

Monilähteisen VMI:n ongelmia ja ratkaisuja

Kai Mäkisara

Luonnonvarakeskus

Muita Luke:sta viime aikoina projektissa mukana olevia/olleita:
Erkki Tomppo, Helena Henttonen, Matti Katila, Juho Pitkänen,
Jouni Peräsaari.



Käsiteltäviä aiheita

- Mikä on monilähteinen VMI (MVMI)?
- historiaa ja nykytilanne
- kuvamateriaaleista
- aikaperspektiivi
- ik-NN-menetelmästä
- laadunvarmistuksesta

Mikä on monilähteinen VMI (MVMI)?

- tarvittiin tuloksia pienemmille alueille kuin mitä VMI-koealoista pystyttiin luotettavasti laskemaan
- kehitettiin VMI-koealoja, satelliittikuvia, digitaalisia karttoja ja maastomallia käyttävä menetelmä
 - tuottaa esim. kuntakohtaisia tilastoja ja rasterikarttoja (nykyisin yli 40 teemaa)
 - rasterikartoista voi laskea tuloksia halutuille alueille (yksittäiset pikselit eivät kovin tarkkoja)
- menetelmä jakaa maastokoealoilla mitatut suureet spatiaalisesti satelliittikuvien ja muiden aineistojen mukaisesti
 - kokonaistaso vaikea saada oikeaksi ilman maastoaineistoa
- varsin yksityiskohtaista kvantitatiivista ja kvalitatiivista tietoa
- melko pienen budjetin työtä

Teemoista

- nykyisin 44 kpl
- keskitilavuus yhteensä sekä puulajeittain (mänty, kuusi, koivu, muu lehtipuu) ja puutavaralajeittain (kaikki, tukkipuu, kuitupuu) (13 kpl)
- pohjapinta-ala, keskipituus, keskiläpimitta, ikä
- maaluokka, kasvupaikka, kasvupaikan päätyyppi
- kasvupaikka FAO:n FRA-määritelmän mukaan (vuodesta 2009)
- latvuspeitto, lehtipuiden latvuspeitto
- biomassa puulajeittain (mänty, kuusi, lehtipuut) ja ositteittain (juuret, kanto, elävät oksat, kuolleet oksat, kuorellinen runkopuu, hukkapuu) (21 kpl)
- lisäksi tietolähdeindeksi (vapaasti saatavissa teemoissa on pilviä "paikattu"vanhemmalla aineistolla)

Historiaa ja nykytilanne

- ensimmäiset operatiiviset tulokset 1990 Pohjois-Savosta
- ensimmäiset koko maan tulokset julkaistu 1998 (maastotyö VMI8 ja VMI8+)
- toiste koko maan tulokset julkaistu 1990 (Springer) (VMI9, maastotyö metsäkeskuksittain 1996-2003)
- kolme rasteriteemaa tehty CORINE2006 varten (keskipituus, latvuspeitto ja lehtipuiden latvuspeitto)
 - INSPIRE-dataa vuodesta 2010, WMS and www.paikkatietoikkunassa.fi
- VMI10:stä lähtien maastotyöt vuosittain koko maassa (1/5th koealoista) => sama ajankohta koko maahan
 - vuoden 2005 tulokset, julkaisu 2009 (Metla MWP111)
 - vuoden 2007 tulokset, julkaisu 2012 (Metla MWP227)
 - vuoden 2009 tulokset, julkaisu 2013 (Metla MWP273)
 - ensimmäinen vapaasti saatava MVMI-tuote, 11/2012
 - vuoden 2011 tulokset, julkaisu 2014 (Metla MWP319)
 - toinen vapaa MVMI-tuote, 6/2013
 - vuoden 2013 tulokset tekeillä

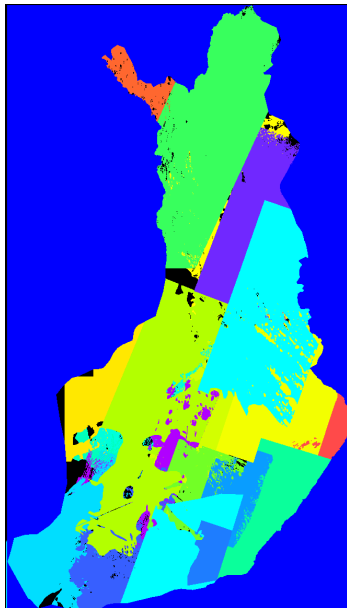
Kuvamateriaaleista

- koko maa pystyttävä kattamaan siten, että kustannukset eivät nouse liian suuriksi (kuvat, työ)
 - pilvet rajoittavat optisten kuvien käyttöä huomattavasti, SAR-kuvilla ei riittävää tarkkuutta
- riittävän tarkka ilmakehäkorjaus ei (vieläkään) onnistu => vain saman jonon kuvia voidaan yhdistää
- kerrallaan käsiteltävälle alueelle mahdollista riittävästi koealoja (=> 1000)

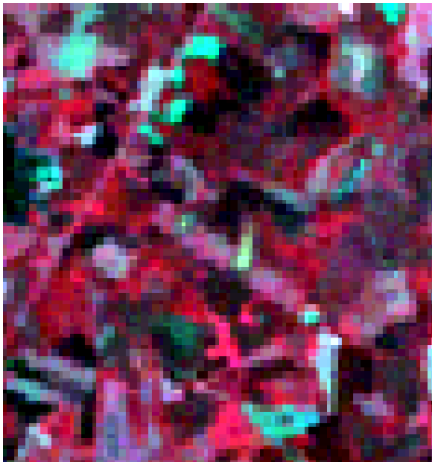
=>

- enimmäkseen käytetty Landsat-kuvia, lisäksi IRS-, Spot- ja ALOS-kuvia
- koko Suomen kuvapeitto yhdeltä kesältä ei usein onnistu
 - Sentinel-2 satelliitit lisäävät kuvien tarjontaa
- instrumenttien paraneminen parantaa vähitellen tuloksia

MVMI-2013 kuvapeitto



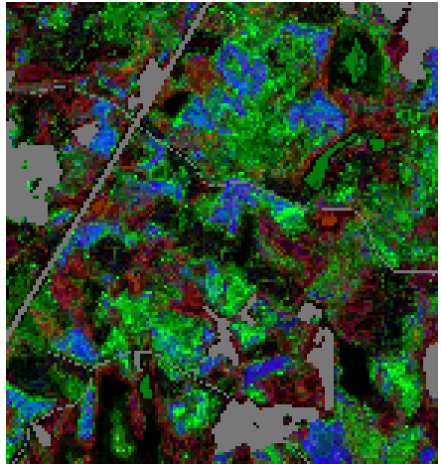
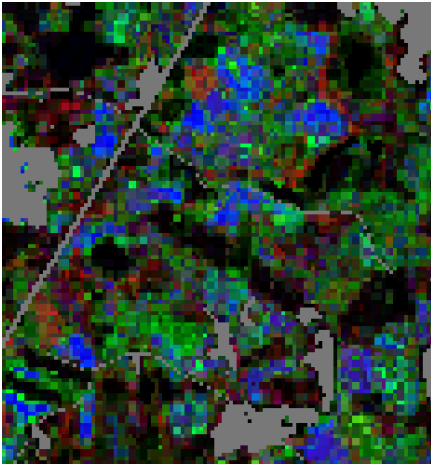
Landsat-5-kuva



Landsat-8-kuva



Landsat-5, kokonaistilavuus Landsat-8, kokonaistilavuus



Kuvien esikäsittely

- työlästä, mutta välttämätöntä
- toivottavasti ei aina tarvitse tehdä itse
- oikaisu tehtävä itse
 - toimitettujen kuvien tarkkuus paranee, mutta ei vieläkään ole riittävä
- pilvimaskit
 - algoritmeja on, mutta tarkkuus ei ole riittävä (reunat, varjot)
 - käsityötä kyllä helpottaa kun tuntee mosaikoinnin

Aikaperspektiivi

Kuvat, maastoaineistot ja muut digitaaliset aineistot eri päiviltä/vuosilta

=> Mikä on analysointituloksen ajankohta?

Osittaisratkaisuja MVMI:ssä:

- päätetään tuotteen tavoiteajankohta (31.7.xxxx)
- valitaan kuvat mahdollisimman läheltä ajankohtaa; ei paljon tehtävissä
- valitaan muut digitaaliset aineistot mahdollisimman läheltä ajankohtaa; ei paljon tehtävissä
- päivitetään maastoaineistot laskennallisesti tavoiteajankohtaan
 - päätehakkuiden yms. suurien muutosten toteaminen ja koealamuuttujien korjaus
 - muiden koealojen päivitys kasvumalleilla siten, että kokonaistilavuus kuvan alueella täsmää maastoestimaattiin (regressio vuosittaisista)

ik-NN-menetelmän ydinkohtia

- jokaisessa kuva-alkiossa piirrevektori, joka koostuu satelliittikuvakanavista ja niistä lasketuista funktioista sekä suuraluutilavuuskuvan kanavista
- jokaiselle piirrekanavalle lasketaan optimaalinen paino geneettisellä algoritmilla
- lasketaan analysoitavan kuva-alkion piirrevektorin ja maastokoealojen kohdilta laskettujen piirrevektrien euklidiset etäisyydet:

$$d_{p_j,p}^2 = \sum_{l=1}^{n_f} \omega_{l,f}^2 (f_{l,p_j} - f_{l,p})^2 + \sum_{l=1}^{n_g} \omega_{l,g}^2 (g_{l,p_j} - g_{l,p})^2$$

- jokaisessa kuva-alkiossa k pienimmästä etäisyydestä lasketaan painot (summa 1)

$$W_{i,p} = \frac{1}{d_{p_i,p}^t} \bigg/ \sum_{j \in \{i_1(p), \dots, i_k(p)\}} \frac{1}{d_{p_j,p}^t}, \quad \text{if and only if } i \in \{i_1(p), \dots, i_k(p)\}$$

$$= 0 \quad \text{otherwise.}$$

- rasterituotetta varten jokaisessa kuva-alkiossa lasketaan teemalle y ennuste

$$\tilde{y}_p = \sum_{i \in i_h} w_{i,p} y_i,$$

- toisaalta alueen u kaikkien kuva-alkiokohtaisten kertoimien avulla voidaan laskea referenssikoealalle kokonaispaino

$$c_{i,u} = \sum_{p \in u} w_{i,p}$$

Keskiennuste alueelle saadaan nyt vastaavasti

$$\hat{y}_u = \frac{\sum_{i \in I} c_{i,u} y_i}{\sum_{i \in I} c_{i,u}}$$

Keskiennusteen voi laskea myös teemarasterin alueen keskiarvona.

Miten tämä poikkeaa edellisestä?

- ei mitenkään, jos lasketaan normaali keskiarvo
- eroa syntyy jos jaetaan arvoalue osiin ja lasketaan summaennuste jollekin arvovälille; esim. paljonko on vähintään 300 m³/ha metsiä esimerkkialueella, jolla on metsää 10 500 km²
 - rasteriteemasta 2,5 %
 - kerrointiedostosta 8,6 %
 - maastoaineistosta 8,1 %

=> jakaumasta riippuvat ennusteet alueille olisi parempi laskea kertoimien kautta.

k-NN-menetelmän etuja ja haittoja

Etuja:

- mallin valinta kuva-alkiokohtaisesti helppoa
 - rajoituksina esim. maantieteellinen etäisyys, korkeusero
- eksplisiittistä mallia ei tarvita
- hyviä tuloksia suuremmilla alueilla

Kyseenalaisia asioita:

- hyviä tuloksia kuva-alkioittain, mutta mahdollisesti suuri ero tuloksissa vaikka ero spektriarvoissa pieni
- ei ekstrapoloi, tulosjakauma kapenee
 - ongelma erityisesti jos vähän koealoja (ei MVMI:n ongelma)

Taulukkotulosten korjaus

Muita Lukessa kertoimia käytettäessä tehtäviä korjauksia

- ns. karttakorjaus
- pinta-alakorjaukset (karttaprojektio ym.)

=> älkää hämmästykö vaikka kartoista laskemalla saatte erilaisia tuloksia kuin Metla on raportoinut

Estimaattien laadun varmistus

- ristiinvalidointi geneettisen optimoinnin yhteydessä
- kuntaryhmävertailut
 - kuntaryhmä kuvan alue, jolla riittävä määrä koealoja (> 1000)
 - lasketaan maastoestimaatteja virharvioineen
 - verrataan MVMI-tulokseen vastaavalla alueella
 - selvitetään, onko estimaattien taso oikea kuvan eri alueilla
- joissakin kuvaikkunoissa joutuu tekemään aika paljonkin työtä ennen kuin tulos on tyydyttävä/niin hyvä kuin voi saada

Yhteenveto

- käytännön työhön liittyy monenlaisia käytännön asioita
- kuvamateriaali vaihtelee, kohteen ominaisuudet vaihtelevat
- kaikkiin ongelmiin löydettävä jonkinlainen ratkaisu
- mukava saada tuote valmiiksi (erityisesti kun sillä on kysyntää :-)