Metsätehon raportti 252 LIITE
19.6.2019

UUSIEN TIETOLÄHTEIDEN SOVITUS METSÄTIEDON PALVELUALUSTAN HILARESOLUUTIOON

Metsätehon raportti 252: Metsätiedon palvelualustan konseptitestaus

Prof. Risto Ritala
DI. Teijo Juntunen
Tampereen yliopisto

ISSN 1796-2374 (Verkkojulkaisu)

METSÄTEHO OY
Vernissakatu 1
01300 Vantaa

www.metsateho.fi

UUSIEN TIETOLÄHTEIDEN SOVITUS METSÄTIEDON PALVELUALUSTAN HILARESOLUUTIOON

Metsätehon raportti 252: Metsätiedon palvelualustan konseptitestaus

Metsätehon raportti 252 LIITE
19.6.2019

ISSN 1796-2374 (Verkkojulkaisu)

© Metsäteho Oy

SISÄLLYS

ESIPUHE **3**

1 JOHDANTO **4**

2 HILARUUTU JA MITTAUSRESOLUUTIO **5**

3 KOORDINAATTIMUUNNOKSET JA KOORDINAATTIKONVERSIOT **6**

4 MITTAUSTA LÄHINNÄ OLEVAN HILASOLUN MÄÄRITTÄMINEN **7**

5 TIETOLÄHTEIDEN RESOLUUTIOIDEN SUHDE HILARESOLUUTIOON **8**

5.1 Mittausresoluutio huomattavasti tiheämpi kuin hilaresoluutio 8

5.1.1 Käyttötapaus 1: Puuston pienoiskopterimittaukset 9

5.1.2 Käyttötapaus 2: Maaperän kivisyysmittaukset 11

5.2 Mittausresoluutio lähes sama kuin hilaresoluutio 12

5.2.1 Käyttötapaus 3: Satelliittimittaus kasvillisuusindeksistä 12

5.3 Mittausresoluutio huomattavasti harvempi kuin hilaresoluutio 15

5.3.1 Käyttötapaus 4: Säätiedot, lämpötila 16

6 YHTEENVETO **19**

LÄHTEET **20**

ESIPUHE

Raportissa tarkastellaan eri resoluution tietolähteiden yhteensovittamisen mahdollisuuksia metsätiedon palvelualustaan. Metsätiedon palvelualusta on Metsätehon ja maa- ja metsätalousministeriön yhteishankkeessa kehitetty pilottivaiheen testausjärjestelmä heterogeenisten metsätietojen jalostamiseen ja välittämiseen sovelluksille, ja sitä on jatkossa ehdotettu tuotteistettavaksi Suomen metsäkeskuksen tarjoaman avoimen metsätiedon välityskanavaksi (tarkempi esittely lähteissä Hämäläinen ym. 2019a ja Hämäläinen ym. 2019b. Tampereen teknillinen yliopisto (nykyisin Tampereen yliopisto) toimi hankkeessa asiantuntijana datakyselyihin ja -fuusioon liittyvissä kysymyksissä.

Raportin ovat kirjoittaneet Prof. Risto Ritala ja DI. Teijo Juntunen Tampereen yliopistosta. Raportin on toimittanut Kirsi Riekki Metsätehosta.

1 JOHDANTO

Metsätiedon palvelualustan perusyksikköinä ovat 16 x 16 m hilaruudut ja perinteiset metsävarakuviot. Tässä työssä keskitytään hilaruutuihin, jotka peittävät koko Suomen pinta-alan sisältäen tietoja peittämänsä maa-alan ominaisuuksista. Työn tarkoituksena on antaa esimerkkitapauksia ja yleisohjeistusta hilaresoluutiosta poikkeavan mittausinformaation sisällyttämiseksi hilaruutuihin siten, ettei mittausdatan informaatiosisältö merkittävästi muuttuisi.

Ohjeessa esitellään käsitteet mittausresoluutio, hilaresoluutio, mittauspeitto ja mittauksen vaikutussäde. Tämän jälkeen eritellään kolme mittausresoluutioltaan poikkeavaa tapausta: mittausresoluutio on tiheämpi kuin hilaresoluutio, mittausresoluutio on lähes sama kuin hilaresoluutio sekä kun mittausresoluutio on huomattavasti harvempi kuin hilaresoluutio.

Suunniteltujen datan sovitusmenetelmien toimivuutta testattiin Metsäteho Oy:ltä, VTT:ltä, TTY:ltä ja Ilmatieteen laitokselta saatua mittausdataa. Käyttötapauksina olivat puuston pienoiskopterimittaukset, maaperän kivisyysmittaukset, kasvillisuusindeksin satelliittimittaukset ja ilmatieteen laitoksen avoimen säädatan lämpötilamittaukset.

2 HILARUUTU JA MITTAUSRESOLUUTIO

Suomen metsäkeskuksen metsävaratiedossa käytettävä hilaruutu on määritelty 16 x 16 m kokoiseksi alueeksi, jonka lounaiskulman koordinaatti on määritetty ETRS-TM35FIN-tasokoordinaatistossa (Holopainen ym. 2013). Tällaisten vierekkäisten hilojen muodostama ruudukko kattaa koko Suomen pinta-alan ja jokaiseen hilaruutuun pyritään tallentamaan kyseiselle alalle osuvaa metsävaratietoa. Hilamuotoista dataa voidaan edelleen jakaa ja käsitellä metsätiedon palvelualustan kautta.

Hilaresoluutio on näin ollen 16 m, kun taas metsävaratietoihin kohdistuvien mittausten spatiaaliset resoluutiot voivat olla merkittävästi hilaresoluutiota pienempiä tai suurempia. Tässä dokumentissa määritellään käyttötapausten avulla menetelmiä mittausten sovittamiseksi valittuun hilaresoluutioon.

Jokaisen tietolähteen kohdalla on aluksi arvioitava kullekin mittaukselle vaikutusala, jonka sisällä yhdessä pisteessä otettu mittausarvo pysyy kuvaavana. Esimerkiksi yhden puun paksuuden mittaus pysyy kuvaavana vain kyseisen puun halkaisijan määrittämällä alalla, mutta lämpötilamittauksen voidaan katsoa pysyvän kuvaavana useiden kilometrien matkan halutusta tarkkuudesta riippuen.

3 KOORDINAATTIMUUNNOKSET JA KOORDINAATTIKONVERSIOT

Mittaukseen liittyvä sijaintitieto ja käytetty koordinaatisto tulee tuntea, jotta mittaus saadaan kohdistettua oikeaan hilaruutuun. Hilakoordinaatistoksi on valittu ETRS-TM35FIN-tasokoordinaatisto, jota on käytetty vuodesta 2005 suomalaisten maakarttojen tasokoordinaatistona. ETRS-TM35FIN-tasokoordinaatiston geodeettinen datum on EUREF-FIN ja karttaprojektio on poikittainen Mercatorin lieriöprojektio.

Koordinaattikonversiot ovat saman datumin (esim. Kartastokoordinaattijärjestelmä, EUREF-FIN, N60) sisäisiä matemaattisia muunnoksia. Niihin luetaan myös karttaprojektioiden väliset siirtymät ja tasokoordinaatistojen väliset muunnokset. Esimerkki koordinaatistokonversiosta on maantieteellisten EUREF-FIN-koordinaattien projisointi ETRS-TM35FIN-tasokoordinaatistoon.

Koordinaattimuunnoksessa puolestaan muunnetaan koordinaatteja kahden eri datumiin perustuvan koordinaatiston välillä. Koordinaattimuunnos suoritetaan molemmissa koordinaatistoissa mitattujen yhteisten pisteiden avulla. Esimerkki koordinaattimuunnoksesta on, kun KKJ-tasokoordinaatistosta muunnetaan koordinaatteja ETRS-TM35FIN-tasokoordinaatistoon.

Koska tasokoordinaatistot ovat aina vääristyneitä yrittäessään kuvata pallomaista pintaa, tasomuunnoksia saa suorittaa vain saman tyypin tasokoordinaatistojen välillä. Esimerkiksi Suomessa käytettävät tasokoordinaatistot ovat yleisesti poikittaisia lieriöprojektioita. Tasomuunnos voidaan tällöin suorittaa, jos lieriöprojektioilla on yhtenevä keskimeridiaani. (Häkli ym. 2009)

Koordinaattimuunnokset ja -konversiot voidaan laskea Julkisen hallinnon suosituksen 154 perusteella (JHS-Suositukset 2008), jotka perustuvat korjattuina lähteeseen Hirvonen (1972). Geodeettisella laitoksella on myös muunnospalvelu osoitteessa http://coordtrans.fgi.fi/.

4 MITTAUSTA LÄHINNÄ OLEVAN HILASOLUN MÄÄRITTÄMINEN

Määritettäessä kunkin yksittäisen mittauksen lähintä hilasolua, vastaus on yleensä aina yksiselitteinen. Ainoastaan mittauksen osuessa tarkalleen kahden tai neljän hilan välille tai neljän hilan yhteiseen pisteeseen päätetään, että ratkaisuksi valitaan jokin yhtä läheisistä hilasoluista. Mittausten koordinaattitiedot ilmoitetaan yleensä sellaisella tarkkuudella, ettei valinnalla ole käytännössä merkitystä: Hilojen rajat ovat tasametrein ja ETRS-TM35FIN-koordinaatit on usein ilmoitettu millimetrin tarkkuudella. Näin ollen vain yksi tuhannesta hilarajan läheisyyteen osuvasta mittauksesta saa tarkalleen koordinaattiarvon hilarajalta. Neljän hilan yhteispisteen läheisyyden mittauksista, joita on jo lähtökohtaisesti huomattavasti vähemmän kuin rajamittauksia, vain yksi mittaus miljoonasta saa tarkalleen arvon yhteispisteestä.

Rajatapauksissa olisi suositeltavaa käyttää lähimmän hilasolun valinnassa sitä hilaa, missä on eniten suoritettuja mittauksia. Näin ollen suositaan mittausalueen keskushiloja reunahilojen kustannuksella. Tällä vältetään tilanteet, joissa yksittäinen reunamittaus osoitettaisiin muutoin tyhjään hilaan, vaikka toisessa yhtä mahdollisessa hilassa olisi puolestaan runsaasti suoritettuja mittauksia. Tilanteen harvinaisuuden vuoksi voi kuitenkin olla riittävää suosia aina esimerkiksi koordinaatiston positiivisien akselisuuntien hilasoluja, mikä mahdollistaa tavanomaisen pyöristysfunktion käyttämisen lähimmän hilasolun etsimiseksi.

5 TIETOLÄHTEIDEN RESOLUUTIOIDEN SUHDE HILARESOLUUTIOON

Tietolähteiden resoluutiot määräytyvät käytetyn mittausmenetelmän tai mittausdatan tallennusmenetelmän mukaan. Tietolähteen mittausresoluutiolla tarkoitetaan tässä keskimääristä mittauspisteiden lukumäärää pinta-alayksikköä kohti: suuri resoluutio tarkoittaa siis tiheää mittausta. Sovitusohjeistuksen säännönmukaistamiseksi jaetaan tietolähteet kolmeen eri pääryhmään mittausresoluution ja hilaresoluution suhteen perusteella:

- Mittausresoluutio >> hilaresoluutio
- Mittausresoluutio ≈ hilaresoluutio
- Mittausresoluutio << hilaresoluutio

Seuraavaksi esitetään käytännön esimerkkien kautta sopiviksi katsotut menetelmät mitatun datan muuntamiseksi hilaresoluutioon.

5.1 Mittausresoluutio huomattavasti tiheämpi kuin hilaresoluutio

Kun jokaiseen hilaan voidaan kohdistaa useita mittauksia, on tehtävänä yhden kuvaavan arvon löytäminen koko hilaruudun mittausjoukolle. Tällöin on syytä tutkia mittausten spatiaalista jakaumaa hilaruudun sisällä. Tämä jakaa tapaukset karkeasti kahteen ryhmään: tasaiseen mittauspeittoon ja epätasaisesti hilaruudun sisällä jakautuneeseen mittauspeittoon. Yleisesti pyritään tasaiseen mittauspeittoon, mutta mitatun alueen reunoilla osa hilaruudusta jää yleensä mittauksetta. Tämän lisäksi esimerkiksi maaston muodot ja heikko kuljettavuus voivat aiheuttaa epätasaisuutta mittauspeittoon.

Tasaisesti hilasoluihin jakautuneille mittausarvoille on yleensä syytä määrittää todennäköisyysjakauma. Jos on syytä olettaa, että mittausten jakauma on esim. keskeisen raja-arvolauseen mukaisesti normaalijakautunut, voidaan mitatun hilan attribuutteina tuottaa mittauksista estimoidut normaalijakauman parametrit: odotusarvo ja varianssi. Muussa tapauksessa voidaan käyttää ilmaisuvoimaisempaa Weibull-jakaumaa (Weibull 1951), jonka parametrit estimoidaan hilan mittauksista ja annetaan hilan attribuutteina palvelualustalla.

Kaksiparametrisen Weibull-jakauman tiheysfunktio on muotoa:

$$f(x|\alpha,\beta) = \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^{\beta}}, \alpha > 0 \; ja \; \beta > 0, \qquad (1)$$

missä α on skaalausparametri ja β muotoparametri. Muotoparametrin avulla Weibull-jakauma voi redusoitua tasajakaumaksi, Rayleigh-jakaumaksi, eksponenttijakaumaksi, normaalijakaumaksi tai näiden välimuodoiksi. Tämän vuoksi se sopii standardijakaumatyypiksi palvelualustan monimuotoisille tietolähteille. Yhtälö 1 tuottaa Weibull-jakauman mukaisen todennäköisyystiheyden, kun mittausarvot ovat ≥ 0. Hilan mittausarvoista laskettava suurimman uskottavuuden estimaatti Weibull-jakauman parametreille löytyy esimerkiksi lähteestä Scholz (2008). Käytetty jakaumatyyppi tulee kertoa tietolähteen dokumentaatiossa.

Kun mittauspeitto on epätasainen, täytyy toimintamalli päättää tietolähdekohtaisesti. On kuitenkin yleensä hyvä ilmaista, kuinka suuri osuus hilan pinta-alasta kuuluu mittausalueeseen. Tämän

mittauspeiton määrittäminen vaatii tiedon mittauksen keskipisteen koordinaateista ja arvion yksittäisen mittauksen vaikutussäteestä.

Yksittäisen mittauksen mittauspeitto M_i kun $i = 1 \dots n$, saadaan laskemalla ympyrän ala mittauksen valitusta vaikutussäteestä r_m. Vaikutussäteen suuruus täytyy päättää mittausmenetelmän perusteella. Yksi määritystapa on esitetty käyttötapauksessa 1 (luku 5.1.1). Hilan sisältämien yksittäisten mittausten unionista lasketaan leikkaus hilan pinta-alan kanssa yhtälön 2 mukaisesti:

$$P_H = A_H \cap (M_1 \cup M_2 \cup \dots M_{n-1} \cup M_n), \quad (2)$$

missä A_H on hilan ala ja P_H on mittausten vaikutusympyröiden yhteispeitto hilan alueella. Tästä voidaan edelleen laskea yhtälön 3 mukaan Q, eli mitatun alueen osuus koko hilan alasta

$$Q = \frac{P_H}{A_H} \quad (3)$$

Jos hilaruudun mittausosuus ei ole 100 %, mittausosuus on syytä ilmaista palvelualustalla hilaruudun attribuuttina, jotta vältetään näiden hilojen käsittely samanarvoisina täydellisesti mitattujen hilojen kanssa esim. runkolukulaskennassa.

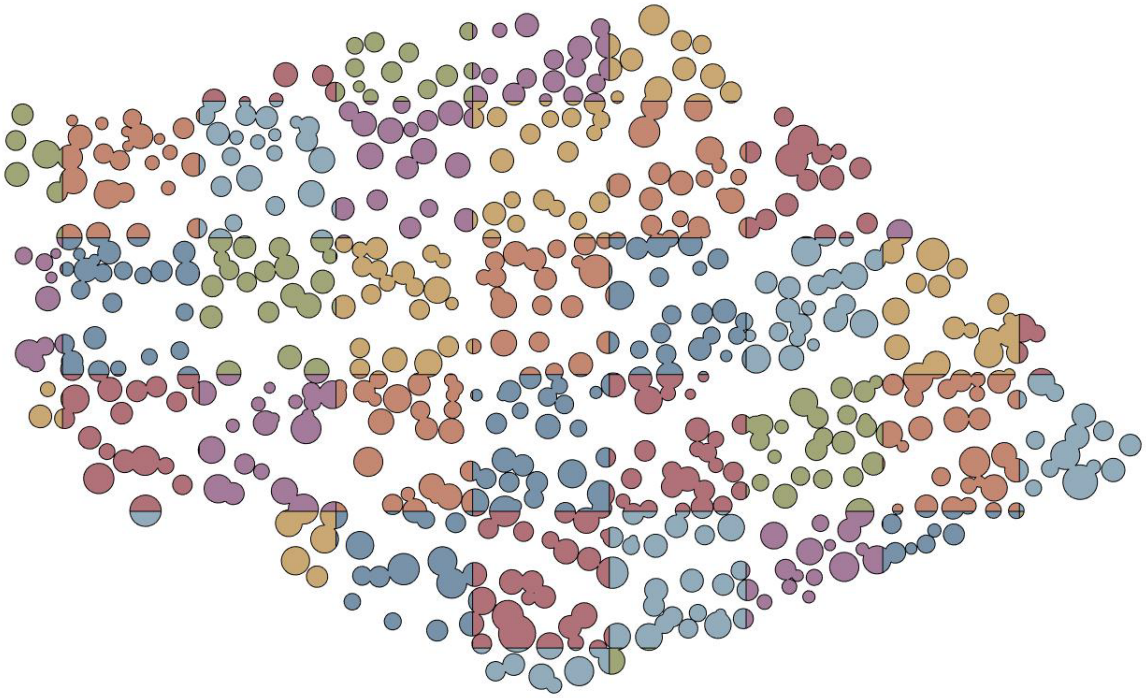
5.1.1 Käyttötapaus 1: Puuston pienoiskopterimittaukset

Mika Vahtila suoritti puuston pienoiskopterimittauksia Lieksassa Pankakoskella ja Sokovaaralla opinnäytetyöhönsä kesällä 2018. Kopterilla kuvattiin mittausalueen puusto, minkä jälkeen Maanmittauslaitoksen Paikkatietokeskus (FGI) paikansi puut kuvien avulla. Paikkatietoihin yhdistettiin Vahtilan mittaamat läpimitat ja osaan puustosta myös hypsometrillä mitatut pituudet sekä latvusrajat. Kuvasta 1 nähdään, että mittausresoluutio on ollut selkeästi hilaresoluutiota tiheämpi. Pankakoskella, missä puiden keskihalkaisija oli mittausalueella hieman Sokovaaraa suurempi, mittauksia suoritettiin täysin mitattua hilaa kohti keskimäärin n. 20 kappaletta.

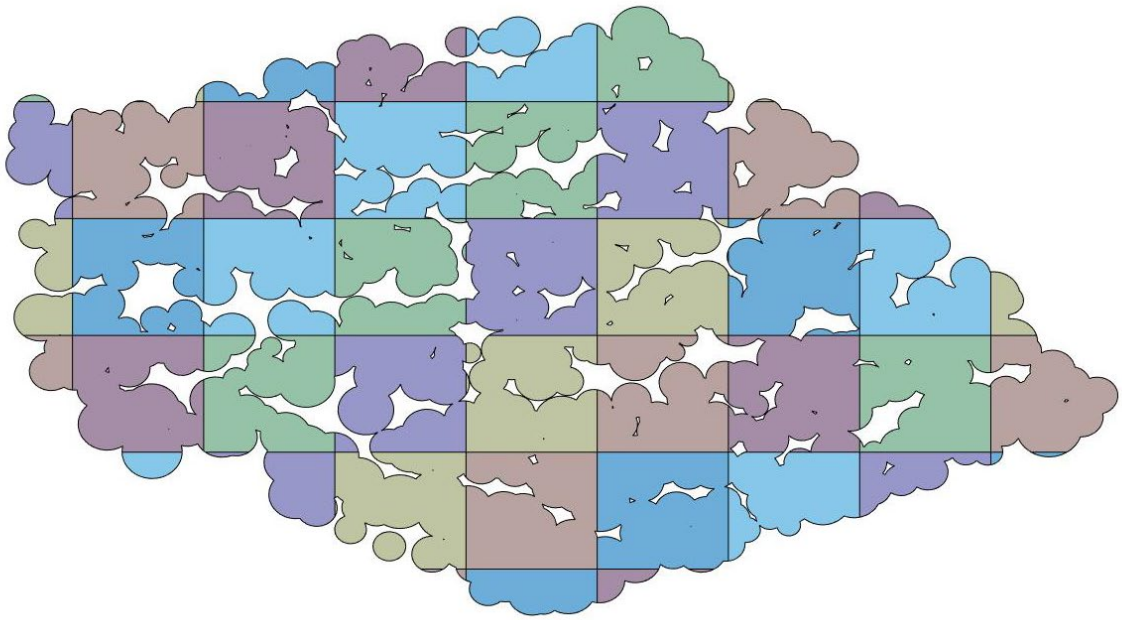
Mittausten vaikutussäde määritettiin mitatun keskihalkaisijan skaalausluvulla. Kun oletetaan, että keskihilojen mitattu osuus on 100 %, määritetään pienin mahdollinen skaalausluku siten, että skaalattujen mittausten alueiden unionin leikkaus mitatusta hilasta on vähintään 95 %. Keskihiloiksi määritetään ne hilat, joiden jokaisessa neljässä pääilmansuunnan naapurihilaruudussa on vähintään yksi mittaus.

Skaalauslukua kasvatetaan niin kauan, kunnes keskihilojen mittauspeitto Q on keskiarvoisesti yli 95% kunkin keskihilan pinta-alasta. Kuvista 1–3 nähdään, kuinka Pankakosken mittausten skaalausluku määräytyi ja kuinka skaalausluvun kasvatus vaikuttaa mittauspeittoon. Kuvassa 1 käytettiin skaalauslukua 15, jolloin mittauspeittojen keskiarvo keskihiloissa oli n. 36 %. Vastaavasti kuvassa 2 skaalausluvulla 30 mittauspeittojen keskiarvo keskihiloissa oli 87 %. Lopulta skaalausluvulla 50 mittauspeittojen keskiarvoksi saatiin keskihilojen osalta 99,5 %. Kuvasta 3 nähdään, että keskihilojen määritelmä saattaa sulkea pois hiloja, joiden mittauspeitto on 100 %. Tämä ei ole kuitenkaan ongelma, niin kauan kuin keskihiloja on koko mittausalueella olemassa.

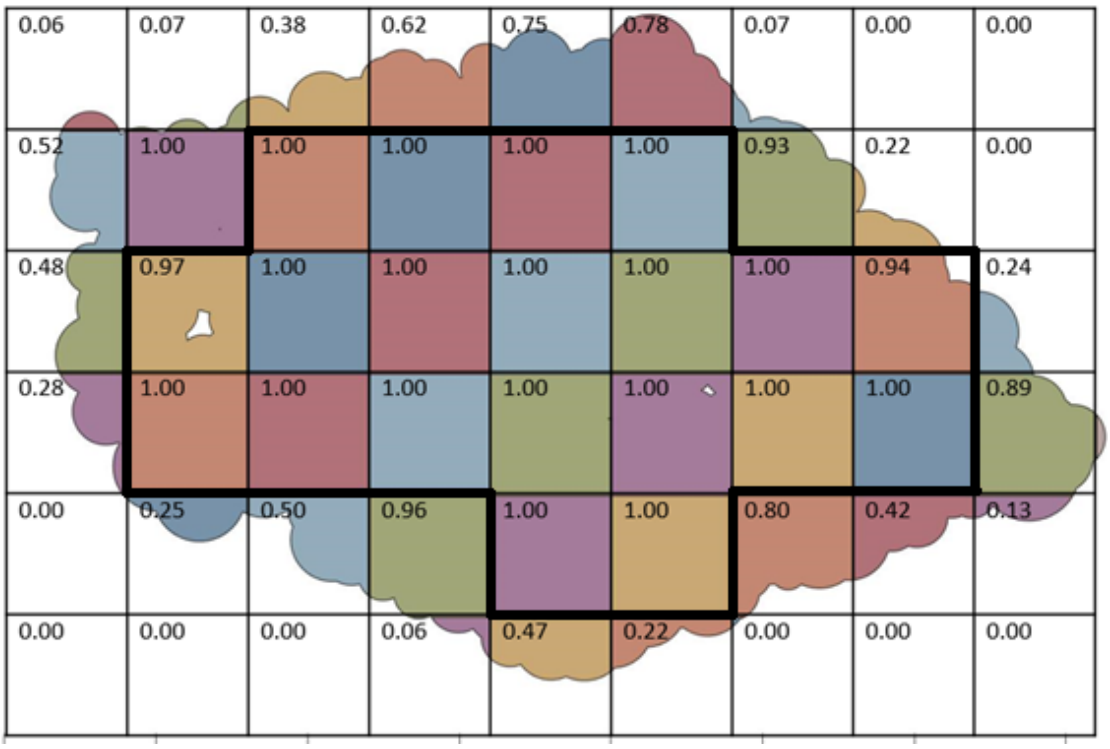
Palvelualustalle hilaruudun attribuuteiksi annetaan skaalausluku, ruudun mittauspeitto sekä kunkin mitatun ominaisuuden jakauman parametrit. Tietolähteen dokumentaatiossa kerrotaan, kuinka skaalausluku on muodostettu ja mitä jakaumatyyppiä on käytetty.



Kuva 1. Pankakosken mittausten mittauspeitto skaalausluvulla 15. Eri hiloihin kuuluvia mittauksia on korostettu eri värein. Keskihilojen mittauspeitto on keskimäärin 36 %.



Kuva 2. Pankakosken mittausten mittauspeitto skaalausluvulla 30. Keskihilojen mittauspeitto on keskimäärin 87 %.

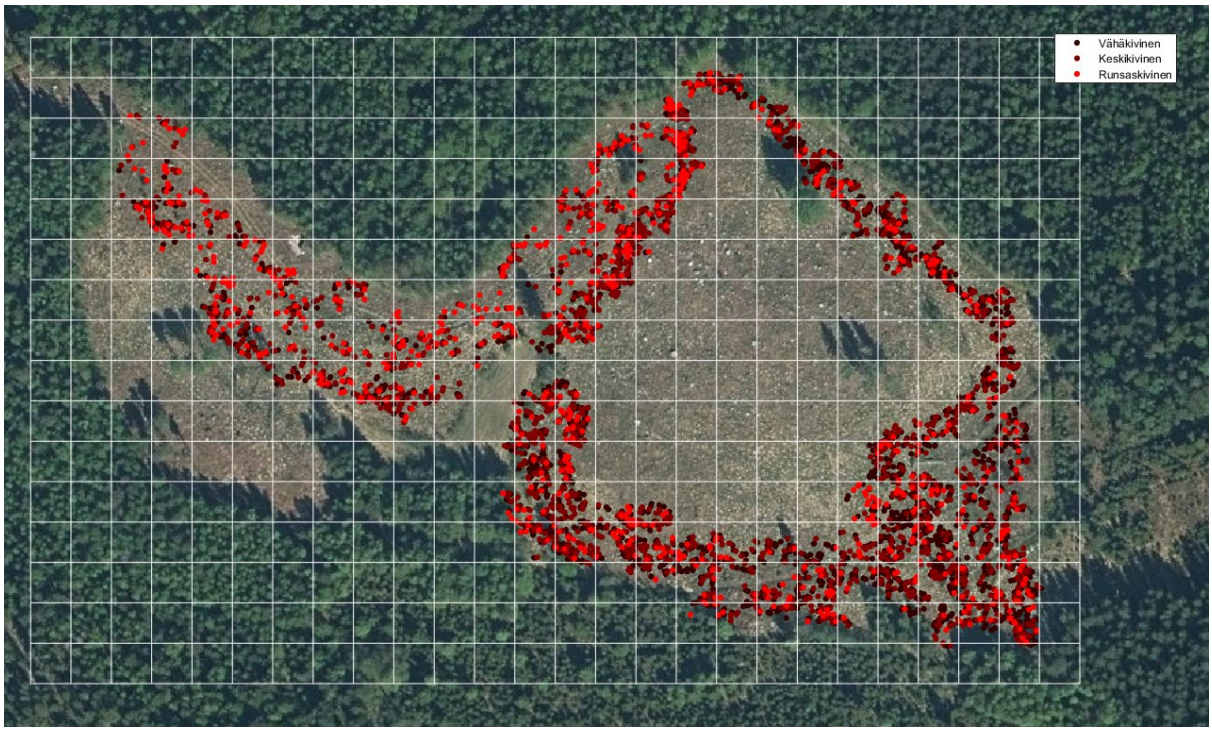


Kuva 3. Pankakosken pienoiskopterimittausten hilakohtaisen mittauspeiton arviointi skaalausluvulla 50. Mustalla paksunnetulla viivalla rajattujen keskihilojen mittauspeitto on nyt keskimäärin 99,5 %.

5.1.2 Käyttötapaus 2: Maaperän kivisyysmittaukset

Maaperän kivisyysmittauksia suoritetaan yhä painamismenetelmällä (Viro 1952). Metsäteho ja TTY tutkivat, voisiko maaperän kivisyyttä arvioida analysoimalla kaivinkoneen kauhaan maansiirtotyössä kohdistuvia voimia. Näytteitä otettiin n. 70 000 kappaletta kolmelta työalueelta. Ne luokiteltiin voimien perusteella kolmeen kivisyysluokkaan: vähäkivinen (maan kivisyysaste on 0–30 %), keskikivinen (30–60 %) ja runsaskivinen (yli 60 %). (Melander ym. 2019)

Kuvasta 4 nähdään, kuinka mittauspaikat osuvat 16 x 16 m hilarakenteeseen. Osaan hiloista saadaan hyvinkin kattava otos kivisyysmittauksia, mutta useissa hiloissa mittauksia on vain alueen reunoilla. Tässäkin tapauksessa on siis syytä ilmaista hilan attribuuttina mittausalueen kattavuus luonnehtimaan hila-arvojen kuvaavuuden ja siten luotettavuuden eroja. Tulosluokkiin on liitetty luokkien todennäköisyys- ja entropiatiedot, jotka voidaan myös tuottaa hilan attribuutteina. Yksinkertaisimmillaan jokaiseen hilaan voidaan tuottaa siihen osuneiden mittaustulosten moodi ja lukumäärä.



Kuva 4. Kivisyysmittaukset 16 x 16 m hilaruudukossa. Kirkkaampi punainen tarkoittaa kivisempää mittausarvoa.

5.2 Mittausresoluutio lähes sama kuin hilaresoluutio

Mittausresoluution voidaan katsoa olevan lähes saman kuin hilaresoluution silloin, kun hilaruutuun osuu keskimäärin $\frac{1}{3} < M_n < 3$ mittausta. Tällöin yleensä riittää, että lähimmän hilaruudun mittausarvo tuotetaan tarkasteltavan hilaruudun attribuuttina yhdessä kyseisen mittauksen ETRS-TM35FIN tasokoordinaattien kanssa.

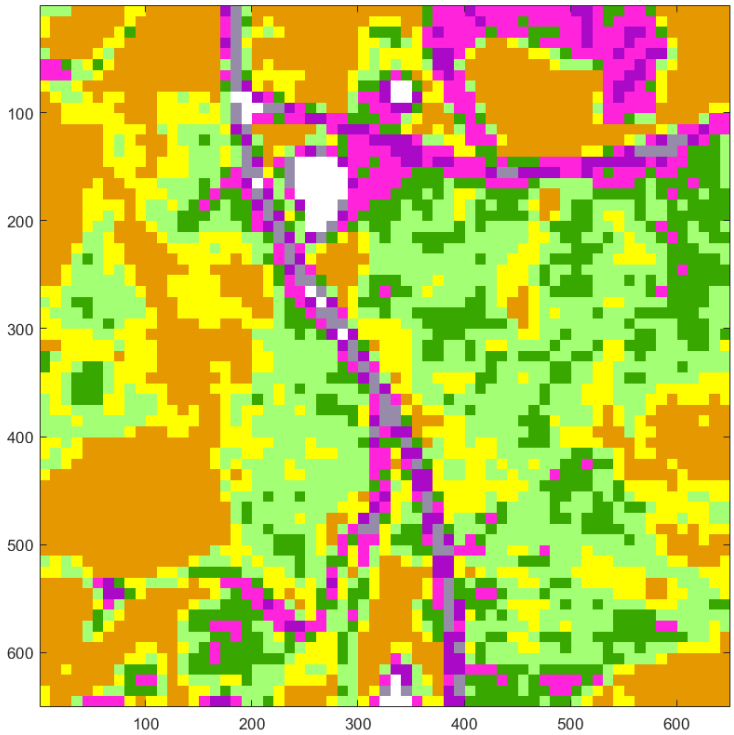
Jos mittauksia on hilassa enemmän kuin yksi 1, näistä voidaan laskea keskiarvo ja keskiarvon vaikutuskohta, joka vastaa mittauksen koordinaatteja. Tapauskohtaisesti riittävä tarkkuus voidaan saavuttaa myös tallentamalla vain hilan keskipistettä lähin mittausarvo koko hilan mittausarvoksi, kuten käyttötapauksessa 3 tehtiin.

5.2.1 Käyttötapaus 3: Satelliittimittaus kasvillisuusindeksistä

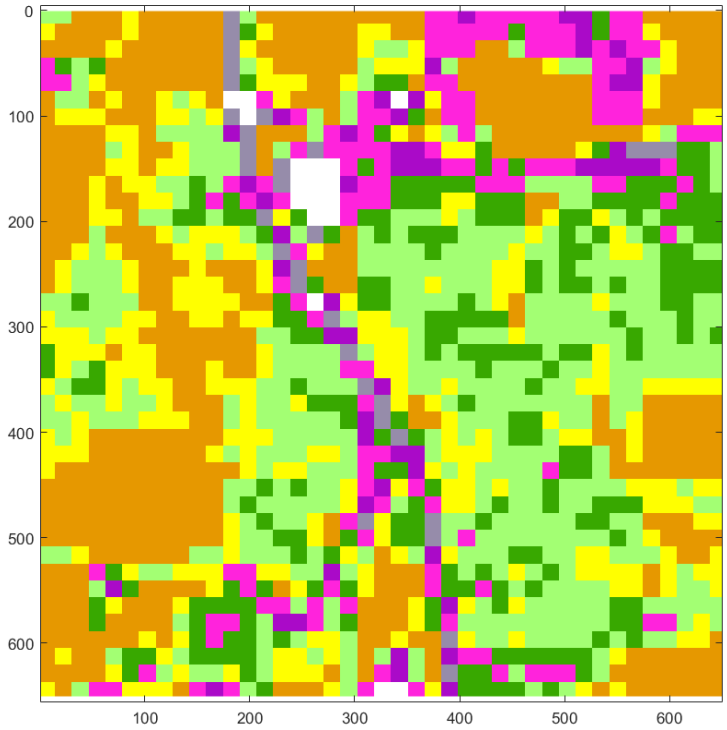
Forestry Thematic Exploitation Platform (F-TEP) on satelliittidatan palvelualusta (Häme ym. 2017), joka mahdollistaa satelliittimittausten haun, lataamisen ja prosessoinnin www-selaimen kautta. Tässä käyttötapauksessa alustalta ladattiin kasvillisuusindeksimittauksia Itä-Suomen alueelta. Mittausdata oli GeoTIFF-muotoinen yksikanavainen kuva, joka sisälsi tiedot jokaisen pikselin RGB-värikoodista visualisointia varten, kuvan koordinaatistotiedot ja NDVI-kasvillisuusindeksiin perustuvan luokittelun vesakoitumisasteen arvosta. Pohjana oleva NDVI on ensi vaiheessa laskettu Sentinel-2-satelliitin optisen instrumentin datasta (Forestry tep: Vegetation indices, 2019). Datan mittausresoluutio on tässä tapauksessa sama kuin kuvan pikselitiheys: 10 x 10 m, jolloin yhteen 16 x 16 m -hilaan tulee keskimäärin 2,56 mittausta.

F-TEP palvelualustan GeoTIFF-kuvan koordinaatisto on EPSG:32635, eli UTM kaista 35N, jonka ero ETRS-TM35FIN-koordinaatistoon on merkityksetön. Molemmat ovat poikittaisia lieriöprojektioita, joilla on sama keskimeridiaani. Ainoana erona edellisen vertausellipsoidi on WGS84 ja jälkimmäisen GRS80. Näiden vertausellipsoidien isoakselit ovat saman pituiset ja pikkuakselit eroavat toisistaan 0,2 mm (JHS-Suositukset 2008). Käytännön sovelluksissa ellipsoideja pidetäänkin yhtenevinä.

Kuvassa 5 nähdään alkuperäisen datan perusteella piirretty kuva, ja kuvassa 6 nähdään puolestaan hilaresoluutiolla piirretty kuva samalta alueelta, kun jokaisen hilan keskipistettä lähinnä oleva kasvillisuusindeksipikselin arvo asetetaan koko hilan alueen arvoksi. Kuvan 6 vastaavuus kuvan 5 kanssa on tällöin hieman yli 38 %.



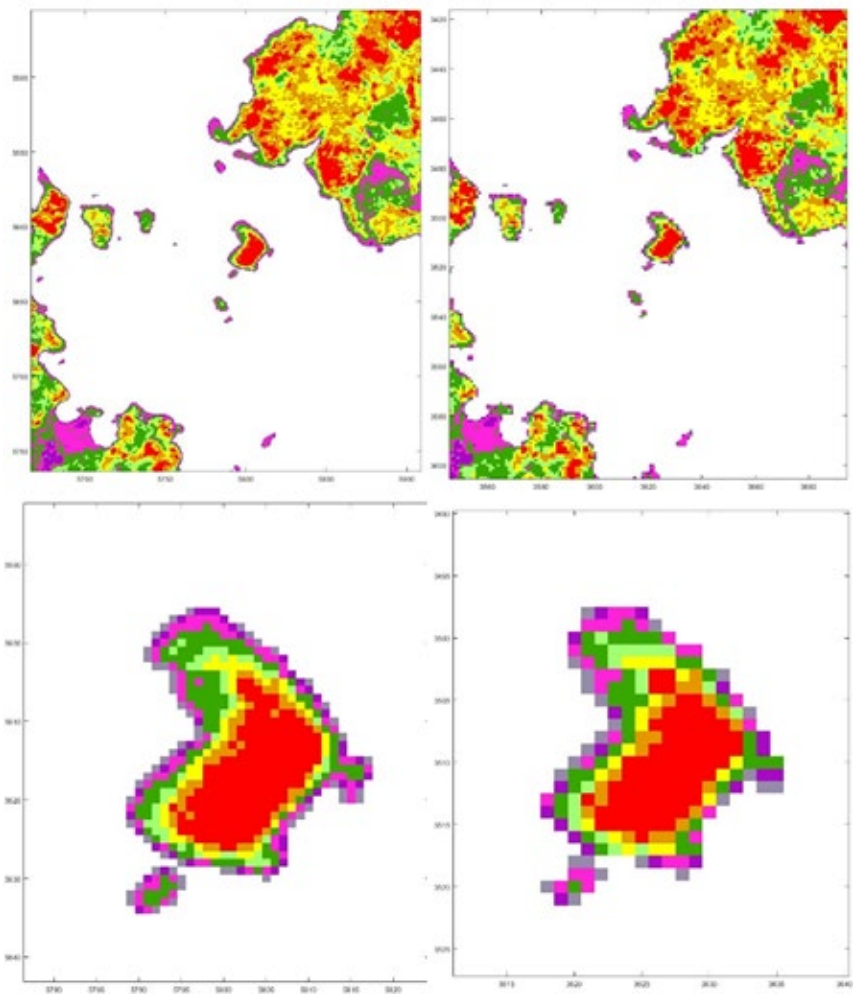
Kuva 5. Alkuperäinen 10 x 10 m -kasvillisuusindeksin mukainen kuva.



Kuva 6. Hilan keskipistettä lähimmän kasvillisuusindeksin perusteella piirretty kuva samalta alueelta kuin kuva 5.

Vastaavuus laskettiin vertaamalla kuvien 5 ja 6 kasvillisuusindeksejä jokaisen neliömetrin alalta. Näin verrattujen kasvillisuusindeksien absoluuttisten virheiden keskiarvo on kuitenkin vain 0,399 kasvillisuusluokkaa. Tämä kuvastaa menetelmän käyttökelpoisuutta tällaisissa tilanteissa, joissa vierekkäiset mittaukset eivät ole riippumattomia toisistaan.

Kuvassa 7 nähdään eri suuruisten alueiden alkuperäisen mittausresoluution ja hilaresoluutioon piirrettyjen kuvien välisiä eroja, kun hilan arvoksi on evaluoitu lähimmän mittapisteen arvo.



Kuva 7. Erisuuruisten alueiden hilaresoluutioon piirrettyjen kuvien silmämääräisen samankaltaisuuden arviointiin tarkoitettuja kuvia. Vasemmalla ovat alkuperäiset kuvat ja oikealla hilaresoluutioon piirretyt kuvat.

5.3 Mittausresoluutio huomattavasti harvempi kuin hilaresoluutio

Kun tietolähteen mittaustiheys on harva, on kyseessä yleensä hitaasti paikan funktiona muuttuvasta mittausarvosta. Tällöin sanotaan, että ilmiö on spatiaalisesti autokorreloitunut ja mittauksen lähellä olevien ilmiöiden arviointi mittauksen perusteella on mielekästä. Tällaisessa tapauksessa voidaan toimia esimerkiksi seuraavilla tavoilla:

Menetelmä 1: Asetetaan hilan keskipistettä lähinnä oleva mittausarvo hilan arvoksi.

Menetelmä 2: Lasketaan kolmen lähimmän arvon etäisyyspainotettu keskiarvo hilan arvoksi.

Menetelmä 3: Muodostetaan mittausalueelle interpoloimalla jatkuva arvopinta, josta evaluoidaan arvot hilakoordinaatteihin.

Ensimmäinen menetelmä on kevein ja sopii käytettäväksi esimerkiksi nopeasti päivittyviin mittauksiin, joiden esiprosessoinnin tulee olla tämän vuoksi vähäistä. Tämä menetelmä tunnetaan myös lähimmän naapurin interpolointina.

Toinen menetelmä olettaa, että hilan keskipistettä lähempänä olevien mittausten painoarvo hilan mittausarvoa estimoitaessa on suurempi kuin kauempana olevien mittausten painoarvo. Jos etäisyyden katsotaan vaikuttavan lineaarisesti mittauksen painoarvoon, voidaan hilalle määrittää estimaatti $\widehat{m}_{hila}$ seuraavasti:

$$\widehat{m}_{hila} = \sum_{i=1}^{n} \frac{\frac{1}{d_i} m_i}{\frac{1}{d_1} + \frac{1}{d_2} + \cdots + \frac{1}{d_n}},$$

missä m_i on mittausarvo pisteessä i, d_i ovat vastaavien mittapisteiden etäisyydet hilan keskipisteeseen ja n on lähimpien mittapisteiden valittu lukumäärä. Tarkempia malleja voidaan tarvittaessa määrittää suoraan mittauspaikkojen mittaushistorioiden ja/tai mahdollisten tarkistusmittausten perusteella. Lähimpien mittausten painoarvoa voidaan tarvittaessa muuttaa korottamalla etäisyysmuuttujat d_i potenssiin p. Korkeilla p:n arvoilla menetelmä 2 muuttuu menetelmäksi 1, eli lähintä naapuria painotetaan muiden mittausten kustannuksella.

Kolmas menetelmä pyrkii ottamaan huomioon koko mittausalueen kaikki mittaukset ja muodostaa näistä interpolaatiopinnan, joka arvioi mittausarvon muuttumista mittapisteiden välillä. Interpolaatiopinnan laskennassa käytettävä interpolointimenetelmä valitaan hiloille evaluoitujen arvojen riittävän kuvaavuuden ja käytettävissä olevan prosessointikapasiteetin perusteella.

5.3.1 Käyttötapaus 4: Säätiedot, lämpötila

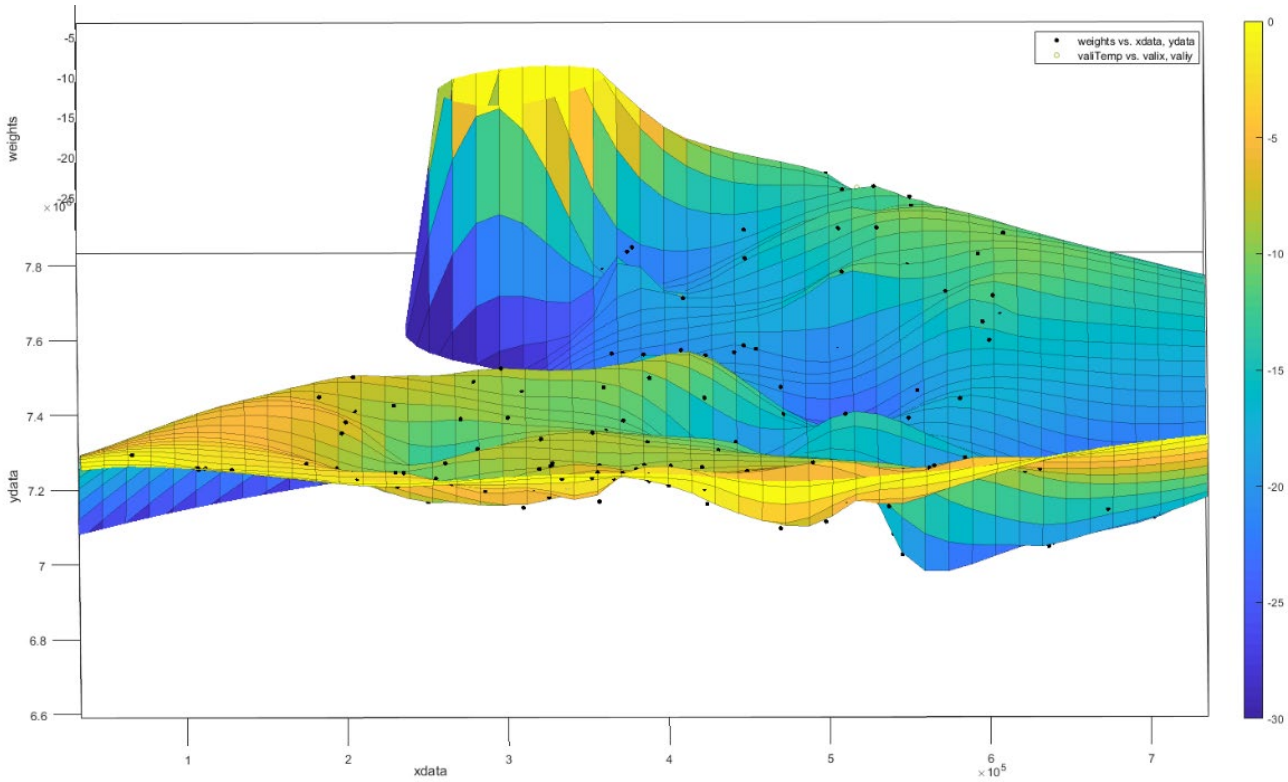
Ilmatieteen laitokselta on saatavissa Suomessa säätiedot julkisesti. Havaintoasemia on noin 200 kappaletta kautta Suomen. Havaintoasemien välien säätiedot estimoidaan Ilmatieteen laitoksen käyttämän mallin perusteella (Sääennustedata). Tässä esimerkissä käytetään havaintoasemilta kerättyä lämpötiladataa esittelemään luvun 5.3 kolmen menetelmän eroja. Eri interpolaatiomenetelmien hyvyyttä alueelliselle säädatalle on vertailtu myös Rooman yliopiston tutkimuksessa (Attorre ym. 2007). Kyseisessä tutkimuksessa parhaat tulokset saavutettiin universaalilla Kriging-interpolaatiolla.

Menetelmien eroja lämpötiladatan osalta testattiin seuraavasti: Kerättiin 182 havaintoasemalta lämpötilamittaukset samalta ajan hetkeltä. Näistä poimittiin kymmenen lämpötilamittausta, jotka poistettiin muiden havaintojen joukosta. Jäljelle jääneiden 172 lämpötilahavainnon perusteella muodostettiin luvun 5.3 menetelmien mukaiset estimaattorit, joista evaluoitiin estimaatit kymmeneen tarkistuskoordinaattiin.

Ensimmäisessä menetelmässä poimittiin kunkin tarkastelukohdan lähin mittapiste estimoimaan tarkastelukohdan lämpötilaa (Nearest Neighbour). Toisessa menetelmässä valittiin kolme lähintä mittapistettä, joiden arvoja painotettiin etäisyyksien käänteisarvojen suhteilla (Inverse Distance Weighting, IDW).

Kolmannen menetelmän interpolaatiopintoina testattiin neljää erilaista metodia: lineaarinen interpolointi, kuutiollinen interpolointi, biharmoninen interpolointi ja Thin plate spline -interpolointi. Parhaimmat tulokset saavutettiin yleensä biharmonisella interpoloinnilla, joten se valittiin käytetyksi metodiksi.

Kuvassa 8 on kyseisellä metodilla muodostettu interpolaatiopinta kuvattuna etelästä katsottuna. Mustat pisteet ovat säähavaintoasemien sijainneissa olevia lämpötilamittausten arvoja. Kuva havainnollistaa kuinka pintaan voi tulla kaukana mittapisteiden ulkopuolella epäjatkuvuuskohtia, mutta pisteiden lähellä pinta seuraa näitä hyvin.



Kuva 8. Thin plate spline -menetelmällä luotu interpolaatiopinta Suomen säähavaintoasemien lämpötiloista eräänä ajan hetkenä tammikuussa 2019.

Taulukossa 1 nähdään eri metodien tarkistuskohtien lämpötilaestimaattien root mean square error -luvut (RMSE), joiden perusteella Thin plate spline valittiin lopulliseen vertailuun.

Taulukko 1. Interpolaatiopintojen vertailu.

Interpolaatiometodi	Tarkistuskohdan estimaatin RMSE
Biharmoninen	0,2156
Thin plate spline	0,4948
Lineaarinen	0,8638
Kuutiollinen	1,0633

Taulukossa 2 nähdään menetelmien estimaattien keskiarvoiset virheet käytetyssä kymmenen tarkistuskohdan aineistossa.

Taulukko 2. Kappaleen 5.3 menetelmien estimaattivirheet testiajossa.

Menetelmä	Estimaattivirhe
1. Nearest Neighbour	1,68 °C
2. Inverse Distance Weighting	1,54 °C
3. Biharmoninen interpolaatio	0,09 °C

Lähimmän naapurin menetelmä ja IDW voivat olla monessa tapauksessa riittävän tarkkoja estimaatteja, mutta laskentakapasiteetin salliessa keskimäärin parempia tuloksia saadaan erilaisilla pintasovitteilla, kuten biharmonisella interpolaatiolla. Tätäkin parempia tuloksia saadaan edellä mainitulla Kriging-interpolaatiolla ja monikerroksisella neuroverkolla (Attorre ym. 2007).

6 YHTEENVETO

Raportissa tutkittiin tapoja tallentaa metsävaratietoja 16 x 16 m hiloihin siten, että se säilyttäisi tarkoituksenmukaisen määrän mittausinformaatiota. Tapauksissa, joissa mittaustiheys oli hilatiheyttä harvempi, voi riittää siirtää lähimmän mittauspisteen arvo hilan arvoksi ja ilmaista toisena hilan attribuuttina kyseisen mittapisteen ETRS-TM35FIN-koordinaatit. Jos tämän menetelmän tarkkuus ei ole riittävä, voidaan käyttää laskennallisesti raskaampia interpolaatiomenetelmiä etenkin voimakkaasti spatiaalisesti autokorreloivissa aineistoissa.

Hilan sisältäessä useita mittauksia, näistä on suositeltavaa muodostaa valitun todennäköisyysjakauman parametrit ja ilmaista ne hilan attribuuttina. Tämän lisäksi tulisi ilmaista kuinka suuren osuuden hilasta mittaukset peittivät. Raportissa esiteltiin tähän eräs mittausten arvioituihin peittosäteisiin perustuva menetelmä.

LÄHTEET

Attorre, F., Alfo, M., De Sanctis, M., Francesconi, F. & Bruno, F. 2017. Comparison of interpolation methods for mapping climatic and bioclimatic variables at regional scale. International journal of climatology, vol. 27, pp. 1825-1843.

Forestry tep: Vegetation indices, 2019. Forestry TEP, 2019. [Verkkojulkaisu.] Saatavissa: https://f-tep.com/?q=content/vegetationindices. [Viitattu 15.2.2019].

Hirvonen, R.A. 1972. Matemaattinen geodesia. Helsinki. Teknillisen korkeakoulun ylioppilaskunta.

Holopainen, M, Hyyppä J. & Vastaranta, M. 2013. Laserkeilaus metsävarojen hallinnassa. Helsingin yliopiston Metsätieteiden laitoksen julkaisuja, nro 5, p. 75, 2013.

Häkli, P., Puupponen, J., Koivula, H. & Poutanen, M. 2009. Suomen geodeettiset koordinaatistot ja niiden väliset muunnokset. Geodeettinen laitos.

Häme, T., Tergujeff, R., Rauste, Y., Farquhar, C., van Zetten, P., Kershaw, P., de Groof, A., Hämäläinen, J., van Bemmelen, J. & Seifert, F.M. 2017. Forestry-TEP responds to user needs for sentinel data value adding in cloud, in 2017 conference on Big Data from Space: BiDS'17, Toulouse, 2017.

Hämäläinen, J., Räsänen, T., Riekki, K. Sorsa, J-A. & Ritala, R. 2019a. Metsätiedon palvelualustan konseptitestaus. Metsätehon raportti 252. Saatavissa: http://www.metsateho.fi/metsatiedon-palvelualustan-konseptitestaus-raportti. [Viitattu 24.6.2019].

Hämäläinen, J., Räsänen, T., Riekki, K. Sorsa, J-A. & Ritala, R. 2019b. Metsätiedon palvelualustan konseptitestaus. Metsätehon tuloskalvosarja 7/2019. Saatavissa: http://www.metsateho.fi/metsatiedon-palvelualustan-konseptitestaus-tuloskalvosarja/. [Viitattu 24.6.2019].

JHS-Suositukset 2008. Julkisen hallinnon tietohallinnon neuvottelukunta, 22.10.2008. [Verkkojulkaisu.] Saatavissa: http://www.jhs-suositukset.fi/web/guest/jhs/projects/jhs154. [Viitattu 15.2.2018]

Melander, L., Ritala, R. & Strandström, M. 2019. Classifying soil stoniness based on the excavator boom vibration data in mounding operations. Silva Fennica vol. 53 no. 2 article id 10068. Saatavissa: https://doi.org/10.14214/sf.10068. [Viitattu 24.6.2019].

Scholz, F. 2008. Inference for the Weibull Distribution. In Stat 498B Industrial Statistics, Washington, Washington university, 2008, pp. 8–17.

Sääennustedata. Ilmatieteenlaitos, avoin rajapinta. Saatavissa: https://ilmatieteenlaitos.fi/avoin-data-saaennustedata-hirlam [Aineisto haettu 28.2.2019].

Viro, P.J. 1952. Kivisyyden määrittämisestä. Metsätieteellisen tutkimuslaitoksen julkaisuja, vol. 40, no. 3, pp. 1–23, 1952.

Weibull, W. 1951. A Statistical Distribution Function of Wide Applicability. ASME Journal of Applied Mechanics, pp. 293-297, September 1951.