

Maanmittauslaitoksen korkeusmallin haltuunotto

Jukka Rahkonen, <http://latuviitta.org>

Lisenssi [Creative Commons Attribution 3.0 Unported](http://creativecommons.org/licenses/by/3.0/)

Viimeksi muokattu 4. maaliskuuta 2013

Johdanto

Avoimena datana oleva Maanmittauslaitoksen 10 metrin korkeusmalli on jaettu 6 x 6 kilometrin karttalehtiin, ja koko maan kattavaan aineistoon kuuluu 11003 zip-pakattua ASCII XYZ -muotoista tiedostoa. Aineisto on saatavilla Maanmittauslaitoksen tiedostolatauspalvelun kautta, mutta jos joku sattuu olemaan kiinnostunut tekemään jotain koko maan laajuista näiden korkeusmallitiedostojen avulla, niin aineiston tehokkaan haltuunoton esteenä on kolme ongelmaa:

- 11033 tiedoston lataaminen MML:n latauspalvelusta on järjetön urakka.
- Zip-arkistot on purettava, ennenkuin niihin kuuluvia tiedostoja voidaan käyttää (tämä on onneksi väärä luulo, kuten kohta todistetaan).
- XYZ-tiedostomuoto ei todennäköisesti sovi suoraan loppukäyttöön, vaan tiedostot on ensin muunnettava johonkin muuhun muotoon.

Jos puuhastellaan yhden tai kymmenen tiedoston kanssa (tai tuntiveloituksella), niin menetelmien tehokkuudella ei ole juuri väliä. Mutta joskus ajalla on merkitystä. Jokainen sekunti, joka käytetään yhden tiedoston käsittelemiseen, merkitsee esimerkiksi tämän esimerkin korkeusmalliaineistolla 11000 sekunnin, eli 183 minuutin, eli hieman yli kolmen tunnin ajankäyttöä.

Tässä esitetyllä menetelmällä tiedostojen lataaminen onnistuu automaattisesti yhdessä yössä. Kaikkien 11003:n zip-tiedostoista löytyvien XYZ-korkeusmallitiedostojen muuntaminen toiseen tiedostomuotoon onnistuu puolestaan hyvällä, mutta kuitenkin ihan tavallisella kannettavalla tietokoneella alle tunnissa, mikä merkitsee 0,33 sekuntia korkeusmallin palasta kohti.

Menetelmät

Korkeusmalliaineiston lataus

Maanmittauslaitoksen korkeusmalliaineisto kannattaa hankkia Kapsin tiedostopalvelusta rsync-ohjelmalla. Aineistoluettelo ja tietoa Kapsista löytyy osoitteesta <http://kartat.kapsi.fi/>

Rsync-lataamisohjeet Windows-koneille löytyy Latuviitan aikaisemmasta ohjeesta

http://latuviitta.org/documents/MML-taustakarttasarja_WMS-palveluksi.pdf

Alla olevalla komennolla haetaan 10 metrin korkeusmallin tiedostot ja tallennetaan ne hakemistoon d:\korkeusmalli_10m

```
rsync -arPv tiles.kartat.kapsi.fi::mml/korkeusmalli_10m/kaikki/etrs89/ascii_xyz/  
/cygdrive/d/korkeusmalli_10m/
```

Zip-tiedostojen lukeminen suoraan GDAL:lla ilman erillistä purkamista

Riittävän uudet GDAL-versiot tukevat tiedotojen lukemista vsizip-menetelmällä, mikä tarkoittaa käytännössä sitä, että zip-tiedostoja ei tarvitse purkaa ollenkaan

Esimerkiksi tämä gdalinfo-komento tutkii zip-tiedoston ja listaa sen sisältä löytämänsä tiedostot.

```
D:\korkeusmalli_10m>gdalinfo /vsizip/D:\korkeusmalli_10m\K2\K23\K2344H.zip  
ERROR 4: `/vsizip/D:\korkeusmalli_10m\K2\K23\K2344H.zip' not recognised as a sup  
ported file format.  
gdalinfo failed - unable to open '/vsizip/D:\korkeusmalli_10m\K2\K23\K2344H.zip'  
.Unable to open source `/vsizip/D:\korkeusmalli_10m\K2\K23\K2344H.zip' directly.  
The archive contains 2 files:  
    /vsizip/D:\korkeusmalli_10m\K2\K23\K2344H.zip/K2344H.xyz  
    /vsizip/D:\korkeusmalli_10m\K2\K23\K2344H.zip/K2344H.log
```

Yhden arkistosta löytyvän tiedoston yleistiedot saadaan listatuksi antamalla arkistoidun tiedoston nimi gdalinfo-ohjelmalle yllä olevan tulostuksen mukaisesti.

```
D:\korkeusmalli_10m>gdalinfo /vsizip/D:\korkeusmalli_10m\K2\K23\K2344H.zip/K2344  
H.xyz  
Driver: XYZ/ASCII Gridded XYZ  
Files: /vsizip/D:\korkeusmalli_10m\K2\K23\K2344H.zip\K2344H.xyz  
Size is 602, 602  
Coordinate System is `'  
Origin = (109995.0000000000000000,6618005.0000000000000000)  
Pixel Size = (10.0000000000000000,-10.0000000000000000)  
Corner Coordinates:  
Upper Left ( 109995.000, 6618005.000)  
Lower Left ( 109995.000, 6611985.000)  
Upper Right ( 116015.000, 6618005.000)  
Lower Right ( 116015.000, 6611985.000)  
Center ( 113005.000, 6614995.000)  
Band 1 Block=602x1 Type=Float32, ColorInterp=Undefined
```

XYZ-tiedostojen muuntaminen toiseen tiedostomuotoon

Yksi zip-arkiston sisältä löytyvä korkeusmallitiedosto muunnetaan toiseen tiedostomuotoon gdal_translate-ohjelmalla käyttämällä tällaista komentoa.

```
gdal_translate -of GTiff -a_srs epsg:3067 -co compress=deflate  
/vsizip/d:\korkeusmalli_10m\K2\K23\K2344H.zip/K2344H.xyz K2344H.tif
```

Yksitoistatuhatta pakattua korkeusmallitiedostoa muuttetaan toiseen tiedostomuotoon tekemällä komentojono, joka käy ne kaikki läpi. Tällaisen komenjonon teon voi aloittaa tallentamalla hakemistolistauksen tiedostoon

```
dir /s /b *.zip >muunna_korkeusmalli.bat
```

Tiedostolistausksen alku näyttää tältä

```
d:\korkeusmalli_10m\K2\K23\K2344H.zip
d:\korkeusmalli_10m\K2\K24\K2433D.zip
d:\korkeusmalli_10m\K2\K24\K2433E.zip
d:\korkeusmalli_10m\K2\K24\K2433F.zip
d:\korkeusmalli_10m\K2\K24\K2433G.zip
...
```

Tätä tiedostolistausta muokataan lisäämällä joka riville haluttu GDAL-komento. Tässä tapauksessa muunnetaan XYZ-tiedosto GeoTIFF-muotoon, liitetään niihin tieto siitä, että ne käyttävät ETRS-TM35FIN-koordinaattijärjestelmää, ja pakataan tiff-tiedostot tilan säästämiseksi deflate-menetelmällä. Tulostiedostot kirjoitetaan hakemistoon ”dem_tif”.

```
gdal_translate -of GTiff -a_srs epsg:3067 -co compress=deflate
/vsizip/d:\korkeusmalli_10m\K2\K23\K2344H.zip/K2344H.xyz dem_tif\K2344H.tif

gdal_translate -of GTiff -a_srs epsg:3067 -co compress=deflate
/vsizip/d:\korkeusmalli_10m\K2\K24\K2433D.zip/K2433D.xyz dem_tif\K2433D.tif

gdal_translate -of GTiff -a_srs epsg:3067 -co compress=deflate
/vsizip/d:\korkeusmalli_10m\K2\K24\K2433E.zip/K2433E.xyz dem_tif\K2433E.tif

gdal_translate -of GTiff -a_srs epsg:3067 -co compress=deflate
/vsizip/d:\korkeusmalli_10m\K2\K24\K2433F.zip/K2433F.xyz dem_tif\K2433F.tif

gdal_translate -of GTiff -a_srs epsg:3067 -co compress=deflate
/vsizip/d:\korkeusmalli_10m\K2\K24\K2433G.zip/K2433G.xyz dem_tif\K2433G.tif
...
```

Yksityiskohtaisempi ohje vastaavasta hakemistopuuhun ripoteltujen lukuisten tiedostohedelmien tehokkaasta mehustamisesta komentojonon avulla on aikaisemmassa Latuviitan ohjeessa

http://latuviitta.org/documents/MML_peruskarttarastereiden_haltuunotto.pdf

Korkeusmallin palasten käsittely yhtenä aineistona

GDAL-virtuaalimosaiikin teko gdalbuildvrt-ohjelmalla

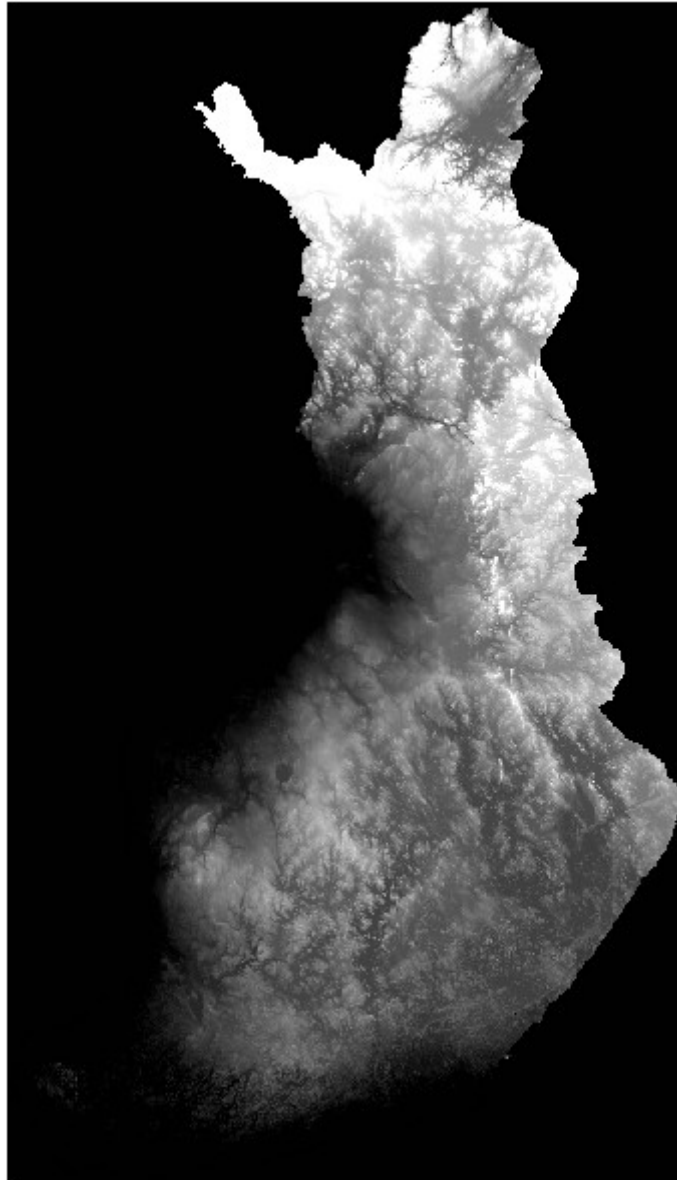
Muunnoksen jälkeen meillä on edelleen käsissä 11000 erillistä korkeusmallitiedostoa. GDAL-ohjelmiin kuuluu kuitenkin virtuaalitiedostoksi kutsuttu mekanismi, jonka avulla niitä kaikkia voidaan kuitenkin käsitellä ikään kuin ne olisivat yksi ainoa tiedosto.

GDAL-virtuaalitiedosto voidaan luoda gdalbuildvrt-apuohjelmalla, ja tässä tapauksessa komento on hyvin yksinkertainen.

```
gdalbuildvrt korkeusmalli_10m.vrt *.tif
```

Virtuaalitiedoston teko kestää hetken, ja kun se on valmistunut, niin se voidaan avata suoraan esimerkiksi Quantum GIS -ohjelmalla. Tämä ei kuitenkaan ole erityisen järkevää, koska

virtuaalitiedosto on vain hakemisto, joka yhdistää 11000 alkuperäistä kuvaa. Siitä syystä kun kuva avataan sellaisessa mittakaavassa, että koko Suomi mahtuu kartalle, niin Quantum GIS joutuu avaamaan kaikki 11000 tiedostoa. Jos tähän operaatioon menee esimerkiksi 0,1 sekuntia/kuva, niin virtuaalitiedoston avaamiseen kuluu aikaa 20 minuuttia. Alla oleva kuva osoittaa, että tämä on mahdollista, mutta järkevä tapa maastomalli-vrt-tiedoston käyttämiseksi Quantum GIS -ohjelmassa on avata ensin joku muu aineisto ja kohdistaa kartta niin pienelle alueelle, että sen peittämiseen tarvitaan vain muutama tai muutama kymmen korkeusmallitiedostoa. Tällaisessa mittakaavassa VRT-tiedoston käyttäminen on nopeaa. Karttatason ominaisuuksien asetuksista kannattaa laittaa korkeusmallitasolle mittakaavaraja, niin ettei Qgis rupea piirtämään sitä epäkäytännöllisen pienessä mittakaavassa.



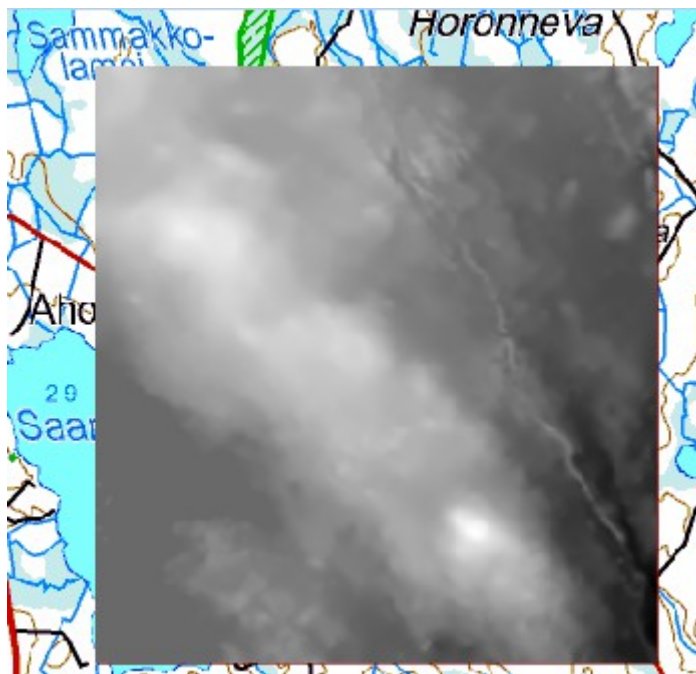
VRT-tiedosto sopii myös erityisen hyvin käytettäväksi muiden GDAL-apuohjelmien lähtötiedostona. Monet ohjelmat eivät esimerkiksi pysty avaamaan kuin pakkaamattomia tiff-kuvia. Gdal_translate-ohjelmalla voidaan kätevästi leikata koko maan korkeusmallista pala halutusta kohdasta ja kirjoittaa siitä pakkaamaton GeoTIFF tällaisia ohjelmia varten.

Leikataan esimerkin vuoksi VRT-korkeusmallista pala, joka on merkitty korostettuna alla olevaan karttaan.

```
gdal_translate -of GTiff -projwin 412214 6964358 415189 6961188 korkeusmalli_10m.vrt valittu_pala.tif
```



Avataan tiedosto "valittu_pala.tif" peruskartan päälle ja todetaan, että hyvältä näyttää.



Shapefile-muotoisen hakemisto teko gdaltindex-ohjelmalla

GDAL-apuohjelma gdaltindex on kehitetty varsinaisesti Mapserver-karttapalvelinta varten, mutta se on usein hyödyllinen muihinkin tarkoituksiin. Ohjelma selvittää sille annettujen kuvien maantieteelliset ulottuvuudet, jalanjäljet, ja tekee niistä vektoritiedoston shapefile-muodossa. Komento on perusmuodossaan

```
gdaltindex hakemisto_shapefile.shp *.tif
```

Alla oleva kuva näyttää tulostiedoston ”hakemisto_shapefile.shp” avattuna karttaohjelmalla. Vaikuttaa siltä, että kaikki korkeusmallin palat ovat tallessa.



Nopeammin, paremmin ja halvemmalla

Tämän mallisuorituksen arvioinnissa voidaan keskittyä nopeuteen, sillä yksinkertaisessa tiedostomuodon muunnoksessa lopputuloksen laatu on täydellinen, ellei ohjelmassa satu olemaan suoranaista karkeaa ohjelmointivirhettä, eikä sellaisen olemassaolosta ole mitään viitteitä. Hinnan arviointi on myös yksinkertaista, sillä GDAL-ohjelmien käyttö on vapaata ja ilmaista mihin tahansa tarkoitukseen.

Nopeuden arvioimisen voi aloittaa muutamista mittaustuloksista.

Ensimmäiset mittaustulokset on saatu kannettavalla tietokoneella, joka jo vuonna 2006 uutena ollessaan oli keskivertoa hitaampi. Tietokoneen prosessorin on Intel Core2 Duo U7600 1,2 GHz. Sekä alkuperäiset että muunnetut kuvat olivat samalla ulkoisella USB2-kovalevyllä, ja muunnos yhdessä komentoikkunassa ajettuna kesti 9 tuntia 55 minuuttia. Yhtä muunnosta kohti aikaa kului siis 3,3 sekuntia, mikä ei tunnu kovin hitaalta.

Seuraavat mittaustulokset saatiin kannettavalla, joka kuului varsin nopeaan luokkaan vuonna 2012. Prosessori on 5755G Intel core i7-2670QM 2.2GHz. Prosessorissa on monta ydintä ja gdal_translate käyttää vain yhtä ydintä, joten oli syytä olettaa, että kokonaisvauhti paranisi, jos muunnoksia ajettaisiin useassa komentoikkunassa yhtä aikaa. Siitä syystä tehtiin pieni esikoe, jossa muunnettiin 100 tiedostoa yhdessä, kahdessa, kolmessa ja neljässä ikkunassa yhtä aikaa.

100 XYZ-tiedoston muunnos yhdessä ikkunassa **72 sek./100 tiedostoa**

2 x 100 XYZ-tiedoston muunnos kahdessa ikkunassa **77 sek./200 tiedostoa**

3 x 100 XYZ-tiedoston muunnos kolmessa ikkunassa **81 sek./300 tiedostoa**

4 x 100 XYZ-tiedoston muunnos neljässä ikkunassa **92 sek./400 tiedostoa**

Oletus piti selvästi paikkansa, ja gdal_translate:a kannattaa tuolla koneella ajaa ainakin neljässä eri ikkunassa yhtä aikaa. Lopullinen kellotus tehtiin jakamalla 11000 kuvan urakka neljään 2750 kuvan erään. Tällä tavalla koko korkeusmalliaineiston muuntaminen kesti **57 minuuttia**. Prosessia kohti muunnosnopeus oli 1,24 sekuntia/tiedosto, mutta kokonaisurakkana laskettuna vain **0,31 sekuntia/tiedosto**. Tämä tuntuu kohtuullisen hyvältä vauhdilta, mutta varmaan parempaankin voi vielä päästä.