

Hierarkkinen Bayes-malli tunnelin vaurioiden
ennustamiseen

Tilastotieteen pro-gradu

Teemu Julkunen

Matematiikan ja tilastotieteen laitos

Jyväskylän yliopisto

12. joulukuuta 2021

JYVÄSKYLÄN YLIOPISTO

Matematiikan ja tilastotieteen laitos

Teemu Julkunen: Hierarkkinen Bayes-malli tunnelin vaurioiden ennustamiseen

Tilastotieteen pro gradu -tutkielma, 32 sivua, 3 liitettä (8 sivua)

12. joulukuuta 2021

Tiivistelmä

Ennakoiva kunnossapito säästää aikaa ja rahaa. Tässä työssä sovitetaan kallioon louhitun tunnelin vaurioaineistoon hierarkkinen Bayes-malli, jolla vaurioiden kehittymistä voi ennustaa. Ennusteiden avulla kunnossapitoa voi kohdentaa tarvealueille ennakoivasti ja ehjäksi ennustetuilla alueilla tarkastusväliä voi harventaa. Aineistona tässä työssä käytetään käytetyn ydinpolttoaineen loppusijoituslaitoksen ONKALO[®]n ajotunnelin vaurioita, joita on kerätty vuosina 2016–2020.

Yleisesti tunnelissa voi esiintyä useita eri vauriotyyppejä. Valittu hierarkkinen Bayes-malli huomioi spatiaalisen riippuvuuden lisäksi myös vauriotyyppien väliset riippuvuudet satunnaisvaikutusten korreloituneisuudella. Ajallinen kehitys puolestaan mallinnetaan satunnaisvaikutusten odotusarvon kautta autoregressiivisesti. Bayesiläinen lähestymistapa mallintamisessa mahdollistaa mallin sovituksen, vaikka aineistossa on puuttuvaa tietoa, kunhan tietyt edellytykset ovat voimassa.

Avainsanat: Bayes-tilastotiede, MCMC, Hierarkkinen malli, Spatiaalinen tilastotiede, Autoregressiivinen malli, Puuttuva tieto, Ennakoiva kunnossapito

Sisällys

Tiivistelmä

1 Johdanto	1
2 Aineisto	3
2.1 Vaurioluokitus	5
2.2 Aineiston luotettavuus	6
2.3 Vasteen muotoilu	8
2.4 Korjaustiedot	11
2.5 Spatiaalinen rakenne	12
2.6 Moranin I	14
3 Menetelmät	17
3.1 Hierarkkinen Bayes-malli	17
3.2 Kovarianssimatriisin laskeminen	20
3.3 Mallin sovitus	21
4 Tulokset	23
4.1 Mallin validointi	27
5 Pohdinta	29
Viitteet	31
Liitteet	33

1 Johdanto

Maanalainen rakentaminen on ollut viime vuosina kasvussa, ja tämän kehityksen uskotaan jatkuvan [Lehtonen J., 2011]. Tämä kasvu luo kasvavaa kysyntää myös kalliotekniselle kunnossapidolle. Tämän työn tarkoituksena on esittää tapa mallintaa eri vauriotyyppeiden kehitystä, kun vaurioiden välillä on spatiaalista riippuvuutta. Tavoitteena on kyetä tunnistamaan tunnelista alueita, joissa on kohonnut riski vaurioiden laajenemiselle lähitulevaisuudessa. Myöhemmin mallia on tarkoitus hyödyntää osana päätöksentekoprosessia, jotta tarkastuksia ja korjauksia on mahdollista kohdentaa riskialueisiin ennakoivasti. Ennakoivan kunnossapidon pääasiallinen tavoite on vähentää kunnossapidon kokonaiskustannuksia ja korjauksiin kuluva-aikaa.

Tässä työssä esitetään malli, jota sovelletaan Posiva Oy:n Olkiluodossa sijaitsevaan käytetyn ydinpolttoaineen loppusijoituslaitos ONKALO[®]n ajotunneliin. Työ toteutetaan työsuhteessa rakennuttajakonsultointipalveluita tarjoavassa yrityksessä Welado Oy. Erityisesti ONKALO[®]n ajotunnelissa on kriittistä pyrkiä minimoimaan kulujen ohella korjaustöiden aiheuttamien käyttökatkojen kestoja, sillä ajotunnelin käyttökatkot aiheuttavat käyttökatkoja koko toiminnalle. Tämä on luonnollisesti epäedullista.

Yleisimmin kallioon louhituissa tunneleissa esiintyy kolmenlaista pintaa: paljasta kalliota, sekä ruiskubetonilla tai metalliverkolla lujitettua pintaa [Paalumäki, T., Lappalainen P. & Hakapää A., 2015]. Paljaan kallion vauriot tarkoittavat käytännössä irtoilevaa kalliota. Näitä vaurioita on usein pyritty ennakoivasti ehkäisemään lujittamalla pintaa ruiskubetonilla tai asentamalla kallioon metalliverkoja. Ruiskubetonivauriot ovat puolestaan luonteeltaan tartuntavaurioita eli ruiskubetoni ei ole tarttunut kalliopintaan odotusten mukaisesti. Verkolla lujitetun pinnan vauriot tarkoittavat käytännössä irronnutta tai rikkiäistä verkkoa. Edellisten vauriotyyppeiden lisäksi tunneleissa voi esiintyä pohjavesien vuotoja. Tässä työssä keskitytään irtoilevan kallion vaurioihin, ruiskubetonin tartuntavaurioihin

ja vesivuotoihin, sillä käytössä olevassa aineistossa esiintyy vain näitä vauriotyyppejä.

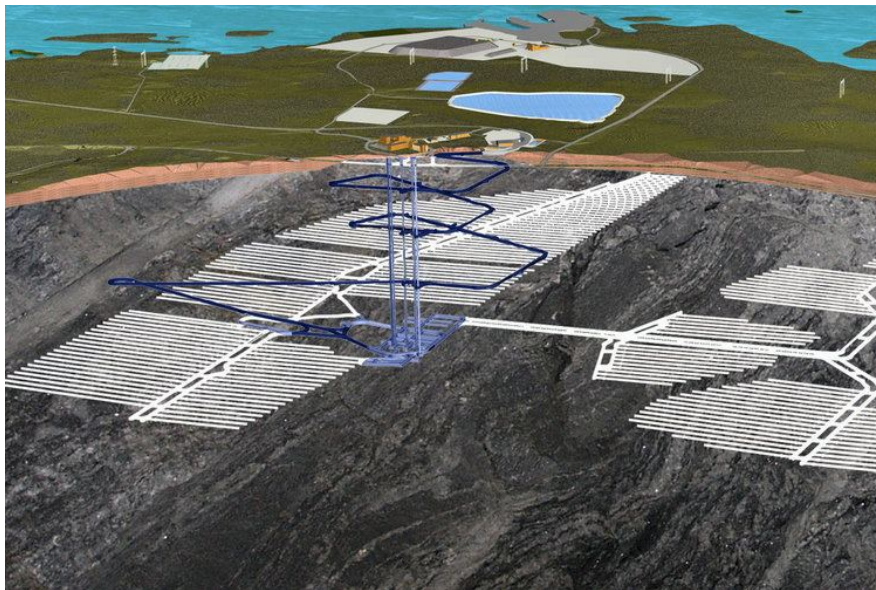
Luvussa 2 esitellään käytettävissä oleva aineisto. Vaurioiden luokitteluperusteet käydään läpi luvussa 2.1. Näiden jälkeen pohditaan aineiston keruuseen ja luotettavuuteen liittyviä haasteita luvussa 2.2. Vaurioiden kehityksen mallintamista varten johdetaan lopulta alueelliset vasteet luvussa 2.3 ja kovariaatteina käytetään tietoa tehdyistä korjauksista, luku 2.4. Alueellisten vasteiden välillä esiintyy spatiaalista korrelaatiota, joka huomioidaan määrittelemällä alueiden välille naapurustorelaatio ja tätä naapurustoa vastaava naapurimatriisi luvussa 2.5. Spatiaalisen riippuvuuden voimakkuutta eri naapurustoilla ja vauriotyypeillä verrataan luvussa 2.6.

Alueellinen vaste on vektoriarvoinen ja vektorin kukin komponentti kuvaa kyseisen alueen vaurioituneisuutta vauriotyypikohtaisesti. Myös vauriotyypit ovat mahdollisesti korreloituneita, joten kaikkia vauriotyyppejä on kyettävä mallintamaan yhtäaikaan. Luvussa 3.1 esiteltävä malli huomioi spatiaalisen korrelaatorakenteen lisäksi vauriotyypien välisen korrelaation, ja temporaalinen korrelaatio mallinnetaan satunnaisvaikutusten autoregressiivisenä rakenteena. Luvussa 3.2 tarkastellaan mallin sovitukseen liittyviä laskennallisia haasteita.

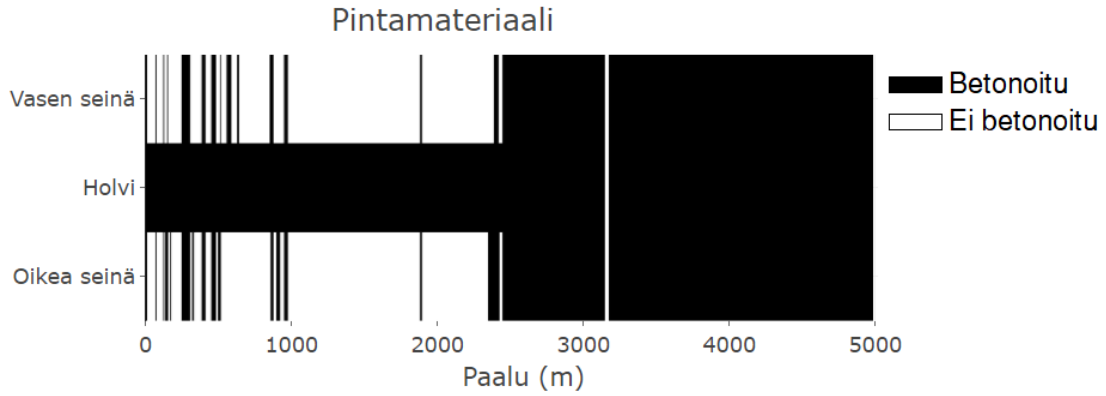
Luvussa 3.3 esiteltävä R-ohjelmiston [R Core Team, 2017] paketti CAR-BayesST [Lee, D., Rushworth, A. & Napier, G., 2018] tarjoaa Bayesiläisen lähestymistavan luvun 3.1 mallin sovitukseen. Bayesiläinen lähestymistapa mahdollistaa puuttuvien havaintojen käsittelyn tietyissä perustapauksissa, joita aineistossa on nyt ja joita toisaalta on tulevaisuudessa odotettavissa, mikäli tarkastuksia ei toteuteta koko tunnelin alueella joka vuosi. Luvussa 4 esitetään mallin tulokset ja varmistetaan mallin sopivuus aineistoon. Lopulta luvussa 5 pohditaan keinoja mallin parantamiselle ja esitetään ideoita mallin hyödyntämismahdollisuuksille kunnossapidollisten resurssien allokoinnissa.

2 Aineisto

Tutkielman aineistona on Posiva Oy:n käytetyn ydinpolttoaineen loppusijoitusta varten rakennetun tunneliverkoston ajotunnelin vauriokirjaukset vuosilta 2016–2020. Loppusijoituslaitoksen havainnekuva on esitetty kuviossa 1, jossa ajotunneli on korostettuna tumman sinisellä. Ajotunneli on 4986 metriä pitkä ja ulottuu noin 450 metrin syvyyteen merenpinnan alapuolelle. Ajotunnelin paaluluku on juokseva metriluku alkaen ylhäältä. Oikeasta ja vasemmasta puolesta puhuttaessa oletetaan kulkusuunnan olevan ylhäältä alaspäin. Ajotunnelin holvi ja paalun 2450 m alapuolella olevat seinät on muutamia poikkeuksia lukuunottamatta ruiskubetonoitua pintaa. Tämän lisäksi joitain seinäkohtia paalun 2450 m yläpuolelta on ruiskubetonoitu. Betonoidut alueet on esitetty kuviossa 2.



Kuvio 1: Havainnekuvio Olkiluodon ONKALO[®]sta, jossa ajotunneli korostettuna tumman sinisellä [Posiva Oy].



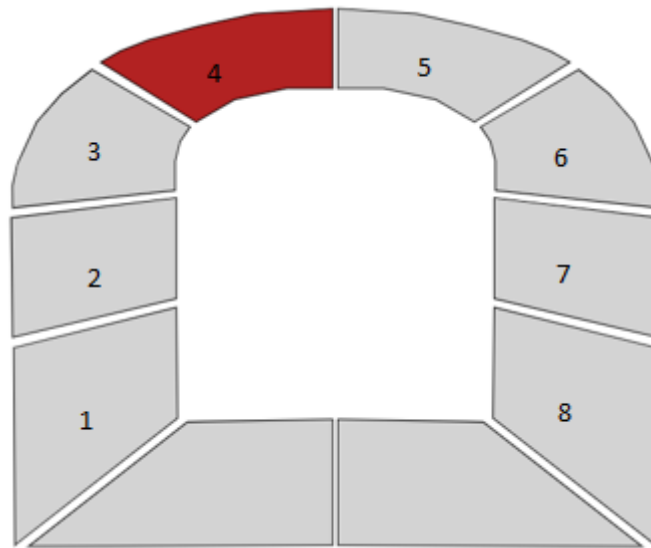
Kuvio 2: Ajotunnelin betonoidut alueet.

Ajotunnelissa irtoilevan kallion, ruiskubetonin tartunnan ja vesivaurioiden havaintoja on kirjattu tarkasteluvälillä yhteensä 3905 kappaletta. Aktiivisten vaurioiden lukumäärät vauriotyypeittäin ja vuosittain on esitetty taulukossa 1. Vauriohavainnoista on raportoitu seuraavat tiedot:

- Vaurion alkupiste metrin tarkkuudella
- Vaurion loppupiste metrin tarkkuudella
- Vaurion sijainti tunnelin profilissa, kuvio 3
- Vaurion luokka, luku 2.1
- Vaurion tyyppi (irtoileva kallio, tartuntavaurio tai vesivaurio)
- Vaurion tila (aktiivinen tai korjattu)

Taulukko 1: Aktiivisten vaurioiden lukumäärät vuosittain ja vauriotyypeittäin.

		Vauriotyyppi			
		Kallio	Tartunta	Vesi	Yhteensä
Vuosi	2016	160	0	3	163
	2017	538	0	24	562
	2018	566	19	32	617
	2019	654	371	322	1347
	2020	706	289	223	1216
	Yhteensä	2622	679	604	3905



Kuvio 3: Vaurion sijainti tunnelin profilissa: 1. Vasen alaseinä 2. Vasen yläseinä 3. Vasen kainalo 4. Vasen holvi 5. Oikea holvi 6. Oikea kainalo 7. Oikea yläseinä 8. Oikea alaseinä.

2.1 Vaurioluokitus

Vauriot on luokiteltu vaurion ominaisuuksien mukaan neljään luokkaan. Luokat 1–4 kuvaavat vaurion laajuutta ja tyyppiä pienimmästä suurimpaan. Luokitus ei itsessään huomioi vaurion merkitsevyyttä kunnossapidon kannalta, vaan perustuu yksinomaan vaurion teknisiin ominaisuuksiin seuraavasti:

- Irtoilevan kallion vaurioluokituksen määritelmät:
 1. Varisevia pieniä kallionsirpaleita, halkaisija alle 10 cm.
 2. Irtoilevia kappaleita, halkaisija vähintään 10 cm.
 3. Paikalleen jääviä lohkareita, jotka eivät irtoa karistelemalla.
 4. Itsestään pudonnut lohkare tai muu vakava stabiliteetin menetys.

- Ruiskubetonin tartuntavaurioiden luokituksen määritelmät:

1. Irti olevan alueen koko alle 0.25m^2 .
2. Irti olevan alueen koko $0.25 - 1\text{m}^2$.
3. Irti olevan alueen koko $1 - 2\text{m}^2$.
4. Irti olevan alueen koko yli 2m^2 .

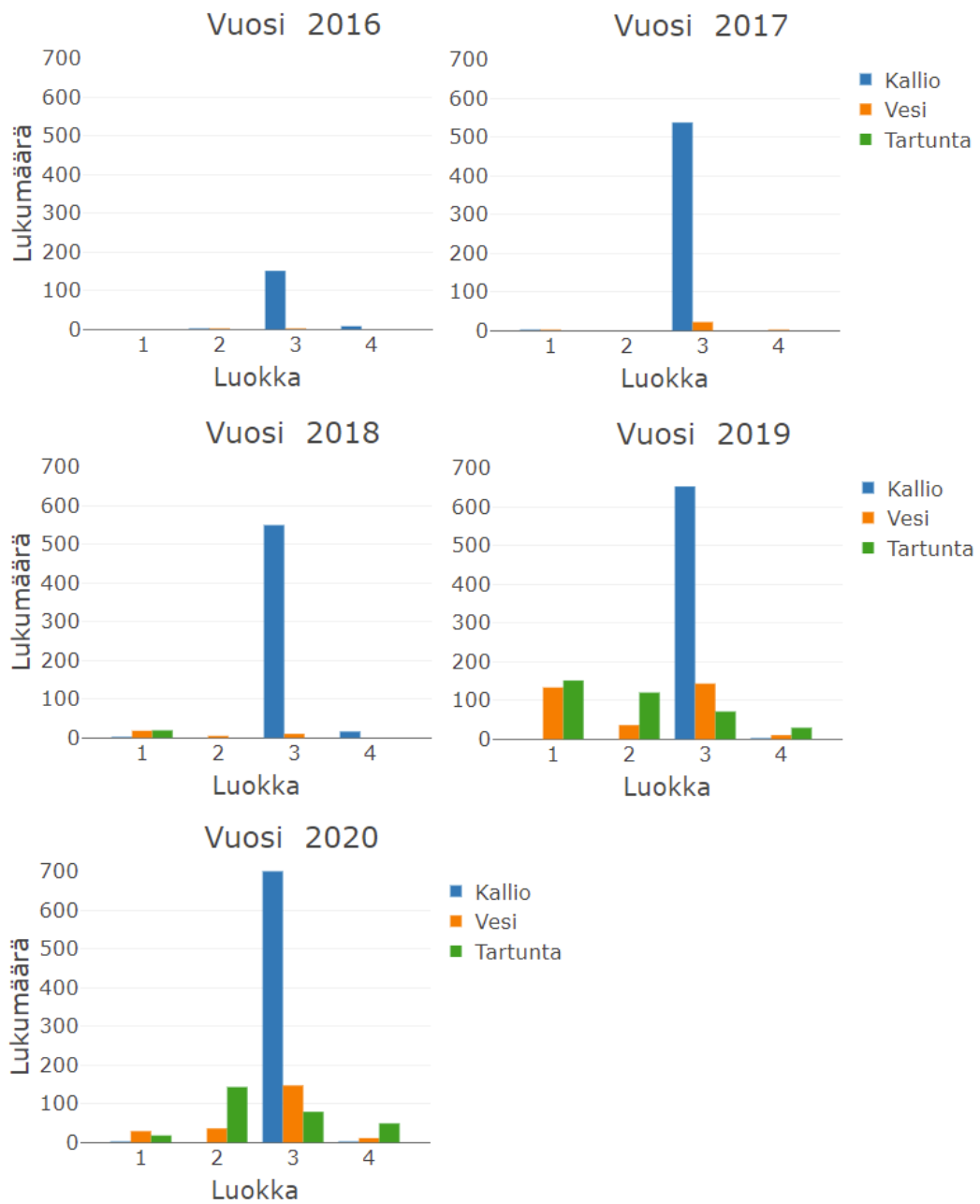
- Vesivaurioiden luokituksen määritelmät:

1. Pinta on kostea tai siinä on kalkkihärmää.
2. Pinnan raossa on vesivuoto, josta ei kuitenkaan tipu vettä.
3. Pinnan raossa on vesivuoto, josta tippuu vettä.
4. Valuva tai muu vakava vesivuoto.

Vaurioiden lukumäärät luokittain ja vuosittain on esitetty kuviossa 4.

2.2 Aineiston luotettavuus

Posiva Oy:n loppusijoituslaitoksen kunnossapidon vastuuhenkilöiden kanssa 2021 käytyjen keskustelujen perusteella tiedetään, että ajotunneli on tarkastettu systemaattisesti kokonaan vain vuosina 2019 ja 2020. Tämä selittää osaltaan vähäisiä vauriokirjauksia taulukossa 1 aikaisemmilta vuosilta. Intuitiivinen oletus, että vuosilta 2016–2018 vain pahimmat vauriot olisi havaittu, ei kuitenkaan kuvion 4 perusteella saa tukea, sillä vauriokirjauksia ei ole vakavammissa vaurioluokissa mitenkään yliedustetusti. Toisaalta ei ole myöskään mitään perusteltua syytä olettaa, että pahimmat vauriot olisi jätetty kirjaamatta tai vaurioita olisi kirjattu vain tietystä osasta tunnelia. Tästä johtuen puuttuvuutta pidetään täysin satunnaisena. Vuosien 2016–2018 osalta käytetään vaurioiden kohdalla luokitusta niin kuin se aineistossa esiintyy, mutta vauriottomat kohdat pidetään täydellisesti satunnaisesti puuttuvana tietona. Oletetaan siis, että vuosina 2019 ja 2020 kaikki vauriot on saatu kirjattua, mutta vuosina 2016–2018 ei ole havaittu yhtään ehjää aluetta, vaan näiden alueiden vaurioituneisuus mallinnetaan.



Kuvio 4: Vaurioiden lukumäärät luokittain ja vuosittain.

Tulevaisuuden kannalta tavoitteena on jättää ehjät alueet säännöllisen tarkastuksen ulkopuolelle ja tarkastaa koko tunneli harvemmin. Tällä tavoin generoituva puuttuvuus ei ole täysin satunnaista vaan perustuu havaintoihin ja malliin.

Luvun 3 mallin sovituksen yhteydessä puuttuva aineisto paikataan datan augmentointia käyttäen [Tanner, M. & Wong, W., 1987]. Tapa on validi molemmissa tämän luvun tilanteissa, sillä voidaan ajatella, että puuttuvuus ei ole kytköksissä vasteen senhetkiseen arvoon vaan se riippuu vain havaitusta aineistosta ja mallista. Bayes-menetelmien yhteydessä puuttuvuus ei näissä tilanteissa teknisesti eroa parametrien estimoinnista, vaan aineistoa voidaan laajentaa iteratiivisesti. Käytämällä yleispätevästi parametreista merkintää θ , havaittusta aineistosta Y_{obs} ja puuttuvasta aineistosta Y_{mis} voidaan siis toistaa seuraavia vaiheita iteratiivisesti:

1. Generoidaan $Y_{mis}^{(t+1)}$ jakaumasta $p(Y_{mis}|Y_{obs}, \theta^{(t)})$.
2. Generoidaan $\theta^{(t+1)}$ jakaumasta $p(\theta|Y_{obs}, Y_{mis}^{(t+1)})$.

2.3 Vasteen muotoilu

Vaurioiden sijainti on kirjattu aineistossa pituussuunnassa metrin tarkkuudella ja tunnelin profiilin suunnassa jako on tehty kuvion 3 mukaisesti kahdeksaan lohkokoon. 4986 metrin pituisessa ajotunnelissa tämä johtaisi $8 \cdot 4986 = 39888$ erilliseen alueeseen. Tämä puolestaan johtaisi mallinnuksen kannalta tietokoneen muistin suhteen laskennalliseen haasteeseen, jota on kuvattu tarkemmin luvussa 3.2. Tästä laskennallisesta seikasta johtuen alueita on yhdistetty seuraavasti:

- Vasen seinä käsittää sekä ala- että yläseinän (kuvio 3 profiilin osat 1 ja 2).
- Oikea seinä käsittää sekä ala- että yläseinän (kuvio 3 profiilin osat 7 ja 8).
- Holvi käsittää molempien holvin osien lisäksi molemmat kainalot (kuvio 3 profiilin osat 3,4,5 ja 6).

Edellisten lisäksi yhdistetään alueita tunnelin pituussuunnassa siten, että kukin alue on pituudeltaan 5 metriä. Pituussuunnassa valintaa perustellaan lasken-

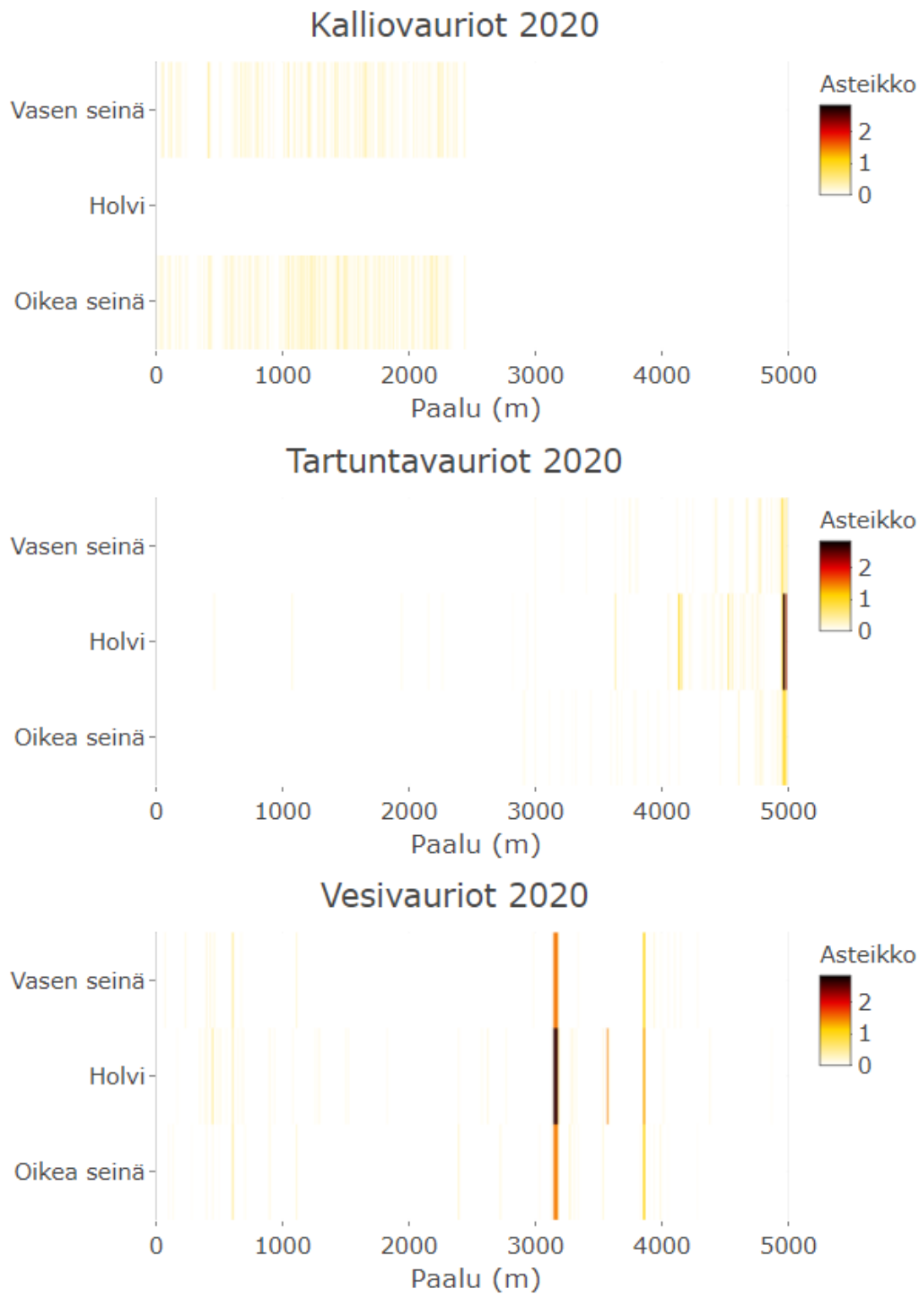
nallisten haasteiden helpottamisen lisäksi sillä, että näin muodostetut alueet ovat pituus- ja leveysuunnassa lähes samansuuruiset. Ajotunnelin seinän korkeus vaihtelee noin 5.3 metristä 5.9 metriin ja tunnelin leveys noin 5.5 metristä 8.5 metriin. Näillä valinnoilla muodostettuja erillisiä alueita on 2994 kappaletta, ja kunkin alueen pinta-ala on arvioitu rakennuspiirrustuksista.

Edellä muodostetuille erillisille alueille johdetaan vaste. Vasteessa esiintyy vaurioluokitusten summa, joten vasteen muodostamista varten vaaditaan oletus, että vaurioluokitus on välimatka-asteikollinen. Tämä oletus ei ole itsestään selvä, mutta jos alkuperäistä luokitusta ajatellaan järjestyslukuasteikollisena approksimaationa jatkuvasta muuttujasta ja luokkia on riittävästi, välimatka-asteikon käyttäminen ei yleensä aiheuta ongelmia [Johnson, D.R. & Creech, J.C., 1983]. Riittävänä määränä edellisessä pidetään vähintään viittä luokkaa. Robitzsch [2020] esittää jopa, että järjestyslukuasteikollisen muuttujan mallintaminen jatkuvana on lähes aina hyväksyttävää.

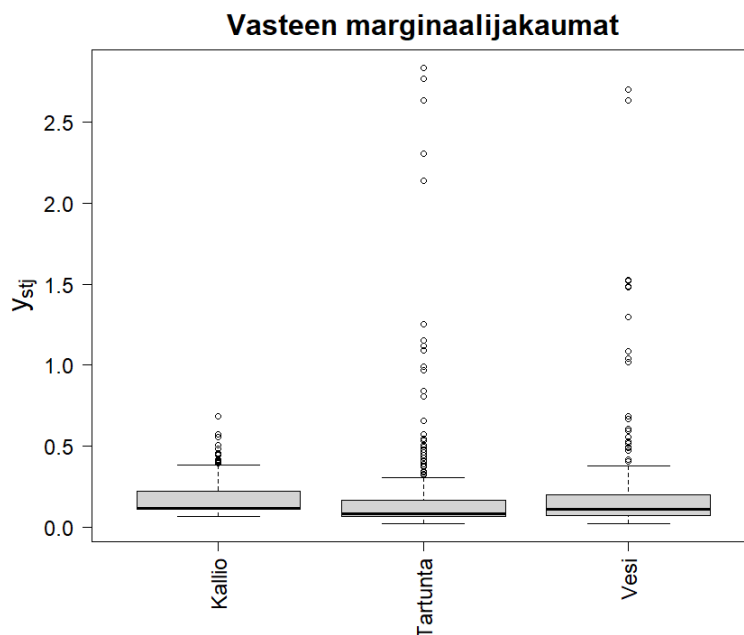
Viitataan merkinnällä V_{stj} joukkoon alueen s tyyppin j aktiivisia vaurioita ajanhetkellä t . Merkitään yksittäistä vauriota kirjaimella v ja tuon vaurion luokasta käytetään merkintää l_v . Vaste lasketaan nyt kaavalla

$$Y_{stj} = \frac{\sum_{v \in V_{stj}} l_v}{A_s},$$

jossa A_s on alueen s pinta-ala. Tällä tavoin johdetut alueelliset vasteet on vuoden 2020 osalta esitetty kallio-, tartunta- ja vesivaurioille kuviossa 5. Vasteen vaihteluväli koko aineistossa on $[0, 2.83]$ ja suurin osa vasteiden arvoista on nolli. Nollasta eroavien vasteenarvojen marginaalijakaumat kullekin vauriotyypille alueiden ja ajan yli on havainnollistettu viiksilaatikoilla kuviossa 6.



Kuvio 5: Ajotunnelin vaurioituneisuus vuonna 2020.



Kuvio 6: Kunkin vauriotyyppin marginaalijakaumat alueiden ja ajan yli. Huomaa, että vauriottomat alueet on jätetty pois.

2.4 Korjaustiedot

Edellä luvussa 2.3 johdettua vastetta selitetään tehdyillä korjauksilla. Korjattujen vaurioiden lukumäärät vuosittain ja vauriotyypeittäin ovat taulukossa 2.

Taulukko 2: Korjattujen vaurioiden lukumäärät vuosittain ja vauriotyypeittäin.

		Vauriotyyppi			
		Kallio	Tartunta	Vesi	Yhteensä
Vuosi	2017	1	0	0	1
	2018	42	0	0	42
	2019	133	23	2	158
	2020	12	0	0	12
	Yhteensä	188	23	2	213

Korjaustiedoista kullekin alueelle s , vauriotyypille j ja ajanhetkelle t johdetaan selittäjäksi tieto siitä, kuinka paljon alueessa on tehty korjauksia. Laskennallinen määrittely on johdettu korjatuista vaurioista samalla tavalla, kuin vaste on määritetty luvussa 2.3 eli myös korjatun vaurion luokitus tiedetään ja tietoa hyödynnetään. Täsmällisemmin, jos merkinnällä \mathcal{K}_{stj} viitataan joukkoon alueen s vauriotyyppin j vuonna t korjattuja vaurioita, korjaustietojen kovariaati on

$$K_{stj} = \frac{\sum_{v \in \mathcal{K}_{stj}} l_v}{A_s},$$

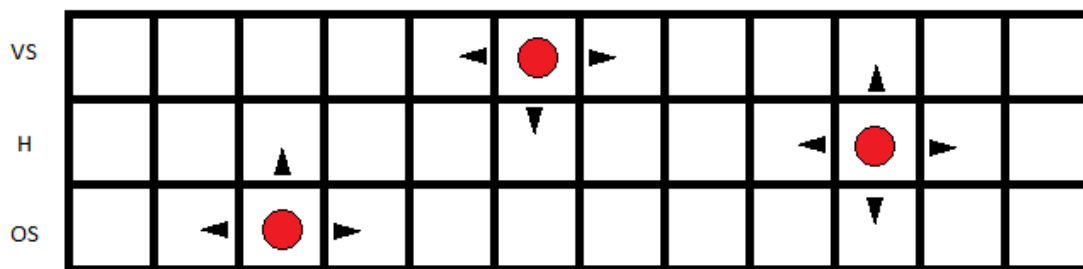
jossa A_s on alueen s pinta-ala.

2.5 Spatiaalinen rakenne

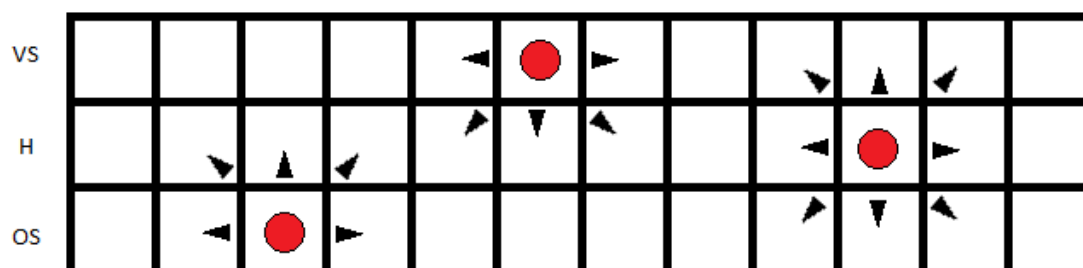
Edellä luvussa 2.3 esitettiin alueellinen vaste. On luontevaa olettaa, että nämä vasteet ovat spatiaalisesti riippuvia. Tämän riippuvuuden mallintamiseksi määritetään naapurustorelaatio ja vastaava naapurimatriisi \mathbf{W} . Yleisesti tämän naapurimatriisin alkiot w_{kj} kuvaavat alueiden k ja j spatiaalista läheisyyttä siten, että suuri luku viittaa läheisyyteen ja pieni luku tai nolla alueisiin, jotka ovat kaukana toisistaan. Huomioitavaa on kuitenkin, että $w_{kk} = 0$, eli alue ei ole itsensä naapuri [Lee, D., Rushworth, A. & Napier, G., 2018].

Teknisesti naapurimatriisin \mathbf{W} määrittelyä rajoittavat vain seuraavat tekijät: symmetrisyys, ei-negatiivisuus ja rivisummien positiivisuus [Lee, D., Rushworth, A. & Napier, G., 2018]. Usein käytetään kuitenkin binääristä matriisia kuvaamaan naapurustoa. Tyypillisiä valintoja on käyttää 1- tai $\sqrt{2}$ -naapurustoa, joita havainnollistetaan kuvioissa 7 ja 8 vastaavasti. Tässä työssä esitetään sovellettavaksi myös kolmas vaihtoehto, jonka havainnollistus on esitetty kuviossa 9. Tätä valintaa voidaan perustella sillä, että samalla paalumetrillä vasemmalla ja oikealla alaseinällä olevat vauriot ovat etäisyydeltään itseasiassa lähempänä toisiaan kuin vaikkapa vasemman alaseinän vaurio ja oikean holvin vaurio.

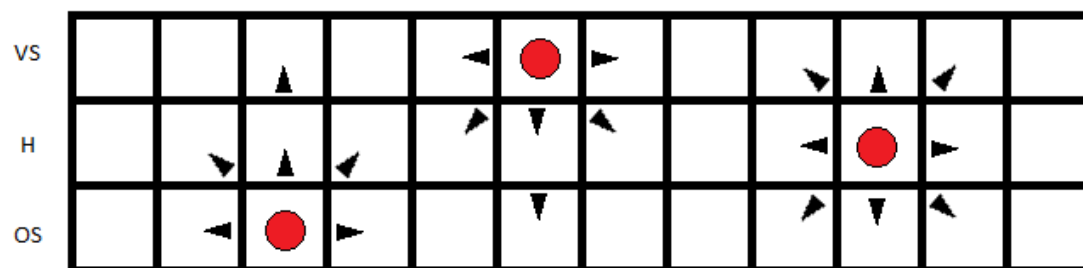
Kunkin tässä työssä käytetyn naapuruston tapauksessa naapurimatriisi $\mathbf{W}_{2994 \times 2994}$ on harva eli alkioista suurin osa (99.98%) on nollia. Lisäksi matriisi on nauhamatriisi eli nollasta poikkeavat alkiot on lähellä diagonaalia. Esitellyissä naapurustoissa naapurimatriisin nauhanleveys $k = 4$, kun käytössä on $\sqrt{2}$ -naapurusto ja $k = 3$ kahden muun naapuruston tilanteessa. Naapurimatriisin alkiot ovat siis $w_{ij} = 0$, jos $j < i - k$ tai $j > i + k$.



Kuvio 7: Havainnollistava kuvio 1-naapurustosta aineiston tilanteessa.



Kuvio 8: Havainnollistava kuvio $\sqrt{2}$ -naapurustosta aineiston tilanteessa.



Kuvio 9: Havainnollistava kuvio itse määritetystä naapurustosta aineiston tilanteessa.

2.6 Moranin I

Luvussa 2.3 johdettujen aluemuuttujien spatiaalista riippuvuutta voidaan kuvata Moranin [1950] esittelemällä tunnusluvulla. Olkoon alueen i aluemuuttujan arvo Y_i ja naapurimatriisin alkiot w_{ij} , tällöin tunnusluku on

$$I = \frac{\sum_{i=1}^S \sum_{j=1}^S w_{ij}^* (Y_i - \bar{Y})(Y_j - \bar{Y})}{\frac{1}{S} \sum_{i=1}^S (Y_i - \bar{Y})^2},$$

jossa $\bar{Y} = \frac{1}{S} \sum_{i=1}^S Y_i$ ja $w_{ij}^* = w_{ij} / (\sum_{k \neq l} w_{kl})$. Tämä voidaan tulkita alueiden väliseksi painotetuksi korrelaatiokertoimeksi, jossa läheiset alueparit painottuvat voimakkaammin. Tunnusluku saa arvoja väliltä $[-1, 1]$, jossa luku -1 viittaisi säännöllisesti hajallaan esiintyviin arvoihin, nolla satunnaisesti esiintyviin arvoihin ja 1 rypäinä esiintyviin arvoihin. Moranin I-tunnusluvulle tunnetaan odotusarvo

$$E[I] = \frac{-1}{S-1}$$

ja varianssi

$$Var[I] = \frac{SM_4 - M_3M_5}{(S-1)(S-2)(S-3) \left(\sum_{i=1}^S \sum_{j=1}^S w_{ij} \right)^2},$$

jossa

$$\begin{aligned}
M_1 &= \frac{1}{2} \sum_{i=1}^S \sum_{j=1}^S (w_{ij} + w_{ji})^2 \\
M_2 &= \sum_{i=1}^S \left(\sum_{j=1}^S w_{ij} + \sum_{j=1}^S w_{ji} \right) \\
M_3 &= \frac{\frac{1}{S} \sum_{i=1}^S (Y_i - \bar{Y})^4}{\left(\frac{1}{S} \sum_{i=1}^S (Y_i - \bar{Y})^2 \right)^2} \\
M_4 &= (S^2 - 3S + 3)M_1 - SM_2 + 3 \left(\sum_{i=1}^S \sum_{j=1}^S w_{ij} \right)^2 \\
M_5 &= (S^2 - S)M_1 - 2SM_2 + 6 \left(\sum_{i=1}^S \sum_{j=1}^S w_{ij} \right)^2 .
\end{aligned}$$

Odotusarvon kaavasta nähdään suoraan, että odotusarvo lähestyy nollaa, kun alueiden määrä kasvaa ja tämän aineiston tilanteessa siis $E[I] = -0.0003$.

Moranin I on laskettu vuosien 2019–2020 vaurioaineistosta eri naapurustoille ja kullekin vauriotyypille. Vertailtavat naapurustot on havainnollistettu luvussa 2.5 ja laskennan tulokset ovat taulukossa 3. Tuloksista nähdään, että kaikissa tapauksissa $I > 0$ eli aineistossa on spatiaalista riippuvuutta kaikkien vauriotyyppien kohdalla. Vauriot esiintyvät tyypillisimmin rypäinä, ja vahvin spatiaalinen riippuvuus on vesivaurioilla.

Taulukko 3: Moranin I-tunnusluku ja hajonta eri naapurustoille ja vauriotyypeille vuosina 2019–2020.

Naapurusto	Vauriotyyppi	Vuosi	Havaittu I	sd(I)
1	Kallio	2019	0.295	0.014
$\sqrt{2}$	Kallio	2019	0.159	0.011
ehdotettu	Kallio	2019	0.317	0.013
1	Kallio	2020	0.290	0.014
$\sqrt{2}$	Kallio	2020	0.155	0.011
ehdotettu	Kallio	2020	0.318	0.013
1	Vesi	2019	0.847	0.014
$\sqrt{2}$	Vesi	2019	0.831	0.010
ehdotettu	Vesi	2019	0.836	0.013
1	Vesi	2020	0.848	0.014
$\sqrt{2}$	Vesi	2020	0.832	0.010
ehdotettu	Vesi	2020	0.836	0.013
1	Tartunta	2019	0.238	0.014
$\sqrt{2}$	Tartunta	2019	0.153	0.010
ehdotettu	Tartunta	2019	0.233	0.012
1	Tartunta	2020	0.520	0.013
$\sqrt{2}$	Tartunta	2020	0.445	0.010
ehdotettu	Tartunta	2020	0.500	0.012

3 Menetelmät

Käytettävän mallin pohja on yleistetyssä lineaarisessa mallissa, jota sovelletaan moniulotteiseen vasteeseen. Kuten edellä luvussa 2.6 osoitettiin, aineistossa esiintyy selvästi spatiaalista riippuvuutta. Tämä sekä temporaalinen ja vauriotyyppien välinen korrelaatio huomioidaan laajentamalla malli hierarkkiseksi Bayes-malliksi, jolla on tilanteeseen soveltuva rakenne. Malli sovitetaan Markovin ketju Monte Carlo- eli MCMC-menetelmällä, joka on kuvattu esimerkiksi artikkelissa [van Ravenzwaaij, D., Cassey, P. & Brown, S.D., 2018].

Mallintamista varten tarkastelu on rajattu $s = 1, \dots, S$ erilliseen alueeseen. Näiltä alueilta on havainnot ajanhetkiltä $t = 1, \dots, N$ kustakin vauriotyypistä $j = 1, \dots, J$. Moniulotteisen vasteen tapauksessa vaste on siis vektori $\mathbf{Y}_{st} = [Y_{s,t,1}, \dots, Y_{s,t,J}]^T$, jonka alkio kuvaa vauriotyyppin j vaurion vakavuutta alueessa s ajanhetkellä t luvussa 2.3 esitetyllä tavalla.

3.1 Hierarkkinen Bayes-malli

Edellä esitetyillä merkinnöillä voidaan kirjoittaa yleistetty lineaarinen Bayes-malli, jossa on mukana paikka-, aika- ja vauriotyypikohtainen satunnaisvaikutus. Malli kunkin vauriotyyppin j vaurioituneisuudelle alueessa s ajanhetkellä t on siis muotoa

$$Y_{stj} \sim N(\mu_{stj}, \nu_j^2) \quad (1)$$

$$\mu_{stj} = \mathbf{X}_{st}^T \beta_j + \phi_{stj} \quad (2)$$

$$\beta_j \sim N(\mathbf{0}, 100000\mathbf{I}). \quad (3)$$

Kunkin vauriotyyppin j vaurioituneisuus alueessa s ajanhetkellä t oletetaan siis normaalijakautuneeksi odotusarvolla μ_{stj} ja vauriotyypikohtaisella varianssilla ν_j^2 , lauseke (1). Tarkalleen ottaen alueellinen vaste voi saada vain diskreettejä

arvoja luvun $1/A_s$ välein. Tätä ei kuitenkaan pidetä ongelmallisena, sillä malli on hierarkkinen. Vasteen odotusarvolle jokaiselle alueelle, ajanhetkelle ja vauriotyypille kirjoitetaan lausekkeen (2) mukainen lineaarinen malli, jossa β_j on vauriotyypikohtaisesti estimoitava vektori regressioparametreille ja ϕ_{stj} on satunnaisvaikutus. Regressiokertoimet oletetaan lausekkeen (3) mukaisesti normaalijakautuneiksi epäinformatiivisella priorilla $N(0, 100000\mathbf{I})$. Tässä työssä selittäjinä käytetään luvussa 2.4 esitettyjä korjaustietoja eli $\mathbf{X}_{st} = [K_{s,t,1}, \dots, K_{s,t,j}]^T$.

Mallin satunnaisvaikutukset ϕ_{stj} ovat keskenään korreloituneita spatiaalisesti, temporaalisesti ja vauriotyyppien välillä. Näiden korrelaatioiden huomioimiseksi muodostetaan vektori $\phi_t = [\phi_{1t}, \dots, \phi_{St}]^T$, jossa kaikki satunnaisvaikutukset ajanhetkellä t on järjestetty siten, että $\phi_{st} = [\phi_{st1}, \dots, \phi_{stJ}]$ on ajanhetken t kaikki satunnaisvaikutukset paikassa s . Näillä merkinnöillä AR(1)-rakenteen malli voidaan lähteitä [Leroux B., Lei, X. & Breslow, N., 2000] ja [Lee, D., Rushworth, A. & Napier, G., 2018] mukaillen esittää muodossa

$$\phi_t | \phi_{t-1} \sim N(\rho_T \phi_{t-1}, [\mathbf{Q}(\mathbf{W}, \rho_S) \otimes \Sigma^{-1}]^{-1}), \quad t=2, \dots, N \quad (4)$$

$$\phi_1 \sim N(\mathbf{0}, [\mathbf{Q}(\mathbf{W}, \rho_S) \otimes \Sigma^{-1}]^{-1}) \quad (5)$$

$$\Sigma \sim \text{Inverse-Wishart}(d, \Omega) \quad (6)$$

$$\rho_S \sim \text{Tas}(0, 1) \quad (7)$$

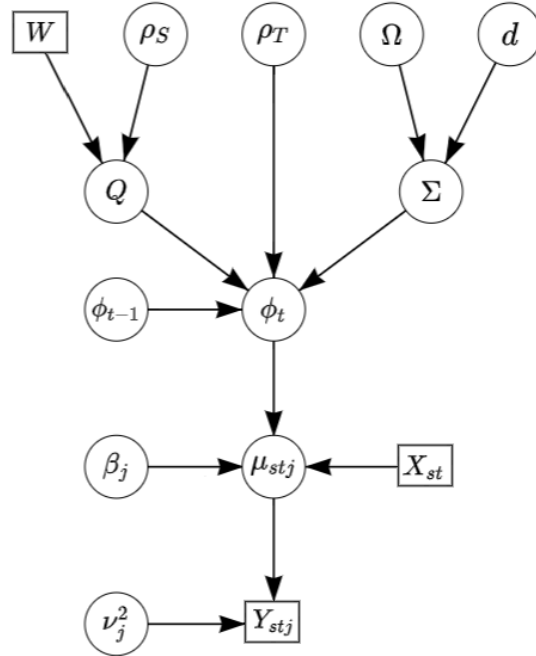
$$\rho_T \propto 1, \quad (8)$$

jossa operaattori \otimes viittaa Kroneckerin tuloon ja tarkkuusmatriisi $\mathbf{Q}(\mathbf{W}, \rho_S)$ määräytyy kaavalla

$$\mathbf{Q}(\mathbf{W}, \rho_S) = \rho_S [\text{diag}(\mathbf{W}\mathbf{1}) - \mathbf{W}] + (1 - \rho_S)\mathbf{I}.$$

Malli huomioi vauriotyyppien välisen korrelaation matriisin Σ kautta satunnaisvaikutusten korreloituneisuudella lausekkeiden (4) ja (5) mukaisesti. Tämä matriisi Σ estimoidaan aineistosta ja prioriksi asetetaan käännteinen Wishart-

jakauma epäinformatiivisilla hyperparametreilla $d = J + 1 = 4$ ja $\Omega = 0.01\mathbf{I}$. Käänteinen Wishart-jakauma on konjugaattinen moniulotteisen normaalijakauman kovarianssimatriisille. Spatiaalinen korrelaatio tulee huomioiduksi tarkkuusmatriisiin $\mathbf{Q}(\mathbf{W}, \rho_S)$ kautta satunnaisvaikutusten korreloituneisuudella, lausekkeet (4) ja (5). Naapurimatriisin \mathbf{W} määrittely on käsitelty tarkemmin luvussa 2.5. Parametri ρ_S kuvaa yleistä spatiaalista riippuvuutta. Arvo nolla viittaisi riippumattomuuteen ja arvo yksi vahvaan spatiaaliseen riippuvuuteen. Priorijakomaksi parametrille ρ_S asetetaan lausekkeen (6) mukaisesti tasajakauma välille $[0,1]$. Satunnaisvaikutusten ajallinen riippuvuus mallinnetaan satunnaisvaikutusten odotusarvon kautta autoregressiivisesti, lauseke (4). Parametri ρ_T on autoregressiivisen rakenteen parametri, jolla edellisen ajanhetken satunnaisvaikutukset odotusarvoisesti kertautuvat lausekkeen (4) mukaisesti. Mallin rakennetta ja parametrien välisiä riippuvuuksia on havainnollistettu suunnatulla syklittömällä graafilla kuviossa 10.



Kuvio 10: Mallin rakenne ja parametrien väliset riippuvuudet esitettynä suunnatulla syklittömällä graafilla.

3.2 Kovarianssimatriisin laskeminen

Kuten tämä työn luvussa 2.5 todettiin naapurimatriisi $\mathbf{W}_{2994 \times 2994}$ on harva eli alkioista suurin osa on nollia. Matriisin \mathbf{W} harvuudesta seuraa suoraan, että myös matriisi $\mathbf{Q}(\mathbf{W}, \rho_S)$ ja edelleen 8982×8982 tarkkuusmatriisi $\mathbf{Q}(\mathbf{W}, \rho_S) \otimes \Sigma^{-1}$ ovat harvoja ja voidaan tallentaa tehokkaasti tietokoneen muistia säästäten. Ongelmana kuitenkin on, että harvan matriisin käänteismatriisi ei välttämättä ole harva. Tämä voidaan osoittaa esimerkiksi niin sanotun nuolimatriisin

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

avulla. Selvästi \mathbf{A} on harva, mutta

$$\mathbf{A}^{-1} = \begin{bmatrix} -\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \\ \frac{1}{6} & -\frac{1}{6} & \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \\ \frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \\ \frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \\ \frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} & -\frac{1}{6} & -\frac{1}{6} \\ \frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} & -\frac{1}{6} \\ \frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{5}{6} \end{bmatrix}$$

ei ole. Tästä johtuen kovarianssimatriisin $[\mathbf{Q}(\mathbf{W}, \rho_S) \otimes \Sigma^{-1}]^{-1}$ tallentaminen yleisessä tilanteessa vaatii runsaasti muistia ja asettaa rajoitteet alueiden ja/tai

vauriotyyppien määrille. Tämän laskentateknisen seikan vuoksi alueita yhdistettiin luvussa 2.3. Approksimatiivisia muistia säästäviä menettelyjä harvan tarkkuusmatriisin kääntämiselle on tutkittu ja eräs ratkaisu on esitetty esimerkiksi artikkelissa [Sidén, P., Lindgren, F., Bolin, D. & Villani, M., 2017]. Tässä työssä käytettävää aluejakoa pidetään kuitenkin riittävän tiheänä ja mallin sovituksessa hyödynnetään valmista R-ohjelmistolle [R Core Team, 2017] tehdyn paketin CARBayesST [Lee, D., Rushworth, A. & Napier, G., 2018] funktiota MVST.CARar, joka laskee kovarianssimatriisin eli tarkkuusmatriisin käänteismatriisin tarkasti.

3.3 Mallin sovitus

Luvun 3.1 mallin sovitus on toteutettu R-ohjelmistolla [R Core Team, 2017] paketin CARBayesST [Lee, D., Rushworth, A. & Napier, G., 2018] funktiolla MVST.CARar. Funktiolle syötetään mallin kaava siten, että vasteena on $(SN) \times J$ matriisi, jossa S on alueiden määrä, N on ajanhetkien määrä ja J on vauriotyyppien määrä. Matriisin alkiot järjestetään siten, että jokaisen sarakkeen S ensimmäistä alkiota vastaavat kyseisen vauriotyyppin vaurioituneisuutta ajanhetkellä $t = 1$ kaikilla alueilla $s = 1, \dots, S$ ja seuraavat S alkiota vaurioituneisuutta ajanhetkellä $t = 2$ jne. Kovariaatit ovat $SN \times 1$ vektoreita ja alkioiden järjestys on vastaava: Ensin selittäjän arvot kaikille $s = 1, \dots, S$ alueille ajanhetkellä $t = 1$ ja tämän jälkeen ajanhetkellä $t = 2$ jne.

Edellisten lisäksi funktiolle on annettava vasteen jakaumaoletus, AR-mallin aste, naapurimatriisi, sekä MCMC-ketjun ja lämmittelyjakson pituudet. Vaihtoehtoina vasteen jakaumalle on normaalijakautunut vaste identtisellä linkillä, binomivaste logistisella linkillä tai Poisson-jakautunut vaste logaritmilinkillä. AR-malli voidaan sovittaa 1 tai 2 asteisena. Tässä työssä oletetaan vasteelle normaalijakauma, kuten luvussa 3.1 on todettu ja sovitetaan AR(1)-malli, sillä ajan suhteen aineistoa on niukasti ja AR(2)-mallin estimointia ei pidetä luotettavana.

Naapurimatriisin määrittelyä on käsitelty luvussa 2.5 ja mallin sovitus on toteutettu kaikilla kolmella tuossa luvussa esitellyllä naapurustolla. Mallin sovituksessa käytetään MCMC-ketjun pituutena 200 000 ja lämmittelyjaksona ketjujen alusta 100 000 havaintoa on jätetty pois.

Muita mahdollisia säätöjä mallin sovituksessa ovat MCMC-ketjujen ohennus, priorijakaumien parametrien säädöt parametreille β_j , ν^2 ja Σ sekä parametrien ρ_S ja ρ_T asettaminen vakioarvoon. Tässä työssä MCMC-ketjuja on ohennettu muistin säästämiseksi siten, että vain joka 40. havainto tallennetaan. Priorijakaumien parametrit ovat oletusasetusten mukaiset ja esitetty luvussa 3.1. Parametrit ρ_T ja ρ_S estimoidaan aineistosta.

Funktio `MVST.CARar` palauttaa parametrien estimaatit (posteriorimediaanit), sovitteiden arvot (posterioriodotusarvot), sekä tässäkin työssä luvussa 4 esiteltävät konvergenssin tunnusluvut. Lisäksi malli laskee informaatiokriteerit DIC [Spiegelhalter, D., Best, N., Carlin, B. & Van Der Linde, A., 2002] ja WAIC [Watanabe, S., 2010] sekä logaritmoidun uskottavuusfunktion. Näitä kolmea tunnuslukua on mahdollista hyödyntää mallien vertailussa, mutta näitä ei käytetä tässä työssä. Mainitaan kuitenkin, että näiden lukujen perusteella paras malli on sellainen, joka maksimoi logaritmoidun uskottavuusfunktion ja minimoi molemmat informaatiokriteerit. Edellä mainittujen lisäksi funktio palauttaa kaikkien parametrien, sekä puuttuvien vasteen arvojen otospolut kokonaisuudessaan.

4 Tulokset

Tässä luvussa tarkastellaan saatuja tuloksia, kun käytössä on ollut tässä työssä ehdotettu naapurusto, jota on havainnollistettu luvun 2.5 kuviossa 9. Muiden naapurustojen tulokset on taulukoitu liitteissä A ja B. Käytetyn MCMC-ketjun pituus on 200 000 ja lämmittelyjaksona on ketjujen alusta puolet havainnoista jätetty pois. Lisäksi muistin säästämiseksi ketjuja on ohennettu tallentamalla vain joka 40. havainto. Näin ollen kunkin parametrin posteriorijakaumista on 2500 havaintoa. Parametrien estimaatit, 95 %:n posterioriväli, efektiivinen otoskoko [Kass, R., Carlin, B., Gelman, A. & Neal, R., 1998] ja Geweken [1991] konvergenssin tunnusluku G ovat taulukossa 4. Efektiivinen otoskoko määritellään

$$n_{eff} = \frac{n}{1 + 2 \sum_i^{\infty} \rho(i)},$$

jossa $\rho(i)$ on kiinnostuksen kohteena olevan parametrin MCMC-ketjun autokorrelaatio viiveellä i . Gewegen konvergenssissa puolestaan verrataan kiinnostuksen kohteena olevan parametrin MCMC-ketjun ensimmäisen 10 %:n keskiarvoa viimeisen 50 %:n keskiarvoon ja taulukossa 4 esitetään $N(0,1)$ jakautunut tunnusluku G . Tämän tunnusluvun perusteella ketjuja voidaan siis pitää konvergoituneina mikäli $G \in [-1.96, 1.96]$.

Taulukko 4: Mallin posteriorijakaumien tunnusluvut, efektiiviset otoskoot ja Geweken konvergenssin tunnusluvut. Regressiokertoimet, joiden 95 %:n posterioriväli ei sisällä lukua 0 on merkitty symbolilla *.

Parametri	Mediaani	2.5%	97.5%	n_{eff}	G
*Kallio - β_0	0.027	0.026	0.027	146.1	0.8
*Kallio - $\beta_{\text{kallion k.}}$	-0.246	-0.275	-0.221	621.1	0.4
Kallio - $\beta_{\text{tartunnan k.}}$	-0.001	-0.034	0.034	1191.1	0.5
Kallio - $\beta_{\text{vesi k.}}$	-0.003	-0.579	0.568	1502.2	-0.1
*Tartunta - β_0	0.010	0.008	0.013	33.1	-1.1
Tartunta - $\beta_{\text{kallion k.}}$	0.000	-0.049	0.050	597.4	-1.6
*Tartunta - $\beta_{\text{tartunnan k.}}$	-0.355	-0.432	-0.275	534.3	-1.1
Tartunta - $\beta_{\text{vesi k.}}$	-0.004	-1.478	1.412	646.2	0.6
*Vesi - β_0	0.027	0.024	0.032	17.8	-0.6
Vesi - $\beta_{\text{kallion k.}}$	-0.018	-0.062	0.029	410.1	-1.2
Vesi - $\beta_{\text{tartunnan k.}}$	-0.010	-0.071	0.048	485.3	-1.2
Vesi - $\beta_{\text{vesi k.}}$	0.803	-0.178	1.827	468.2	-1.8
ν_{kallio}^2	$6.4 \cdot 10^{-5}$	$5.9 \cdot 10^{-05}$	$7.0 \cdot 10^{-05}$	471.8	0.1
ν_{Tartunta}^2	$1.1 \cdot 10^{-3}$	$1.0 \cdot 10^{-03}$	$1.2 \cdot 10^{-03}$	97.2	0.1
ν_{Vesi}^2	$9.4 \cdot 10^{-5}$	$8.2 \cdot 10^{-05}$	$10.7 \cdot 10^{-05}$	570.3	-0.5
Σ_{11}	0.002	0.002	0.002	1114.0	-1.0
Σ_{22}	0.002	0.002	0.002	89.3	-0.1
Σ_{33}	0.005	0.004	0.005	295.8	-0.3
ρ_S	0.926	0.918	0.934	127.2	-1.9
ρ_T	1.004	0.997	1.012	165.5	0.3

Taulukosta 4 nähdään, että vaikka mallien sovituksessa käytettiin pitkiä MCMC-ketjuja ja ohennusta, osa efektiivisistä otoskoista on silti melko pieniä. Vaikka pienet efektiiviset otoskoot viittaavatkin ketjun huonoon sekoittumi-

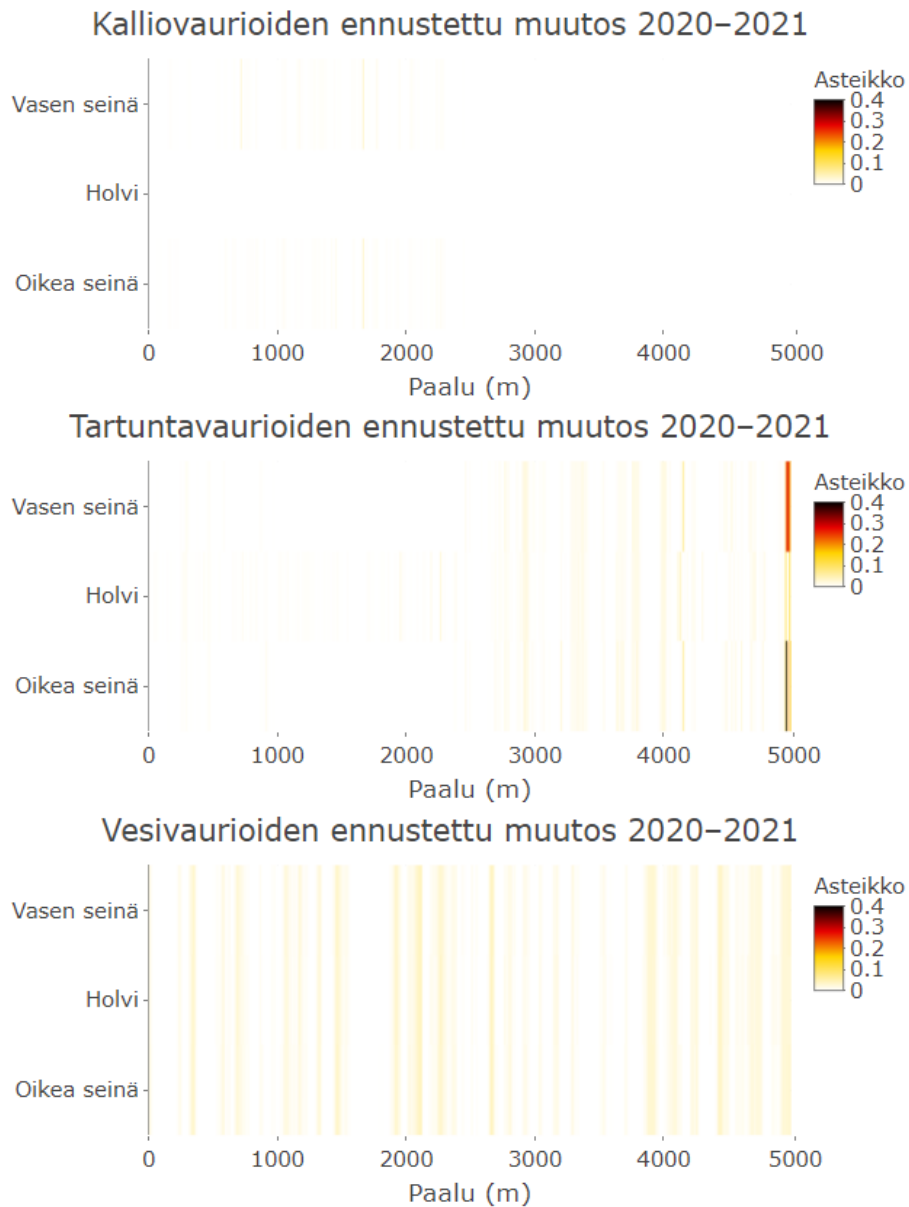
seen eli autokorrelaatioon otoksessa, ei tämä mallin toimivuuden kannalta ole välttämättä kriittistä; esimerkiksi kirjan [Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. & Rubin, D.B., 2013] mukaan jo $n_{eff} \geq 10$ on usein riittävä takaamaan MCMC-ketjun stabiilisuuden. Huomataan myös, että Geweken konvergenssin perusteella kaikki ketjut näyttäisivät supenneen. Lisäksi todetaan, että haasteellisimmat parametrit (β_0 tartunta- ja vesivaurioille) eivät ole tulkinallisesti kiinnostavia. Kaikkien parametrien otospolut ja posteriorijakaumien estimaatit on esitetty liitteessä C.

Taulukossa 4 esitetyt estimaatit ovat pääasiassa järkeenkäyviä ja odotettuja. Kalliovaurioiden korjaukset laskevat kallion vaurioituneisuutta ja muiden vaurioiden korjauksilla ei käytännössä ole vaikutusta, vaan kerroin on lähes nolla. Vastaavasti tartuntavaurioiden korjaukset laskevat betonin vaurioituneisuutta ja muiden vauriotyyppien korjauksella ei ole vaikutusta betonoidun alueen kuntoon. Vesivaurioiden osalta estimaatti $\hat{\beta}_{\text{vesi k.}} = 0.803$ ei ole uskottava. Toisaalta aineiston vesivaurioista vain kaksi on korjattu, joten tarkka estimointi ei siksi olekaan luotettavaa. Tämä näkyy myös tuon parametrin estimaatin leveässä posteriorivälissä.

Selittämättömän vaihtelun parametrit ν_j^2 , ovat hyvin pieniä. Suurin vaihtelu jää tartuntavaurioihin. Parametri $\rho_s \in [0, 1]$ kuvaa mallinnetun spatiaalisen riippuvuuden voimakkuutta ja tulosten perusteella riippuvuus on hyvin voimakasta. Parametri ρ_T on autoregressiivinen parametri, joka kuvaa ajallista riippuvuutta. Kun $\rho_T > 1$, seuraavan ajanhetken kunto on odotusarvoisesti heikompi kuin nykyinen. Odotetusti parametrin estimaatti on kuitenkin hyvin lähellä lukua 1 eli ajallinen muutos on hidasta.

Kuviossa 11 on esitetty kallio-, tartunta- ja vesivaurioiden ennustetut muutokset. Kunkin alueen ennustetusta vasteen arvosta (posterioriennustejakauman mediaani) vuonna 2021 on vähennetty havaittu vasteen arvo vuonna 2020. Kuvioista nähdään esimerkiksi, että ennustettu kalliovaurioiden kehitys on hyvin maltit-

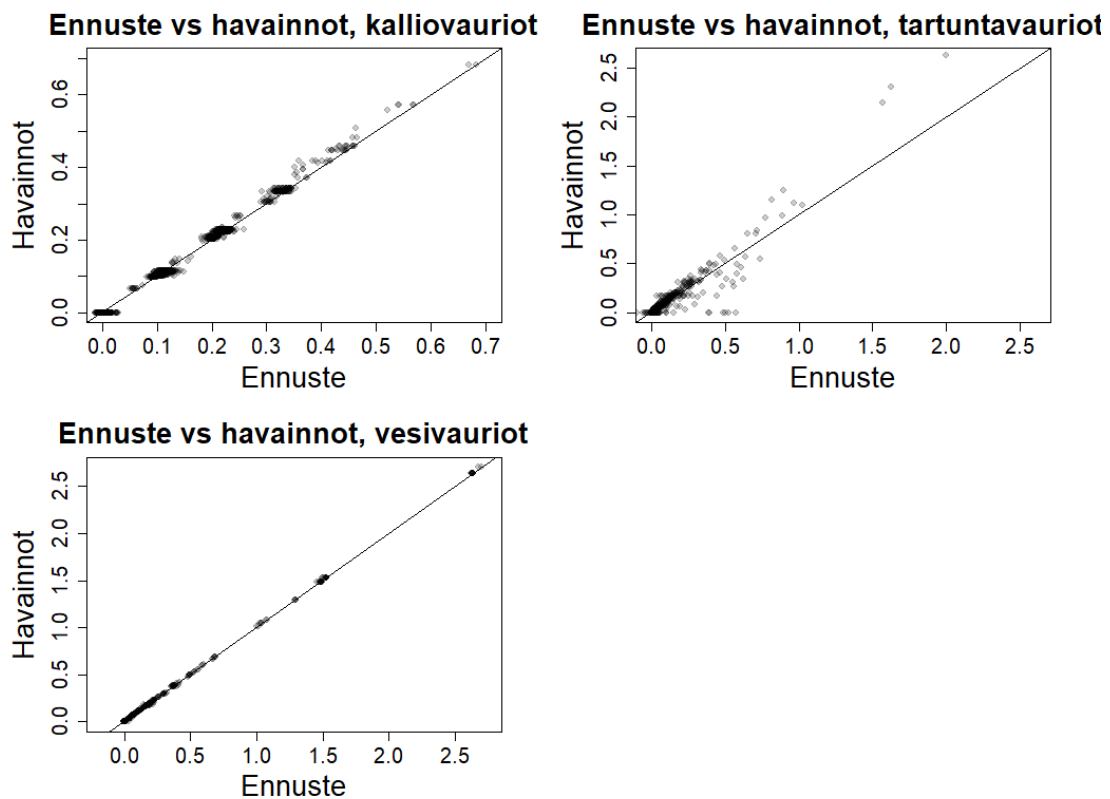
lista. Lisäksi, jos kuvion 11 tulosta verrataan kuviossa 5 esitettyyn nykytilaan, niin ennusteen mukaan olemassa olevat vesivauriot vaikuttaisivat leviävän herkästi viereisiin alueisiin. Tämä tulos on linjassa myös luvussa 2.6 esitetyn tuloksen kanssa, jonka mukaan spatiaalinen riippuvuus on voimakkainta vesivaurioilla.



Kuvio 11: Ennustettu muutos aikavälillä 2020–2021 vauriotyypeittäin.

4.1 Mallin validointi

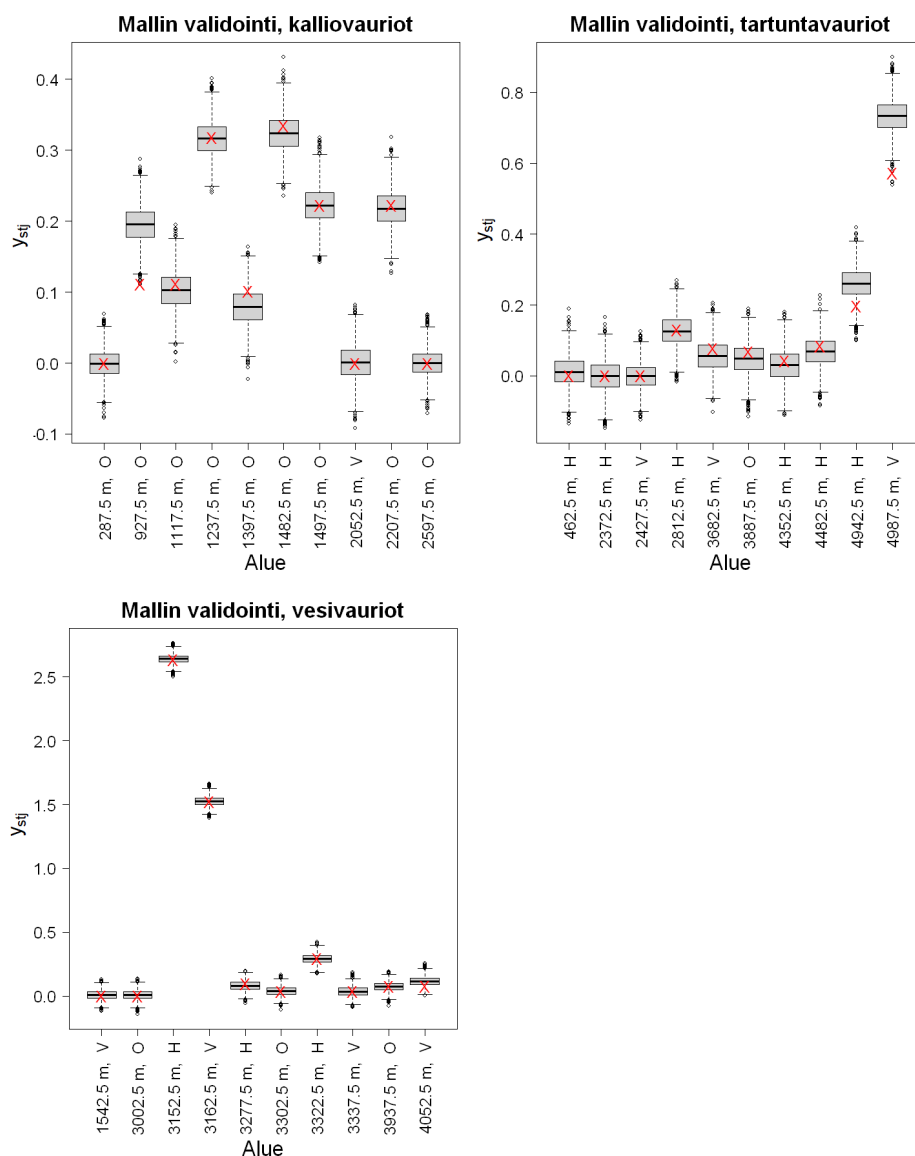
Tarkastellaan tässä luvussa mallin sopivuutta aineistoon. Kuviossa 12 on esitetty todelliset vasteen arvot ennusteita vastaan vauriotyypeittäin eriteltynä. Havaitaan, että kallio- ja vesivaurioilla pisteet asettuvat referenssisuoralle melko hyvin. Tartuntavaurioiden osalta huomataan, että vaihtelu on suurempaa ja pahasti vaurioituneiden alueiden estimaatit jäävät hieman alhaisiksi. Näitä pahasti vaurioituneita alueita onkin aineiston kokoon nähden hyvin vähän.



Kuvio 12: Ennustetut vasteen arvot havaittuja vasteen arvoja vastaan vauriotyypeittäin.

Mallin suunniteltu käyttö edellyttää mallilta puuttuvan aineiston käsittelyn toimivuutta. Tätä on tarkasteltu seuraavasti: Luodaan aineisto, josta poistetaan vuoden 2020 havaituista vasteen arvoista kunkin vauriotyyppin osalta 10 havaintoa. Alueet, joihin poisto kohdistuu, on arvottu siten, että kolme poistettavaa vasteen arvoa on ehjästä alueesta ja lopuissa alueista on havaittu vähintään yksi

vaurio. Alueet voivat olla sijainniltaan poikkeavat eri vauriotyypeille. Muutoin ai-
neisto pidetään ennallaan ja mallin sovitus on toistettu tässä työssä esitellyllä ja
kuviossa 9 havainnollistetulla naapurustolla. Tämän validoinnin tulokset on esi-
tetty kuviossa 13. Tulokset ovat vastaavat kuin kuviossa 12; kallio ja vesivaurioi-
den estimaatit ovat pääasiassa hyvin lähellä todellisia arvoja, mutta pahimpien
tartuntavaurioiden estimaatit jäävät alhaisiksi.



Kuvio 13: Kunkin validointiaineistoon päätyneen alueen vaurioituneisuuden posterioriennustejakauman viiksilaatikko, sekä todellinen vasteen arvo kyseisessä alu-
eessa (punainen X). Alueen sijainti ilmoitettu paalumetreinä ja kirjaintunnuskel-
la: O=Oikea seinä, H=Holvi ja V=Vasen seinä.

5 Pohdinta

Tässä työssä esitettiin luvun 2 vaurioaineiston pohjalta malli tunnelin vaurioiden kehittymiselle. Tavoitteena oli kyetä ennustamaan vaurioiden kehittymistä ja tunnistaa tunnelista alueita, joihin kunnossapidollisia toimia tulisi painottaa ja vastaavasti alueita, jotka eivät ole lähitulevaisuudessa kunnossapidon tarpeessa. Oletuksena on, että ennakoiva kunnossapito vähentää kustannusten lisäksi myös korjaustöiden kestoa, ja näin toimintaa tunnelissa ei tarvitse keskeyttää tarpeettoman pitkiksi ajoiksi.

Lähestymistavaksi valikoitui hierarkkinen Bayes-malli, joka mahdollistaa spatiotemporaalisen korrelaation huomioimisen mallissa. Lisäksi esitetty malli mahdollistaa useiden erityyppisten vaurioiden kehityksen mallintamisen samanaikaisesti. Tällöin myös vauriotyyppien väliset korrelaatiot on mahdollista huomioida ja ennusteet voidaan esittää vauriotyyppikohtaisesti. Lähestymistavan erityisenä vahvuutena voidaan pitää sitä, ettei aineiston tarvitse jatkossakaan olla täysin ehjää, vaan puuttuvuus tulee huomioiduksi oikein, kun tarkastuspäätös perustuu ainoastaan tehtyihin havaintoihin ja malliin. Tällöin tässä työssä esitettyä lähestymistapaa voidaan käyttää jatkossakin, vaikka mallin avulla tunnistettuja ehjiä alueita ei yhtä tiheästi tarkastettaisi kuin muuta tunnelia.

Aineistoa ei ole vielä kerätty kovin pitkältä ajalta, joten tässä työssä luvussa 3.1 esitetty malli huomioi riippuvuuden vain edellisestä ajanhetkestä AR(1)-rakenteisella mallilla, mutta tulevaisuudessa voisi olla mielekäästä harkita vakavammin myös AR(2)-rakenteista mallia.

Kovariaatteina tässä työssä hyödynnettiin vain korjaustietoja. Tulevaisuudessa voisi olla hyödyllistä seurata erityisesti betonoinnin aikana kerättyjen tietojen yhteyttä vaurioiden kehittymiseen. Tiedetään esimerkiksi, että ruiskubetonoinnin reseptiikka on ollut erilaista eriosissa tunnelia ja ruiskubetonoinnin paksuudessakin on vaihtelua. Teoriassa näidenkin seikkojen huomioiminen kovariaatteina on mahdollista, kunhan tieto on esitettävissä aluekohtaisesti. Näiden selittäjien

huomioiminen voisi vähentää tartuntavaurioiden estimointiin liittyvää korkeampaa epävarmuutta.

Mallin ennusteille saadaan laskettua myös aluekohtainen hajonta. Tämänkin tiedon hyödyntämisen mahdollisuutta on syytä pohtia tulevassa päätöksenteossa. Erityisesti tarkastustiheyttä pohdittaessa voisi olla syytä painottaa alueita, joissa ennusteen epävarmuus on suurta.

Viitteet

- Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. & Rubin, D.B. Bayesian data analysis (3rd ed.). 2013.
- Geweke, J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. *Federal Reserve Bank of Minneapolis, Research Department*, 148, 1991.
- Johnson, D.R. & Creech, J.C. Ordinal measures in multiple indicator models: A simulation study of categorization error. 48:398–407, 1983.
- Kass, R., Carlin, B., Gelman, A. & Neal, R. Markov chain Monte Carlo in practice: a roundtable discussion. *The American Statistician*, 52(2):93–100, 1998.
- Lee, D., Rushworth, A. & Napier, G. Spatio-temporal areal unit modeling in R with conditional autoregressive priors using the CARBayesST package. *Journal of Statistical Software*, 84:1–39, 2018.
- Lehtonen J. Maanalainen asemakaavoitusmenettely ja 3d-kiinteistönmuodostus. *Metropolia Ammattikorkeakoulu, Insinööriyö*, 2011.
- Leroux B., Lei, X. & Breslow, N. Estimation of disease rates in small areas: A new mixed model for spatial dependence. *Statistical Models in Epidemiology, the Environment, and Clinical Trials*, pages 179–191, 2000.
- Moran, P. A. P. Notes on continuous stochastic phenomena. *Biometrika*, 37: 17–23, 1950.
- Paalumäki, T., Lappalainen P. & Hakapää A. Kaivos- ja louhintatekniikka. 2015.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017. URL <https://www.R-project.org/>.

- Robitzsch, A. Why ordinal variables can (almost) always be treated as continuous variables: Clarifying assumptions of robust continuous and ordinal factor analysis estimation methods. *Frontiers in Education*, 5:177, 2020.
- Sidén, P., Lindgren, F., Bolin, D. & Villani, M. Efficient covariance approximations for large sparse precision matrices. *Journal of Computational and Graphical Statistics*, 27, 2017.
- Spiegelhalter, D., Best, N., Carlin, B. & Van Der Linde, A. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B*, 64(4):583–639, 2002.
- Tanner, M. & Wong, W. The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82:528–540, 1987.
- van Ravenzwaaij, D., Cassey, P. & Brown, S.D. A simple introduction to Markov chain Monte–Carlo sampling. *Psychonomic Bulletin and Review*, 25:143–154, 2018.
- Watanabe, S. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11, 2010.

Liitteet

Liite A

Taulukko: Mallin posteriorijakaumien tunnusluvut, efektiiviset otoskoot ja Gewerken konvergenssin tunnusluvut, kun käytössä on 1-naapurusto. Regressiokertoimet, joiden 95 %:n posterioriväli ei sisällä lukua 0 on merkitty symbolilla *.

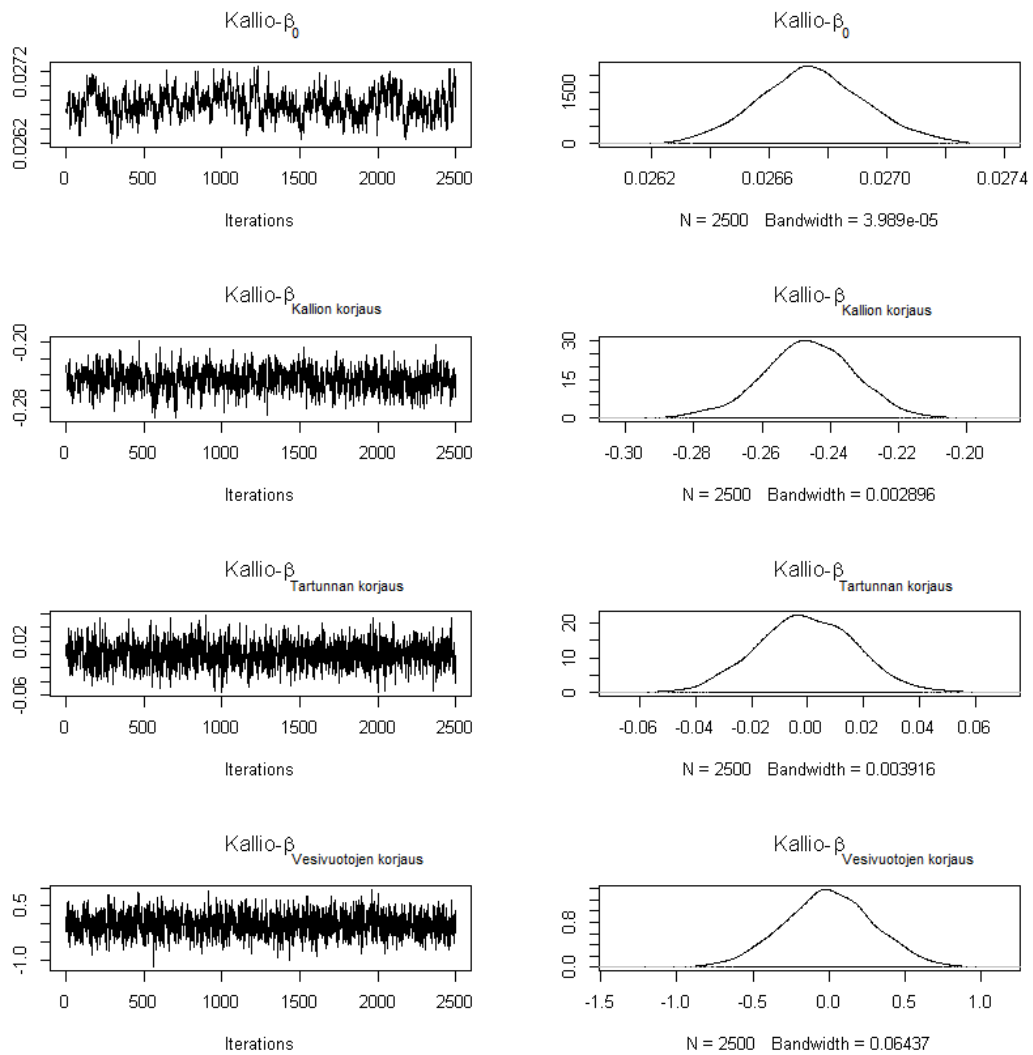
Parametri	Mediaani	2.5%	97.5%	n_{eff}	G
*Kallio - β_0	0.0264	0.0260	0.0267	151.8	-0.4
*Kallio - $\beta_{kallion\ k.}$	-0.2516	-0.2769	-0.2231	597.3	0.6
Kallio - $\beta_{tartunnan\ k.}$	0.0007	-0.0291	0.0310	1361.7	-1.5
Kallio - $\beta_{vesivuotojen\ k.}$	-0.0053	-0.5181	0.5091	1586.6	-1.5
*Tartunta - β_0	0.0106	0.0086	0.0129	28.6	-2.7
Tartunta - $\beta_{kallion\ k.}$	-0.0003	-0.0514	0.0510	539.3	0.4
*Tartunta - $\beta_{tartunnan\ k.}$	-0.2990	-0.3838	-0.2166	460.5	-2.9
Tartunta - $\beta_{vesivuotojen\ k.}$	-0.0021	-1.5070	1.4493	489.3	0.5
*Vesi - β_0	0.0251	0.0209	0.0293	19.8	0.1
Vesi - $\beta_{kallion\ k.}$	-0.0155	-0.0653	0.0348	239.7	0.3
Vesi - $\beta_{tartunnan\ k.}$	-0.0121	-0.0715	0.0509	493.0	0.9
Vesi - $\beta_{vesivuotojen\ k.}$	0.8102	-0.2350	1.7644	707.2	-0.1
ν_{kallio}^2	$5.4 \cdot 10^{-5}$	$4.9 \cdot 10^{-5}$	$5.9 \cdot 10^{-5}$	603.3	1.2
$\nu_{Tartunta}^2$	$1.3 \cdot 10^{-3}$	$1.3 \cdot 10^{-3}$	$1.4 \cdot 10^{-3}$	164.8	-4.4
ν_{Vesi}^2	$1.1 \cdot 10^{-4}$	$9.1 \cdot 10^{-5}$	$1.2 \cdot 10^{-4}$	472.0	-0.1
Σ_{11}	0.0017	0.0016	0.0017	1277.1	-1.9
Σ_{22}	0.0017	0.0016	0.0018	96.2	4.0
Σ_{33}	0.0043	0.0042	0.0045	286.9	0.9
ρ_S	0.9398	0.9329	0.9463	160.0	-2.1
ρ_T	1.0107	1.0039	1.0184	131.5	-2.0

Liite B

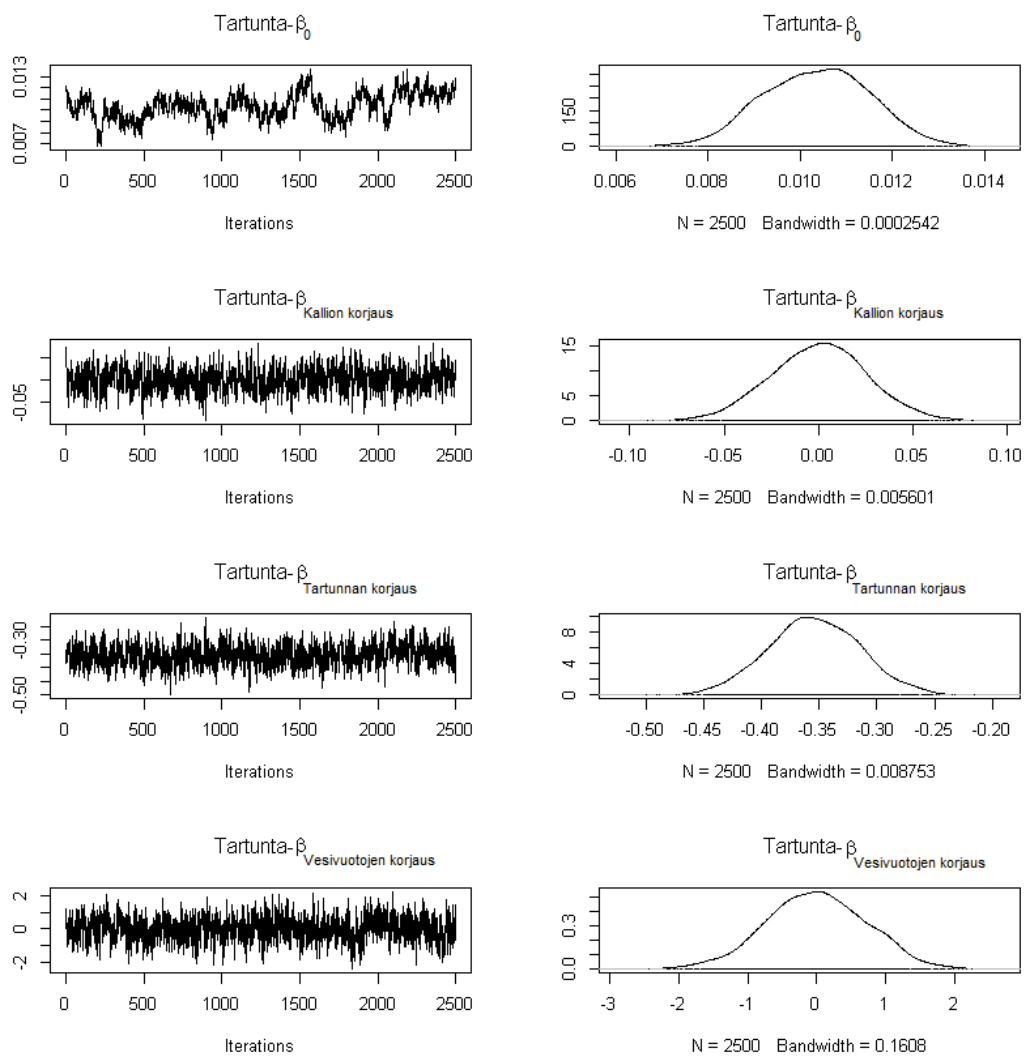
Taulukko: Mallin posteriorijakaumien tunnusluvut, efektiiviset otoskoot ja Gewe-
ken konvergenssin tunnusluvut, kun käytössä on $\sqrt{2}$ -naapurusto. Regressiokerto-
imet, joiden 95 %:n posterioriväli ei sisällä lukua 0 on merkitty symbolilla *.

Parametri	Mediaani	2.5%	97.5%	n_{eff}	G
*Kallio - β_0	0.0258	0.0255	0.0261	172.4	-0.5
*Kallio - $\beta_{kallion\ k.}$	-0.2609	-0.2902	-0.2328	486.5	-0.3
Kallio - $\beta_{tartunnan\ k.}$	-0.0004	-0.0300	0.0294	1374.6	0.2
Kallio - $\beta_{vesivuotojen\ k.}$	0.0058	-0.5096	0.5406	1675.8	-2.2
*Tartunta - β_0	0.0106	0.0090	0.0127	27.4	-0.4
Tartunta - $\beta_{kallion\ k.}$	0.0006	-0.0543	0.0574	439.5	1.1
*Tartunta - $\beta_{tartunnan\ k.}$	-0.2386	-0.3179	-0.1609	553.1	-0.1
Tartunta - $\beta_{vesivuotojen\ k.}$	0.0184	-1.5520	1.4757	485.2	0.7
*Vesi - β_0	0.0267	0.0215	0.0303	16.7	0.4
Vesi - $\beta_{kallion\ k.}$	-0.0093	-0.0660	0.0458	301.9	-0.7
Vesi - $\beta_{tartunnan\ k.}$	-0.0158	-0.0688	0.0413	588.7	-0.3
Vesi - $\beta_{vesivuotojen\ k.}$	0.8319	-0.2239	1.8875	550.6	-1.4
ν_{kallio}^2	$5.5 \cdot 10^{-5}$	$5.0 \cdot 10^{-5}$	$6.0 \cdot 10^{-5}$	536.4	-1.5
$\nu_{Tartunta}^2$	$1.6 \cdot 10^{-3}$	$1.5 \cdot 10^{-3}$	$1.6 \cdot 10^{-3}$	124.0	0.3
ν_{Vesi}^2	$1.1 \cdot 10^{-4}$	$9.7 \cdot 10^{-5}$	$1.3 \cdot 10^{-4}$	441.1	0.6
Σ_{11}	0.0032	0.0032	0.0033	685.7	-2.7
Σ_{22}	0.0029	0.0028	0.0031	90.0	0.8
Σ_{33}	0.0091	0.0087	0.0094	256.2	-0.5
ρ_S	0.8843	0.8703	0.8970	95.4	-4.6
ρ_T	1.0182	1.0114	1.0258	120.3	-1.5

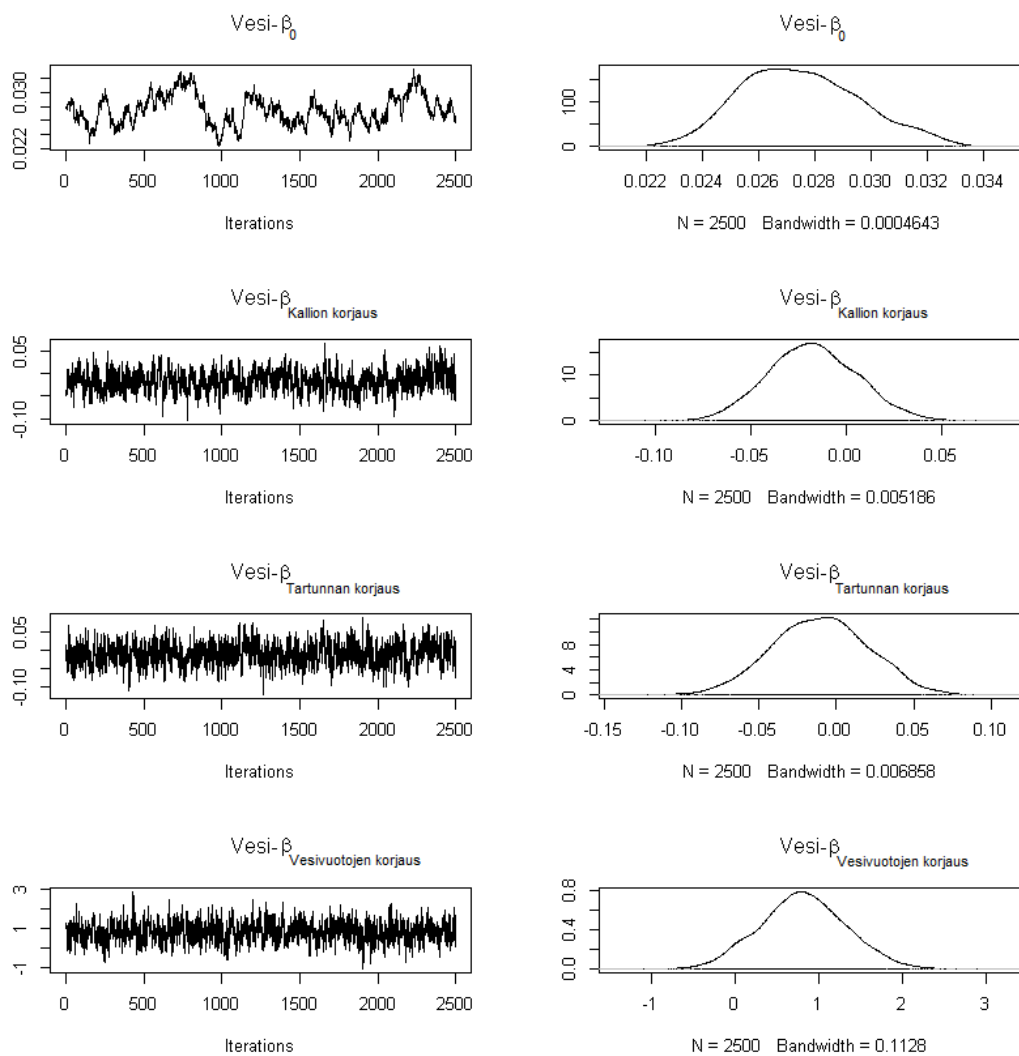
Liite C



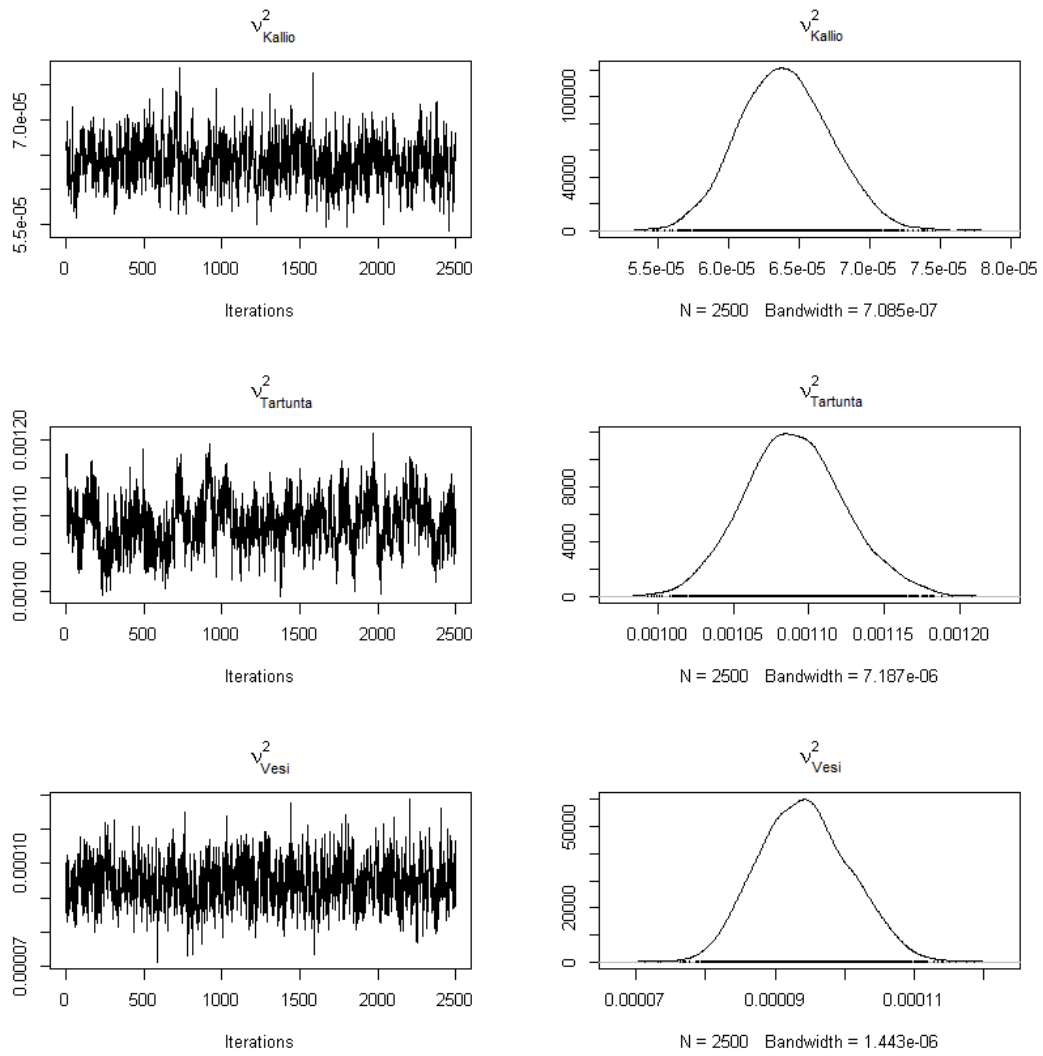
Kuva: Otospolut ja posteriorijakaumien estimaatti parametreille β_j , kun vasteena on kallioauriot.



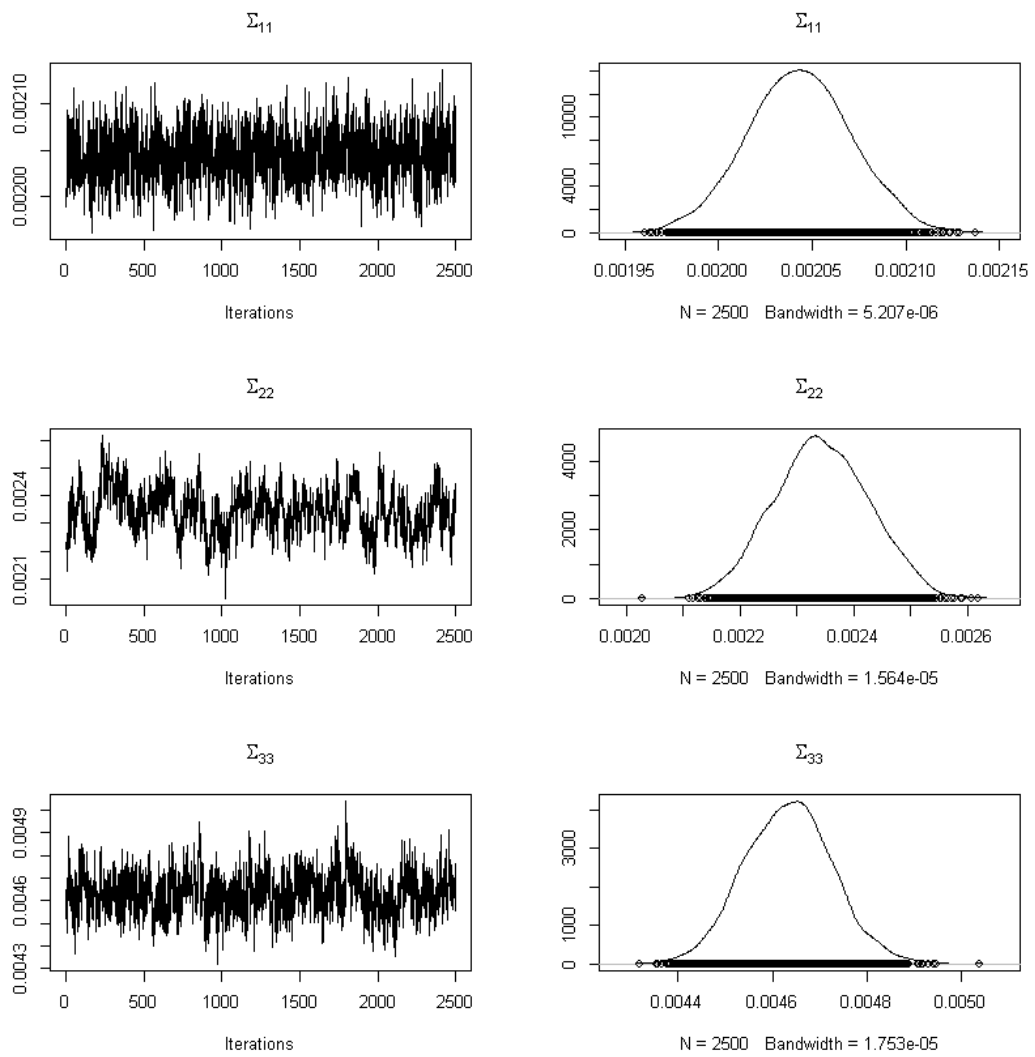
Kuva: Otopolut ja posteriorijakaumien estimaatti parametreille β_j , kun vasteena on tartuntavauriot.



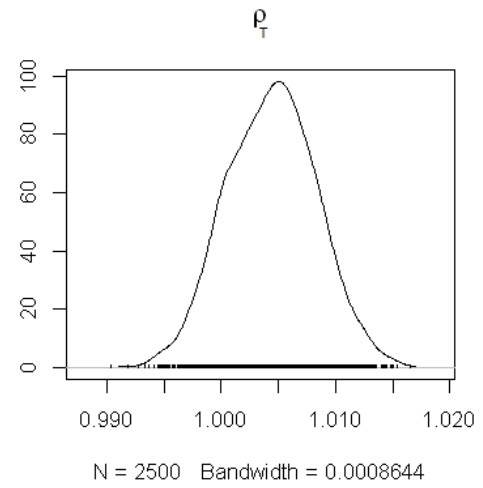
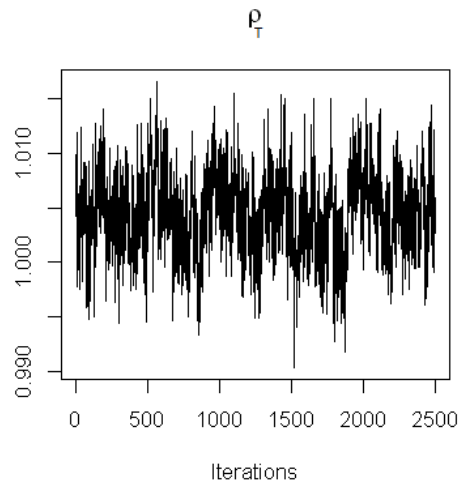
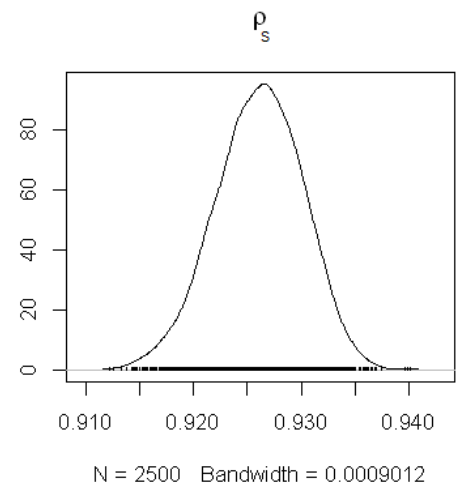
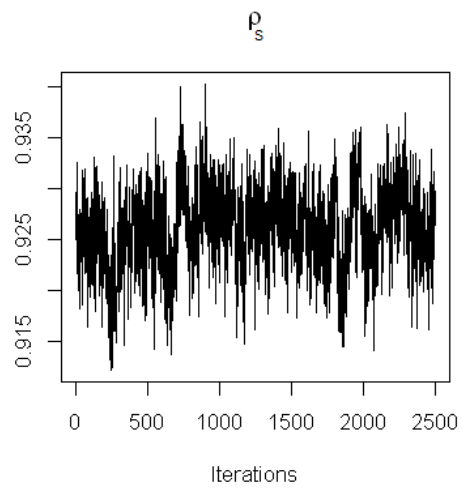
Kuva: Otospolut ja posteriorijakaumien estimaatti parametreille β_j , kun vasteena on vesivauriot.



Kuva: Otopolut ja posteriorijakaumien estimaatti parametreille ν_j .



Kuva: Otospolut ja posteriorijakaumien estimaatti parametreille Σ_{11} , Σ_{22} ja Σ_{33}



Kuva: Otospolut ja posteriorijakaumien estimaatti parametreille ρ_S ja ρ_T .