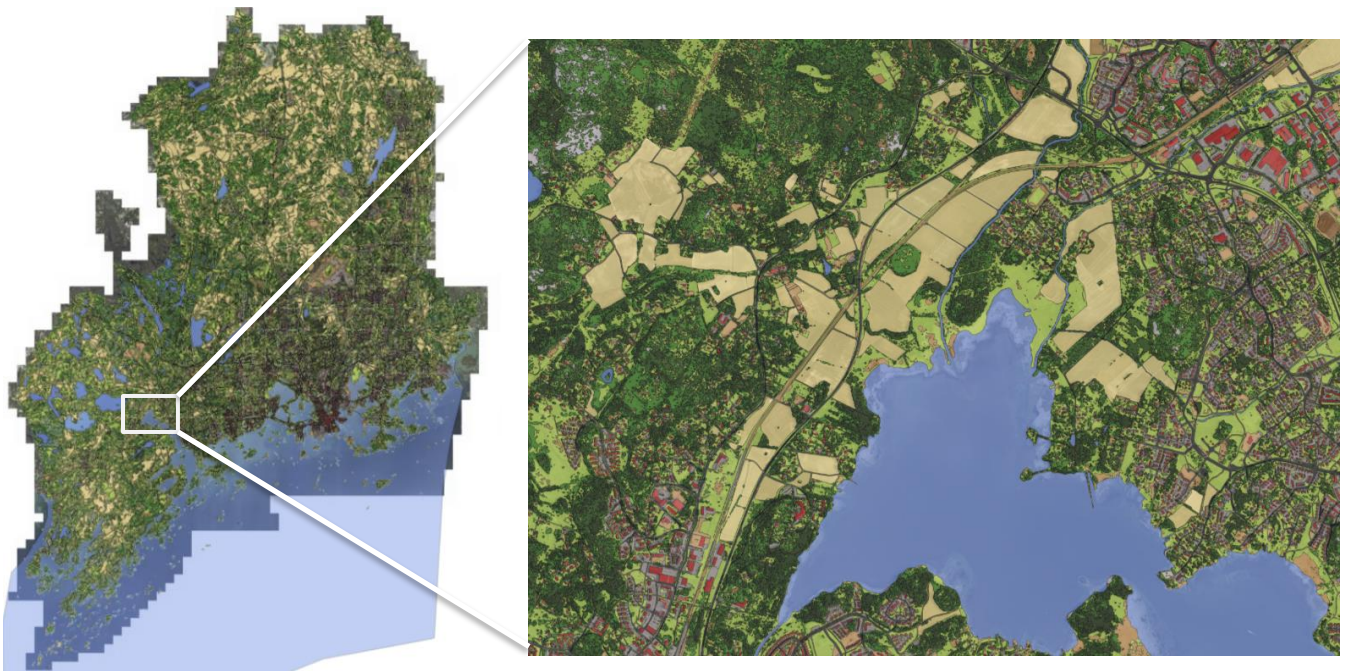


# Helsingin seudun maanpeiteaineisto 2024

Loppuraportti



## Sisällysluettelo

1	Johdanto .....	3
1.1	Projektin kuvaus .....	3
1.2	Tausta .....	3
1.3	Aikataulu ja työvaiheet .....	5
1.4	Projektin vaiheet .....	6
1.5	Lopputuote .....	7
2	Aineistot .....	9
3	Maanpeiteluokittelu .....	10
3.1	Tekoälymallinnus .....	10
3.2	Pistepilviluokitus .....	11
3.2.1	Voimalinjapisteet .....	11
3.2.2	Maanpinnan suodatus pistepilviaineistosta .....	12
3.3	Yhdistysvaihe .....	12
3.3.1	Perustoimitus (Basic product) .....	12
3.3.2	Leikkaamaton viherpinta (Uncut green) .....	12
3.4	Vektorisointi .....	12
4	Validointi .....	13
	Liite 1 Yhdistysvaihe .....	15
	Liite 2 Segmenttien yhdistäminen .....	18

# 1 Johdanto

Tässä raportissa kuvataan "Helsingin seudun maanpeiteaineisto 2024" -hankkeen lähtötiedot, menetelmät, syöttötiedot, sekä tulokset ja niiden validointi. Hankintasopimus laadittiin HSY:n ja yhteistyökuntien sekä Scalgon kanssa, ja tämä raportti kuvaa hankkeen ja sen vaatimusten mukaisten aineistojen kokoamiseen sekä tulosten tuottamiseen liittyvät tiedot.

## 1.1 Projektin kuvaus

Hankkeen tavoitteena on ollut tuottaa ilma- ja pistepilviaineistosta tekoälytulkintainen maanpeiteaineisto pääkaupunkiseudun yhteistyökunnille, joihin lukeutuu Espoo, Helsinki, Kauniainen, Vantaa, Järvenpää, Kirkkonummi, Kerava, Tuusula ja Nurmijärvi, mukaan lukien saaret, joiden pinta-ala on Espoon ja Kirkkonummen alueella vähintään 1 ha ja Helsingin alueella vähintään 0,3 ha. Projektialueen kokonaispinta-ala on 1849,2 km<sup>2</sup>.

Projektissa tuotettuun 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 2023 ilmakuvia ja pistepilviaineistoja hyödyntäen, sisältäen lähi-infrapunakanavan. Mahdollisuuksien mukaan pistepilviaineistoja on kerätty kunnilta ja jos kunnilla ei ole ollut omaa pistepilviaineistoa, on alueella käytetty Maanmittauslaitoksen 5 p/m<sup>2</sup> aineistoa.

Lisäksi HSY toimitti Scalgolle 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

enimmäistä kertaa Tanskaan valtakunnallisen veden läpäisemättömyyttä kuvaavan aineiston, josta esimerkki kuvassa 1.

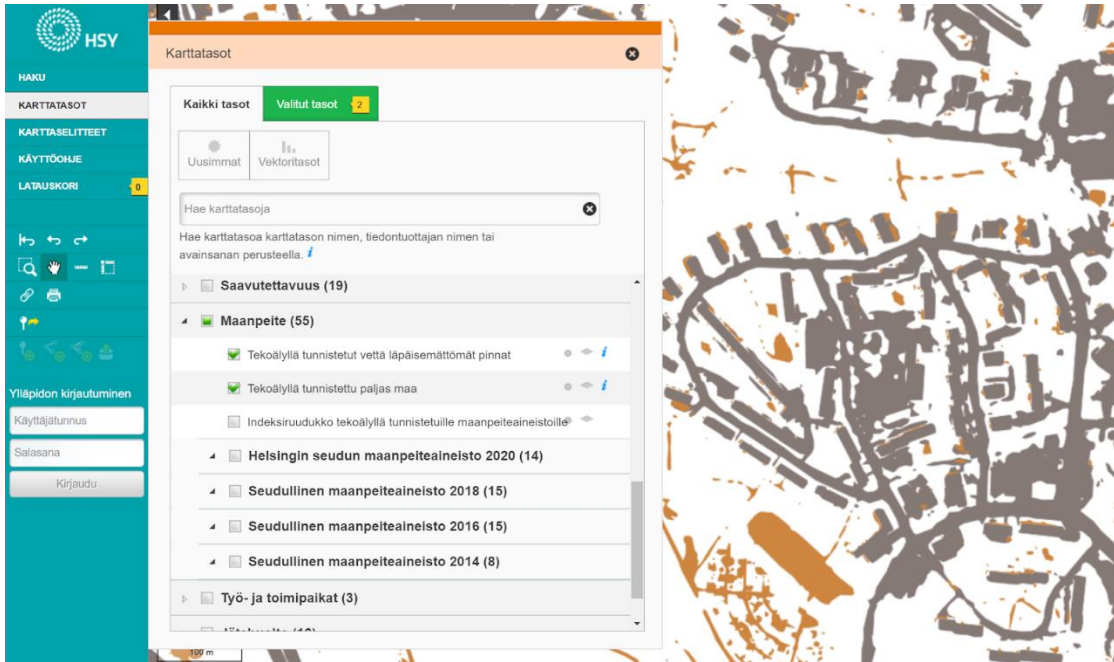


Kuva 1. Scalgon tuottama veden läpäisemättömyyttä kuvaava aineisto Tanskassa, ensimmäinen versio tuotettu vuonna 2019.

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

- Maa- ja metsätalousministeriön rahoittamassa **LaserVesi-hankkeessa** Scalgon UNET-mallin avulla tunnistettiin läpäisemättömät pinnat pääkaupunkiseudulla. Hanke tehtiin yhteistyössä HSYn, Helsingin kaupungin ja SYKEN kanssa.
- Jatkona LaserVesi-hankkeelle HSY ja Scalgo testasivat, toimiiko menetelmä myös paljaan maan tunnistamiseksi. UNET-mallia muokattiin ja opetettiin tunnistamaan myös hiekka / sora.
- Vuoden 2022 **Seudullinen maanpeiteaineisto -hankkeessa**, jossa tuotettiin pääkaupunkiseudun alueen kunnille entistä tarkempi maanpeitemalli. Hankkeen vaiheissa mallia iteroitiin projektiryhmän toimesta, mikä paransi mallin luotettavuutta.

Aineistot ovat vapaasti saatavilla [HSY:n karttapalvelussa](#), ks. kuva 2.



Kuva 2. Aineisto vuodelta 2024 esitettynä karttatasoikkunassa HSY:n karttapalvelussa.

### 1.3 Aikataulu ja työvaiheet

Projekti käynnistettiin toukokuussa 2024 kick off -tapaamisella (Scalgo ja HSY). Työn edistymistä seurattiin sekä erikseen sovittujen kokouksien että aineiston väliaikatietojen jakamisen pohjalta. Tilaaja antoi kommentit väliaikavaiheissa tuotettuun aineistoon, jonka hyväksymisen jälkeen edettiin seuraavaan työvaiheeseen. Hanke jaettiin tarjousvaiheessa esitetysti neljään erilliseen työvaiheeseen työn edistymisen osalta:

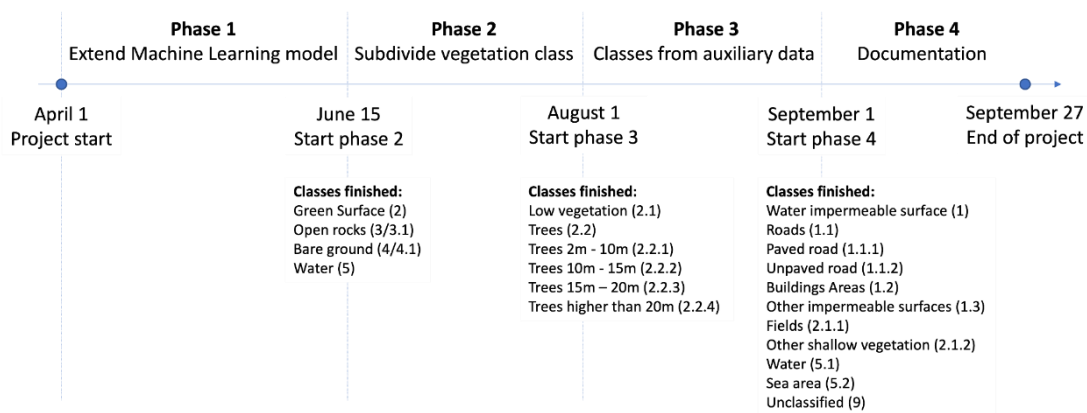
**Vaihe 1:** Scalgon olemassa olevaa tekoälyluokitusta hienosäädettiin, jotta lähtökohtainen maanpeitteen luokittelu olisi entistä tarkempaa. Vuoden 2022 tehdyn Seudullinen maanpeiteaineisto hankkeen lisäksi malli on saanut paljon mallin opetukseen käytettävää testiaineistoa Ruotsin, Norjan ja Tanskan maanpeiteaineistoista. Loppuvaiheessa jalostettiin koneoppimismalli, jota voitiin hyödyntää aineiston luokitteluun.

**Vaihe 2:** Ns. vihreät pinnat jaettiin yhä edelleen kuuteen kasvillisuusluokkaan niiden HAG-mallin perusteella. Tässä prosessin vaiheessa sellaiset korkeuspisteet, jotka eivät vastanneet kasvillisuutta (kuten sähkölinjat) poistettiin vihreitä pintoja kuvaavasta aineistosta. Jäljelle jääneestä pisteaineistosta luotiin pintamalli, jossa kasvillisuus oli luokiteltu korkeuden mukaan. Metsäalueet

segmentointiin ei-metsäalueista klusterointi-algoritmeja käyttäen (esim. DBSCAN). Yhtenäiset metsäalueet luokiteltiin HAG-mallin avulla, jotta eri keilausaineistojen datan resoluutio voitiin asettaa toisiaan vastaaviksi.

**Vaihe 3:** Apuaineisto (rakennukset, tiet, vesistöt ja kentät) yhdistettiin vaiheessa 2 tuotettuun malliin. Aineistojen väliset ristiriidat tarkastettiin, ja sen jälkeen tuotettiin lopullinen luokitteluaineisto. Pikselöity malli ”pehmennettiin”, jotta mallista tulee visuaalisesti parempi.

**Vaihe 4:** Aineisto toimitettiin tilaajille sekä tuotettiin tämä loppuraportti, jossa kuvataan aineisto, menetelmät ja tulokset.



Kuva 2. Tarjousvaiheen alustava aikataulu sekä työn vaiheistus.

## 1.4 Projektin vaiheet

Maapeitteen tunnistaminen ja luokitettu tehtiin tässä projektissa neljässä eri työvaiheessa:

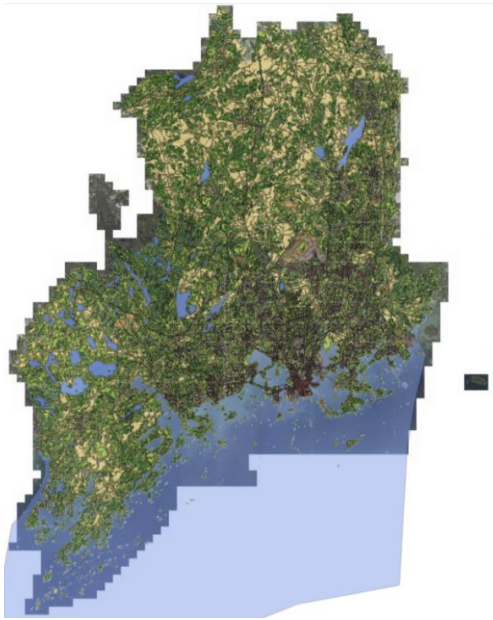
- Tekoälyluokittelu:** Lämpisemättömien pintojen, paljaan maan, paljaan kallion, matalan kasvillisuuden, tiheän kasvillisuuden ja veden luokittelu suoritettiin UNET-neuroverkkomallinnuksella. Malli tunnistaa luokat ilmakuvista.
- Pistepilviluokittelu:** Kasvillisuuden korkeus, eli maanpeitetyypin "tiheä kasvillisuus" alaluokitus, tuotettiin pistepilviaineistosta.
- Yhdistysvaihe:** Vektoritiedot (apuaineistot) yhdistettiin maanpeiteaineistoon.
- Vektorisointivaihe:** Rasteritulokset vektorisoitiin

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

## 1.5 Lopputuote

Lopputuotteena toimitetaan alueen kattava maanpeitemalli vaatimusten mukaisissa formaateissa sekä vaadituilla tarkkuustasoilla sekä loppuraportti.

Alla esitetyissä kuvissa 3 ja 4 on koko hankealueen maanpeitemalli sekä tarkastelu erilaisista pinnoista, ensin ilmakuvana ja sitten maanpeitteen tulkintana.



*Kuva 3. Maanpeitemalli koko projektialueesta, merimaskit ja saaret mukaan lukien.*



*Kuva 4. Ortokuva, ortokuva-maanpeitemalli-yhdistelmä sekä maanpeitemalli esimerkkikohteesta (tässä järjestyksessä).*

Tulokset ovat saatavilla sekä rasteri- että vektorimuodossa. Lopputuote sisältää seuraavat taulukossa 1 esitetyt 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>5 Vesi</b>			
5.1	Vesistöt (järvet ja joet)	1 Tekoäly / 3 Yhdistys	100, 149, 237
5.2	Merialue	1 Tekoäly / 3 Yhdistys	100, 149, 237

\*) Työvaiheet kuvattuna kohdassa 1.4 Projektin vaiheet

Lopputuote sekä formaatit toimitetaan tilaajan vaatimusmäärittelyn mukaisesti (Liite 2, Helsingin seudun maanpeiteaineisto 2024: Vaatimusmäärittely).



## 2 Aineistot

Projektissa hyödynnetyt aineistot on vastaanotettu HSY:ltä. Projektissa käytettiin seuraavia tietolähteitä:

Taulukko 2. Projektissa hyödynnetyt aineistot

Aineisto	Kommentti	Vastaanotettu HSY:ltä
Ilmakuvat	Tulkittavien ilmakuvien määrä oli 2492 kpl	maalis-huhtikuu 2024
Pistepilviaineistot	Ks. Taulukko 3.	maalis-huhtikuu 2024
Apuaineisto - polygonit	Päällystetyt tiet Päällystämättömät tiet Rakennukset Vesistöt Meri Paljas kallio	elokuu 2024
Apuaineisto - viivat	Voimalinjat Teiden keskiviivat	helmikuu 2024

Pistepilviaineistoa kunnilla oli saatavasti seuraavasti taulukossa 3 esitetysti:

Taulukko 3. Pistepilviaineisto kunnittain

Kunta	Käytettävä pistepilvi	Vuosi	Pistetiheys p/m <sup>2</sup>	Luokiteltu
Helsinki	Kunnan oma	2023	60	kyllä
Espoo ja Kauniainen	Kunnan oma	2021	20–40	kyllä
Vantaa	Kunnan oma	2022	5	kyllä
Kirkkonummi	MML	2020 / 2021	5	kyllä
Järvenpää	Kunnan oma	2021	30	kyllä
Kerava	Kunnan oma	2021	31	kyllä
Tuusula	Kunnan oma + MML	2022, MML 2020 ja 2023	30, MML 5	kyllä
Nurmijärvi	Kunnan oma + MML	2022, MML 2020 ja 2023	2022 Klaukkala, muut 5	kyllä

Tilaja toimitti seuraavat apuaineistot palvelun tuottajalle (ei omia apuaineistoja):

Taulukko 4. Tilaajan toimittama apuaineisto ja lähde

Apuaineisto	Lähde
tiestö päällystetietoineen	kunnat ja Digiroad
rakennukset	kunnat
sisävesialueet	MML maastotietokanta ja kunnat
pellot	MML maastotietokanta
suurjännitelinjat	MML maastotietokanta ja kunnat

## 3 Maanpeiteluokittelu

Maanpeiteaineisto valmistetaan neljässä työvaiheessa.

- 1) **Tekoölymallinnus.** Tekoölyn 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. Height Above Ground-malli (HAG) joka jakaa tiheän 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 Tekoölymallinnus

UNET-tyyppinen neuroverkkomalli on alun perin Scalgon 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 Scalgon kehittämässä validointiohjelmassa. Jokaiselle polygonille merkittiin maanpeite ja polygonit hyödynnettiin suoraan mallin opettamiseksi.
- Opetusaineistoa tuotettiin myös Scalgon 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 Voimalinjalisteet

Pistepilviaineiston jokainen piste luokiteltiin kahteen luokkaan riippuen siitä, onko kyseessä voimalinja vai ei ja voimalinjalisteet 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 voimalinjalisteiksi, jos  $HAG > 0,5$ , tiheys  $< 6$  ja arvo on 1 tai 2.

### 3.2.2 Maanpinnan suodatus 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>. Vuoden 2022 hankkeesta poiketen, tässä hankkeessa ei tuotettu rasteria ilman segmenttien yhdistämistä.

### 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 HSY:n toimittaman ruudukon mukaisesti. Tarvittaessa segmentit yhdistetään (jos alle 25m<sup>2</sup>).
- 7) Kuntarajat liitetään aineistoon ja vaaditut attribuutit lisätään.

## 4 Validointi

Aineisto validoidaan käyttäen 16 polygonia, jotka on piirretty Scalgon tai HSY:n toimesta. Yhteensä nämä polygonit sisältävät 1 889 822 pikseliä. Arvioimme tekoälymallinnuksen tuottamia luokkia lo-pullisten 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>. Tekoälymallin suorituskyky on yleisesti ottaen hyvä, selvästi 80 %:n rajan yläpuolella. Tekoälymalli tuotti seuraavat luokat:

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

	IV	BL	SV	DV	BR
IV	<b>0.998375</b>	0.000041	0.001161	0.000423	0.00000
BL	0.000000	<b>1.000000</b>	0.000000	0.000000	0.00000
SV	0.000180	0.000000	<b>0.999820</b>	0.000000	0.00000
DV	0.000000	0.000000	0.000865	<b>0.999135</b>	0.00000
BR	0.000000	0.000000	0.000000	0.000220	<b>0.99978</b>

<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 sulautettu suurempiin osiin. Sulauttaminen on tehty HSY:n vaatimusten mukaisesti. Siinä ei huomioida RGB-kuvia, jolloin lopputuloksen tarkkuus kärsii eteenkin maanpeiteluokissa, joissa on paljon pieniä segmenttejä, esimerkiksi puusto. Vuoden 2022 hankkeessa tämä huomioitiin tuottamalla rasteri ilman segmenttien yhdistämistä (katso kappale 4 ja liite 3), mutta vuoden 2024 hankkeessa yhdistysvaiheeseen on lisätty vaihe, jossa alle 240 m<sup>2</sup> kokoiset metsäalueet yhdistettiin yhtenäisempien metsäalueiden tuottamiseksi (liite 1, kohta 10).

## 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:
 
$$aux_{mask} = aux_{paved} \text{ tai } aux_{unpaved} \text{ tai } auc_{building} \text{ tai } aux_{fresh-water} \text{ tai } aux_{sea}$$
 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 % ympärysmitasta on jaettu *other-impervious*-luokan kanssa ja päinvastoin.
8. Lisäämme apuaineistot:
  - a. Pellot: Määritämme että  $p = field$  jos pikseli ei ole jo merkitty puustoksi (eli 2.2\*)
  - b. 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.
10. Ylimääräisenä vaiheena kooltaan alle 240 m<sup>2</sup> kokoiset metsäalueet sulautettiin yhtenäisempien metsäalueiden tuottamiseksi; matala kasvillisuus, paljas maa ja avokallio, jos kyseisen segmentin vieressä on puustoksi luokiteltava segmentti. Jos kyseessä on useampi kuin yksi rinnakkainen puustoksi luokiteltava segmentti, ne sulautetaan siihen, jolla on näistä pisin yhteinen raja.



## 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 yhdistämisen vaiheen, jossa varmistamme, ettei rasterissa ole alle 10 m<sup>2</sup>:n segmenttiä. Tämän vaiheen yksityiskohdat on kuvattu Liitteessä 2.
4. 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:

$$\text{jos } k = True^2 \text{ ja } (p = 2.1.2 \text{ tai } p = 2.2.1) \text{ niin } p = \text{nodata}$$

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

$$\text{jos } field = True^3 \text{ ja } (p = 2.1.2 \text{ tai } p = \text{nodata}) \text{ niin } p = 2.1.1$$

---

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

<sup>3</sup>  $field = True$  tarkoittaa sitä, että analysoituva 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) Toistetaan vaiheet 1–3, kunnes  $s$ :tä pienempien segmenttien määrä ei muutu.