

Tautikartoitus CAR- ja partitiomalleilla

Esimerkkeinä sydän- ja verisuonitaudit sekä keuhkosityöpä

Aki Havulinna^{1,2}

¹Lääketieteellisen tekniikan ja laskennallisen tieteen laitos, TKK

²Terveyden ja hyvinvoinnin laitos (THL)

Terveydenhuollon uudet analyysimenetelmät (TERANA)
Loppuseminaari 1.4.2009



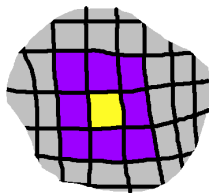
HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Biomedical Engineering and Computational Science



TERVEYDEN JA
HYVINVOINNIN LAITOS

Tautikartoitus CAR-mallilla

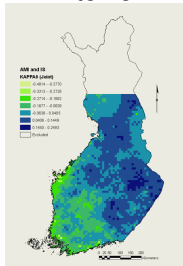
- CAR (conditional autoregression) on perinteisesti käytetty tasoitusmalli tautikartoituksessa (Besag, York & Mollié 1991).
- CAR-priorioletus: kunkin alueen lähimmät naapurit antavat lisätietoa alueen tautiriskistä
- Bayes-mallissa priorin painoarvo, ts. tasoituksen määrä, estimoidaan datasta.



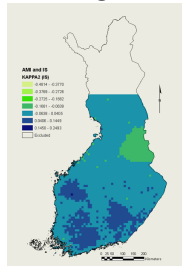
CAR – Esimerkki 1: aivohalvaukset ja sydänkohtaukset 1991-2003

Iskeemisten aivohalvauksien (IS) ja akuuttien sydäninfarktien (AMI) ilmaantuvuuden yhteinen ja tautispesifi komponentti

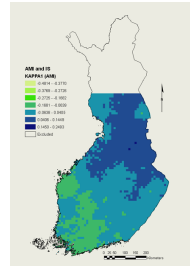
Yhteinen



IS



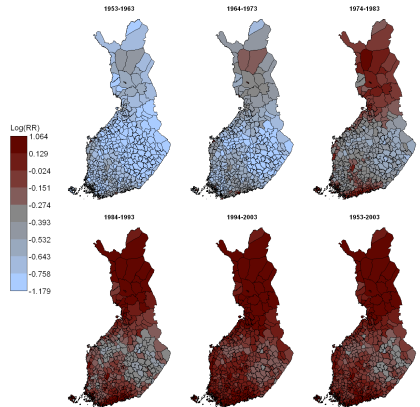
AMI



Viite: Havulinna, Pääkkönen, Salomaa, Karvonen (2008)

CAR – Esimerkki 2: Naisten keuhkosyöpäkuolleisuus 1953-2003

- Mallissa huomioidaan spatiaalinen, temporaalinen ja spatiotemporaalinen komponentti
- Alinna oikealla spatiaalinen komponentti, muissa kartoissa kaikki komponentit yhdessä
- Malli: Knorr-Held (2000)



Kuntarajat © Affecto Finland Oy, Karttakeskus, Helsinki

CAR – edut ja ongelmat

Etuja

- Tunnettu perinteinen malli, toteutettu myös WinBUGSissa
- Melko nopea mallinnus, skaalautuu hyvin

Ongelmia

- Joustamattomuus tilanteissa, joissa tautiriskin alueellisen vaihtelun suuruus ei ole vakio eri puolilla tutkimusaluetta

Tässä projektissa

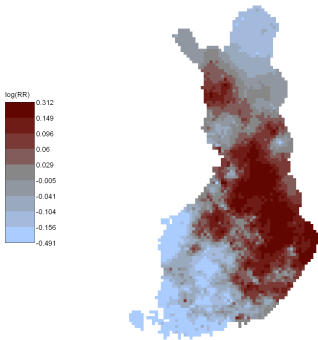
- Tutkittiin CAR-mallin laajennuksia
- Kehitettiin nopea C++ -pohjainen algoritmi mallien estimointiin ⇒ tarvetta etenkin spatiotemporaalimallinnuksessa

Partitiomallit ja Reversible jump MCMC

- Partitiomalleissa tutkimusalue ositetaan klustereihin, joiden sisällä tautiriski on vakio
- Jako perustuu klusterikeskuksiin; kuhunkin klusteriin kuuluvat ne alueet, joita lähinnä ko. klusterikeskus on (Voronoi-tesselaatio)
- Reversible jump MCMC (Green, 1995) mahdollistaa otannan, jossa klusterikeskusten paikkaa ja määrää voidaan vaihdella \Rightarrow mallien keskiarvoistaminen
- Kaksi tunnettua mallia: Knorr-Held & Raßer 2000 (Log-normal–Poisson), Denison & Holmes 2001 (Gamma–Poisson)

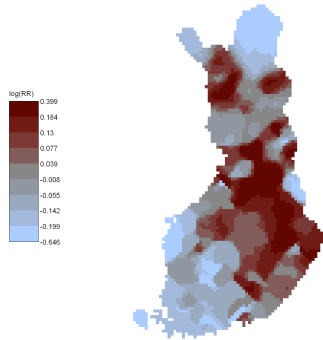
Partiomallit – Esimerkki: Akuutit sydänkohtaukset 1991-2003

CAR –malli



15 000 iter., 2 min. DIC 15 695

Knorr-Held & Raßer –malli



5 000 000 iter., 20 min., DIC 14 943

Partitiomallit – edut ja ongelmat

Etuja

- Hyvin joustava alueellisten erojen suhteen
- Laskentanopeus verrattain kohtuullinen, skaalautuu hyvin

Ongelmia

- Käyttökelpoinen laskentanopeus vaatii hyvin optimoidun C++ -algoritmin
- Etäisyysmatriisin etukäteen laskemiseen kuluu 3 tuntia C++:lla

Tässä projektissa

- Verrattiin Knorr-Held & Raßer –mallia perinteiseen CAR –malliin
- Kehitettiin nopea C++ -pohjainen algoritmi mallin estimointiin \Rightarrow Projektin alussa toteutetulla algoritmilla (R:llä) 5 000 000 iteraation läpikäyminen olisi vaatinut viikon laskenta-aikaa

Yhteenveto

- Tämän projektin aikana on kehitetty nopeita laskenta-algoritmeja sekä verrattu CAR-mallien ja partitiomallien keskenäistä paremmuutta
- Jatkokehityksenä:
 - Partitiomallin spatiotemporaaliversio on tekeillä
 - Algoritmien julkaiseminen yleiseen käyttöön
 - Tautien ilmaantuvuuden muutosten ennustaminen

Viitteet

Besag J, York J, Mollié, A (1991). Bayesian image restoration, with two applications in spatial statistics (with discussion). *Ann. Inst. Statist. Math.*, 43, 1-59

Denison DGT, Holmes CC (2001). Bayesian Partitioning for Estimating Disease Risk *Biometrics* 57(1): 143-149

Havulinna AS, Pääkkönen R, Karvonen M, Salomaa V (2008). Geographic patterns of incidence of ischemic stroke and acute myocardial infarction in Finland during 1991-2003. *Ann Epidemiol.* 18(3): 206-13

Green PJ (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* 82(4):711-732

Knorr-Held L (2000). Bayesian modelling of inseparable space-time variation in disease risk. *Statistics in Medicine*, 19, 2555-2567

Knorr-Held L, Raßer G (2000). Bayesian Detection of Clusters and Discontinuities in Disease Maps. *Biometrics*, 56(1): 13-2