

# Helsingin seudun maanpeiteaineisto 2022

Menetelmäkuvaus ja tulokset





## Sisältö

1	Johdanto .....	4
1.1	Projektin kuvaus .....	4
1.2	Tausta.....	4
1.3	Projektin vaiheet .....	6
1.4	Lopputuote.....	6
2	Aineistot .....	8
3	Maanpeiteluokittelu .....	9
3.1	Tekoälymallinnus .....	9
3.2	Pistepilviluokitus .....	10
3.2.1	Voimalinjapisteiksi luokitellut pisteet suodatetaan.....	10
3.2.2	Maanpinta suodatetaan pistepilviaineistosta .....	11
3.3	Yhdistysvaihe.....	11
3.3.1	Perustoimitus (Basic product) .....	11
3.3.2	Leikkaamaton viherpinta (Uncut green).....	11
3.4	Vektorisointi .....	11
4	Rasteri ilman segmenttien yhdistämistä .....	12
5	Validointi.....	13
	Liite 1 Yhdistysvaihe .....	16
	Liite 2 Segmenttien yhdistäminen .....	19
	Liite 3 Rasteri ilman segmenttien yhdistämistä .....	20

# 1 Johdanto

Tämä raportti kuvaa HSY:n tilaamaa ja SCALGOn toteuttamaa hanketta "Helsingin seudun maanpeiteaineisto 2022, TP 20/22 ". Raporttiin on koottu kuvaus menetelmistä, syöttötiedoista, tuloksista ja tulosten validoinnista.

## 1.1 Projektin kuvaus

Hankkeen tavoitteena oli tuottaa yksityiskohtainen maanpeiteaineisto pääkaupunkiseudulle, mukaan lukien kunnat Espoo, Helsinki, Kauniainen, Järvenpää, Kirkkonummi ja Vantaa. Projektialueen kokonaispinta-ala on 1147,7 km<sup>2</sup>.

Seudullinen maanpeiteaineisto tuotetaan tässä projektissa hyödyntäen tekoälyyn perustuvaa menetelmää maanpeitteen tunnistamiseksi. Menetelmän on alun perin kehittänyt SCALGO yhteistyössä Århusin yliopiston kanssa vuosina 2019-2020. Tässä projektissa menetelmää on edelleen kehitetty vastaamaan tämän projektin tarpeita.

Luokittelu on tehty vuoden 2021 ilmakuvia ja pistepilviaineistoja hyödyntäen. Mahdollisuuksien mukaan pistepilviaineistoja on kerätty kunnilta ja jos kunnilla ei ole ollut omaa pistepilviaineistoa, on alueella käytetty Maanmittauslaitoksen 5p/m<sup>2</sup> aineistoa.

Lisäksi HSY toimitti SCALGO:lle teiden, rakennusten, vesien ja peltojen vektoriaineistoja, joiden avulla tekoälyn tuloksia on tarkennettu kattamaan kaikki projektissa vaaditut maanpeiteluokat.

## 1.2 Tausta

Maanpeiteaineisto on suurelta osin tuotettu UNET "Convolutional Neural Network" -mallia hyödyntäen. UNET-mallin vahvuus on erityisesti sen kyky erotella vettä läpäisemättömät pinnat (kuten asfaltti) ja paljas maa (sora ja hiekka). Nämä maanpeiteluokat ovat perinteisillä menetelmillä vaikea erotella.

Alkuperäinen, SCALGOn ja Århusin yliopiston kehittämä, malli osasikin erotella juuri vettä läpäisemättömät pinnat läpäisemättömistä pinnoista. Mallin opetusaineistona käytettiin muun muassa Skanderborg vesilaitoksen manuaalisesti tuotettua maanpeiteaineistoa sekä monen tanskalaisen konsulttitoimiston tuottamia aineistoja pienemmiltä valuma-alueilta. Vuonna 2019 SCALGO tuotti ensimmäistä kertaa Tanskaan valtakunnallisen veden läpäisemättömyyttä kuvaavan aineiston, josta esimerkki kuvassa 1.



*Kuva 1. SCALGO:n tuottama veden läpäisemättömyyttä kuvaava aineisto Tanskassa, ensimmäinen versio tuotettu vuonna 2019.*

UNET-mallia on parannettu ja muokattu Suomen olosuhteisiin sopivaksi eri yhteyksissä:

- Maa- ja metsätalousministeriön rahoittamassa **LaserVesi-hankkeessa** SCALGO:n UNET-mallin avulla tunnistettiin läpäisemättömät pinnat pääkaupunkiseudulla. Hanke tehtiin yhteistyössä HSY:n, Helsingin kaupungin ja SYKE:n kanssa.
- Jatkona LaserVesi-hankkeelle HSY and SCALGO testasivat, toimiiko menetelmä myös paljaan maan tunnistamiseksi. UNET-mallia muokattiin ja opetettiin tunnistamaan myös hiekka / sora.

Molemmat aineistot ovat vapaasti saatavilla HSY-karttapalvelussa.



### 1.3 Projektin vaiheet

Maapeitteen tunnistaminen ja luokitettu tehtiin tässä projektissa neljässä eri työvaiheessa:

1. **Tekoälyluokittelu:** Lämpäsemättömien pintojen, paljaan maan, paljaan kallion, matalan kasvillisuuden, tiheän kasvillisuuden ja veden luokittelu suoritettiin UNET-neuroverkko-mallinnuksella. Malli tunnistaa luokat ilmakuviista.
2. **Pistepilviluokittelu:** Kasvillisuuden korkeus, eli maanpeitetyypin "tiheä kasvillisuus" alaluokitus, tuotettiin pistepilviaineistosta.
3. **Yhdistysvaihe:** Vektoritiedot (apuaineistot) yhdistettiin maanpeiteaineistoon.
4. **Vektorisointivaihe:** Rasteritulokset vektorisoitiin

Ensimmäisessä työvaiheessa (tekoälyluokittelu) tuotettiin useita iteraatioita. Jokainen iteraatio jaettiin HSYn kanssa, jossa projektiryhmä antoi palautetta ja auttoi tunnistamaan väärin luokiteltuja alueita. Iteraatioiden tarkastelu tuotti lisää harjoitusaineistoa, joka auttoi parantamaan mallin suorituskykyä.

### 1.4 Lopputuote

Maanpeiteaineisto kattaa kuvassa 2 (vasemmalla) esitetyn alueen. Kuvassa 2 (oikealla) on esimerkki siitä, miten aineistossa on eroteltu paljas maa (vaaleanruskea) vettä lämpäsemättömistä pinnoista (harmaa)



Kuva 2. Vasemmalla projektialue, oikealla esimerkki paljaan maan ja lämpäsemättömien pintojen luokittelusta tekoälyn avulla.

Tulokset ovat saatavilla sekä rasteri että vektorimuodossa. Lopputuote sisältää seuraavat luokat (lisäksi mainittuna työvaihe, jossa luokka on tuotettu):

Taulukko 1. Maanpeiteaineiston luokat

Luokka nro	Luokka	Työvaihe *)	Värit (RGB)
<b>1 Vettä läpäisemättömät pinnat</b>			
<b>1.1 Tiet /kadut</b>			
1.1.1	Päällystetty tie	3 Yhdistys	0, 0, 0
1.1.2	Päällystämätön tie	3 Yhdistys	130, 69, 19
1.2	Rakennukset	3 Yhdistys	216, 0, 0
1.3	Muu vettä läpäisemätön pinta	1 Tekoäly	133, 121, 118
<b>2 Viherpinta</b>			
<b>2.1 Matala kasvillisuus</b>			
2.1.1	Pellot	1 Tekoäly / 3 Yhdistys	255, 217, 128
2.1.2	Muu matala kasvillisuus	1 Tekoäly	178, 223, 67
<b>2.2 Puusto</b>			
2.2.1	2m-10m	1 Tekoäly / 2 Pistepilvi	50, 160, 40
2.2.2	10m-15m	1 Tekoäly / 2 Pistepilvi	50, 130, 40
2.2.3	15m-20m	1 Tekoäly / 2 Pistepilvi	50, 100, 40
2.2.4	>20m	1 Tekoäly / 2 Pistepilvi	50, 70, 40
<b>3 Avokalliot</b>			
3.1	Avokalliot	1 Tekoäly	208, 202, 208
<b>4 Paljas maa</b>			
4.1	Paljas maa	1 Tekoäly	205, 133, 63
<b>4 Vesi</b>			
5.1	Vesi	1 Tekoäly / 3 Yhdistys	100, 149, 237
5.2	Merialue	1 Tekoäly / 3 Yhdistys	100, 149, 237

\*) Työvaiheet kuvattuna kohdassa 1.3 Projektin vaiheet

## 2 Aineistot

Projektissa hyödynnetyt aineistot on vastaanotettu HSYltä. Projektissa käytettiin seuraavia tietolähteitä:

Taulukko 2. Projektissa hyödynnetyt aineistot

Aineisto	Kommentti	Vastaanotettu HSYltä
Ilmakuvat, CIR ja RGB	CIR ja RDG hödynnetään NDVIn tuottamisessa. RGB käytetään tekoälymallinnuksessa.	25 huhtikuuta 2022
Pistepilviaineistot	katso seuraava taulukko	25 huhtikuuta 2022
Apuaineisto - polygonit	Päällystetyt tiet Päällystämättömät tiet Rakennukset Vesistöt Meri Paljas kallio	1 heinäkuuta 2022
Apuaineisto - viivat	Voimalinjat Teiden keskiviivat	1 heinäkuuta 2022 15 kesäkuuta (Digiroad *)

\*) Käytetty UNET-mallin opetukseen

Pistepilviaineistoa kunnilla oli saatavasti seuraavasti:

Taulukko 3. Pistepilviaineisto kunnittain

Kunta	Käytettävä pistepilvi	Vuosi	Pistetiheys p/m2
Helsinki	Kunnan oma	2021	40
Espoo ja Kauniainen	Kunnan oma	2021	20-40
Vantaa	MML	2022	5
Kirkkonummi	MMI	2022, 2021	5 *)
Järvenpää	Kunnan oma	2021	30

\*) Pienessä osassa Kirkkonummea ei ollut saatavilla 5 p/m2 aineistoa. Tällöin käytettiin vanhempaa 0.5 p/m2



## 3 Maanpeiteluokittelu

Maanpeiteaineisto valmistetaan neljässä työvaiheessa.

- 1) **Tekoälymallinnus.** Tekoälyn avulla tuotetaan seuraavat maanpeiteluokat:
  1. "läpäisemätön pinta" (IV)
  2. "paljas maa" (BL)
  3. "paljas kallio" (BR)
  4. "matala kasvillisuus" (SV)
  5. "tiheä kasvillisuus" (DV)
  6. "vesi" (WA)
- 2) **Pistepilviluokitus.** Tässä työvaiheessa tuotetaan ns. High Above Ground-malli (HAG) joka jakaa tiheään kasvillisuuden, eli puuston (2.2), eri alaluokkiin (2.2.1-2.2.4).
- 3) **Yhdistysvaihe.** Aiempien työvaiheiden tulokset yhdistetään apuaineistojen kanssa (katso taulukko 2), jolloin saadaan rasteriaineisto, joka sisältää kaikki vaaditut luokat (katso taulukko 1)
- 4) **Vektorisointi.** Tuotettu rasteriaineisto vektoroidaan ja tuloksena olevat segmentit merkitään vaadituilla attribuuteilla.

Jokainen vaihe on kuvattu yksityiskohtaisesti alla.

### 3.1 Tekoälymallinnus

UNET-tyyppinen neuroverkkomalli on alun perin SCALGO:n ja Aarhusin yliopiston kehittämä. Tässä projektissa mallin suorituskykyä parannettiin edelleen tiiviissä yhteistyössä HSY:n kanssa.

UNET-mallin opetus vaatii suuret määrät opetusaineistoa. Suomen olosuhteisiin sopivaa mallia kehitettiin projektissa tuottamalla uutta opetusaineistoa seuraavasti:

- HSY ja kuntien edustajat tuottivat aineistoa piirtämällä polygoneja SCALGO:n kehittämässä validointiohjelmassa. Jokaiselle polygonille merkittiin maanpeite ja polygonit hyödynnettiin suoraan mallin opettamiseksi.
- Opetusaineistoa tuotettiin myös SCALGO:n sisäisiä, semi-automatisoituja prosesseja hyödyntäen.

UNET-malli analysoi RGB-ilmakuvia ja arvioi millä todennäköisyydellä jokainen pikseli on maanpeiteluokka:

1. "läpäisemätön pinta" (IV)
2. "paljas maa" (BL)
3. "paljas kallio" (BR)
4. "matala kasvillisuus" (SV)
5. "tiheä kasvillisuus" (DV)
6. "vesi" (WA)

Mallin lopputulos on näin ollen 6 todennäköisyyttä jokaista pikseliä kohden. Malli hyödyntää myös teiden keskiviivoja (Digiroadista). Kesiviivat auttavat mallia arvioimaan onko tie päällystetty vai päällystämätön.

UNET-mallin tulosta käytettiin edelleen seuraavissa työvaiheissa.

## 3.2 Pistepilviluokitus

HAG (Height Above Ground) -malli tuotettiin kahdessa eri vaiheessa.

### 3.2.1 Voimalinjapisteiksi luokitellut pisteet suodatetaan.

Pistepilviaineiston jokainen piste luokiteltiin kahteen luokkaan riippuen siitä, onko kyseessä voimalinja vai ei ja voimalinapisteet poistettiin pistepilviaineistosta. Tämä tehtiin seuraavasti:

- HSYn toimittamien voimalinjojen sijaintia kuvaavien linjojen avulla tunnistettiin XY-tasolla kaikki pisteet, jotka ovat alle 20 m päässä voimalinjasta.
- Merkityille pisteille annettiin 3 arvoa:
  - Tiheys: Pistetiheys valitun pisteen ympärillä (maksimi 1 m valitusta pisteestä)
  - Arvo: Lähipisteiden kovarianssimatriisin arvioitu arvo. Esimerkiksi, jos arvo on 1, se tarkoittaa, että pisteen ympäristö voidaan ilmaista viivana. Jos arvo on 2, paikallinen naapurusto voidaan ilmaista tasona jne.
  - HAG: Korkeus maanpinnasta merkitylle pisteelle.
- Pisteet luokiteltiin voimalinjapisteiksi, jos  $HAG > 0,5$ , tiheys  $< 6$  ja arvo on 1 tai 2.

### 3.2.2 Maanpinta suodatetaan pistepilviaineistosta

- Pistepilviaineiston pisteet jaettiin kahteen ryhmään: Luokan 2 maapisteet vain viimeisillä palautuksilla ja muut luokkien 2-11 ja 20-23 pisteet vain ensimmäisillä palautuksilla.
- Molemmat ryhmät rasteroitiin ja "maarasteri" vähennettiin toisesta rasterista lopullisen HAG-mallin tuottamiseksi.

## 3.3 Yhdistysvaihe

Yhdistysvaiheessa tuotettiin kaksi rasteriaineistoa:

### 3.3.1 Perustoimitus (Basic product)

Koko tulkinta-alueen ilmakeu- ja pistepilvitulkinta ja merimaski Taulukon 1 mukaisesti luokiteltuna. Pienin kartoitettava yksikkö on 25 m<sup>2</sup>.

### 3.3.2 Leikkaamaton viherpinta (Uncut green)

Viherpinnan luokittelu kokonaisuudessaan siten, että sitä ei ole leikattu tasohierarkian mukaisesti katu- ja tiealueilla eikä rakennuspolygoneilla. Leikkaamattomaan viherpintaan pätee sama luokitus ja spatiaalinen tarkkuus sekä toimitusformaatit kuin perustoimitukseen. Pienin kartoitettava yksikkö on 10 m<sup>2</sup>.

Aineistot tuotettiin yhdistämällä apuaineistot tekoälymalliin tuloksiin ja HAG-malliin tietyssä ennalta määritellyssä järjestyksessä. Menetelmä on dokumentoitu Liitteessä 1.

## 3.4 Vektorisointi

Vektoriaineisto tuotettiin perustoimituksesta apuaineistojen alkuperäisiä vektoreja hyödyntäen. Vektorisointi sisälsi seuraavat vaiheet:

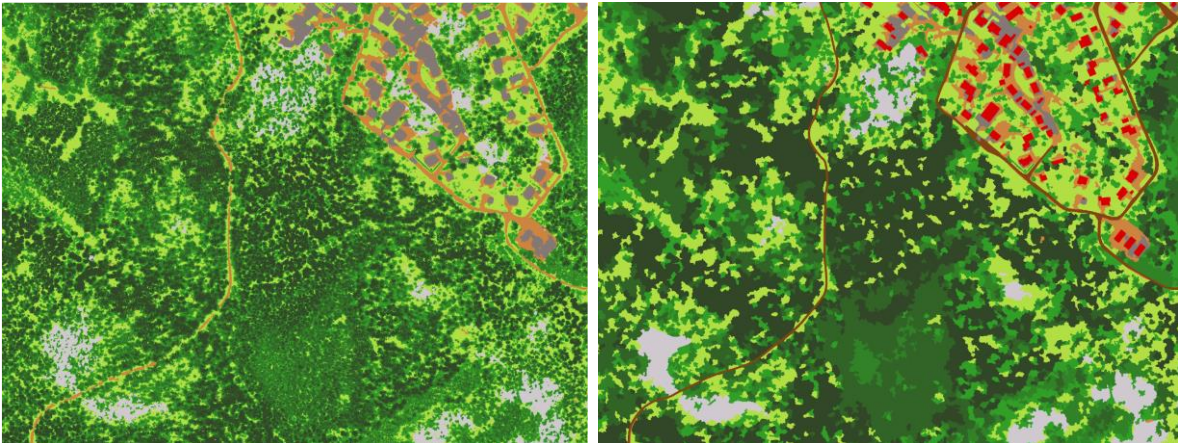
- 1) Perustoimituksesta poistetaan apuaineistoja vastaavat luokat koska rasterisointi on muokannut näitä luokkia. Kyseessä on seuraavat luokat: päällystetty tie, päällystämätön tie, meri ja vesistö.
- 2) Jäljellä olevia segmenttejä laajennetaan hieman. Eli, jos segmentin vieressä on *nodata* pikseleitä, segmenttiä laajennetaan näihin pikseleihin. Tällöin varmistamme, että kun lisäämme apuaineistot yksinkertaistamisen jälkeen, vektoriaineistoon ei tule aukkoja.

- 3) Rasteriaineisto vektoroidaan ja yksinkertaistetaan niin, että tulos on tasainen ilman turhia, rasterista tulevia kulmia. Tässä hyödynnetään SCALGOn omaa metodia (ei avoimesti saatavilla).
- 4) Itseen leikkaavista polygoneista poistetaan tarvittavat leikkauspisteet.
- 5) Rasterointi yksinkertaistaminen muuttaa polygonien geometriaa niin, että ne mahdollisesti menevät päällekkäin apuaineistojen kanssa. Mahdolliset päällekkäisyydet poistetaan ennen apuaineistojen lisäämistä.
- 6) Tiepolygonit leikataan sopiviksi HSYn toimittaman ruudukon mukaisesti. Tarvittaessa segmentit yhdistetään (jos alle 25m<sup>2</sup>).
- 7) Kuntarajat liitetään aineistoon ja vaaditut attribuutit lisätään.

## 4 Rasteri ilman segmenttien yhdistämistä

Osana toimitusta tuotettiin rasteri ilman segmenttien yhdistämistä, koska yhdistäminen suosii suuria segmenttejä ja tämä poistaa tietoa luokista, joissa on paljon pieniä segmenttejä, esimerkiksi puusto (yksittäiset puut) jota ympäröi matala kasvillisuus.

Kuvassa 3 esimerkki siitä, miten segmenttien yhdistäminen vaikuttaa lopputulokseen.



*Kuva 3. Vasemmalla: Alkuperäiset kasvillisuussegmentit ovat hyvin pieniä. Oikealla: Segmenttien yhdistäminen yksinkertaistaa kasvillisuuden kartoitusta.*

Liite 3 sisältää yksityiskohtaisen menetelmäkuvauksen tämän (ei-yhdistetyn) rasterin tuottamisesta.

## 5 Validointi

Aineisto validoidaan käyttäen 25 polygonia, jotka on piirretty SCALGOn tai HSYn toimesta. Yhteensä nämä polygonit sisältävät 7 781 452 pikseliä.

Arvioimme tekoälymallinnuksen tuottamia luokkia lopullisten luokkien (ennen yhdistämistä) sijaan koska apuaineistojen tarkkuus parantaisi validointituloksia suhteettoman paljon.

Validoinnin tulokset näkyvät alla. Tulosten kuvaukseen on käytetty ns. Confusion Matrix<sup>1</sup>.

Tekoälymalli tuotti seuraavat luokat:

1. "läpäisemätön pinta" (IV)
2. "paljas maa" (BL)
3. "paljas kallio" (BR)
4. "matala kasvillisuus" (SV)
5. "tiheä kasvillisuus" (DV)
6. "vesi" (WA)

	IV	BL	SV	DV	WA	BR
IV	<b>0.986615</b>	0.012858	0.000000	0.000016	0.000000	0.000511
BL	0.013562	<b>0.974160</b>	0.011492	0.000286	0.000000	0.000500
SV	0.000002	0.004858	<b>0.943068</b>	0.048245	0.000000	0.003827
DV	0.000006	0.000640	0.014371	<b>0.984983</b>	0.000000	0.000000
WA	0.000000	0.187369	0.072335	0.004244	<b>0.736051</b>	0.000000
BR	0.000008	0.000000	0.000000	0.002297	0.000000	<b>0.997695</b>

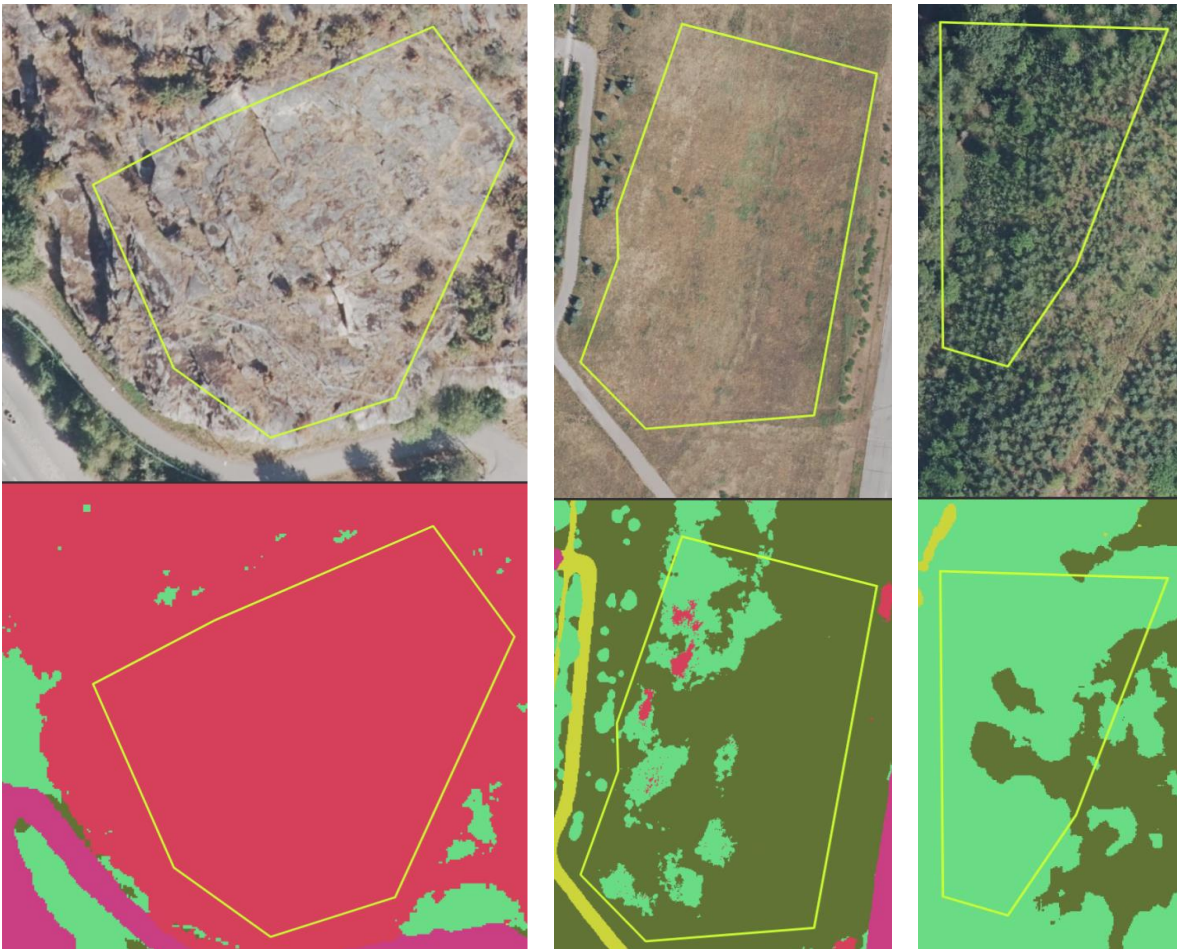
Tekoälymallin suorituskyky on yleisesti ottaen hyvä, paitsi vesiluokissa, joissa mallilla on taipumus ennustaa paljasta maata ja matalaa kasvillisuutta. Tämä on kuitenkin pieni ongelma, koska kaikki virheet ovat vesistö- ja meripolygonien sisällä, osana apuaineistoja, ja ne voidaan helposti korjata. Lopputuotteessa luokittelimme siis alueet vesistöksi /mereksi, vaikka malli ennustaa paljasmaata tai matalaa kasvillisuutta. Tämä tarkoittaa sitä, että vesiluokan todellinen tarkkuus on 0,995755, joka on selvästi 80 %:n rajan yläpuolella.

<sup>1</sup> [https://en.wikipedia.org/wiki/Confusion\\_matrix#Table\\_of\\_confusion](https://en.wikipedia.org/wiki/Confusion_matrix#Table_of_confusion)



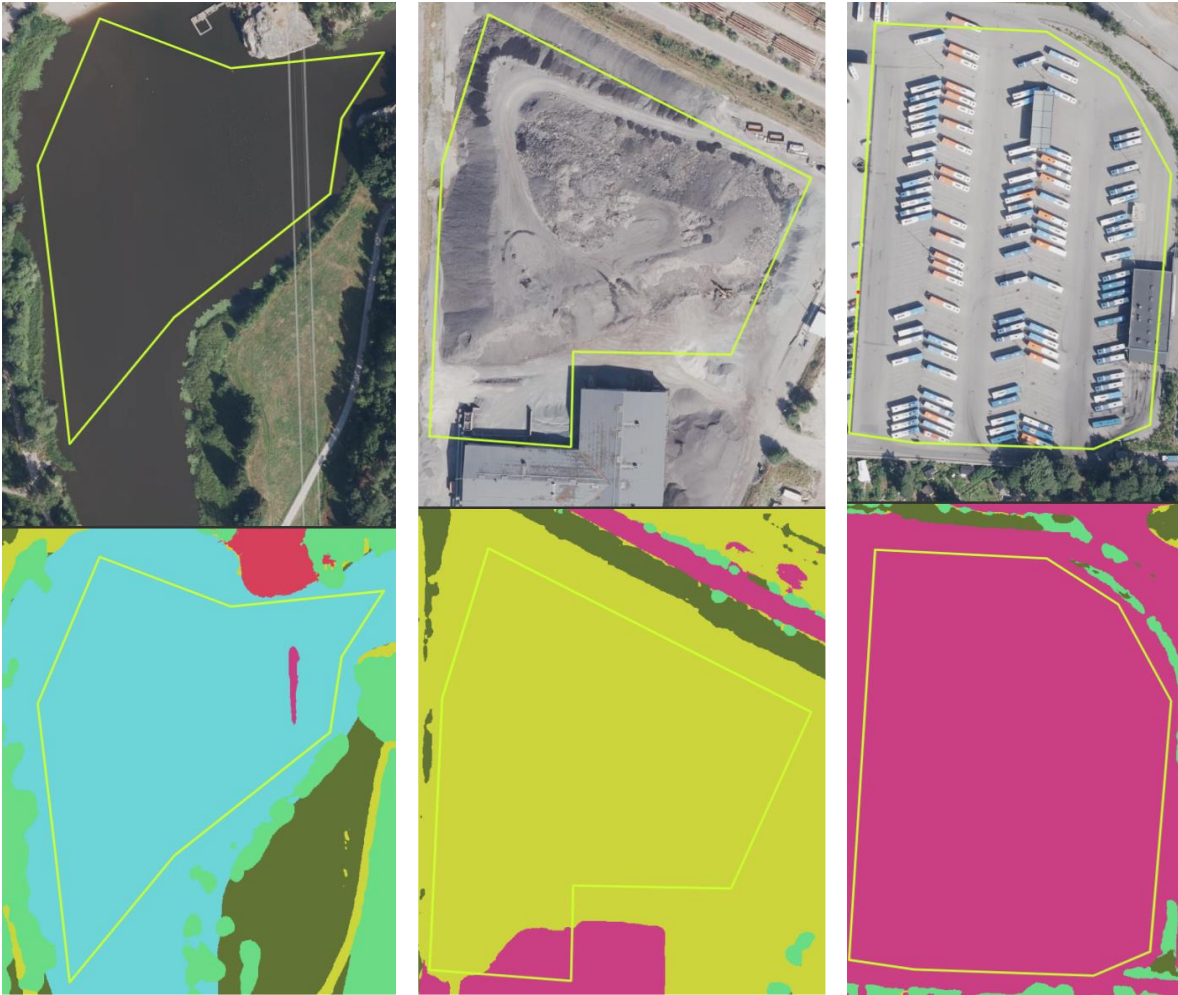
Perustoimituksessa pieniä maanpeitesegmenttejä on yhdistetty suurempiin osiin. Yhdistäminen on tehty HSYn vaatimusten mukaisesti. Yhdistämisessä ei huomioida RGB-kuvia, jolloin lopputuloksen tarkkuus kärsii eteenkin maanpeiteluokissa, joissa on paljon pieniä segmenttejä, esimerkiksi puusto. Emme arvioi miten paljon yhdistäminen vaikuttaa tuloksen tarkkuuteen, mutta silmämääräisesti ero on huomattava. Tämän takia olemme myös tuottaneet rasterin ilman segmenttien yhdistämistä (katso kappale 4 ja liite 3).

Alla on esimerkkejä validoinnissa käytetyistä polygoneista.



Kuva 4. Vasen: paljas kallio, Keskellä: matala kasvillisuus, Oikealla: tiheä kasvillisuus





*Kuva 5. Vasemmalla: Vesi, keskellä: paljas maa, Oikealla: päällystetty pinta*

## Liite 1 Yhdistysvaihe

Perustoimituksen (**Basic product**) yhdistäminen on dokumentoitu yksityiskohtaisesti (katso 1) Basic product). Leikkaamattoman viherpinnan (**Uncut green**) teko seuraa pitkälti perustoimituksen metodia. Erot on kuvattu osiossa "2) Uncut green".

Seuraavassa osiossa "ml" vastaa Tekoälyvaiheen maanpeiteluokkia (ml = machine learning), eli tekoälyvaiheen lopputuloksia jossa arvioitiin todennäköisyys luokille:

1. Lämpäisemätön pinta = other impervious
2. Paljas maa = bare land
3. Paljas kallio = bare rock
4. Matala kasvillisuus = shallow vegetation
5. Tiheä kasvillisuus = dense vegetation
6. Vesi = fresh water or sea

Lyhenne "aux" kuvaa apuaineistoja (aux= auxillary data). Yhdistysvaiheessa käytetyt apuaineistot ovat:

1. Päällystetyt tiet = Paved
2. Päällystämättömät tiet =Unpaved
3. Rakennukset = Buildings
4. Vesistöt = Fresh water
5. Meri = Sea
6. Paljas kallio = Bare rock

### 1) BASIC PRODUCT:

1. Tuotamme rasterin tekoälymallinnuksen lopputuloksesta, eli jokainen pikseli luokitellaan:

$$p = \max(ml_{other-impervious}, ml_{bare-land}, ml_{shallow-vegetation}, ml_{dense-vegetation})$$

2. Jos  $p = ml_{shallow-vegetation}$  tai  $ml_{dense-vegetation}$ , luokittelemme pikselin uudelleen HAG-mallin arvon perusteella, HSY:n määritelmien mukaisesti.

3. Määritämme että  $p = other\_impervious$  jos  $ml_{other-impervious} > 0,5$ , koska haluamme "other impervious"-luokan (muu vettälämpäisemätön) hallitsevan muita luokkia, esimerkiksi parkkipaikan päällä oleva puusto.

4. Otamme käyttöön bare-rock luokan määrittämällä  $p = bare\_rock$  jos  $ml_{bare-rock} > 0,5$  ja  $aux_{bare-rock} = True$

5. Määritämme että:

$$aux_{mask} = aux_{paved} \text{ tai } aux_{unpaved} \text{ tai } aux_{building} \text{ tai } aux_{fresh-water} \text{ tai } aux_{sea}$$

ja että  $p = nodata$  jos pikseli on  $aux_{mask}$ :in sisäpuolella. Tämä tehdään, koska seuraavat segmenttien yhdistämisvaiheet tehdään ilman  $aux_{mask}$ :ia.

6. Suoritamme pääasiallisen segmenttien yhdistämisen, jossa varmistamme, ettei rasterissa ole alle 25 m<sup>2</sup>:n segmenttejä. Tämän vaiheen yksityiskohdat on kuvattu Liitteessä 2.
7. Yhdistämme *bare-land* segmentit *other-impervious* segmenttiin, jos *bare-land* on pienempi kuin 200m<sup>2</sup> ja vähintään 90 % ympärysmistä on jaettu *other-impervious*-luokan kanssa ja päinvastoin.
8. Lisäämme apuaineistot:
- Pellot: Määritämme että  $p = field$  jos pikseli ei ole jo merkitty puustoksi (eli 2.2\*)
  - Vesistö, Meri, Päälystämätön tie, Päälystetty tie, Rakennus: Määritämme että  $p = aux_*$ , jossa \* on jokin seuraavista, *fresh-water*, *sea*, *unpaved*, *paved* ja *building*. Aineistot lisätään juuri tässä järjestyksessä. Tämä on tärkeää, jotta vältämme konfliktit.
9. Apuaineistojen välisen pienen kohdistusvirheen vuoksi rasterointi luo yksittäisiä aukkoja. Yhdistämme nämä viereiseen segmenttiin, jolla on suurin yhteinen ympärysmitta.

## 2) UNCUT GREEN

Leikkaamaton viherpinta (Uncut green) noudattaa samoja vaiheita, kun lopputuote (Basic product) pienin eroin:

1. Tuotamme binäärisen maskin kasvillisuudelle hyödyntäen tekoälymallin tuloksia:

$$p = ml_{shallow-vegetation} \text{ tai } ml_{dense-vegetation}$$

2. Jos  $p$ , luokittelemme pikselin uudestaan HAG-mallin mukaan
3. Suoritamme pääasiallisen segmenttien yhdistämisvaiheen, jossa varmistamme, ettei rasterissa ole alle 10 m<sup>2</sup>:n segmenttiä. Tämän vaiheen yksityiskohdat on kuvattu Liitteessä 2.

- Poistamme väärät positiiviset vesistöissä ja meressä yhdistämällä *fresh-water* ja *sea* maskit, ja tämän jälkeen tuotetaan maskille vaikutusalue, buffer -1. Tällöin tuotamme maskin *k* jolle annetaan sääntö:

*jos k = True*<sup>2</sup> ja (*p = 2.1.2* tai *p = 2.2.1*) niin *p = nodata*

- Lisäämme pellot (*fields*, 2.1.1). Puusto priorisoidaan säännöllä:

*jos field = True*<sup>3</sup> ja (*p = 2.1.2* tai *p = nodata*) niin *p = 2.1.1*

---

<sup>2</sup> *k=True* tarkoittaa sitä, että analysoitava pikseli on vesistö/meri-maskin sisällä.

<sup>3</sup> *field = True* tarkoittaa sitä, että analysoitava pikseli on pelto-polygonin sisällä

## Liite 2 Segmenttien yhdistäminen

Yhdistysalgoritmi toimii rasteritiedoilla ja varmistaa, että mikään yhdistetty segmentti ei ole pienempi kuin jokin koko  $s$ . Algoritmi yhdistää segmentit seuraavasti:

- 1) Etsi rasterista kaikki toisiinsa kytketyt segmentit käyttämällä vain vaaka- ja pystysuuntaista yhteyttä. Toisin sanoen kahta pikseliä pidetään yhdistettynä, jos niillä on sama arvo ja niillä on joko pysty- tai vaakasuuntainen reuna, esim. diagonaaliliitännöjä ei huomioida.
- 2) Valitse kaikki liitetyt komponentit, jotka ovat samassa luokassa ja pienempiä kuin  $s$ . Käsittele niitä kasvavan koon mukaan.
- 3) Jos segmentin koko on edelleen pienempi kuin  $s$ , yhdistä segmentti naapuriin, jonka kanssa segmentillä on pisin yhteinen rajapinta. Jos segmenttiä ei voida yhdistää, esim. segmentti ei jaa rajaa kelvollisen luokan kanssa, segmentti ohitetaan.
- 4) Toista vaiheet 1-3, kunnes  $s$ :tä pienempien segmenttien määrä ei muutu.

## Liite 3 Rasteri ilman segmenttien yhdistämistä

Seuraavassa osiossa "ml" vastaa Tekoälyvaiheen maanpeiteluokkia (ml = machine learning), eli tekoälyvaiheen lopputuloksia jossa arvioitiin todennäköisyys luokille:

1. Läpäisemätön pinta = other impervious
2. Paljas maa = bare land
3. Paljas kallio = bare rock
4. Matala kasvillisuus = shallow vegetation
5. Tiheä kasvillisuus = dense vegetation
6. Vesi = fresh water or sea

Rasteri ilman segmenttien yhdistämistä tuotetaan seuraavasti:

1. Tuotamme rasterin tekoälymallin tuloksista jokaiselle pikselille:

$$p = \max(ml_{other-impervious}, ml_{bare-land}, ml_{shallow-vegetation}, ml_{dense-vegetation}, ml_{bare-rock})$$

2. Jos  $p = ml_{shallow-vegetation}$  tai  $ml_{dense-vegetation}$  luokittelemme pikselin uudelleen HAG-mallin arvon perusteella.
3. Määritämme että  $p = other\_impervious$  jos  $ml_{other-impervious} > 0,5$ , koska haluamme että *other-impervious* dominoi muita luokkia, esimerkiksi parkkipaikan päällä oleva puusto.
4. Määritämme että  $p = bare\_land$  jos  $ml_{bare-land} > 0,5$  ja  $p = shallow\_vegetation$ , koska se antaa kaiken kaikkiaan paremman suorituskyvyn.
5. Jos pikseli on merimaskin sisällä, merkitsemme *nodata*.