kumulatiivinen kustannusfunktio. (mm. Attoh-Okine, 1999.)

Tarkastelemisemme tieverkon rappeutumisen tutkimisessä valvottu oppiminen on todettu sopivammaksi koulutustavaksi. Suosituin algoritmi on selkeästi ollut backpropagation-koulutusalgoritmi. Se on helposti yleistettävissä erilaisiin tapauksiin sekä tarjoaa selkeän matemaattisen konseptin, joka on helppo ohjelmoida (Attoh-Okine, 1999). Algoritmi itsessään voi hyödyntää joko gradienttilaskeutumis- tai liittogradientiaalgoritmeja (gradient descent, conjugate gradient), mutta tämänkin valinta tulee tehdä etsimällä pienintä virhearvoa eri algoritmeilla (Tran ym., 2007). Backpropagation-koulutusalgoritmi käsittelee syötteet verkon läpi alkuperäisiksi valituilla painokertoimilla. Tämän jälkeen lasketaan tavoiteulostulojen ja ennusteiden väliset virheet  $E$ , joiden perusteella algoritmi säätää painoarvoja ( $\Delta w$ ) oikeaan suuntaan taaksepäin edeten. Seuraavaksi esitetään gradienttilaskeutumismenetelmän määritelmä (mm. Lou ym., 2001):

$$E_r = \frac{\sum_r^p \sum_k^m (T_{kr} - O_{kr})^2}{2} \quad (17)$$

$$(\Delta w_{lk})_r = -\eta \frac{\partial E_r}{\partial w_{lk}} + \alpha (\Delta w_{lk})_{(edellinen)} \quad (18)$$

Ensimmäisessä funktiossa  $E_r$  on kaikkien reitin  $r$  ulostulojen virheiden neliö,  $p$  on eri reittien määrä,  $m$  on ulostuloneuronien määrä,  $T_{kr}$  on neuronin  $k$  reitin  $r$  tavoiteulostulo ja  $O_{kr}$  on Sigmoid funktioon perustuva neuronin  $k$  reitin  $r$  ulostulo. Toisessa funktiossa  $(\Delta w_{lk})_r$  on painokertoimen syötoneuronin  $l$  ja piilotetun neuronin  $k$  muutos,  $\eta$  on oppimisnopeuden positiivinen vakio,  $\alpha$  on momenttitermi ja  $(\Delta w_{lk})_{(edellinen)}$  on edellisen edelliseltä kierrokselta saatu muutosarvo.

Koulutusvaiheessa virhe pienenee, kunnes malli saavuttaa tarpeeksi pienen kokonaisvirheen ja pystyy täten ennustamaan datan käyttäytymisen. Virheen pieneneminen voidaan jakaa ajallisesti kolmeen toisistaan erotettavissa olevaan osaan. Ensimmäisessä vaiheessa virhe laskee nopeasti. Tämän jälkeen virhe säilyy kohtuullisen vakiona, kunnes kolmannessa vaiheessa virhe laskee nopeasti pienimpään arvoonsa. (Attoh-Okine, 1999.)

### 7.3.1 Vaihtoehto backpropagation-koulutusalgoritmile

Vaikka tieverkon rappeutumisen mallintamisessa backpropagation-koulutusalgoritmi on ollut selkeästi suosituin, ei se kuitenkaan tuota jokaisessa tapauksessa yksiselitteisesti parhaimpia tuloksia (Tran ym., 2007). Algoritmin vaatimuksena on esimerkiksi vähintään kahden historiallisen mittausdatan saatavuus samasta kohteesta, sekä kaikissa tarkastelemissamme tutkimuksissa historiapistettä on ollut kolmesta viiteen kappaletta. Onneksi kuitenkin tieverkon mallintaminen ei ole erityisen monimutkainen ongelma, sillä esimerkiksi robotiikan alan neuroverkot tarvitsevat paljon vaikeasti saatavaa koulutusdataa. Monimutkaisuus ja suuri datan tarve kasvattavat myös koulutuksen kestoa ja heikentävät tehokkuutta. Mallin heikkouksiksi onkin tunnistettu lineaaristen funktioiden puutteesta johtuva hidas oppimisnopeus ja tehon minimin lähestymisnopeus.

Tutkittaessa viemäriputkien rappeutumista päätettiin vertailla kolmea erilaista lähestymistapaa: backpropagation koulutettu neuroverkko, Markov Chain Monte Carlo -simulaatiolla koulutettu neuroverkko ja perinteisempi lineaarinen monen ryhmän erotteluanalyysi. Tutkimuksessa ei haluttu tyytyä backpropagation-koulutusalgoritmin perustuviin tuloksiin, sillä se koettiin liian epävarmaksi. Backpropagation saattaa pysähtyä lokaaleihin minimeihin, varsinkin jos verkko on liian monimutkainen. Jotta tutkimuksessa saataisiin vähennettyä kompleksisuutta ja kasvatettua yleistettävyyttä, koulutusalgoritmiksi valittiin Markov Chain Monte Carlo -simulaatio (MCMC). Tällä menetelmällä neuroverkko



ennustaa ulostulojen arvot luottamusvälille, kun taas backpropagation menetelmän avulla ennusteet ovat pistearvoja. Näin tutkimuksessa saatiin noin 20 % paremmat ennusteet kuin perinteisellä backpropagation menetelmällä. (Tran ym., 2007.)

MCMC menetelmä perustuu Bayes-teoreemalle, jonka mukaan saadaan painokertoimien seuraava jakauma

$$P(w | y, x) = \frac{P(y | x, w)P(w | x)}{P(y | x)}, \quad (19)$$

jossa  $w$  on painokertoimien vektori,  $y$  on havaintojen lukumäärä,  $x$  muodostaa mallin syötetekijät,  $P(y/x)$  on vakio laskettu  $x$ :n yli ja  $P(w/x)$  on aikaisempi tietämys  $w$ :stä.  $P(y/x, w)$  on todennäköisyysfunktio, jolla havaitaan todellinen  $y$  tiedetyn  $x$  ja etsityn  $w$  avulla:

$$P(y | x, w) = \prod_{i=1}^N P(y_i | x, w) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{T_i - O_i}{\sigma}\right)^2\right]. \quad (20)$$

Tässä  $T_i$  on tavoiteulostuloarvo,  $O_i$  on ulostuloarvo ja  $\sigma$  on mallin residuaalien keskihajonta. Tunnistettuamme vakiot voidaan yksinkertaisemmin merkitä

$$P(w | y, x) \propto P(y | x, w). \quad (21)$$

Kokeilemalla tarpeeksi suuren määrän painokertoimia voidaan kertoimet määrittellä luottamusvälien avulla. Näin ei tarvitse luottaa vain yhteen pistearvoennusteeseen painokertoimista kuten backpropagation mallissa. Ainakin TTK:n Centre of Excellence in Computational Complex Systems Research (COSY) keskuksella on tarjolla MCMC Matlab Toolbox, jota voi hyödyntää tämän mallin rakentamisessa (Vehtari & Vanhatalo, 2006).

#### 7.4 Kokemuksia ja oppeja aikaisemmista tutkimuksista

Tarkastelemisamme tutkimuksissa neuroverkon mallintamisessa oli käytetty enimmäkseen kaupallisia sovelluksia kuten Brainmaker, Matlab ja NeuralWorks. Mitään analyysia ohjelmistojen tehokkuudesta tai soveltuvuudesta tutkimuksien ongelmiin ei esitetty. Kaikki mallinsivat ongelmaa, jossa tarvitsee ennustaa vain yhtä ulostuloa, seuraavan vuoden kuntokerrointa.

Tutkimuksissa käytettiin 7-12 eri syötettä ennustamaan tieverkon kuntoa seuraavana vuonna. Jokaisessa tutkimuksessa mittauspisteestä oli ainakin kaksi edellistä mittaustulosta tiedossa. Backpropagation koulutuksella mallit ovat saavuttaneet kohtuullisen hyviä ennusteita, sillä neliöllinen keskivirhe (RMS) on useimmissa tutkimuksissa ollut parhaimmillaan hieman alle 0,06 ja yhteensopivuusaste noin 0,8-0,9. Tutkimuksissa piilotettuja neuroneita on aina

ollut vain yhdessä kerroksessa ja niitä on ollut noin 1,5 kertaa enemmän kuin syöteneuroneita.

Perusmallin ennustamisen lisäksi tutkimukset ovat tehneet myös muutamia erikoishuomioita suhteessa perinteisiin autoregressiomalleihin:

- Neuroverkkomallien ennustuskyky paranee, jos pyritään ennustamaan useampaa vuotta eteenpäin.
- Kun tie on vanha ja täten rappeutuu todennäköisesti nopeammin, neuroverkko tuottaa selkeästi tarkempia tuloksia.

Kaiken kaikkiaan neuroverkot ovat erittäin lupaavalta kuulostava ennustemenetelmä tieverkon kunnon rappeutumiselle. Koska menetelmän toimivuus on kuitenkin hyvin riippuvainen saatavilla olevasta datasta, ei toimivuutta voi arvioida etukäteen. Tästä syystä teemmekin projektin puitteissa myös pienimuotoisen katsauksen neuroverkkomallin soveltuvuuteen, minkä avulla voidaan päätellä kannattaako neuroverkkojen ennustumahdollisuuksiin tutustua tarkemmin.

### 7.5 Tutkimusasetelma neuroverkkokatsaukselle kirjallisuuskatsauksen perusteella

Katsauksessa on tarkoitus arvioida neuroverkkomallin toimivuutta toimeksiantajan asettamaan ongelmaan. Seuraavissa kappaleissa on käsitelty yksinkertaistetut toimintavaiheet katsastukselle.

Katsauksessa keskitytään ensin backpropagation mallin testaamiseen ja pyritään alustavasti kokeilemaan MCMC-mallia. Katsaus tulee aloittaa datan käsittelyllä, mikä sisältää seuraavat vaiheet

- Yleinen tarkastelu
- Virheellisten datajoukkojen poiskarsinta
- Datajoukoilla oltava saman luokan datapistheet
- Pääkomponenttianalyysi merkityksien tunnistamiseksi
- Normalisointi
- Luokkamuuttujien muuttaminen binäärisiksi
- Jako ryhmiin: koulutus 65 %, testaus 25 % ja vahvistus 10 %

Backpropagation koulutus suoritetaan Matlabin Neural Networks työkalulaatikolla, joka riittää katsastuksen tarpeisiin. Alustavasti verkossa tulee olla yksi piilotettu kerros sekä yksi ulostulo (rakennetta voidaan myöhemmin muokata jos siihen on tarvetta). Alustaviksi lukuarvoiksi malliin syötetään

- Painokertoimet satunnaislukuja väliltä -0,3 – 0,3 keskiarvolla 0
- Momenttitermi väliltä 0,4 – 0,5
- Oppimisenopeus väliltä 0,2 – 0,5

Näitä arvoja voidaan myös hienosäätää tarpeen vaatiessa – oppimisenopeus kannattaa useimmiten kuitenkin pitää mahdollisemman alhaisena. Seuraavaksi tulee määrittää oikea piilotettujen neuronien määrä. Kokemattomuuden takia olisi hyvä aloittaa esimerkiksi kahdella piilotetulla neuronilla. Piilotettujen neuronien määrää kasvatetaan kunnes lisätyillä neuroneilla ei enää pystytä kasvattamaan selitysastetta. Optimaalisen neuronimäärän löydyttyä testataan eri algoritmien paremmuus toisiinsa nähden. Jos valitaan uusi algoritmi, tulee piilotettujen neuronien määrä arvioida varmuuden vuoksi uudestaan.

Kun verkko on saatu valmiiksi, tulee sen soveltuvuus vielä varmistaa erotetulla vahvistusjoukolla. Jos malli kelvollinen, voidaan uusimmasta datasta seuraavaksi muodostaa ennusteet seuraavalle vuodelle.

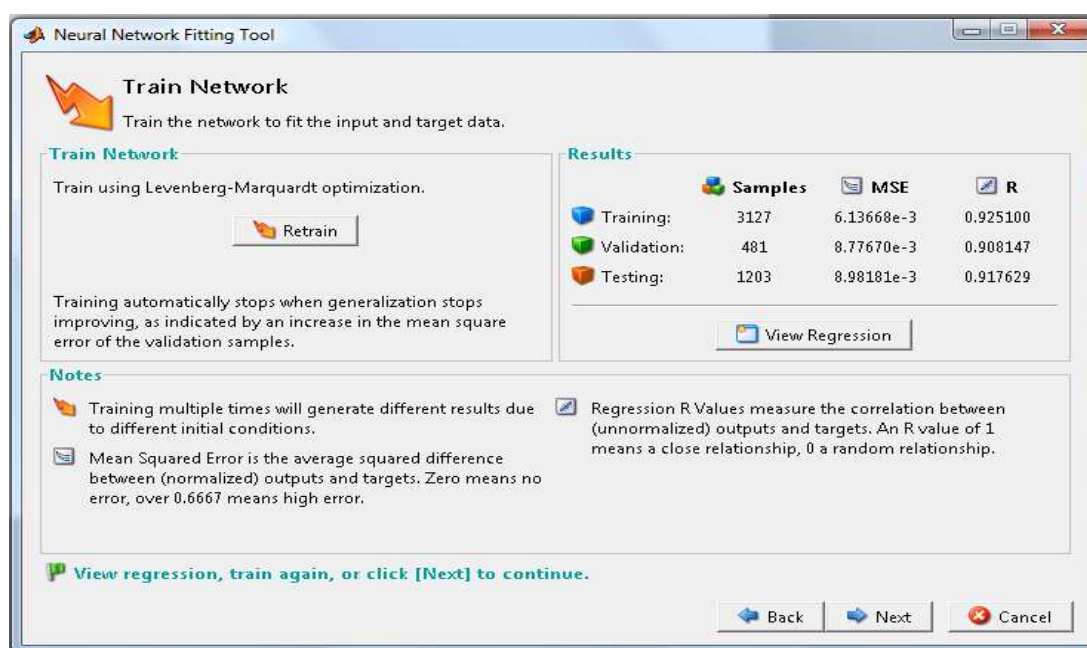
## 8 Alustava tarkastelu neuroverkkomallin käytöstä Suomen tieverkon ennustemallina

Käytimme neuroverkkojen tarkastelussa Matlab-ohjelmiston Neural Network Fitting Toolbox -osaa. Ohjelman käyttäminen oli yksinkertaista. Malli muodostui, kun ohjelmaan lisättiin syöteinä edellisten vuosien (-06 ja -07) mittaustulokset ja valitsemiamme teiden kunnan muutokseen vaikuttavia tekijöitä. Tavoitearvoina käytimme viimeisimmän vuoden (-08) mittaustuloksia. Datan syöttämisen jälkeen ohjelma muodosti datasta sopivimman mallin. Pystyimme itse valitsemaan, kuinka suurta osaa datasta halusimme käyttää kouluttamiseen, vahvistamiseen ja testaamiseen. Päädyimme kirjallisuuskatsauksessa esitettyyn ratkaisuun, jossa kouluttamiseen käytetään 65, vahvistamiseen 10 ja testaukseen 25 prosenttia datasta. Neuroverkko muodostui niin, että virhetermin suuruus oli pienin mahdollinen. Neuronien määrä oli myös itse valittavissa ja useimmissa tapauksissa päädyttiin noin kaksinkertaiseen piilotettujen neuronien määrään syötteiden määrään verrattuna. Eli, mikäli syötteitä oli kolme, niin piilotettuja neuroneja oli yleensä kuusi.

Käytimme malleissa vain osaa kunkin osaverkon datasta. Matlabin neuroverkko-ohjelma tarvitsee toimiakseen täydellistä dataa, jossa ei ole lainkaan tyhjiä paikkoja. Tämän takia jouduimme poistamaan datasta kaikki ne tiedot, joista puuttui yksi tai useampi käyttämämme muuttujan tai mittauksen arvo. Alkuperäisten osaverkkojen datan suuruudet olisivat olleet huomattavasti suuremmat kuin mitä nyt käytettiin. Olisimme myös voineet arvioida puuttuvat mittaustulokset esimerkiksi lineaarisen interpolaation avulla. Koska datan

suuruus oli karsittunakin todella suuri, päädyimme yksinkertaisesti poistamaan tyhjät kohdat, minkä seurauksena dataa oli vähemmän käytettävissä.

Mallien kouluttamisessa ohjelma valitsi satunnaisesti, mitä dataa käytettiin kouluttamiseen, mitä vahvistamiseen ja mitä testaamiseen. Tästä johtuen tulokset eivät olleet jokaisella koulutuskerralla aivan samat, vaikka suuruusluokat säilyivätkin oikeina. Kuvassa 34 on esitetty havainnollistuksen vuoksi Matlabin antama tulostus. Kuvassa 35 on esimerkkitulostus, jossa tarkastellaan AB3000 uria. Lähes kaikki kuvat olivat kyseisenlaisia, olipa kyseessä mikä tahansa osaverkko.

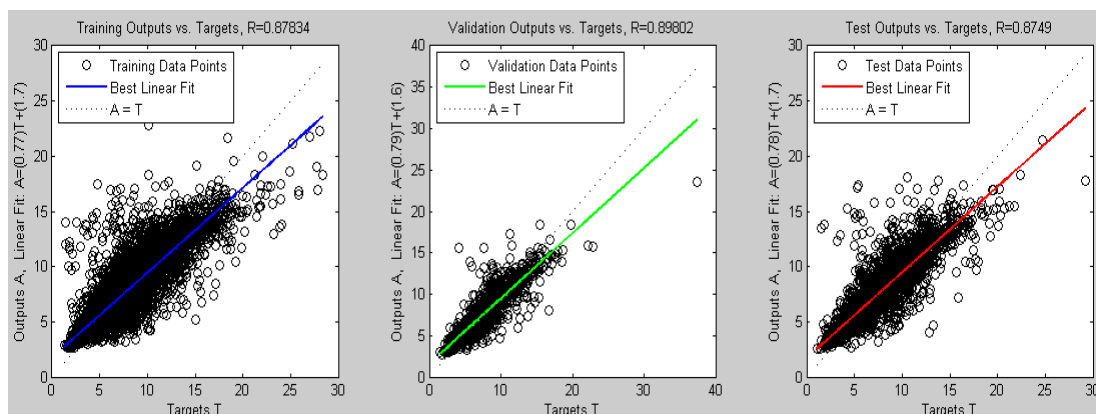


Kuva 34 Matlabin Neural Network Fitting Tool

Tutkimme uria ja tasaisuutta erikseen. Tämä lähinnä sen takia, että tutkimme pääasiassa teiden kunnan huononemista millimetreissä. Näin ollen, koska pääpaino ei ollut kuntoluokissa, oli helpompaa tarkastella kumpaakin, uria ja tasaisuutta erikseen. Tulokset olivat hyvin samantyyppisiä eri osaverkoille. Lähes kaikille AB-osaverkoille saatiin selitysasteeksi 0,82 – 0,95 sekä tasaisuudelle että urille. Pientä poikkeamaa oli joskus havaittavissa, mutta toistettaessa useita kertoja samaa osaverkkoa, saatiin äsken mainitulta väliltä oleva selitysaste. Osaverkolle PAB saatiin selitysasteeksi yli 0,9. Johtopäätöksenä on siis, että neuroverkkojen antamat tulokset teiden kunnan huonontumiselle antavat hyviä tai erittäin hyviä tuloksia, mikäli tarkastellaan huonontumista millimetreinä.

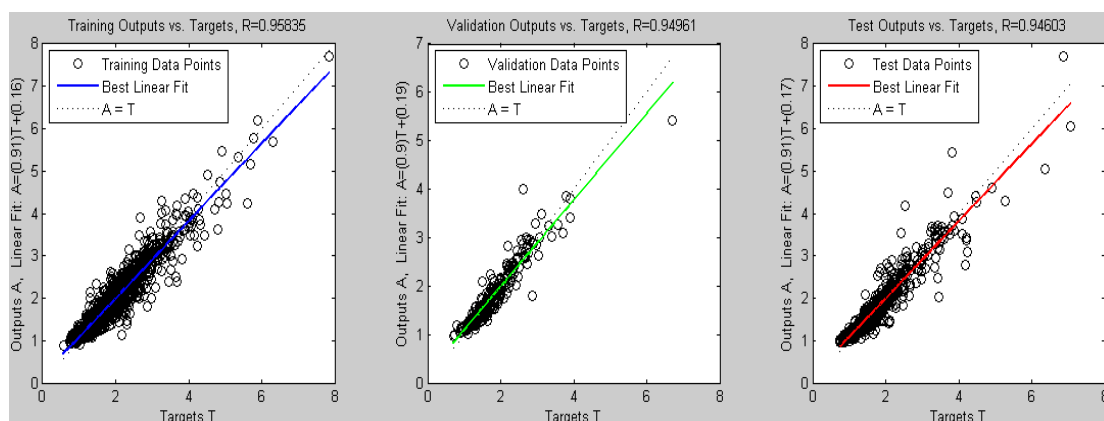
Parhaat tulokset saatiin siten, että syöteinä olivat vuoden 2006- ja 2007 mittaustulokset sekä joitain muita tekijöitä. Lähes kaikissa osaverkoissa muiksi tekijöiksi riitti ikä, mutta kaikkiin sitäkään ei tarvinnut ottaa mukaan. Muiden tekijöiden kuin ikä mukaan ottaminen ei oleellisesti parantanut tuloksia

useimmissa tapauksissa. Muita jossain määrin relevantteja tekijöitä olivat esimerkiksi keskimääräinen vuorokausiliikenne (KVL), toimenpideluokka (tp\_lk) sekä nopeus. Mutta kuten sanottu, ne eivät olleet kovin merkityksellisiä. Yli 0,8 olevaan selitysasteeseen päästiin useimmiten jo siten, että syötteenä on kahden edellisen vuoden havaitut arvot ja tavoitearvona vuoden 2008-havaitut arvot.



Kuva 35 MatLabin tulostus osaverkon AB\_3000 urille. Vaaka-akselilla on datasta saatu tavoite millimetreinä ja pystyakselilla mallin antama ulostulo.

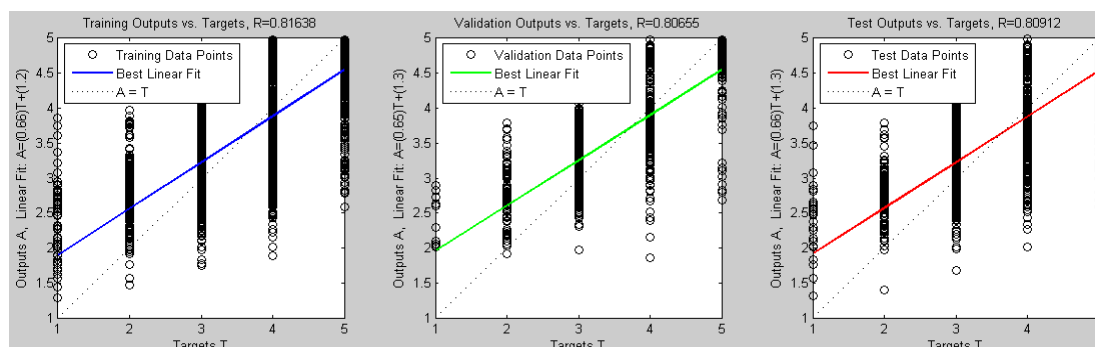
Huomasimme otoskoon vaikutuksen mallin hyvydessä. Esimerkiksi osaverkolle PAB\_350 saimme dataa 2774:stä tieosuudesta, jolloin mallin selitysasteeksi saatiin tasaisuudelle noin 0,95 (Kuva 36).



Kuva 36 MatLabin tulostus osaverkon PAB\_350 tasaisuudelle. Vaaka-akselilla on datasta saatu tavoite millimetreinä ja pystyakselilla mallin antama ulostulo.

Toisin kuin ensisijaisen mallin rakentamisessa, emme olettaneet, että kuntoluokka voi vain pysyä samana tai huonontua yhden kuntoluokan verran vuoden aikana. Oletuksen poisjättämisen takia dataa jäi enemmän tutkittavaksi ja koekelimme neuroverkkojen toimimista myös kuntoluokkien kanssa. Kuten millimetrien kanssa, samat muuttujat (ikä, keskimääräinen vuorokautinen liikennemäärä ja kaksi historiapistettä datasta) selittivät dataa parhaiten. Muiden muuttujien lisääminen ei parantanut mallia ainakaan huomattavasti. Tuloksiksi saatiin hieman huonompia malleja kuin millimetreillä, mikä johtuu suurimmalta osalta siitä, että kuntoluokitus ei ole jatkuva.

Kuntoluokkien käyttäminen neuroverkkojen rakentamisessa huonontaa selvästi mallin selittävyyttä. Kuten datapisteiden levittäytymisestä ja regressiosuorista (Kuva 37) nähdään, havaintopisteet eivät sovi suoralle, koska x-akselilla ovat pelkät kuntoluokat ja kuntoluokkien arvot eivät ole jatkuvia. Y-akselilla on mallin input-datasta lasketmat arvot.



Kuva 37 MatLabin tulostus osaverkon AB\_12000 tasaisuudelle. Vaaka-akselilla on datasta saatu tavoite millimetreinä ja pystyakselilla mallin antama ulostulo.

Kuvassa 37 on esitetty osaverkon AB\_12000 tasaisuuden kuntoluokat muuttujien ja kahden edellisen vuoden kuntoluokkien funktiona. Mallin selitysasteeksi saadaan noin 0,81. Datassa oli mukana yhteensä 14376 samaan osaverkkoon kuuluvaa tienpätkää. Myös muille osaverkoille saadaan samaa suuruusluokkaa olevia selitysasteita, jotka ovat selkeästi huonompia kuin millimetreistä saadut tulokset.

Toimeksiantajan toivomuksena oli, että käyttäisimme teiden kunnan ennustamisessa tielaitoksen jaottelemia kuntoluokkia. Neuroverkoissa millimetrien käyttäminen on kuitenkin järkevämpää, sillä neuroverkot soveltuvat suoran tai jonkun muun riippuvuussuhteen muuttujien välille, toisin kuin muun muassa Markov-malleissa ja käyttämässämme putoamistodennäköisyysmallissa.

## 9 Yhteenveto

Tässä projektityön loppuraportissa käsiteltiin päällystettyjen teiden ja tieverkostojen rappeutumisen matemaattista mallinnusta. Raportin ensimmäisessä osassa rakennetun mallin pohjalta tehtiin riskianalyysi Suomen päällystetyn tieverkon rappeutumisesta. Toiseen osaan sisältyvä kirjallisuuskatsaus keskittyi esittelemään erilaisia lähestymistapoja projektin ongelmaan, jonka pohjalta tehtiin alustava tarkastelu neuroverkkojen soveltuvuudesta kyseiseen tarkoitukseen.

Projektissa rakennettiin ja testattiin Suomen tieverkon tasaisuuden ja uraisuuden kehittymistä kuvaava malli. Alkuperäinen idea Bayes-verkoista jouduttiin hylkäämään ja tilalle kehitettiin kahden mittauskerran dataa

hyödyntävä malli, jossa putoamistodennäköisyydet alempaan kuntoluokkaan lasketaan keräämällä havaintoja siitä, miten monta tienpätkää datassa on pysynyt tietyssä kuntoluokassa määrätyn ikäisinä – ja vastaavasti montako tienpätkää on sen ikäisinä pudonnut alempaan kuntoluokkaan. Laskemalla näiden suhde saadaan laskettua putoamistodennäköisyydet. Datatiedoston lukuun ja havaintojen poimintaan kehitettiin ohjelma Java-ohjelmointikielellä. Suomen päällystetyn tieverkon jokainen mitattu sadan metrin tienpätkä samaistettiin vastaaviin tienpätkiin, jolloin voitiin luoda riittävän laaja joukko eli osaverkko todennäköisyyksien laskemista varten. Luokittelutekijöiksi valittiin päällystetyyppi ja keskimääräinen vuorokausiliikenne sekä tarkentaviksi tekijöiksi nopeusrajoitus ja maantieteellinen sijainti.

Kuntoluokkiin perustuvana ennustemalli ei sovellu hyvin yksittäisten teiden rappeutumisen ennustamiseen. Haluttaessa tarkempaa ennustetta tulisi käyttää kulumisen mittana millimetrejä, jotka ovat karkean kuntoluokituksen taustalla. Myös osaverkkojen luokittelutekijöiden tarkempi syy-seuraus suhde tulisi ilmi paremmin, jos malli perustuisi kuntoluokkien sijasta millimetreihin. Projektissa tehty malli soveltuu laajojen osaverkkojen rappeutumisen ennustamiseen ja se on tehty palvelemaan erityisesti riskianalyysiä, jolloin yksi sen vahvuuksista on korjauspäätösten tukena toimiminen.

Riskianalyysin tavoitteena oli antaa kuva siitä, kuinka suuri joukko tiestöstä voisi huonossa tapauksessa pudota alempaan kuntoluokkaan. Teiden kunnan putoamisten mahdolliset määrät simuloitiin binomijakaumasta ja komponenttien välinen korrelaatio huomioitiin tarkastelemalla ja vertailemalla kolmea eri skenaariota, joista keskityttiin erityisesti vahvan korrelaation tarjoamiin worst-case -skenaarioihin. Pehmeällä asfalttibetonilla päällystettyjen teiden tasaisuuden havaittiin kuluvan liikennemäärän vaikutuksesta asfalttibetoniteitä nopeammin, mutta urasyvyyden kehittymisessä huomattiin enemmän eroavaisuuksia osaverkkojen välillä. Silmiinpistävin havainto oli asfalttibetoniteiden viidennen luokan erittäin nopea putoaminen alempaan kuntoluokkaan.

Kirjallisuuskatsauksessa keskityttiin Bayes-verkkojen ja Markov-mallien esittelyyn teiden rappeutumisen näkökulmasta. Kirjallisuustutkimusta tehtäessä löydettiin myös projektin aihetta sivuavia neuroverkkosovelluksia, joka vaikuttaa lupaavalta mallinnusmenetelmältä. Tästä syystä katsauksessa esiteltiin myös neuroverkkojen teoriaa ja lopuksi tehtiin alustava tarkastelu neuroverkkomallin käytöstä Suomen tieverkon ennustemallina. Saadut tulokset osoittavat, että neuroverkot antavat hyviä tuloksia, mikäli tiekunnan huonontumista tarkastellaan millimetreihin perustuen mutta kuntoluokkia käyttämällä tulokset ovat selkeästi huonompia.

## 10 Viitteet

- AASHO. (1962). "Road test." *Rep. 7, Summary Rep., Spec. Rep. 61G*, Highway Research Board, Washington D.C.
- Abaza, K. A. (2006). "Iterative linear approach for nonlinear nonhomogenous stochastic pavement management models." *J. Transp. Eng., Vol. 132, No. 3, March 1*.
- Abaza, K. A., and Ashur, S. A. (1999). "Optimum decision policy for management of pavement maintenance and rehabilitation." *Transportation Research Record 1655*, Transportation Research Board, Washington, D.C., 8 – 15.
- Abaza, K. A., Ashur, S. A. and Al-Khatib, I. (2004). "Integrated pavement management system with a markovian prediction model." *J. Transp. Eng., Vol. 130, No. 1, January 1*.
- Alsugair, A. M. & Al-Quadrah, A. A. (1998). "Artificial neural network approach for pavement maintenance." *Journal of computing in civil engineering*, 12(4), s. 249-255.
- ASCE. (2000). "Artificial neural networks in hydrology I: Preliminary concepts." *Journal of Hydrologic Engineering*, s. 115-123.
- Attoh-Okine, N.O. (1999). "Analysis of learning rate and momentum term in backpropagation neural network algorithm trained to predict pavement performance." *Advances in Engineering Software*, 30(4), s. 291-302.
- Bandara, N., and Gunaratne, M. (2001). "Current and future pavement maintenance prioritization based on rapid visual condition evaluation." *J. Transp. Eng., Vol. 127, No. 2, March/April*.
- Basheer I.A. & Hajmeer, M. (2000). "Artificial neural networks: fundamentals, computing, design, and application." *Journal of Microbiological Methods*, 43(1), s. 3-31.
- Butt, A. A., Shahin, M. Y., Feighan, K. J., and Carpenter, S. H. (1987). "Pavement performance prediction model using the Markov process." *Transportation Research Record 1123*, Transportation Research Board, Washington, D.C., 12 – 19.
- Calculus, Volume II*, Tom M. Apostol, 2<sup>nd</sup> edition, John Wiley & Sons, 1969
- Charniak, (2002). "Bayesian Networks without Tears."; <http://www.cs.tut.fi/~elomaa/opetus/ai07-7.pdf>
- Durrett, R. (1999). *Essentials of stochastic processes*. Springer-Verlag, New York.
- Dussadee, S., and Pannapa, H. (2007). "Exploring predictive strengths of stochastic pavement deterioration models: a case study of Thailand's highway network." *Asia-Pacific journal of Rural Development*, Vol. XVII, No. 1. July.
- Finn, (1998). "Pavement Management Systems — Past, Present, and Future."
- Fwa, T. F. & Chan, W. T. (1993). "Priority rating of highway maintenance needs by neural networks. *Journal of Transportation Engineering*." 119(3), s. 419-432.



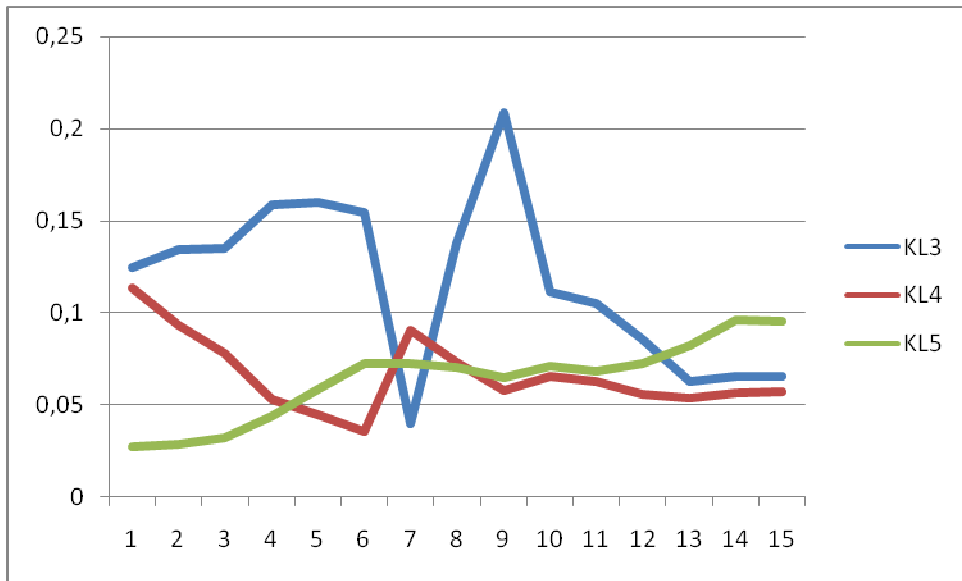
- Golabi & Kulkarni & Way, (1982). "A Statewide Pavement Management System."
- Golabi, K., Kulkarni, R. B., and Way, G. B. (1982). "A state wide pavement management system." *Interface Mag.*, 12(6), 5 – 21.
- Haykin, S. (2008). "Neural Networks: A Comprehensive Foundation."
- Hong, H. P., and Wang, S. S. (2003). "Stochastic modeling of pavement performance." *Int. J. Pavement Eng.*, 4(4), 235 – 243.
- Jorion, P. (1997). "Value at risk: The new bechmark for controlling market risk."
- Li, N., Xie, W.-C., and Haas, R. (1996). "Reliability-based processing of Markov chains for modeling pavement network deterioration." *Transportation Research Record 1524*, Transportation Research Board, Washington, D.C., 203 – 213.
- Looney, C. G. (1996). "Advances in feedforward neural networks: demystifying knowledge acquiring black boxes. *Knowledge Data Engineering*, 8(2), s. 211-226."
- Lou, Z. & Gunaratne, M. & Lu, J. J. & Dietrich, B. (2001). "Application of neural network model to forecast short-term pavement crack condition: Florida case study." *Journal of infrastructure systems*, 7(4), s. 166-171.
- MathWorks. (2009.) "MatLab Neural Network toolbox 6.0.2." Saatavilla osoitteesta <http://www.mathworks.com/products/neuralnet> Haettu 2.4.2009 klo 21:00.
- Ortiz-García, J. J., Costello, S. B., and Snaith, M. S. (2006). "Derivation of transition probability matrices for pavement deterioration modeling." *J. Transp. Eng.*, Vol. 132, No. 2, February 1.
- Roberts, G. A. & Attoh-Okine, N. O. (1998). "A comparative analysis of two artificial neural networks using pavement performance prediction." *Computer-aided civil and infrastructure engineering*, 13(5), s. 339-348.
- Sirvio, K., and Hollmén, J. (2008). "Spatio-temporal road condition forecasting with Markov chains and artificial neural networks." *Hybrid Artificial intelligence systems*, Springer Berlin / Heidelberg, 204 – 211.
- Stulz, R. M. (2001). "Risk management and Derivatives." Thomson Sout-Western.
- Tran, D. H. & Ng, A. W. M. & Perera B. J. C. (2007). "Neural networks deterioration models for serviceability condition of buried stormwater pipes." *Engineering Applications of Artificial Intelligence*, 20(8), s. 1144-1151.
- Wang, K. C. P., Zaniewski, J., Way, G. (1994). "Probabilistic behavior of pavements." *J. Transp. Eng.*, Vol. 120, No. 3, May/June.
- Way, G. B., Eisenberg, J., and Kulkarni, R. B. (1982). "Arizona pavement management system: Phase 2, Verification of performance prediction models and development of data base." *Transportation Research Record 846*, Transportation Research Board, Washington,

D.C., 49 – 55

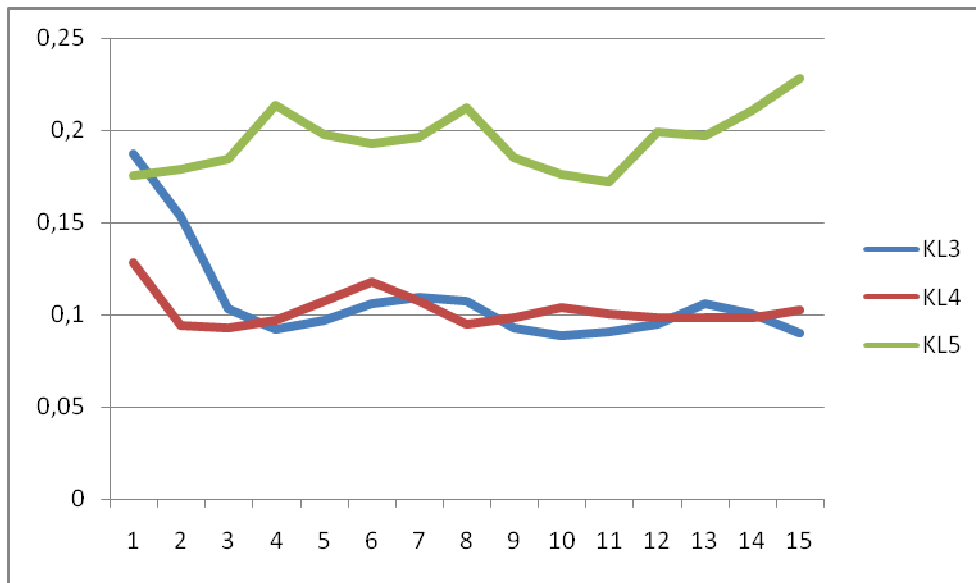
Vehtari, A. & Vanhatalo, J. (2006.) "MCMC Methods for MLP and GP and Stuff (for Matlab) V2.1. COSY." Saatavilla <http://www.lce.hut.fi/research/mm/mcmcstuff> Haettu 2.4.2009

Yang, J. & Lu, J. J. & Gunaratne, M. & Xiang, Q. (2003). "Forecasting Overall Pavement condition with neural network." Transportation Research record, n. 1853, s. 3-12.

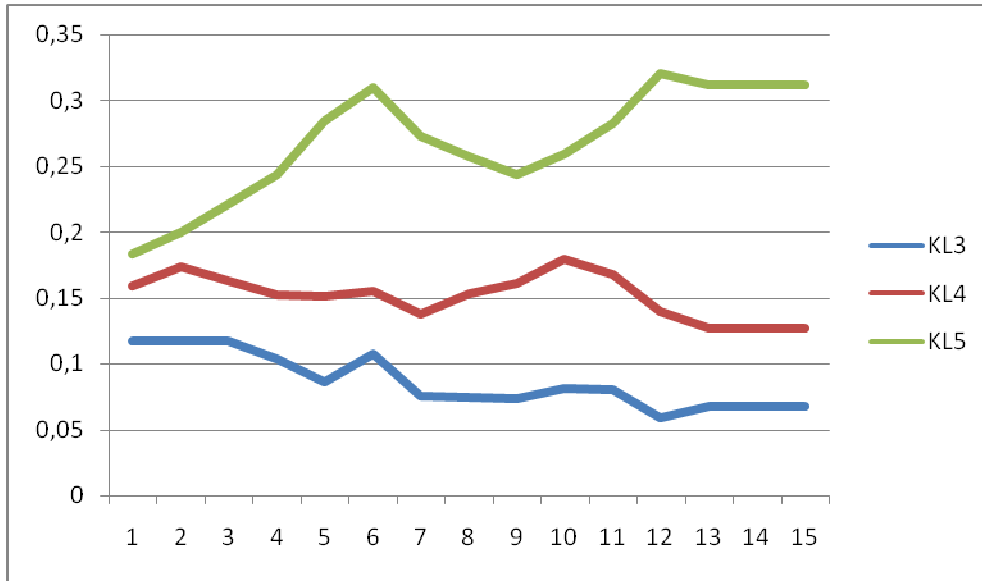
## LIITE 1 – Putoamistodennäköisyyksien kuvaajat



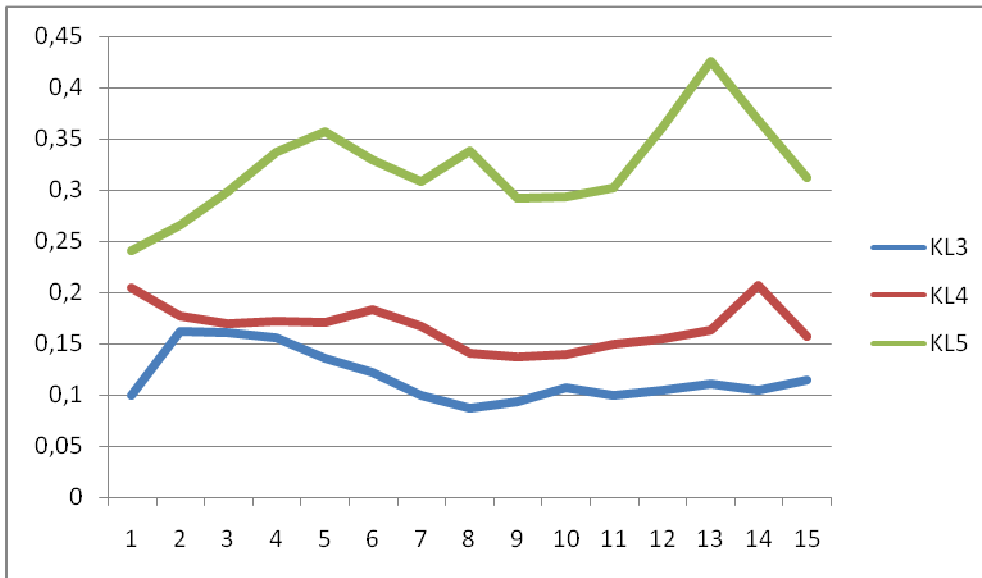
Kuva 38 Ura AB alle 1500, vaaka-akselilla ikä ja pystyakselilla todennäköisyydet yhdelle tiepätkälle pudota yksi kuntoluokka yhden vuoden aikana.



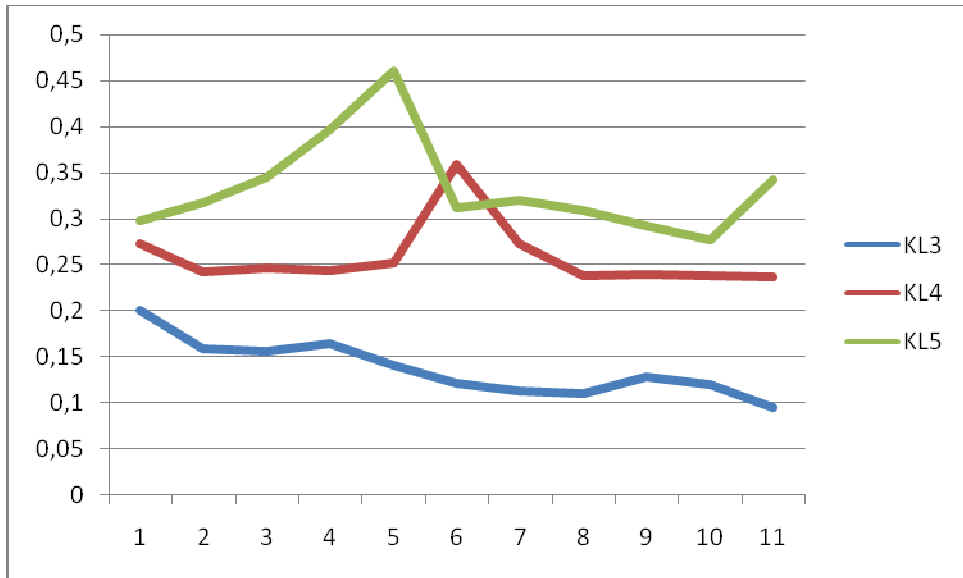
Kuva 39 Ura AB 1000-3000 alle 85 km/h, vaaka-akselilla ikä ja pystyakselilla todennäköisyydet yhdelle tiepätkälle pudota yksi kuntoluokka yhden vuoden aikana.



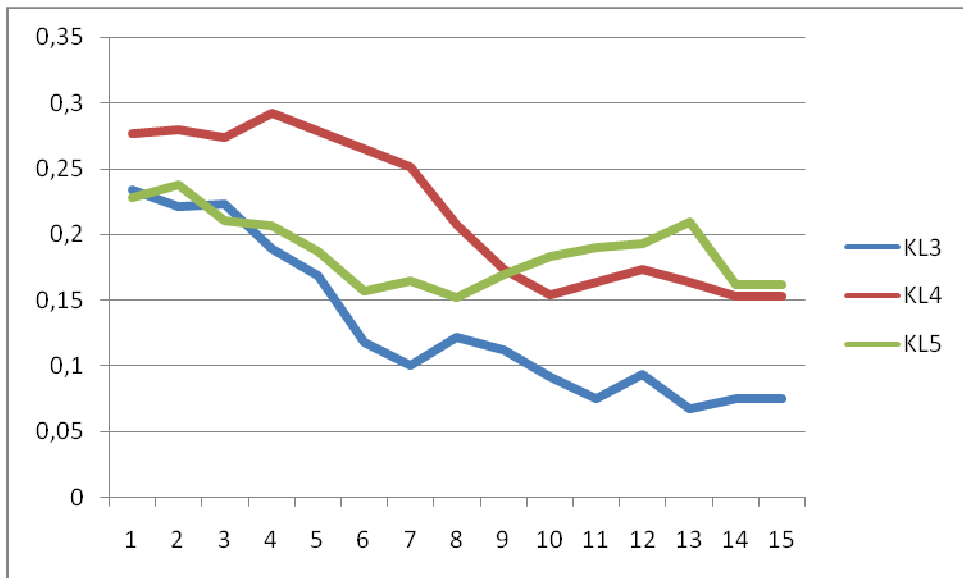
Kuva 40 Ura AB 1000-3000 yli 85km/h, vaaka-akselilla ikä ja pystyakselilla todennäköisyydet yhdelle tiepätkälle pudota yksi kuntoluokka yhden vuoden aikana.



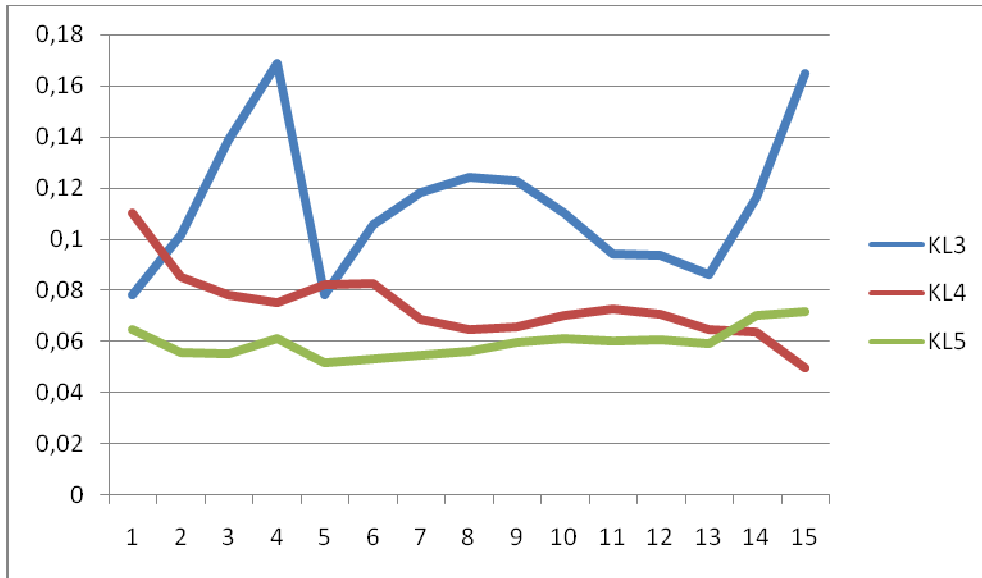
Kuva 41 Ura AB 3000-6000 alle 85 km/h, vaaka-akselilla ikä ja pystyakselilla todennäköisyydet yhdelle tiepätkälle pudota yksi kuntoluokka yhden vuoden aikana.



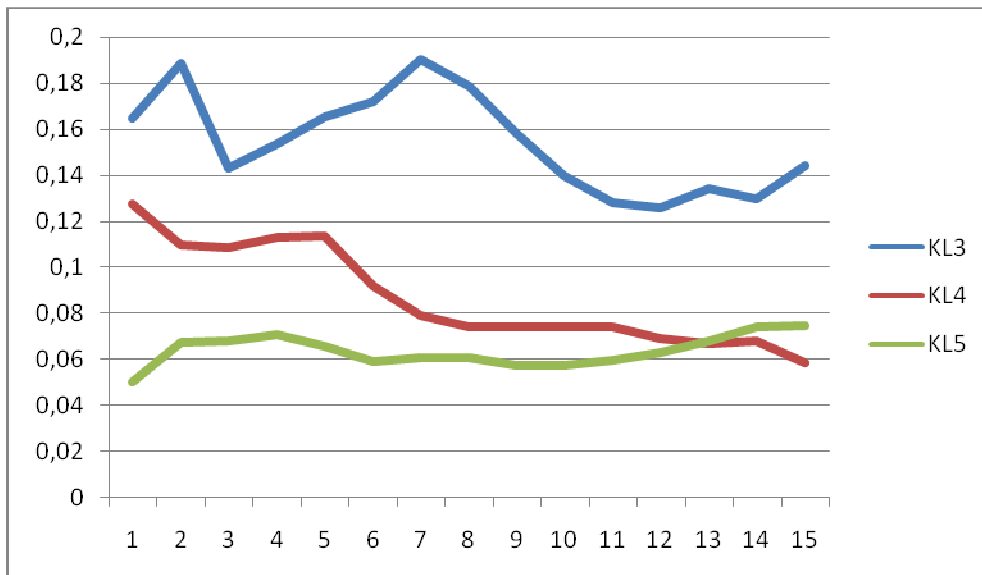
Kuva 42 Ura AB 3000-6000 yli 85 km/h, vaaka-akselilla ikä ja pystyakselilla todennäköisyydet yhdelle tiepätkälle pudota yksi kuntoluokka yhden vuoden aikana.



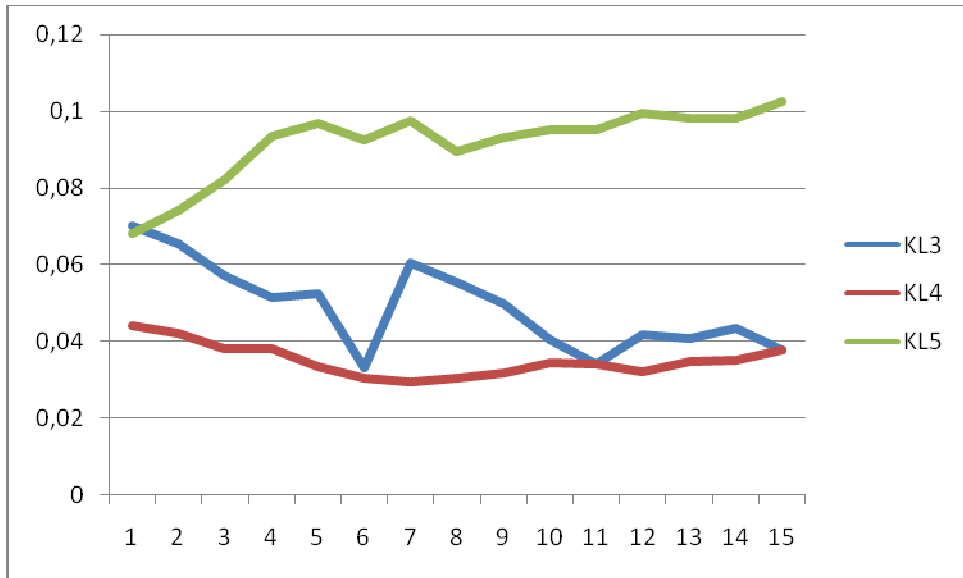
Kuva 43 Ura AB yli 6000, vaaka-akselilla ikä ja pystyakselilla todennäköisyydet yhdelle tiepätkälle pudota yksi kuntoluokka yhden vuoden aikana.



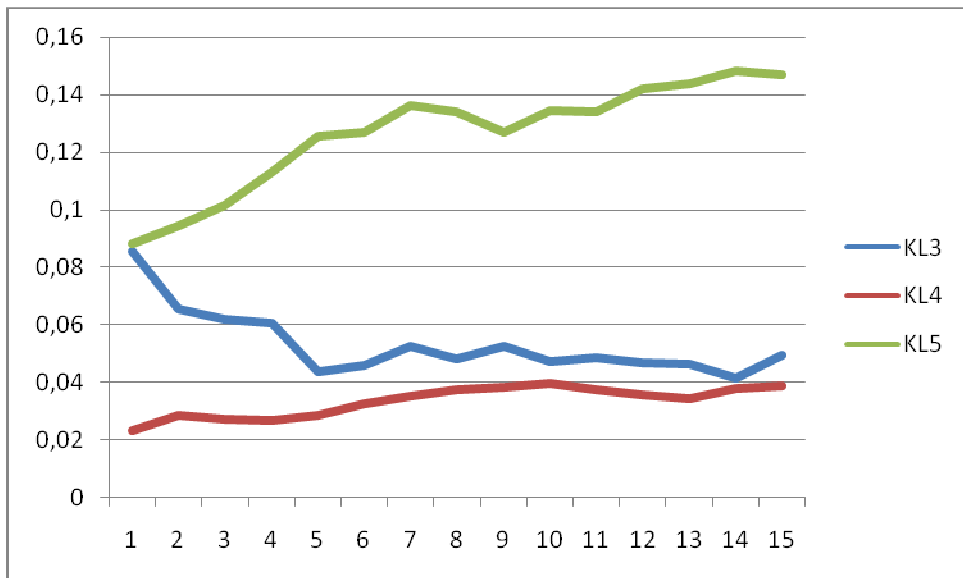
Kuva 44 Ura PAB Oulu ja Lappi, vaaka-akselilla ikä ja pystyakselilla todennäköisyydet yhdelle tiepätkälle pudota yksi kuntoluokka yhden vuoden aikana.



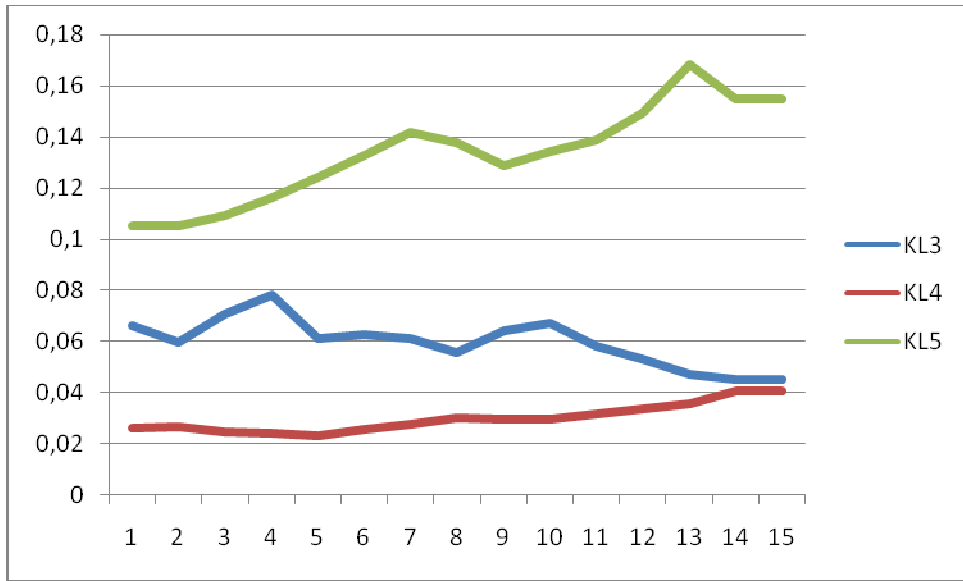
Kuva 45 Ura PAB muu Suomi, vaaka-akselilla ikä ja pystyakselilla todennäköisyydet yhdelle tiepätkälle pudota yksi kuntoluokka yhden vuoden aikana.



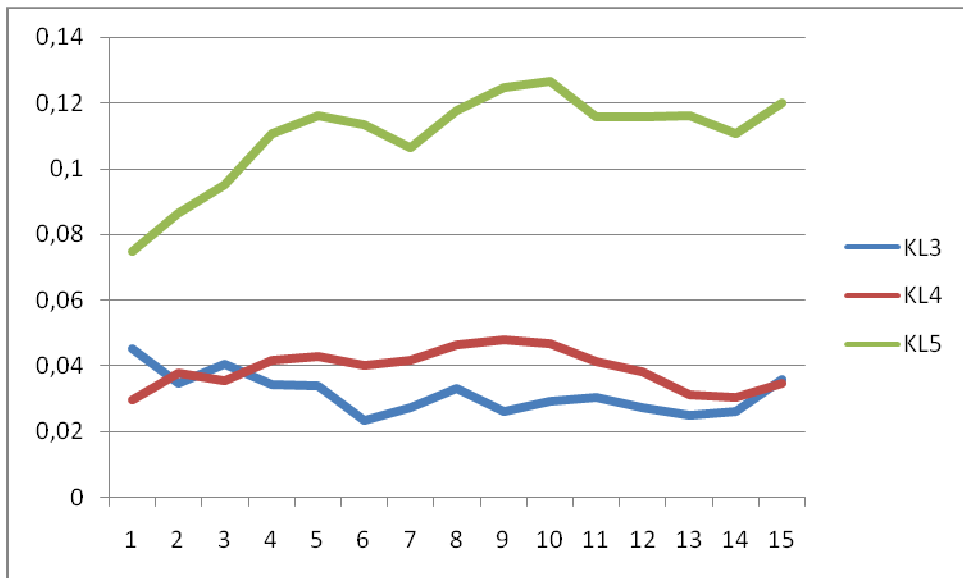
Kuva 46 Tasaisuus AB alle 1500, vaaka-akselilla ikä ja pystyakselilla todennäköisyydet yhdelle tiepätkälle pudota yksi kuntoluokka yhden vuoden aikana.



Kuva 47 Tasaisuus AB 1000-3000, vaaka-akselilla ikä ja pystyakselilla todennäköisyydet yhdelle tiepätkälle pudota yksi kuntoluokka yhden vuoden aikana.

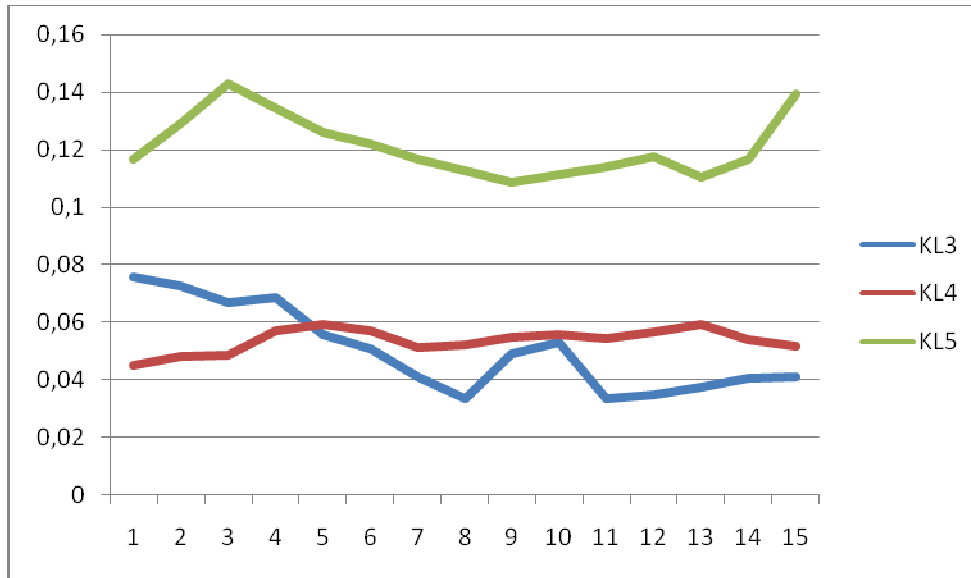


Kuva 48 Tasaisuus AB yli 3000, vaaka-akselilla ikä ja pystyakselilla todennäköisyydet yhdelle tiepätkälle pudota yksi kuntoluokka yhden vuoden aikana.

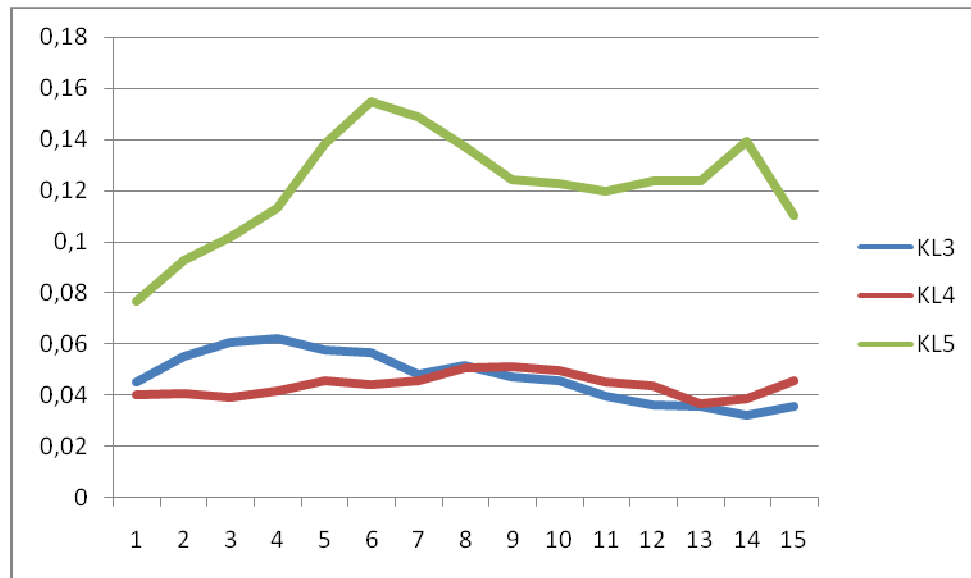


Kuva 49 Tasaisuus PAB alle 350 Oulu ja Lappi, vaaka-akselilla ikä ja pystyakselilla todennäköisyydet yhdelle tiepätkälle pudota yksi kuntoluokka yhden vuoden aikana.

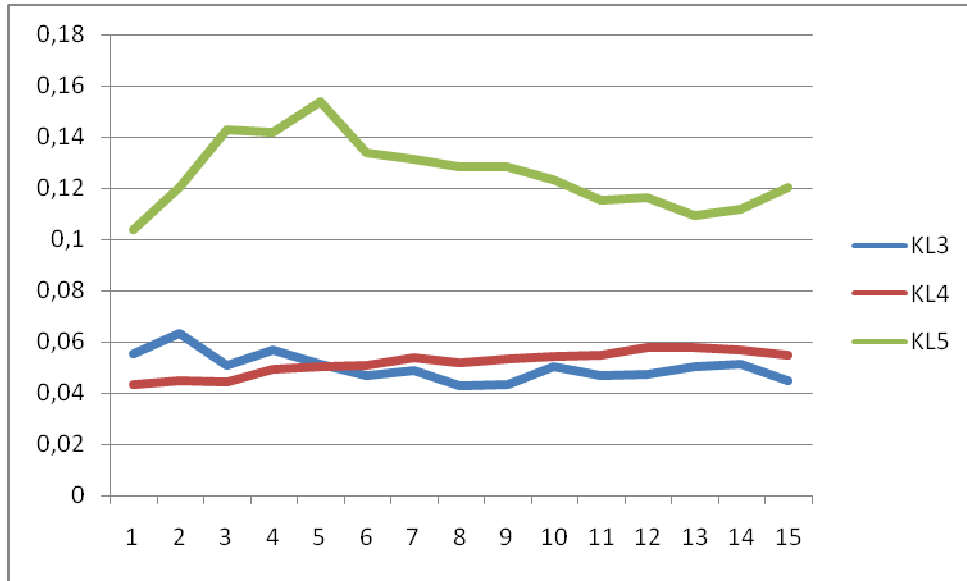




Kuva 50 Tasaisuus PAB alle 350 muu Suomi, vaaka-akselilla ikä ja pystyakselilla todennäköisyydet yhdelle tiepätkälle pudota yksi kuntoluokka yhden vuoden aikana.



Kuva 51 Tasaisuus PAB yli 350 Oulu ja Lappi, vaaka-akselilla ikä ja pystyakselilla todennäköisyydet yhdelle tiepätkälle pudota yksi kuntoluokka yhden vuoden aikana.



Kuva 52 Tasaisuus PAB yli 350 muu Suomi, vaaka-akselilla ikä ja pystyakselilla todennäköisyydet yhdelle tiepätkälle pudota yksi kuntoluokka yhden vuoden aikana.

## LIITE 2 - Java koodi havaintojen keräämiseen

```
import java.io.*;

//Kyseessä on ohjelma, joka lukee samassa kansiossa olevaa CSV-
tiedostosta rivi kerrallaan tiepätkien tiedot, ja mikäli kyseinen tie
täyttää osaverkkoehto, niin tie käsitellään. Lopuksi ohjelma
tallentaa uuteen tiedostoon kerätyt havainnot.

public class kuntokehitysOsaverkolle {

//Matriisit, johon kerätään tiedot havainnoista. Esimerkiksi komento
"tippuu[i][j]=tippuu[i][j]+0.5" lisää matriisiin painoarvoltaa
puolikkaan havainnon, jossa kyseinen tiepätkään ei ole tehty
toimenpiteitä 'j':teen vuoteen, ja havainto on tippuminen
kuntoluokasta 'i' kuntoluokkaan 'i-1'

public static float[][] tippuu=new float[5][40];
public static float[][] pysyy=new float[5][40];

//Rekursiivinen metodi, joka kaikki kaikki kuntoluokan kehityspolut
pisteestä (kuntoluokka,ikä)=( 'polkuPisteenKunto', 'polkuPisteenIka')
pisteeseen ('polkuPisteenKunto'- 'deltaKunto', 'polkuPisteenIka'-
'deltaT'). Samalla se lisää havaintoja matriiseihin 'tippuu' ja
'pysyy', jotta merkittyjen havaintojen yhteenlaskettu painoarvo on 1,
metodi laskee jokaiselle ketityspuun oksalle montako lehteä siinä on,
ja jakamalla argumentilla 'polkuMaara', se saa kyseisen
havaintopisteen painoarvon.

public static int pathFinder (int deltaT, int deltaKunto, int
polkuMaara, int polkuPisteenIka, int polkuPisteenKunto) {

    int solmunPysyyPolut=0;
    int solmunTippuuPolut=0;

//Selvitetään lähtöpisteestä montako lehteä on sen oksan alla, jossa
kuntoluokka ei vähene.

    if ( deltaT>deltaKunto && deltaT>=0) {

        solmunPysyyPolut+=pathFinder (deltaT-
1,deltaKunto,polkuMaara, polkuPisteenIka+1,polkuPisteenKunto);

        pysyy[polkuPisteenKunto-1][polkuPisteenIka-1]+=(float)
solmunPysyyPolut/ (float)polkuMaara;

    }

//Selvitetään lähtöpisteestä montako lehteä on sen oksan alla, jossa
kuntoluokka pienenee yhdellä.

    if (deltaKunto>=1 && deltaT>=0) {

        solmunTippuuPolut+=pathFinder (deltaT-1,deltaKunto-
1,polkuMaara, polkuPisteenIka+1,polkuPisteenKunto-1);

        tippuu[polkuPisteenKunto-1][polkuPisteenIka-1]+= (float)
solmunTippuuPolut/ (float) polkuMaara;

    }

}
```

```
//Metodi laskee montako havaintoa kyseisen oksan alla on, jossa mahdollinen aikaisempi metodi (huom rekursiivisuus) tietää montako lehteä oksassa on.
```

```
    if (solmunPysyyPolut+solmunTippuuPolut<1) {  
        return 1;  
    } else {  
        return solmunPysyyPolut+solmunTippuuPolut;  
    }  
}
```

```
//Ohjelma, jota ajetaan.
```

```
public static void main (String args[]) {  
    int polkuMaara=0;  
    int aikaisempiIka=0;  
    int uudempiIka=0;  
    int aikaisempiKunto=0;  
    int uudempiKunto=0;  
    int deltaT=0;  
    int deltaKunto=0;  
    int rivi=0;  
    int kuntoNouseeIlmanKorjausta=0;  
    int tippuuLiianNopeasti=0;  
    int rivitKasitelty=0;  
    String str1="";  
    String[] strarray;  
    DataInputStream din;  
  
    // Päätetään, mikäli tarkastetaan urien kehityksen, vai tasaisuuden.  
    Urat: "int uratVaiTasaisuus=0;". Tasaisuus: "int  
    uratVaiTasaisuus=1;".  
    int uratVaiTasaisuus=0;  
  
    FileInputStream fin; // Stream to read file  
    FileOutputStream fout; // Stream to write file  
  
    try  
    {  
        fin = new FileInputStream ("TKK_data.txt"); // Avaa pilkulla eroteltun CSV-tiedoston, esim. "TKK_data.txt"  
        din = new DataInputStream(fin);
```

```
fout = new FileOutputStream ("Data_out.txt"); // Tekee tiedoston,
johon ohjelma tallentaa tiedot lopuksi.
```

```
//Luetaan otsikkorivi läpi, ettei sitä käsitellä datarivinä.
```

```
str1=din.readLine(); // Read a line of text
```

```
while (din.available() != 0) {
```

```
    rivi+=1;
```

```
    aikaisempiIka=0; // Nollaa edellisen rivin
```

```
    str1=din.readLine(); //Luetaan seuraava datarivi
```

```
    strarray = str1.split(","); //Pilkulla erotetut datakohdat
erotetaan ja tallennetaan taulukoksi.
```

```
// Joka tuhannennella datarivillä tulostetaan näytölle, rivimäärä,
jotta näkee, että ohjelman etenemisen.
```

```
    if (rivi%1000==0) {
```

```
        System.out.println("Rivi: " + rivi);
```

```
    }
```

```
// Mahdollisuus valita mitkä rivit käsitellään, eli esittää
osaverkkokriteerejä.
```

```
// Esimerkiksi:
```

```
//int kriteeri1=0; // Kriteeri tarkastaa mikäli sarakkeessa 11
(nollasta alkaen) lukee "PAB" tai ei mitään, eli "" ("AB" ei siis
täytä kriteerejä)
```

```
//int kriteeri2=0; // Kriteeri tarkastaa mikäli sarakkeessa 6
(nollasta alkaen) lukee "12" tai "14", elikkä Oulun tai Lapin
maakuntaa
```

```
//int kriteeri3=1; // Kriteeri hylkää tiet jossa KVL (sarake 9,
nollasta alkaen) on tyhjä ("") tai <350
```

```
//if (strarray[11].equalsIgnoreCase("")) ||
```

```
//strarray[11].equalsIgnoreCase("2")) {
```

```
    //    kriteeri1=1;
```

```
    //}
```

```
//if (strarray[6].equalsIgnoreCase("12") ||
```

```
//strarray[6].equalsIgnoreCase("14") ) {
```

```
    //    kriteeri2=1;
```

```
    //}
```

```
//if (strarray[9].equalsIgnoreCase("")) {
```

```

// kriteeri3=0;
//} else if (Integer.parseInt(strarray[9])<350) {
// kriteeri3=0;
//}

//if (kriteeri1==1 && kriteeri2==1 && kriteeri3==1) { // Tiepätkä
(eli datarivi) hyväksytään osaverkkoon kuuluvaksi, mikäli kaikki
ehdot toteutuu.

    int kriteeri1=0; // Kriteeri tarkastaa mikäli sarakkeessa 11
(nollasta alkaen) lukee "PAB" tai ei mitään, eli "" ("AB" ei siis
täytä kriteerejä)

    int kriteeri2=0; // Kriteeri tarkastaa mikäli sarakkeessa 6
(nollasta alkaen) lukee "12" tai "14", eli Oulun tai Lapin
maakuntaa

    int kriteeri3=1; // Kriteeri hylkää tiet jossa KVL (sarake 9,
nollasta alkaen) on tyhjä ("") tai <350

    if (strarray[11].equalsIgnoreCase("") ||
strarray[11].equalsIgnoreCase("2")) {

        kriteeri1=1;

    }

    if (strarray[6].equalsIgnoreCase("12") ||
strarray[6].equalsIgnoreCase("14") ) {

        kriteeri2=1;

    }

    if (strarray[9].equalsIgnoreCase("")) {

        kriteeri3=0;

    } else if (Integer.parseInt(strarray[9])<350) {

        kriteeri3=0;

    }

    if (kriteeri1==1 && kriteeri2==1 && kriteeri3==1) {

//Datasta luetaan läpi myöhäisimmän mittauksen kuntoluokka

        if ( strarray[34+uratVaiTasaisuus].equalsIgnoreCase("5")
|| strarray[34+uratVaiTasaisuus].equalsIgnoreCase("4") ||
strarray[34+uratVaiTasaisuus].equalsIgnoreCase("3") ||
strarray[34+uratVaiTasaisuus].equalsIgnoreCase("2") ||
strarray[34+uratVaiTasaisuus].equalsIgnoreCase("1")) {

            uudempiIka=2008-Integer.parseInt(strarray[18]);

            uudempiKunto=Integer.parseInt(strarray[34]);

        } else if (
strarray[36+uratVaiTasaisuus].equalsIgnoreCase("5") ||
strarray[36+uratVaiTasaisuus].equalsIgnoreCase("4") ||
strarray[36+uratVaiTasaisuus].equalsIgnoreCase("3") ||
strarray[36+uratVaiTasaisuus].equalsIgnoreCase("2") ||
strarray[36+uratVaiTasaisuus].equalsIgnoreCase("1")) {

            uudempiIka=2007-Integer.parseInt(strarray[18]);

```

```

        uudempiKunto=Integer.parseInt(strarray[36]);
    } else if (
strarray[38+uratVaiTasaisuus].equalsIgnoreCase("5") ||
strarray[38+uratVaiTasaisuus].equalsIgnoreCase("4") ||
strarray[38+uratVaiTasaisuus].equalsIgnoreCase("3") ||
strarray[38+uratVaiTasaisuus].equalsIgnoreCase("2") ||
strarray[38+uratVaiTasaisuus].equalsIgnoreCase("1")) {
        uudempiIka=2006-Integer.parseInt(strarray[18]);
        uudempiKunto=Integer.parseInt(strarray[38]);
    } else if (
strarray[40+uratVaiTasaisuus].equalsIgnoreCase("5") ||
strarray[40+uratVaiTasaisuus].equalsIgnoreCase("4") ||
strarray[40+uratVaiTasaisuus].equalsIgnoreCase("3") ||
strarray[40+uratVaiTasaisuus].equalsIgnoreCase("2") ||
strarray[40+uratVaiTasaisuus].equalsIgnoreCase("1")) {
        uudempiIka=2005-Integer.parseInt(strarray[18]);
        uudempiKunto=Integer.parseInt(strarray[40]);
    } else if (
strarray[42+uratVaiTasaisuus].equalsIgnoreCase("5") ||
strarray[42+uratVaiTasaisuus].equalsIgnoreCase("4") ||
strarray[42+uratVaiTasaisuus].equalsIgnoreCase("3") ||
strarray[42+uratVaiTasaisuus].equalsIgnoreCase("2") ||
strarray[42+uratVaiTasaisuus].equalsIgnoreCase("1")) {
        uudempiIka=2004-Integer.parseInt(strarray[18]);
        uudempiKunto=Integer.parseInt(strarray[42]);
    }
}

//Datasta luetaan läpi aikaisimman mittauksen kuntoluokka
    if ( strarray[44+uratVaiTasaisuus].equalsIgnoreCase("5")
|| strarray[44+uratVaiTasaisuus].equalsIgnoreCase("4") ||
strarray[44+uratVaiTasaisuus].equalsIgnoreCase("3") ||
strarray[44+uratVaiTasaisuus].equalsIgnoreCase("2") ||
strarray[44+uratVaiTasaisuus].equalsIgnoreCase("1")) {
        aikaisempiIka=2003-Integer.parseInt(strarray[18]);
        aikaisempiKunto=Integer.parseInt(strarray[44]);
    } else if (
strarray[42+uratVaiTasaisuus].equalsIgnoreCase("5") ||
strarray[42+uratVaiTasaisuus].equalsIgnoreCase("4") ||
strarray[42+uratVaiTasaisuus].equalsIgnoreCase("3") ||
strarray[42+uratVaiTasaisuus].equalsIgnoreCase("2") ||
strarray[42+uratVaiTasaisuus].equalsIgnoreCase("1")) {
        aikaisempiIka=2004-Integer.parseInt(strarray[18]);
        aikaisempiKunto=Integer.parseInt(strarray[42]);
    } else if (
strarray[40+uratVaiTasaisuus].equalsIgnoreCase("5") ||
strarray[40+uratVaiTasaisuus].equalsIgnoreCase("4") ||
strarray[40+uratVaiTasaisuus].equalsIgnoreCase("3") ||
strarray[40+uratVaiTasaisuus].equalsIgnoreCase("2") ||
strarray[40+uratVaiTasaisuus].equalsIgnoreCase("1")) {
        aikaisempiIka=2005-Integer.parseInt(strarray[18]);

```

```

        aikaisempiKunto=Integer.parseInt(strarray[40]);
    } else if (
strarray[38+uratVaiTasaisuus].equalsIgnoreCase("5") ||
strarray[38+uratVaiTasaisuus].equalsIgnoreCase("4") ||
strarray[38+uratVaiTasaisuus].equalsIgnoreCase("3") ||
strarray[38+uratVaiTasaisuus].equalsIgnoreCase("2") ||
strarray[38+uratVaiTasaisuus].equalsIgnoreCase("1")) {
        aikaisempiIka=2006-Integer.parseInt(strarray[18]);
        aikaisempiKunto=Integer.parseInt(strarray[38]);
    } else if (
strarray[36+uratVaiTasaisuus].equalsIgnoreCase("5") ||
strarray[36+uratVaiTasaisuus].equalsIgnoreCase("4") ||
strarray[36+uratVaiTasaisuus].equalsIgnoreCase("3") ||
strarray[36+uratVaiTasaisuus].equalsIgnoreCase("2") ||
strarray[36+uratVaiTasaisuus].equalsIgnoreCase("1")) {
        aikaisempiIka=2007-Integer.parseInt(strarray[18]);
        aikaisempiKunto=Integer.parseInt(strarray[36]);
    }

    deltaT=uudempiIka-aikaisempiIka;
    deltaKunto=aikaisempiKunto-uudempiKunto;

//Varmistetaan, että mittauksia on vähintään kaksi, ja että kehitys
on mallin oletusten mukainen, jolloin siitä kerätään havainnot
kutsumalla metodia pathFinder

        if (uudempiIka<=40 && aikaisempiIka>0 &&
uudempiIka>aikaisempiIka && aikaisempiKunto>=uudempiKunto &&
deltaT>=deltaKunto) {
            rivitKasitelty+=1;
            polkuMaara=1;

            for ( int i = deltaT; i > deltaT-deltaKunto ; i-- )
{
                polkuMaara = polkuMaara * i;
            }
            for ( int i = 1; i <= deltaKunto; i++ ) {
                polkuMaara = polkuMaara / i;
            }

            int i=pathFinder (deltaT, deltaKunto, polkuMaara,
aikaisempiIka, aikaisempiKunto);

// Tarkastetaan montako riviä ei hyväksytä mallin oletuksien takia

```



```

        } else if (aikaisempiKunto<uudempiKunto &&
aikaisempiIka>0 && uudempiIka>aikaisempiIka) {
            kuntoNouseeIlmanKorjausta+=1;
        } else if (deltaT<deltaKunto && aikaisempiIka>0 &&
uudempiIka>aikaisempiIka) {
            tippuuLiianNopeasti+=1;
        }
    }
}

    new PrintStream(fout).println ("Rivit: " + rivi);
    new PrintStream(fout).println ("Määrä rivejä, joita on
hyväksytty: " + rivitKasitelty);

    new PrintStream(fout).println ("Määrä rivejä, jossa kunto nousee
ilman korjausta: " + kuntoNouseeIlmanKorjausta);
    new PrintStream(fout).println ("Määrä rivejä, jossa kunto tippuu
liian nopeasti: " + tippuuLiianNopeasti);
    new PrintStream(fout).println ("Pysyy:");
    for ( int i = 0; i <= 39; i++ ) {
        new PrintStream(fout).println ((pysyy[0][i] + " " + pysyy[1][i]
+ " " + pysyy[2][i] + " " + pysyy[3][i] + " " +
pysyy[4][i]).replace('.',','));
    }

    new PrintStream(fout).println ("Tippuu:");
    for ( int i = 0; i <= 39; i++ ) {
        new PrintStream(fout).println ((tippuu[0][i] + " " +
tippuu[1][i] + " " + tippuu[2][i] + " " + tippuu[3][i] + " " +
tippuu[4][i]).replace('.',','));
    }

    fin.close(); // Close our input stream
    fout.close();

}
catch (IOException e) {
    System.err.println ("Unable to read from or write to file"); //
Catches any error conditions
    System.exit(-1);
}

} //main-ohjelma suljetaan
} //Class suljetaan

```