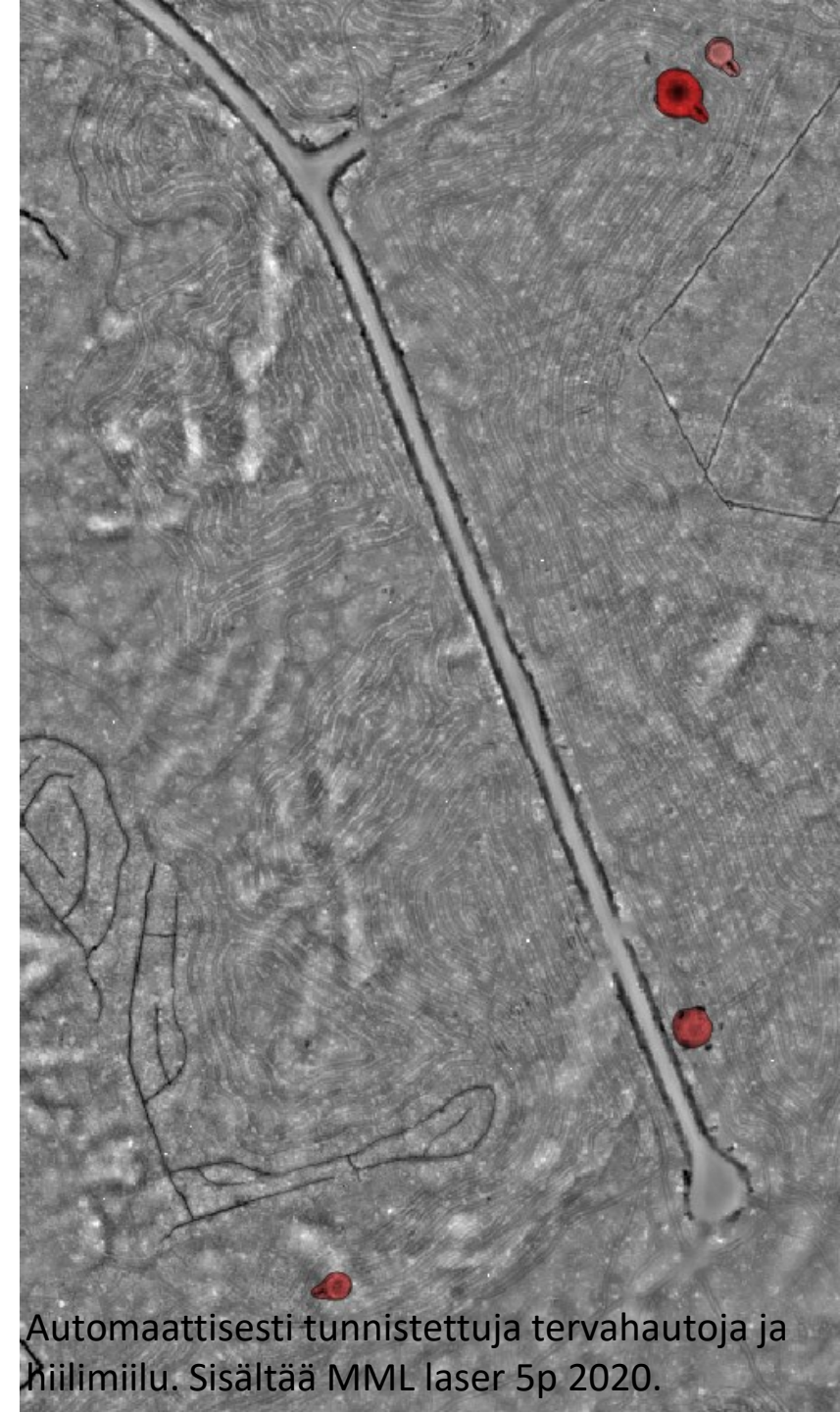


Arkeologisten kohteiden tekoälyavusteinen tunnistus laserkeilausdatasta

Niko Anttiroiko 10.5.2023



Museovirasto
Museiverket



Automaattisesti tunnistettuja tervahautoja ja
hiilimiilu. Sisältää MML laser 5p 2020.

LIDARK-hankkeen taustat

Tavoitteet:

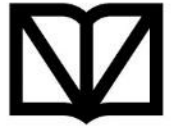
- Kehittää menetelmiä arkeologisten kohteiden automaattiseen tunnistamiseen 5p laserkeilausaineistoista.

Työryhmä:

- Lidark-konsortio: Museovirasto & Oulun yliopisto
- Yhteistyökumppanit: Blom kartta oy, Maanmittauslaitos, Metsähallitus & alueelliset vastuumuseot

Rahoitus:

- 2021-2022 Maa- ja metsätalousministeriö



**Museovirasto
Museiverket**



**OULUN
YLIOPISTO**

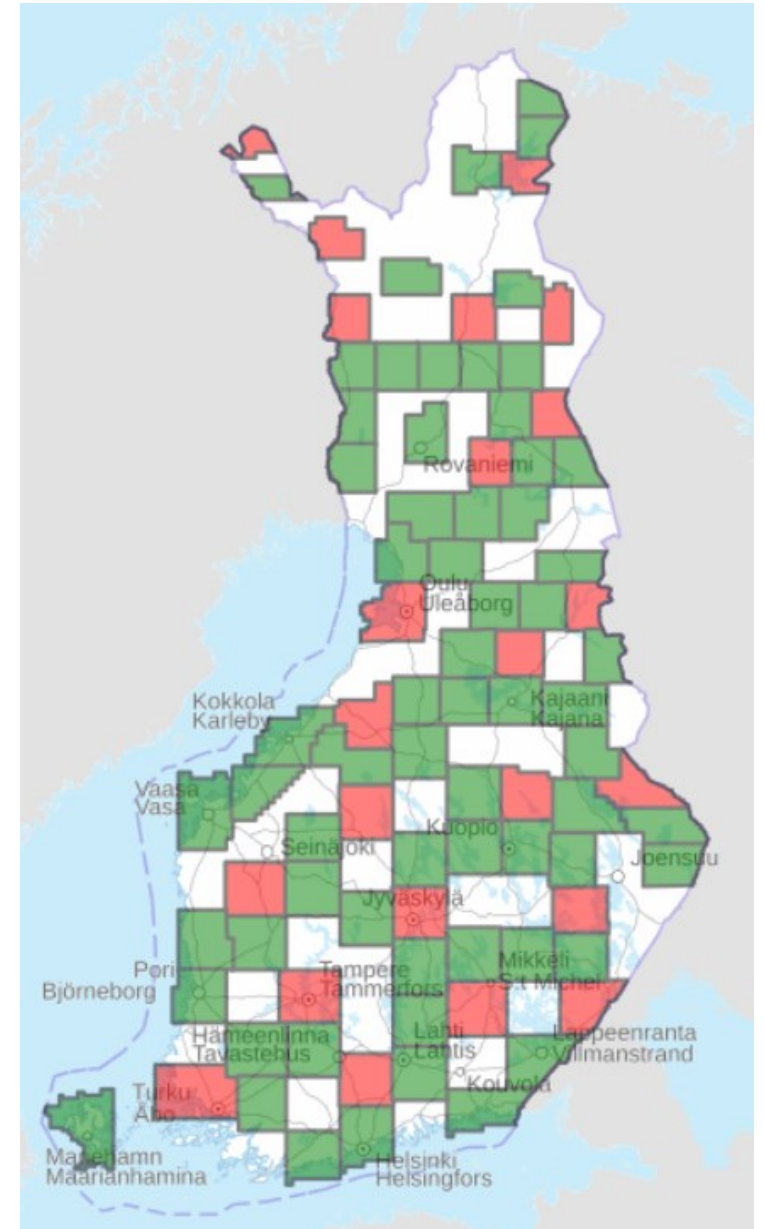
Mihin arkeologisten kohteiden puoliautomaattista tunnistusta tarvitaan?

- Runsaasti arkeologista kulttuuriperintöä sijaitsee alueilla, jotka tapaavat jäädä selvitysten ulkopuolelle.
 - Metsät, erämaa-alueet jne.
 - Tarve kustannustehokkaille kartoitusmenetelmille, joilla voidaan kattaa laajoja alueita.
- Kohteiden sijainnista tarvitaan aiempaa enemmän ja tarkempaa tietoa
 - Helpottaa kohteiden huomioimista maankäytön yhteydessä.
 - Erityisesti aluemaisissa rajauksissa parannettavaa.
- Tiettyjä kohdetyyppejä koskevat tiedot ovat puutteellisia
 - Muinaisjäännösrekisterissä on niukasti tietoja esimerkiksi tietyistä runsaslukuisista kohdetyypeistä (esim. tervahaudat, hiilimiilut)



Laserkeilausaineistot ja kansallinen laserkeilausohjelma 2020-2025

- Maanpintaa ja maanpinnalla olevia kohteita kuvaava pistepilviaineisto
 - Tuotetaan yleensä lentokoneeseen kiinnitetyllä laserkeilaimella
- Kansallisen aineiston tuottamisesta vastaa Maanmittauslaitos
- 2020 alkaen tuotettu aineisto
 - Pistetiheys: 5 pistettä / m²
 - Kattavuus nyt n. 165 000 km²



Laser 5p aineiston kattavuus (vihreä) ja suunnitelma 2023, Maanmittauslaitos

Arkeologisten kohteet laserkeilausaineistossa

- Osaan arkeologisista kohteista liittyy maan pinnalle erottuvia rakenteita - 'kuopat ja kasat'
 - Nämä on ainakin periaatteessa mahdollista havaita laserkeilausaineiston perusteella
- Laserkeilausaineistoja on hyödynnetty tähän tarkoitukseen 2008 alkaen
 - Menetelmä on vakiintunut kiinteäksi osaksi alan 'työkalupakkia'



Taavetin linnoitus LiDAR-visualisoinnissa.
Sisältää Laser 0,5p, maanmittauslaitos

Esimerkkejä tunnistettaviksi sopivista kohteista

Pyyntikuopat

Kuva



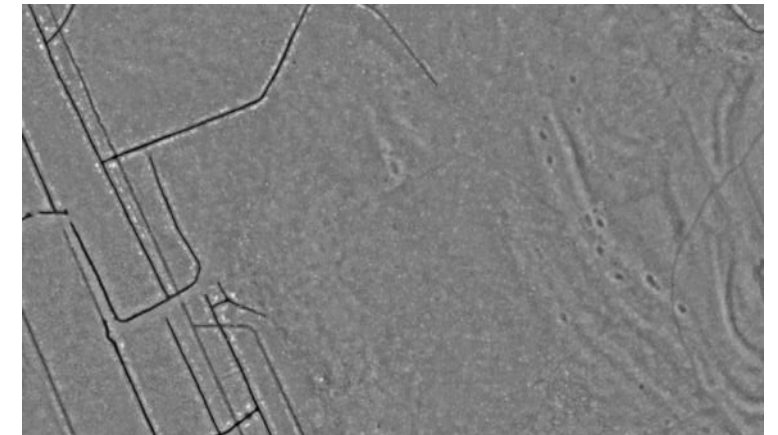
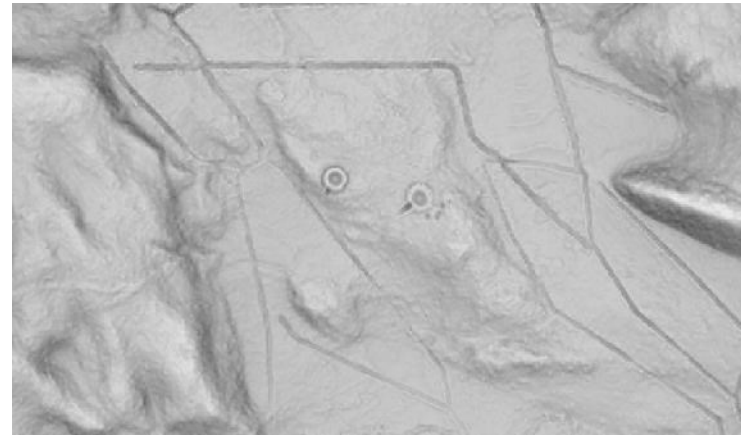
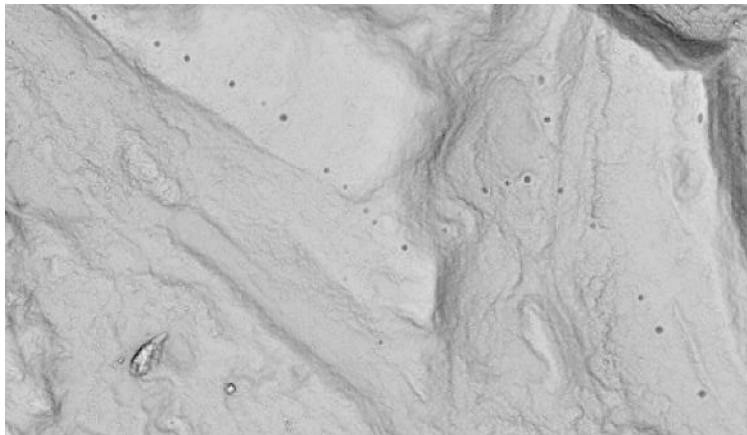
Tervahaudat



Kivikautiset asumuspainanteet



MML / Laser 5p



Menetelmät

- Hankkeessa käytetään U-Net arkkitehtuuriin perustuvaa syväoppivaa mallia
- Malli tuottaa kuvasta semanttisen segmentaation, eli luokittelee kuvan pikselit eri luokkia edustaviin alueisiin
 - Käytössä binaarinen luokittelu – malli tunnistaa vain yhden tyyppisiä kohteita kerrallaan
- Lopputuloksena tunnistettujen kohteiden rajaukset
 - Monipuolisesti käyttökelpoista dataa sellaisenaan tai muokattuna

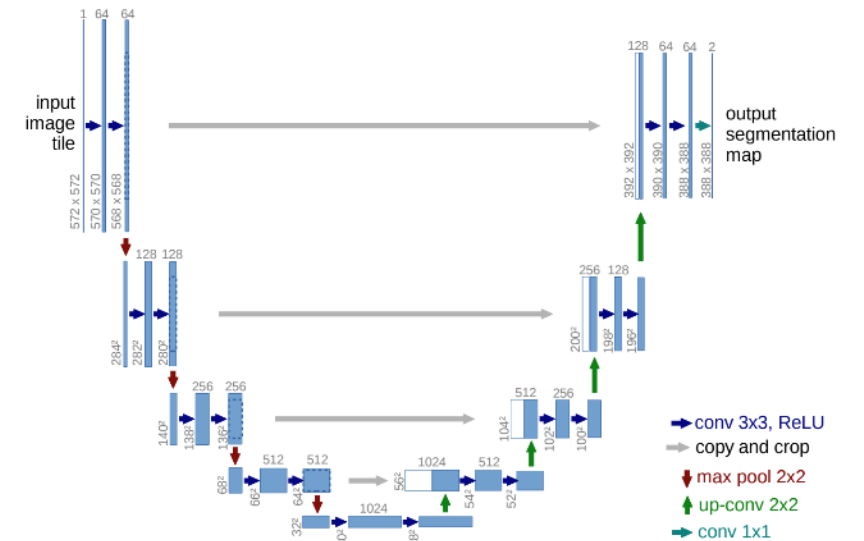


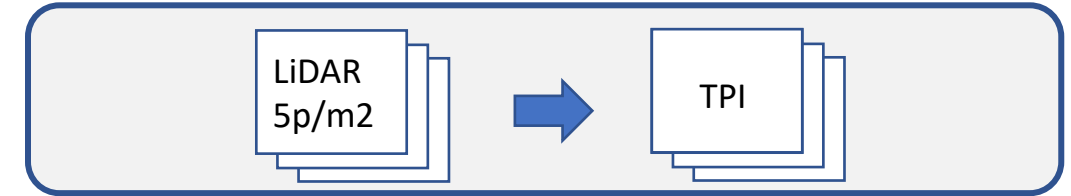
Fig. 1. U-net architecture (example for 32x32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows denote the different operations.

Ronneberger *et al.* 2015

Työnkulku

- Laserkeilausdatan esikäsittely
 - Luodaan halutut rasterimuotoiset visualisoinnit eri tarkoituksiin (U-Net + visuaaliseen tarkasteluun)
- Luodaan opetusaineisto
 - Koostetaan tiedot eri rekistereistä (MV, Maanmittauslaitos, Metsähallitus)
 - Digitoidaan uusia kohteita
- Mallin kouluttaminen ja automaattinen tunnistus
 - Koulutetaan U-Net malli opetusdatan avulla
 - Sovelletaan koulutettua mallia uusien kohteiden tunnistamiseen
- Tulosten käsittely
 - Automaattinen jälkikäsittely (Python)
 - Tulosten tarkastus (GIS + kenttätyöt)
 - Tulosten muokkaus eri käyttötarkoituksiin sopiviksi

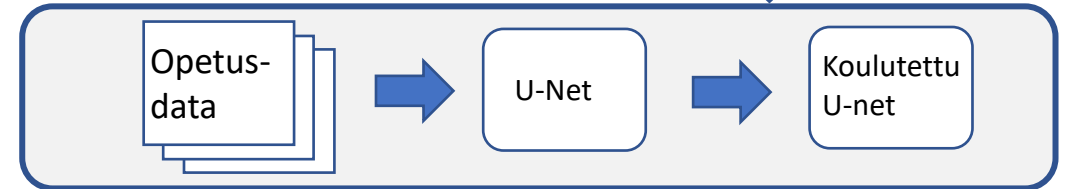
Laserkeilausaineiston esikäsittely



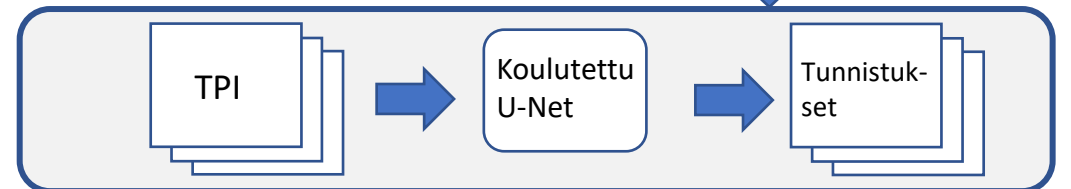
Opetusdatan muodostaminen



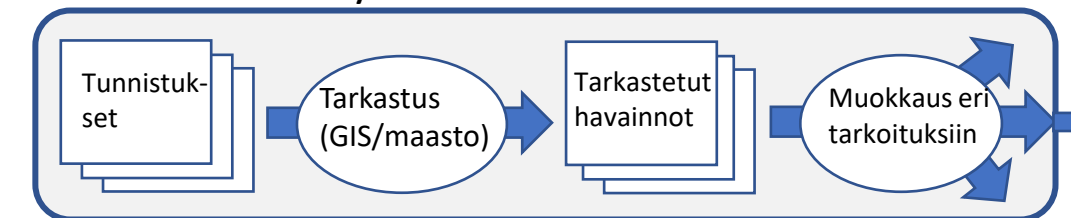
Mallin kouluttaminen



Automaattinen tunnistus

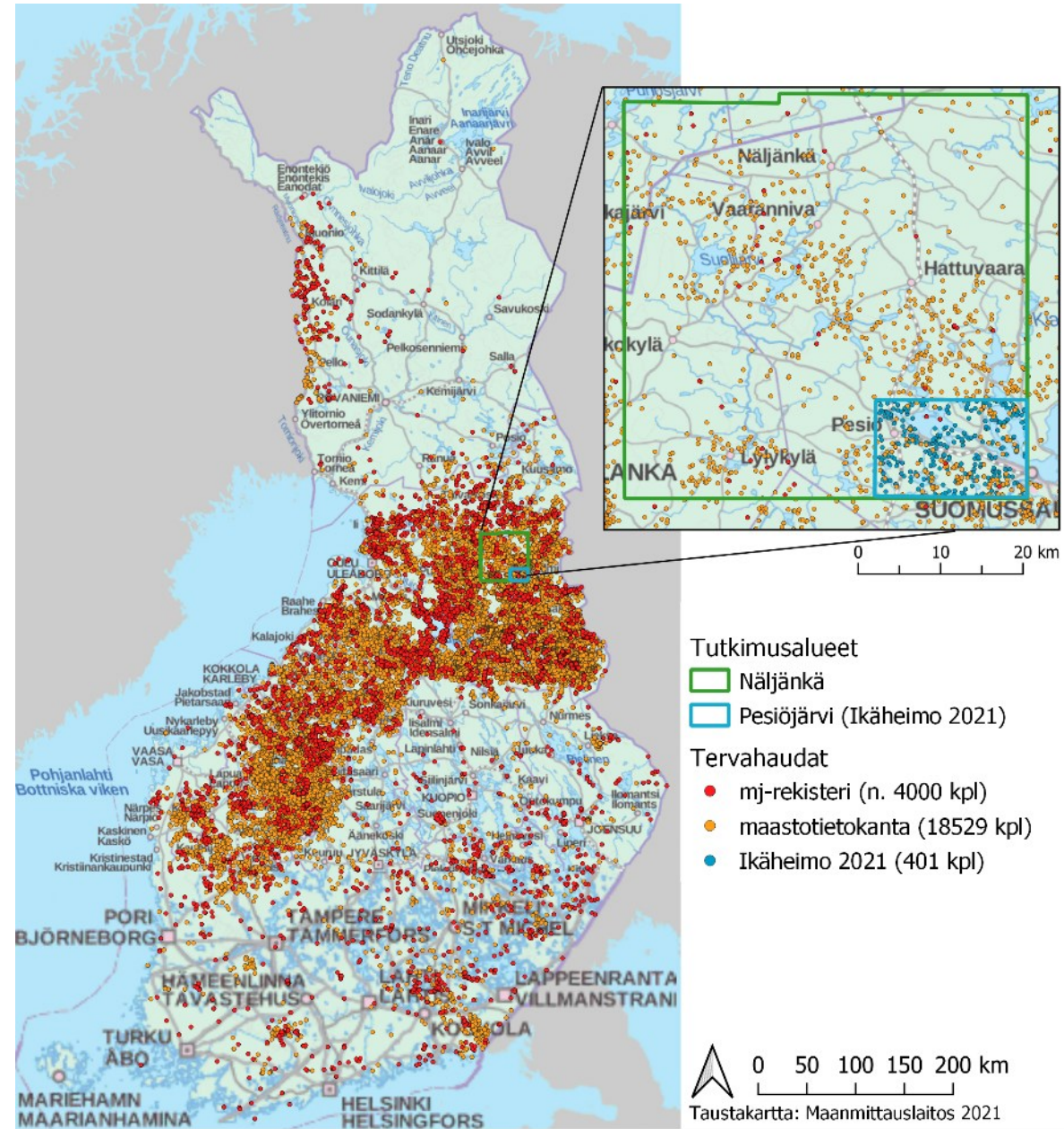


Tulosten käsittely



Case Näljänkä

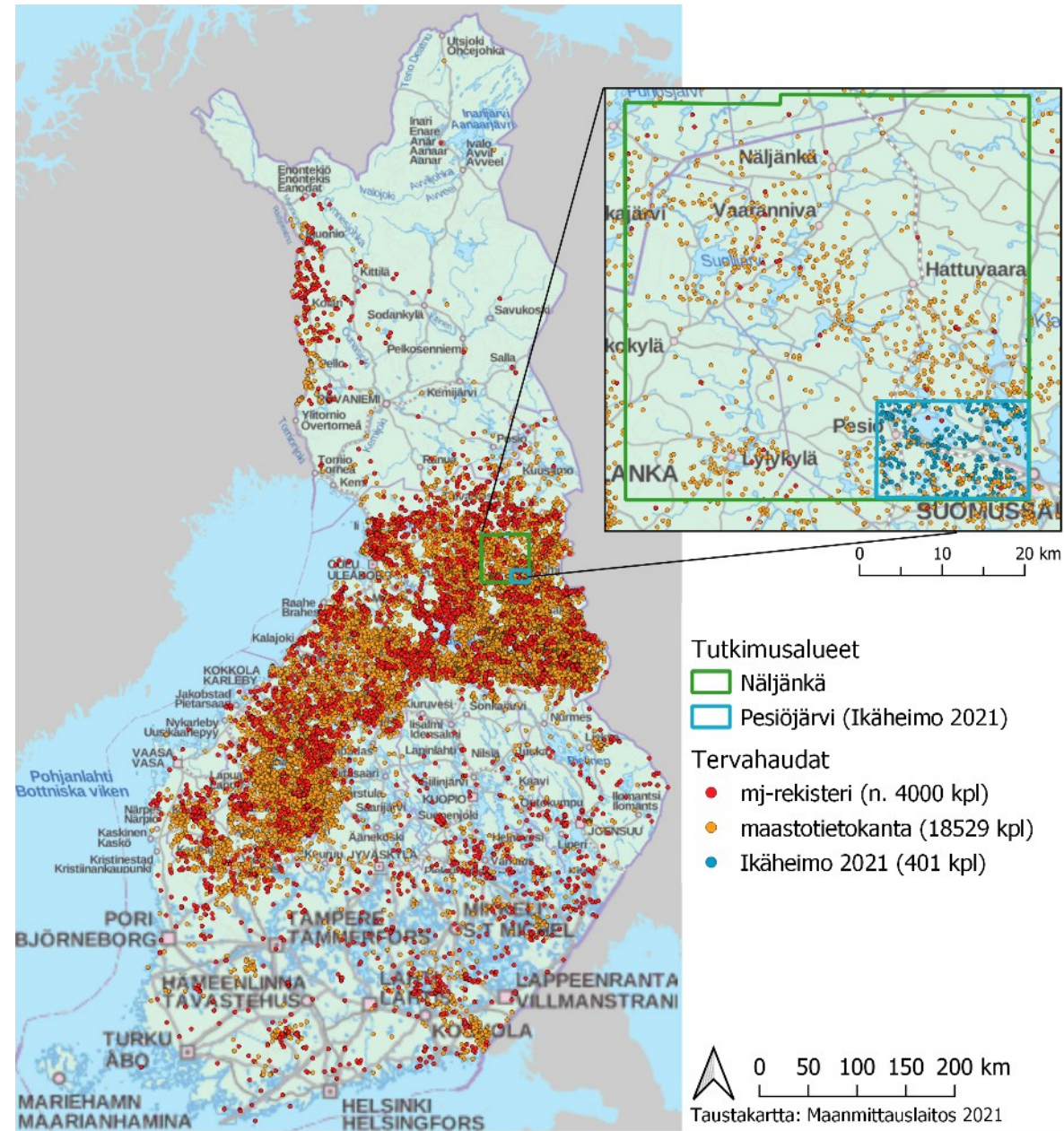
- LIDARK-hankkeen ensimmäinen pilottitutkimus
 - 1 LiDAR 5p tuotantoalue (2304 km²)
 - 1 helposti tunnistettava kohdetyyppi (tervahaudat)
- Fokuksessa menetelmien kehittäminen



Case Näljänkä

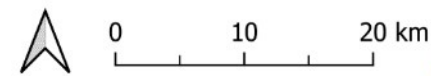
Iteratiivinen lähestymistapa

- Menetelmiä ja opetusdataa parannettiin toistojen välillä.
- Aluksi opetusaineiston ytimen muodostivat Ikäheimon (2021) Pesiöjärven alueelta tunnistamat kohteet (401 kpl)
- Seuraavissa versioissa opetusdataa kasvatettiin työn aikana tunnistetuilla kohteilla.



Case: Näljänkä

- Alueelta tunnistettiin runsaasti uusia kohteita:
 - Ennen: 750-800
 - Jälkeen: 2750
- Tulosten perusteella suurin osa havaittavissa olevista kohteista saatiin tunnistettua
 - Muista lähteistä on tiedossa noin 50 kohdetta, joita ei tunnistettu
 - Tunnistamatta jäivät kohteet yleensä epätyypillisiä, vahingoittuneita tai muusta syystä vaikeasti havaittavia



Tutkimusalueet

aluerajaukset

■ Näljänkä

■ Pesiöjärvi

Tunnetut kohteet

● mj-rekisteri (39 kpl)

● mastotietokanta (757 kpl)

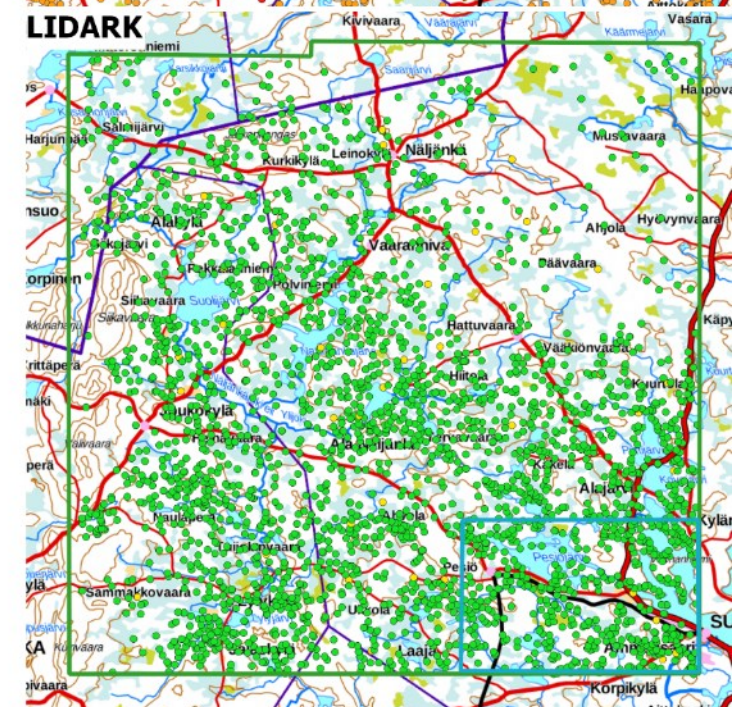
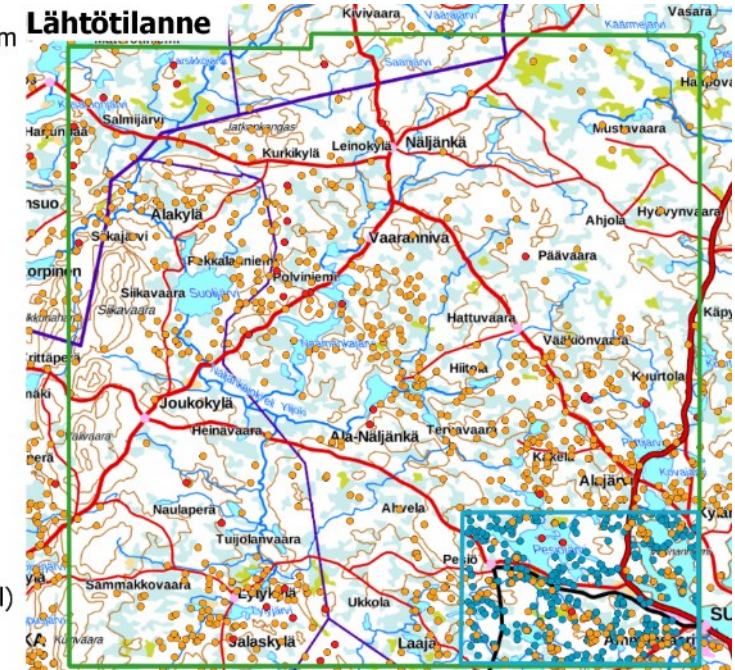
● Ikäheimo 2021 (401 kpl)

LIDARK

Automaattisesti tunnistetut

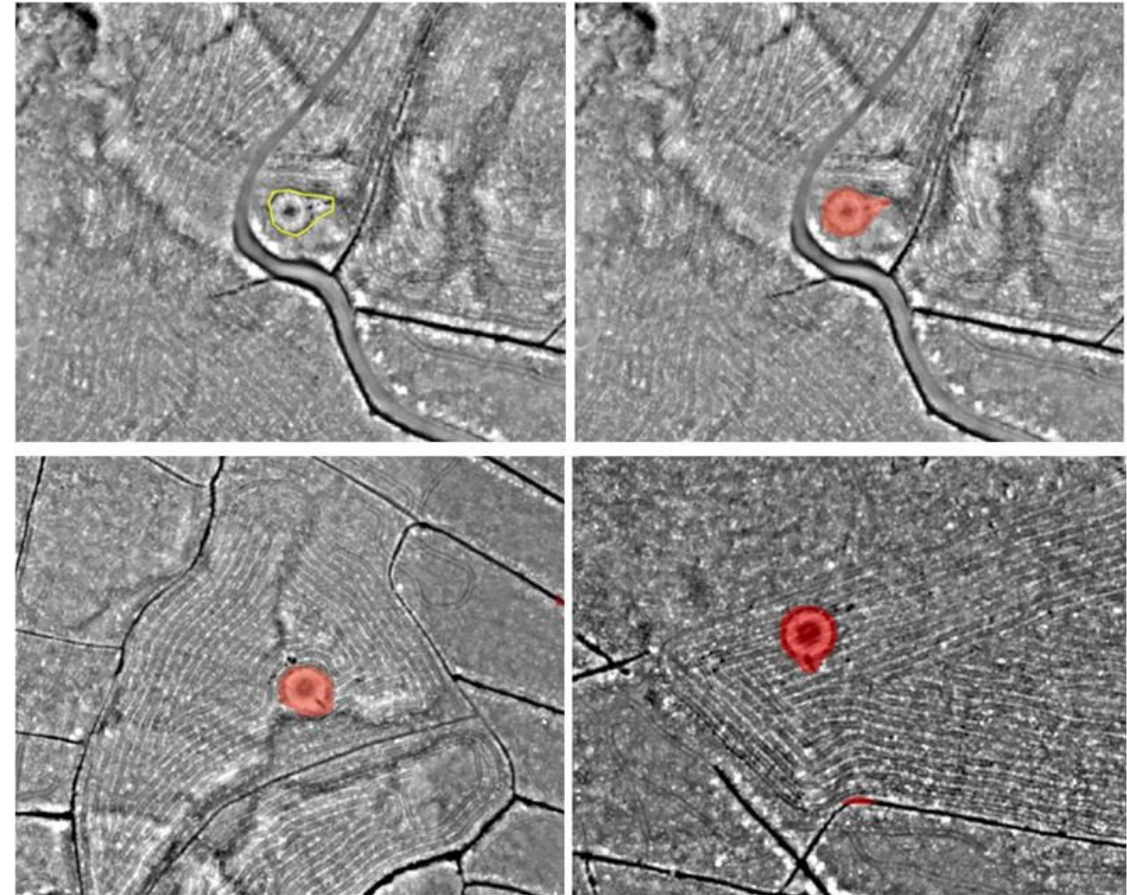
● Tervahauta (n. 2750 kpl)

● mahdollinen tervahauta (40kpl)



Case: Näljänkä

- Mallin tuottamat rajaukset pääosin erittäin tarkkoja
- Vahingoittuneiden tai osittain tuhoutuneiden kohteiden tunnistus onnistuu yleensä ongelmitta



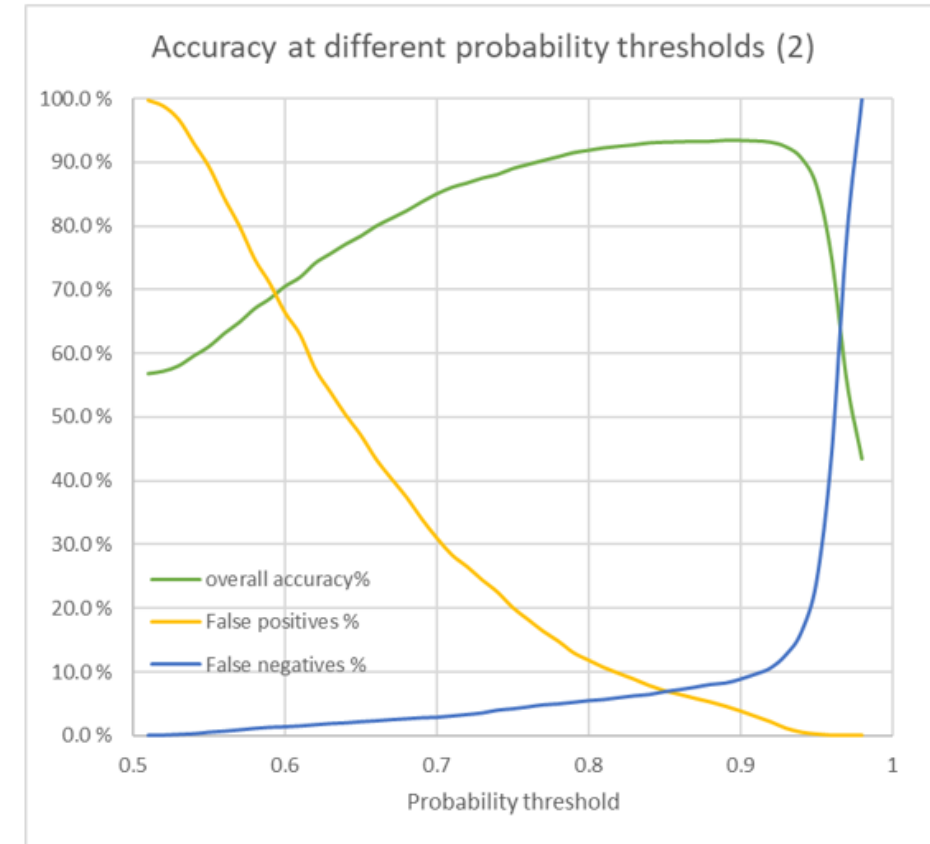
Tervahaudan rajaus opetusdatassa (ylä vasen) ja mallin rajaamia tervahautoja (punaisella)

Kuva: Blom Kartta 2021, sisältää MML LiDAR 5p 2020



Tunnistusten suodattaminen

- Mallit eivät koskaan ole täydellisiä, vaan tuloksena on myös virheitä
 - False positive = tunnistettu kohde ei ole mitä piti
 - False negative = oikea kohde jää tunnistamatta
- Virheisiin voidaan vaikuttaa suodattamalla aineistoa esimerkiksi:
 - Jonkin ehdon perusteella (esim. pinta-ala, p-arvo, muoto, sijainti, korkeus merenpinnasta...)
 - Käymällä aineistoa manuaalisesti läpi
 - Maastotarkastukset

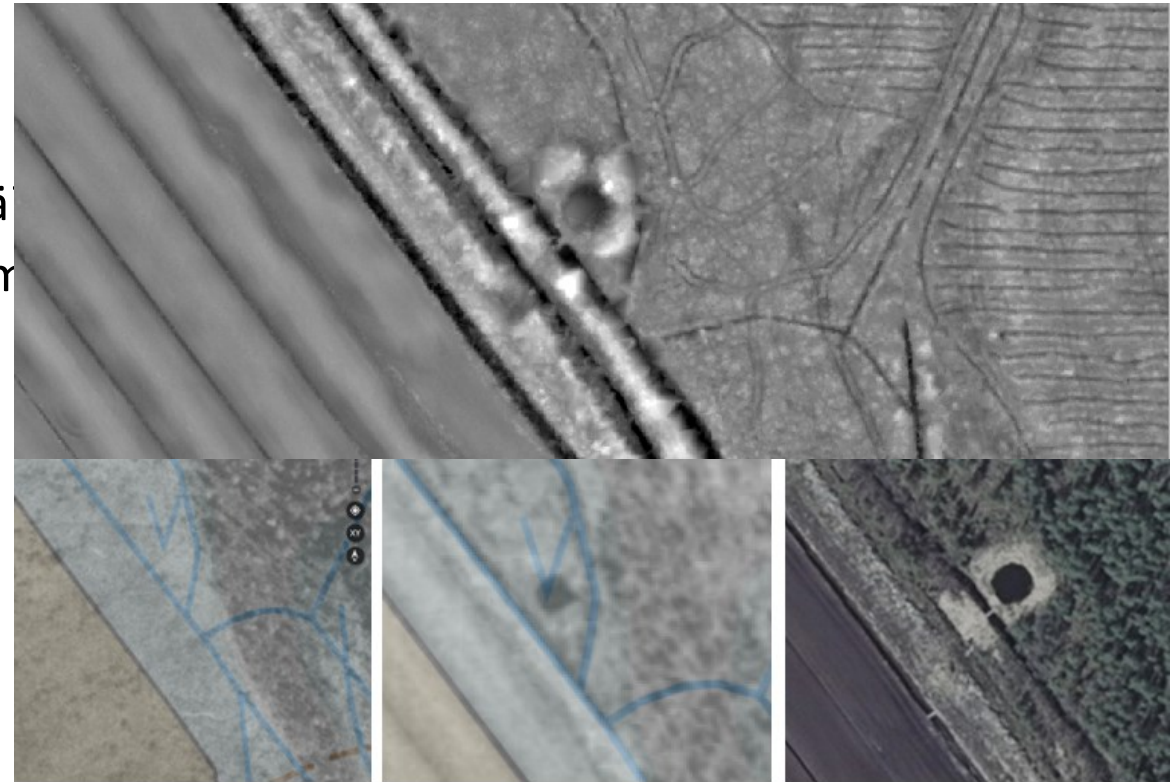


Tunnistuksen todennäköisyyden perusteella suodatuksen vaikutus eri tyyppisten virheiden määriin.



Mallin tuottaman aineiston suodattaminen

- Sopivin lähestymistapa virheiden vähentämiseen on tapauskohtaista
 - Resurssit?
 - Miten paljon erilaisia virheitä voidaan hyväksyä?
 - Voidaanko kohde tunnistaa yksiselitteisesti esim. laserkeilausaineistosta?
- LIDARK-hankkeessa havainnot on pääosin tarkastettu manuaalisesti laserkeilausaineiston perusteella.
 - Osa kohteista tarkastettu maastossa



1958

1996

2013

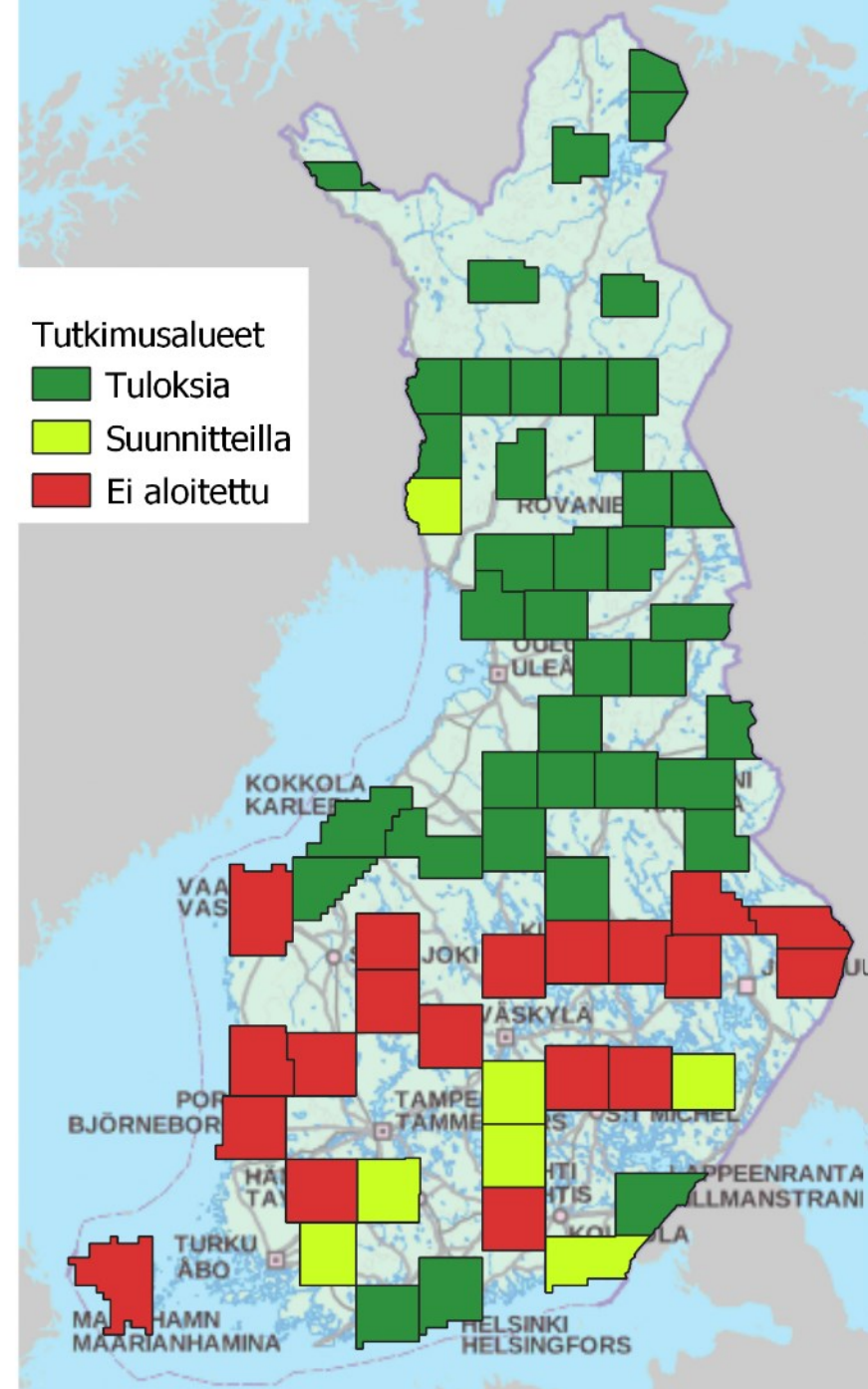
Tervahaudaksi tulkittu suon ojitukseen liittyvä vesiallas.
Sisältää MML laser 5p



Museovirasto

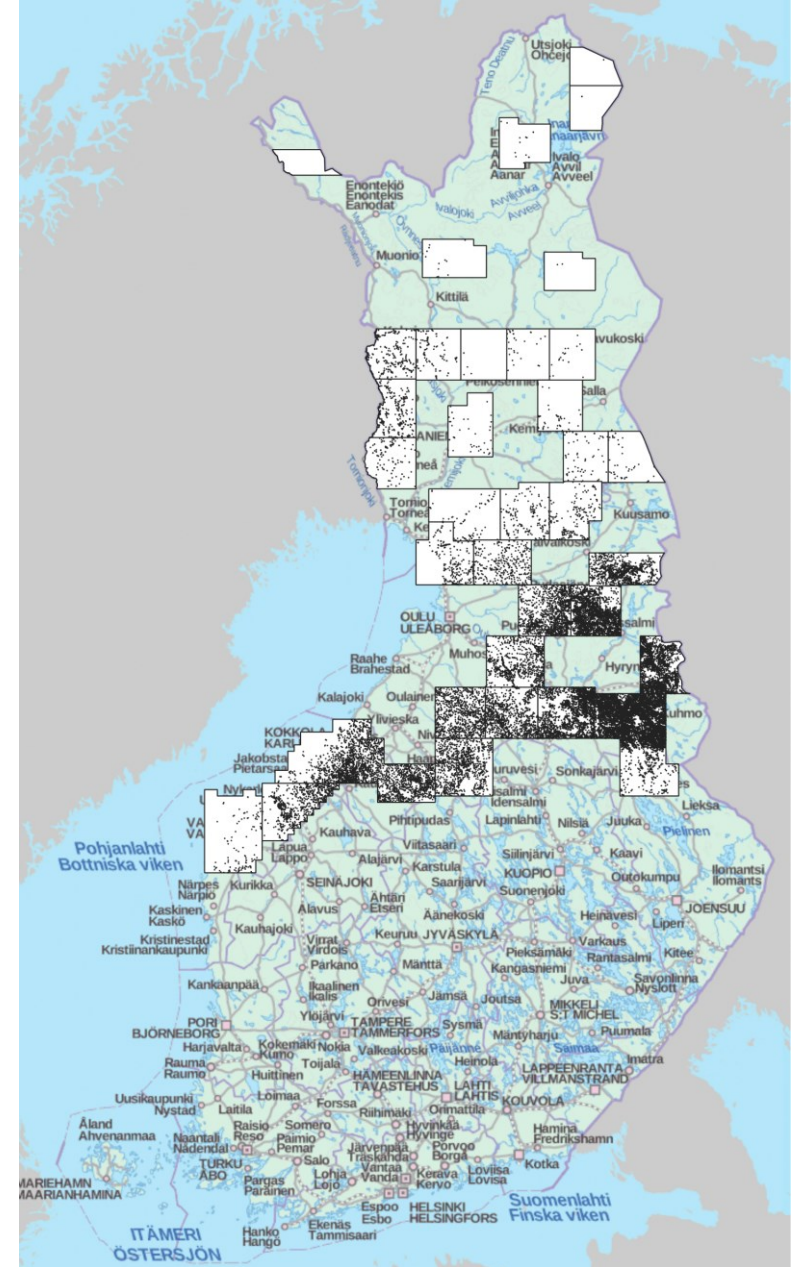
Menetelmien skaalautuvuus

- Valmiita malleja voidaan soveltaa tehokkaasti laajoille alueille
 - LIDARK-hankkessa käsitelty noin 40 laser 5p tuotantoaluetta
 - Tutkittu pinta-ala tähän mennessä noin 90 000 neliökilometriä
- Merkittäviä etuja ovat nopeus ja kustannustehokkuus suhteessa perinteisiin menetelmiin
- Rajoitteena soveltuvuus vain osalle arkeologisista kohteista sekä havaintojen varmentamiseen liittyvät kysymykset



Johtopäätöksiä LIDARK-hankkeesta

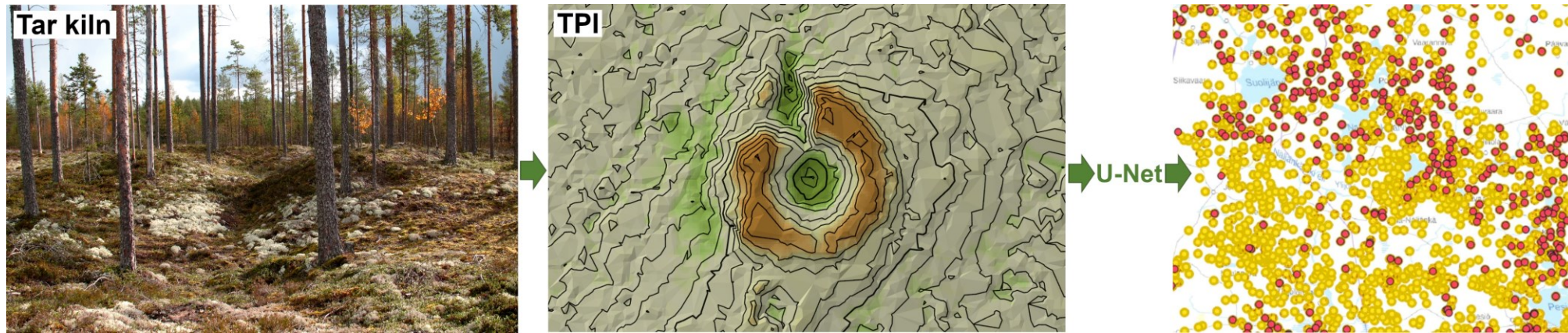
- Toimiva ja kustannustehokas tapa kartoittaa arkeologisesti kiinnostavia ilmiöitä
- Tuotetun aineiston laatu on pääosin hyvä
 - Vaatii vähintään manuaalisen luokittelun
- Seuraava haaste on aineiston saaminen hyötykäyttöön.
 - Luotava käytännöt aineiston hyödyntämiseen arkeologisen kulttuuriperinnön suojelussa
 - LIDARK-hankkeen tervahauta- ja miiluhavainnot ovat Museoviraston ja alueellisten vastuumuseoiden koekäytössä



LIDARK-hankkeen tervahauta- ja miiluaineiston (Ikäheimo & Kotkajärvi 2023) kattavuus



Lisää LIDARK-hankkeen tuloksista:



Anttiroiko N, Groesz FJ, Ikäheimo J, Kelloniemi A, Nurmi R, Rostad S, Seitsonen O. Detecting the Archaeological Traces of Tar Production Kilns in the Northern Boreal Forests Based on Airborne Laser Scanning and Deep Learning. *Remote Sensing*. 2023; 15(7):1799.
<https://doi.org/10.3390/rs15071799>