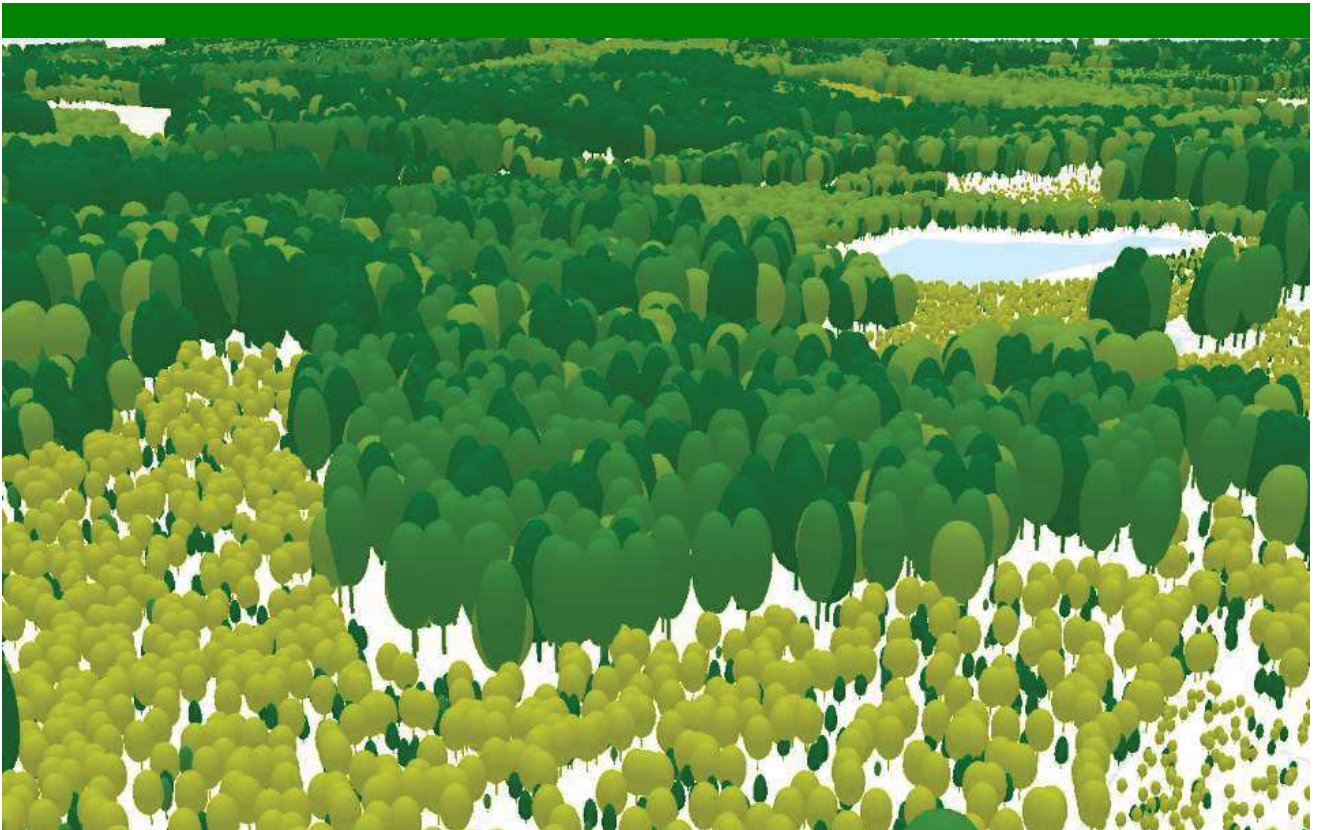




Taimikkotiedon tuottamisen automatisointi koneoppimisen avulla Loppuraportti



Tiivistelmä

Taimikkotiedon tuottamisen automatisointi koneoppimisen avulla

6/2018 - 3/2019

Taimikkotiedon tuottaminen on muodostunut nykyisen metsävaratiedon tuotannon ongelmakohdaksi, koska taimikoiden kattavaan maastoinventointiin ei ole riittävästi resursseja. Toisaalta taimikoiden tiedot ovat useissa tapauksissa pääteltävissä muiden aineistojen perusteella melko luotettavasti, mikäli tavoitteena on oikea toimenpidetarve sekä ajantasaistuksen kannalta riittävä puustotieto. Tässä projektissa kehitettiin uutta toimintatapaa, jolla taimikkotietoa voidaan tuottaa ilman maastokäyntiä automaattisesti eri tietolähteiden avulla koneoppimista hyödyntäen.

Projektin päätuloksena todettiin, että koneoppiminen ei yksinään ratkaise taimikkotiedon tuottamisen automatisointia, mutta saattaa olla osa lopullista ratkaisua. Ongelma on edelleen relevantti, sillä Metsäkeskuksen tuottamaan metsävaratietoon sisältyvä taimikkotieto ei tässäkään projektissa tehdyn analyysin perusteella ole tarpeeksi laadukasta. Projektin aikana Metsäkeskuksen ymmärrys koneoppimisen hyödyntämisen mahdollisuuksista ja vaatimuksista kasvoi. Projektin tulosten avulla voidaan edistää myös metsävaratiedon käytettävyyttä parantavan laatuindeksin kehittämistä.

Projektin tulosten perusteella syntyi useita jatkotoimenpiteitä. Metsäkeskus jatkaa taimikkotiedon tuottamisen automatisointia sekä laatuindeksin kehittämistä osana metsävaratiedon keruun ja ylläpidon kehittämistä. Metsäkeskus pyrkii vähentämään sekä maastotyön että manuaalisen työn määrää ja tarkastelee koneoppimisen hyödyntämistä yhtenä mahdollisuutena Metsäkeskuksen eri tehtävien näkökulmasta.

22.5.2019

Suomen metsäkeskus

Henna Etula henna.etula@metsakeskus.fi

Sisällys

1	Tutkimuksen tavoitteet.....	4
2	Tutkimusosapuolet ja yhteistyö	5
3	Tutkimuksen tulokset	6
3.1	Tutkimusmenetelmät ja aineisto.....	6
3.2	Tutkimustulokset.....	7
3.3	Toteutusvaiheen arviointi.....	8
3.4	Julkaisut	9
4	Tulosten arviointi.....	9
5	Jatkotoimet	10
	Liitteet.....	11

1 Tutkimuksen tavoitteet

Uuden toimintatavan luominen taimikkotiedon tuottamiseen ja ylläpitoon on tarpeen. Tämän projektin aloitushetkellä käytössä tai näköpiirissä ei ollut yhtään menetelmää, jolla pystyttäisiin hankkimaan laadukas taimikkotieto kattavasti tai riittävän kustannustehokkaasti.

Taimikoista saadaan tietoa useista eri tietolähteistä, joita ovat esimerkiksi eri kaukokartoitusaineistoista tuotetut tulkinnot, metsä- ja kemera-lakeihin liittyvät asiakirjat, omaoikeus- tai muu toimenpidetieto sekä maastossa kerätty tieto. Yksittäisen taimikon näkökulmasta näistä lähteistä voidaan tuottaa parhaassa tapauksessa luotettava tieto, mutta pahimmassa tapauksessa ei lainkaan tietoa. Mikäli yksittäisen tietolähteen avulla tuotettu tieto ei vaikuta luotettavalta, voidaan sitä useissa tapauksissa parantaa eri tietolähteitä yhdistämällä. Tavoitteena on tuottaa tieto taimikoiden hoitotarpeesta sekä ajantasaistuksen kannalta riittävä puustotieto.

Tietojen kerääminen eri tietolähteistä on työlästä, eikä lopputulos ole ihmisen tekemänä tasalaatuista. Ihmisen tiedonkäsittelykapasiteetti on rajallinen, joten kaikkien tietolähteiden systemaattinen läpikäynti on vaikeaa. Toisaalta kokenut metsäammattilainen pystyy esimerkiksi ilmakuvan perusteella usein päättelemään taimikon tiedot.

Päätelyt eivät perustu suoriin sääntöihin, vaan tilanteen mukaan on huomioitava eri aineistoja ja arvioitava niiden antamien lähtötietojen luotettavuutta. Tällaista päätelyä on mahdollista automatisoida mallintamalla ihmisen ajattelua tekoälyn avulla. Tehdyistä päätelyistä saatua palautetta on mahdollista hyödyntää uusissa päätelyissä, eli oppivan menetelmän avulla tuotetun tiedon laatu parantuu koko ajan.

Taimikkotiedon tuottamisen automatisointi tukee Suomen metsäkeskuksen strategiassa kuvattuja tavoitteita, joilla pyritään parantamaan metsätalouden kestävyttä ja kannattavuutta sekä asiakkaiden palvelukokemusta. Metsäkeskuksen kaukokartoitusperusteisen metsävaratiedon keruun lähtöajatuksena on ollut tuottaa puustotulkinnan avulla varttuneiden metsien puustotiedot ja inventoida taimikot maastotyönä (kohdennettu maastoinventointi). Niukkojen resurssien vuoksi maastotyötä ei ole pystytty tekemään riittävästi erilaisista priorisointitoimenpiteistä huolimatta. Mikäli taimikkotiedon tuottamista ei pystytä automatisoimaan, tarvitaan maastoinventointiin kymmenien henkilötyövuosien lisäresurssi vuosittain tai vaihtoehtoisesti taimikkotieto rapautuu käyttökelvottomaksi. Jälkimmäinen vaihtoehto on huono metsätalouden kestävyuden tai kannattavuuden näkökulmasta.

Verrattuna tilanteeseen, jossa maastoinventointia ei pystytä resurssisyistä tekemään, asiakkaille projektin tavoitteiden toteutuminen näkyy nykyistä luotettavampana ja ajantasaisempana tasalaatuisena taimikkotietona, jota voidaan käyttää metsätalouden operatiivisessa toiminnassa. Samalla saavutetaan merkittävä kustannussäästö, kun maastotyö voidaan suurelta osin korvata automatisoidulla työllä. Taimikkotiedon tuottamisen automatisointiin ei liity suuria riskejä taimikkotiedon laadun kannalta, koska luotettavaksi arvioitua tietoa ei korvata heikkolaatuisemmalla, vaan tieto voi käytännössä korvautua vain luotettavammalla tiedolla.

Tämän projektin tavoitteena oli automatisoida taimikkotiedon tuottaminen niin, että kullekin taimikkokuviolle pystytään tuottamaan riittävän hyvät tiedot yhdistelemällä eri tietolähteistä saatavaa tietoa automaattisesti. Puustotietojen tuli olla niin tarkat, että niiden perusteella voidaan päätellä oikea hoitotoimenpidetarve. Vaihtoehtoisesti, mikäli jonkin tietolähteen perusteella saadaan tietää luotettavasti toimenpidetarve, tavoitteena oli tuottaa tämän tiedon kanssa ristiriidattomat riittävän tarkat puustotiedot. Mikäli parempien tietojen päättely ei onnistu, jätetään alkuperäiset tiedot ennalleen, ja tällaiset kuviot inventoidaan tarvittaessa maastossa. Taimikkotiedon tuottamisen ohella tavoitteena oli lisätä Suomen metsäkeskuksen osaamista ja kokemusta koneoppimisen hyödyntämisestä paikkatietoaineistojen käsittelyssä mahdollista muuta jatkokehitystä varten.

Työn edetessä havaittiin, aineistojen käsittelyyn kuluva aika ylitti merkittävästi arvioidun ajankäytön. Tämän vuoksi alkuperäiset tavoitteet olivat mahdottomia saavuttaa projektin aikataulun ja budjetin puitteissa. Lopulliset tavoitteet supistuivat valmiin toimintaprosessin tuottamisesta menetelmän testaamiseen, Metsäkeskuksen koneoppiseen liittyvän osaamisen kasvattamiseen ja jatkokehitystarpeiden tunnistamiseen.

2 Tutkimusosapuolet ja yhteistyö

Projekti jakautui Metsäkeskuksen omana työnä tehtävään ja ulkopuolelta ostettavaan osaan. Omana työnä toteutettiin tavoitetilan kuvaus, lähtöaineistojen kuvaus, päättelysääntöjen luominen (osin yhteistyössä valitun yhteistyökumppanin kanssa) sekä menetelmän käyttökelpoisuuden arviointi. Yhteistyökumppanilta ostettiin sovelluksen rakentamiseen liittyvä työ. Mikäli projektin tulokset olisi viety tuotantoon, Metsäkeskuksen vastuulla olisi ollut suunnitelma tuotantokäyttöön otosta sekä tarvittavista tietojärjestelmämuutoksista.

Yhteistyökumppaniksi valittiin Teknologian tutkimuskeskus VTT Oy. VTT on toteuttanut erilaisia projekteja, joissa on käytetty koneoppimista ympäristöön liittyvien muuttujien ennustamisessa. Hankinta tehtiin suoraan hankintana, koska hankintalain kynnysarvo ei ylittynyt.

VTT:ltä projektiin osallistuivat tutkimusprofessori Tuomas Häme, erikoistutkija Heikki Astola ja varttunut tutkija Lauri Seitsonen. Metsäkeskuksessa projektin omistajana toimi metsätiedon palvelupäällikkö Jorma Jyrkilä, projektipäällikkönä kehityspäällikkö Henna Etula sekä projektiryhmän jäseninä metsätietopäälliköt Esko Välimäki ja Jussi Lappalainen ja metsätietoasiantuntijat Mika Rautiainen ja Juha Antinluoma.

Yhteydenpito VTT:n ja Metsäkeskuksen välillä oli säännöllistä ja kumpikin osapuoli oli sitoutunut projektin aikatauluun. Työtä tehtiin yhdessä niin, että kumpikin osapuoli oli tietoinen projektin tilanteesta ja eteen tulevia kysymyksiä ratkottiin nopeasti siten, ettei työn edistyminen kärsinyt.

Metsäkeskuksen näkökulmasta havaittiin, että Metsäkeskuksen metsävaratieto on monimutkaista ja vaikeaselkoista niille, jotka eivät ole käsitelleet sitä aiemmin syvällisesti. Aineiston tuottamiseen ja ylläpitoon liittyy sääntöjä, joiden vaikutuksesta mallinnukseen pitää olla kokemusta (esimerkiksi vuosina 2010-2015 laserpituuden 2 metrin leikkuri, joka saa aikaan yliarvion 2-3 metriä pitkien taimikoiden pituuteen). VTT:n esittämät kysymykset saatiin selvitettyä, mutta mikäli vastaavia projekteja käynnistetään, tulee Metsäkeskuksen asiantuntijoiden käytettävissä olo varmistaa.

Aineistojen käsittelyn ajankäyttö tuli yllätyksenä molemmille osapuolille. Tämän vuoksi projektin rajauksista päätettiin hyvässä yhteistyössä, vaikka nämä rajaukset estivätkin alkuperäisen tavoitteen saavuttamisen.

3 Tutkimuksen tulokset

3.1 Tutkimusmenetelmät ja aineisto

Kehittämistyön aineistona olivat Metsäkeskuksen aineistot, joiden katsottiin olevan hyödyksi taimikkotiedon tuottamisessa. Mallin laadinnassa käytettävissä olivat: maastossa inventoidut taimikkokuvat (referenssiaineisto), maastossa inventoiduille taimikkokuvioille tuotetut

puustotulkintatiedot, kamera-kuvat, metsänkäyttöilmoitukset, taimikon perustamisilmoitukset, toimenpidekuvat, ilmakehän kuva ja kasvillisuuden pintamalli (CHM).

Työ jakautui projektisuunnitelman mukaan neljään vaiheeseen, jotka olivat 1) aineiston hankinta ja määrittelyt, 2) testiympäristön rakentaminen, 3) taimikkotiedon automaattisen käsittelysystemin kehitys ja 4) taimikkotiedon tuottaminen valitulle testialueelle/testialueille. Ensimmäinen vaihe piti sisällään eri formaateissa ja tietomalleissa olevan aineiston muokkaamisen muotoon, jossa sitä voitiin hyödyntää mallinnuksen lähtöaineistona. VTT suoritti aineiston muokkaamisen ja Metsäkeskus tarkensi, mitkä olivat kunkin aineiston oleelliset muuttujat ja antoi lisätietoa aineistojen sisällöstä. Toinen vaihe jätettiin tekemättä, koska työn edistyessä havaittiin, että varsinaiselle testiympäristölle ei ollut tarvetta. Kolmas vaihe tehtiin VTT:n ja Metsäkeskuksen yhteistyönä niin, että VTT esitteli välituloksia ja kysymyksiä Metsäkeskukselle. Työtä jatkettiin yhdessä sovittujen tarkennusten ja rajausten jälkeen. Neljännessä vaiheessa VTT tuotti taimikkotiedot testialueelle.

Tarkemmat kuvaukset aineistosta ja menetelmistä ovat VTT:n asiakasraportissa (liitteenä).

3.2 Tutkimustulokset

Projektin tavoitteita ei onnistuttu saavuttamaan kokonaisuutena. Projektin aikataulu ja budjetti olivat rajattuja. Metsäkeskuksen aineistot osoittautuivat ennakoitua monimutkaisemmiksi ja niiden käsittely selvästi arvioitua työläämmäksi, joten projektin tavoitteita jouduttiin rajaamaan projektin aikana. Keskeisimmät tehdyt rajakset olivat maantieteellinen raja (koko Suomen sijaan testialueena Satakunnan maakunta) ja käytettyjen aineistojen määrä (valittiin vain kattavimmat aineistot). Jälkimmäisen rajauksen vuoksi ei päästy testaamaan tiedon päättelyä useammasta eri lähteestä, vaan menetelmä oli mallinnusta, jossa kaikkien selittävien muuttujien piti olla käytettävissä. Mitään metsävaratiedon tuotantojärjestelmään integroitavissa olevaa toiminnallisuutta ei pystytty toteuttamaan, koska oleellinen osa, eri tietolähteiden vertailu ja tiedon päättely/korjaaminen, jäi toteuttamatta.

Projektin aikana saatiin tuotettua taimikkotiedon estimoinnin malli ja taimikkotiedot testialueelle. Tilastollisesti tarkasteltuna tällä tavoin tuotetut taimikkotiedot ovat puustotulkintaan verrattuna tarkemmalla tasolla valtaosalla kuvioista. Tämän tuloksen vahvistaminen vaatii lisäselvityksiä (ks. luku 4).

Keskeisenä tavoitteena oli testata koneoppimisen tekniikkaa ja saada siitä kokemuksia. Tämä tavoite toteutui. Oleellinen havainto oli, että koneoppiminen ei yksittäisenä tekniikkana pysty vastaamaan alkuperäisen ongelman ratkaisuun. Sen sijaan se voi olla osa ratkaisua tuottamalla lisätietoa taimikkotiedon päättelyyn. Koneoppimisen avulla tehtävässä mallinnuksessa voidaan hakea selittäviä muuttujia tehokkaasti. Esimerkiksi tässä projektissa havaittiin, että puustotunnuksia selittivät useammat muuttujat kuin ennakkoon arveltiin (VTT:n asiakasraportin liite D).

Projektin aikana syntyi myös tuloksia, joita ei ennakoitu projektisuunnitelmassa. Metsäkeskuksen taimikkoaineistoa analysoitiin huolellisesti ja saatiin tutkittua tietoa sen laatutasosta. Tätä tietoa voidaan hyödyntää tiedon tuotannossa ja siinä, mitkä kuviot merkitään epäluotettaviksi/maastotarkastustarpeellisiksi Metsäkeskuksen metsävaratiedon tuotantoprosessissa.

Projektissa kehitettiin menetelmä virheellisten kuvioden löytämiseksi eli ns. kuviotason laatuindeksi. Tämä menetelmä oli tarkoitettu virheellisten taimikkokuvioden löytämiseen. Menetelmässä on potentiaalia myös metsävaratiedon yleisemmän, kaikkia kehitysluokkia koskevan laatuindeksin luomiseen.

Tulokset on kuvattu tarkemmin VTT:n asiakasraportissa (liitteenä).

3.3 Toteutusvaiheen arviointi

Projekti toteutettiin kesäkuun 2018 ja maaliskuun 2019 välisenä aikana.

Lähtöaineistojen kuvaus ja aineistotoimitukset tehtiin kesäkuun 2018 aikana, jotta työt lähtivät käyntiin sujuvasti kesälomakauden jälkeen. Aineistojen käsittely jatkui pitkälle syksyyn. Tämän työn rinnalla tehtiin päättelysääntöjen ja haluttujen toiminnallisuuksien kuvaus. Varsinainen mallinnus aloitettiin marras-joulukuussa 2018 ja sen tulokset valmistuivat tammikuussa 2019. Johtopäätökset ja menetelmän käyttökelpoisuuden arviointi tehtiin helmi-maaliskuussa 2019.

Alkuperäisen aikataulun mukaan mallinnukseen olisi pitänyt alkaa alkusyksystä 2018. Aineistojen käsittelyn venymisen vuoksi mallinnus kyllä saatiin tehtyä, mutta tuloksia olisi ollut hyödyllistä pystyä tarkastelemaan VTT:n ja Metsäkeskuksen yhteistyönä vielä syvällisemmin ja jatkaa mallinnusta iteratiivisesti havaintojen perusteella. Todennäköisesti aineistosta olisi saatu tällä tavoin vielä enemmän tuloksia.

3.4 Julkaisut

Tämän loppuraportin lisäksi projektin työ on dokumentoitu tarkemmin VTT:n kirjoittamaan asiakasraporttiin (liitteenä). Nämä raportit julkaistaan Metsäkeskuksen verkkosivulla. Projektista ei synny muita julkaisuja.

4 Tulosten arviointi

Tilastollisesti tarkasteltuna projektissa tuotetut taimikkotiedot ovat laserkeilaukseen, ilmakuvauskeilaukseen ja koealamittaukseen perustuvaa puustotulkintaa tarkempia. Tulosten tulkinnassa on kuitenkin huomioitava, että tässä projektissa tuotetut tiedot eivät todennäköisesti ole niin selvästi laserkeilaukseen perustuvaa puustotulkintaa parempia kuin hankkeen tulokset osoittavat. Tässä hankkeessa referenssiaineistona olivat maastossa inventoidut taimikkokuviot. Ne olivat valikoituneet maastoinventointiin siksi, että alkuperäisten puustotulkintatietojen on arveltu olevan epäluotettavia. Sen sijaan luotettaviksi arvioitua ja tämän vuoksi ilman maastoinventointia jääneet taimikkotiedot eivät olleet mukana arvioinnissa. Projektissa tuotettuja taimikkotietoja on tarpeen arvioida maastossa.

Taimikkoaineiston läpikäynti tuotti mielenkiintoista tietoa sen laadusta. Metsäkeskuksen metsätietoasiantuntijoilla on vahvaa kokemukseräistä tietoa siitä, mitkä tunnuksat taimikkotiedossa ja minkä tyyppisillä taimikkokuvioilla puustotulkintatiedot ovat kunnossa. Projektin aikana pystyttiin vahvistamaan monia näistä oletuksista. Syntynyttä tietoa voidaan käyttää hyväksi metsävaratiedon tuotantoprosessissa.

Alkuperäisen taimikko-ongelman ratkaisu on edelleen relevantti kysymys metsävaratiedon laadun ja sen tuottamisen kustannusten kannalta. Vaikka tämän projektin tulokset eivät tuoneet siihen valmista ratkaisua, Metsäkeskuksen näkemys tarvittavista jatkotoimenpiteistä vahvistui.

Metsäkeskuksen ymmärrys koneoppimisen hyödyistä, ja sen käyttöön tarvittavista lähtöaineistoista kasvoi. Tätä tietämystä voidaan hyödyntää tulevaisuudessa, kun arvioidaan mahdollisuuksia automatisoida manuaalista työtä Metsäkeskuksen toiminnassa. Manuaalisen työn määrä on merkittävä kustannustekijä Metsäkeskukselle. Sen vähentämiseen tehtävät panostukset ovat tärkeä investointi, jotta henkilötyötä saadaan suunnattua rutiinitöistä ihmisen asiantuntemusta vaativaan työhön.

Yhtenä tärkeänä metsävaratiedon kehittämistoimenpiteenä on tuotetun tiedon laadun esittäminen kuviokohtaisesti. Tässä projektissa tuotettua laatuindeksiä voidaan hyödyntää yhtenä vaihtoehtona laatutunnuksen kehittämiseen. Laatutunnus nostaisi metsävaratiedon käytettävyyttä merkittävästi, kun maastokäynnit vähentyisivät avointa metsätietoa hyödyntävissä eri organisaatioissa.

5 Jatkotoimet

Tulosten perusteella VTT esitti jatkotoimiksi seuraavia:

- 1) Ilmakuviin ja latvusmalliaineistoon perustuvan puustotulkintamenetelmän jatkokehitys.
- 2) Automaattisen taimikkotiedon tuottamisen jatkokehitys.
- 3) Taimikkotiedon laatuindeksin jatkokehitys.
- 4) Taimikkotiedon puustotunnusten estimointi Suomen alueelle VTT:n tässä hankkeessa kehittämällä menetelmillä.

Metsäkeskuksen päätökset jatkotoimista ovat seuraavat:

- 1) Ilmakuviin ja latvusmalliaineistoon perustuvan puustotulkintamenetelmänä kehittämistä ei jatketa. Metsäkeskuksen käytössä oleva inventointimenetelmä on laserkeilaukseen, ilmakuviin ja koealoihin perustuva menetelmä, jonka oletetaan tuottavan taimikoihinkin paremmat puustotiedot tämän hankkeen tuloksista poiketen. Tämän vuoksi taimikoiden tulkintatiedon tuottamiseen liittyvät kehittämispanokset ohjataan käytössä olevan inventointimenetelmän kehittämiseen (tiheämpi laseraineisto, taimikkokoealojen tarkempi/kattavampi mittaus esim. dronella, puustotulkintamenetelmien kehitys).
- 2) Automaattisen taimikkotiedon tuottamisen jatkokehitys. Taimikkotiedon tuottamisen automatisoinnin kehittämistä jatketaan, jotta tämän hankkeen alkuperäinen, edelleen relevantti ongelma saadaan ratkaistua. Tämän hankkeen tulosten perusteella ratkaisu ei löydy pelkästään koneoppimisesta vaan todennäköisemmin sääntöpohjaisen päättelyn ja koneoppimisen yhdistelmästä. Koneoppimisen käytön tarkoituksena on todennäköisesti tuottaa uutta informaatiota, jota voidaan käyttää päättelyn osatekijöinä. Jatkokehitys käynnistetään metsävaratiedon keruun ja ylläpidon menetelmien kehittämiseen sisältyvänä työnä. Aikataulu riippuu Metsäkeskuksen omista käytettävissä olevista asiantuntijaresursseista.
- 3) Taimikkotiedon laatuindeksin jatkokehitys. Hankkeessa toteutetulla laatuindeksillä on jatkokehityspotentiaalia. Metsäkeskus haluaa 2020-luvulla kuvata metsävaratiedon laatua kuviokohtaisella tunnuksella. Jatkokehitys käynnistetään metsävaratiedon keruun ja ylläpidon

menetelmien kehittämiseen sisältyvänä työnä. Samalla selvitetään, voidaanko tässä hankkeessa kehitetyn taimikkotiedon laatuindeksin tuottamisen logiikkaa laajentaa varttuneiden metsien laatuindeksiksi.

- 4) Taimikkotiedon puustotunnusten estimointi Suomen alueelle VTT:n tässä hankkeessa kehittämällä menetelmillä. Metsäkeskus arvioi VTT:n Satakunnan alueelle toimittamia puustoestimaatteja maastotyönä kesän 2019 aikana. Lopullinen päätös koko Suomen puustotunnusten estimoinnista tehdään vasta maastotarkastuksen jälkeen.
- 5) Mahdollisuutta hyödyntää koneoppimista ratkaisuna manuaalisen työn vähentämisessä tarkastellaan Metsäkeskuksen eri tehtävien näkökulmasta. Mikäli koneoppimisesta katsotaan olevan hyötyä, toteutetaan kehitystyötä Metsäkeskuksen digitiekartan mukaisissa tietojärjestelmäprojekteissa tai erikseen käynnistettävässä projektissa.

Liitteet

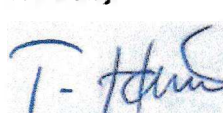
VTT:n asiakasraportti *Taimikkotiedon tuottamisen automatisointi* 4.3.2019

Taimikkotiedon Tuottamisen Automatisointi

Kirjoittajat: Heikki Astola, Lauri Seitsonen

Luottamuksellisuus: Luottamuksellinen



Raportin nimi TAIMIKKOTIEDON TUOTTAMISEN AUTOMATISOINTI		
Asiakkaan nimi, yhteyshenkilö ja yhteystiedot Suomen Metsäkeskus Henna Etula, kehityspäällikkö Suomen Metsäkeskus Huhtalantie 2, 60220 Seinäjoki henna.etula@metsakeskus.fi		Asiakkaan viite 21304/500
Projektin nimi TAIMIKKOTIEDON TUOTTAMISEN AUTOMATISOINTI		Projektin numero/lyhytnimi 119938/SMK Tatuointi
Tiivistelmä Hankkeen tavoitteena oli tuottaa automaattinen menetelmä taimikkotiedon parantamista varten. Menetelmän avulla virheellistä taimikkotietoa voidaan korjata ja paikantaa taimikot, joihin maastotarkastukset tulee kohdentaa. Koska analyysiaineiston esikäsittelyvaihe vei ennakoitua huomattavasti suuremman osan hankkeen resursseista, alkuperäisiä tavoitteita tarkennettiin yhdessä asiakkaan kanssa hankkeen kuluessa. Menetelmäkehityksen aineistolähteiden lukumäärää ja testialueita vähennettiin. Toisaalta hankkeen tehtäviin lisättiin taimikon pituuden estimointi SMK:n ilmakuva- ja latvuston korkeusmalliaineiston avulla. Hankkeessa analysoitiin ensin SMK:n eri lähteistä koottua Satakunnan alueen kuviotietoa aineiston rakenteen, korrelaatioiden ja puustotietojen jakaumien selvittämiseksi. Yksi tämän vaiheen tuloksia oli, että SMK:n toimittamissa taimikkotiedoissa oli huomattavan paljon harhaa. Alustavan analyysin tulosten perusteella estimointiin taimikoiden puustotiedot ilmakuva- ja latvusmallipiirteiden ja maastossa inventoitujen taimikkokuvioiden avulla. Estimoidut puustotiedot olivat taimikon pituus (H), runkoluku (N), keskiläpimitta (D) ja ikä (T) puulajeille mänty ja kuusi. Estimoinnissa käytettiin neuroverkkoihin perustuvia malleja. VTT:n estimoinnin tuloksia verrattiin SMK:n toimittamiin puustotietoihin: yksittäisten kuvioiden tarkastelussa VTT:n estimaatti oli parempi 87%:lla kuvioista, ja SMK:n estimaatti parempi 13%:lla. Kaikkien testattujen puustotietojen suhteellinen keskineliövirhe (RMSE%), suhteellinen harha (BIAS%) ja selitysaste (R^2) olivat VTT:n estimaateilla paremmat kuin SMK:n toimittamilla puustotunnuksilla. Lisäksi hankkeessa kehitettiin menetelmä niiden kuvioiden löytämiseksi, joissa on odotettavissa suurin estimointivirhe (taimikkotiedon laatuindeksi). Menetelmä perustuu ohjatun ja ohjaamattoman oppimisen menetelmien yhdistämiseen. Menetelmää testattiin alustavasti Satakunnan taimikkokuvioilla ja tulokset olivat lupaavia. Menetelmän käyttökelpoisuuden arviointi vaatii kuitenkin jatkokehittelyä ja lisätestausta yhteistyössä SMK:n kanssa. Ehdotukset hankkeen jatkotoimenpiteiksi ovat: 1) ilmakuviin ja latvusmalliaineistoon perustuvan puustotunnusten tulkintamenetelmän jatkokehitys, 2) automaattisen taimikkotiedon tuotantomenetelmän jatkokehitys, 3) taimikkotiedon laatuindeksin jatkokehitys, ja 4) taimikoiden puustotunnusten estimointi Suomen alueelle VTT:n tässä hankkeessa kehittämällä menetelmillä.		
Espoo 7.2.2019		
Laatija  Heikki Astola Erikoistutkija	Tarkastaja  Tuomas Häme Research Professor	Hyväksyjä  Matti Koivu Solution sales lead
VTT:n yhteystiedot Tuomas Häme, tuomas.hame@vtt.fi , 040 587 0631		
Jakelu (asiakkaat ja VTT) Metsäkeskus, VTT		
VTT:n nimen käyttäminen mainonnassa tai tämän raportin osittainen julkaiseminen on sallittu vain Teknologian tutkimuskeskus VTT Oy:ltä saadun kirjallisen luvan perusteella.		

Sisällysluettelo

Sisällysluettelo	3
1 Toimeksiannon kuvaus ja tavoitteet	4
2 Aineisto	5
2.1 SMK:n toimittama aineisto	5
2.2 Testialueet ja aineiston rajausta	6
2.3 Esikäsittely	6
3 Menetelmät	9
3.1 Analyysin rajausta	9
3.2 Aineiston alustava analyysi	11
3.2.1 Eri tietolähteiden määrä Satakunnan aineistossa	11
3.2.2 Jakaumat ja korrelaatiotaulukot	12
3.3 Puustotunnusten estimointi	13
3.3.1 Estimointiaineiston jako	13
3.3.2 Estimointimenetelmät	13
3.4 Menetelmä virheellisten kuvioiden löytämiseksi	14
4 Tulokset	16
4.1 Puustotunnusten jakaumat	16
4.1.1 Aineiston D19_2 vertailu maastossa inventoituun tietoon	16
4.2 Korrelaatiotaulukot	19
4.3 Puustotietojen estimointi	20
4.3.1 Piirrevalintatestit	20
4.3.2 Puustotietojen estimointi yhden piilokerroksen neuroverkolla	20
4.3.3 Puustotietojen estimointi syvällä neuroverkolla	22
4.3.4 Puustotietojen estimointi käyttäen aineistoa D19_2 selittävinä muuttujina (ilmakuvien ja CHM-aineiston lisäksi)	22
4.3.5 VTT:n estimaattien vertailu SMK:n puustotietoihin	24
4.3.6 Yhteenveto estimointimalliennuksen tuloksista ja tulosten vertailu	27
4.4 Virheellisten kuvioiden löytäminen (laatuindeksi)	30
5 Johtopäätökset ja yhteenveto	35
5.1 Johtopäätökset	35
5.2 Yhteenveto ja ehdotus jatkotoimenpiteistä	36
Lähdeviitteet	38
Liitteet /	39

1 Toimeksiannon kuvaus ja tavoitteet

Hankkeen alkuperäisenä tavoitteena oli automatisoida taimikkotiedon tuottaminen niin, että kullekin taimikkokuviolle pystytään tuottamaan riittävän hyvät tiedot yhdistelemällä eri tietolähteistä saatavaa tietoa automaattisesti. Tavoitteena on tuottaa taimikoihin niin tarkat puustotiedot, että niiden perusteella voidaan päätellä oikea hoitotoimenpidetarve [1], [2].

VTT:n projektisuunnitelmassa [3] määriteltiin hankkeessa tehtävän työn sisältö ja tulokset seuraavasti:

Hankkeessa tehtävä työ sisältää prototyypin rakentamisen automaattisen taimikkotiedon tuottamiseen.

Työ on jaettu neljään osatehtävään, jotka ovat:

- Aineiston hankinta ja määrittelyt (Tehtävä 1.)
- Testiympäristön rakentaminen (Tehtävä 2.)
- Taimikkotiedon automaattisen käsittelysysteemin kehitys (Tehtävä 3.)
- Taimikkotiedon tuottaminen valitulle testialueelle/testialueille (Tehtävä 4.)

SMK:lle toimitetaan seuraavat tulokset:

- Automaattisen taimikkotiedon tuotantoprototyyppi, joka sisältää lähtöaineistojen ja päättelysääntöjen kokoelman sekä sovelluksen eri tietolähteiden yhdistämiseksi ja taimikkotiedon tuottamiseksi automaattisesti erillisessä testiympäristössä.
- Lyhyt, korkeintaan 10-sivuinen raportti, joka sisältää kuvauksen käytetyistä aineistoista, menetelmistä ja tuloksista. Sisällöstä sovitaan tarkemmin SMK:n kanssa.

Hankkeen ensimmäistä osatehtävää toteutettaessa selvisi, että valmistelu- ja tarjousvaiheessa kirjattuja tavoitteita ei tulla täysin saavuttamaan. Lähtöaineisto oli hajautettuna SMK:ssa useaan eri tietokantaan, joilla oli erilaiset rakenteet ja tiedostoformaattit. Lisäksi aineisto oli hyvin heterogeenistä: eri puustotunnusten määrä taimikkokuviolta toiselle vaihteli ja eri lähteiden taimikkokuvioiden geometria ei ollut yhteneväinen vaan vaati erikoiskäsittelyä. Aineiston luku SMK:n tietokannoista, esikäsittelyt ja muokkaus analyysin vaatimaan tiedostoformaattiin edellyttivät tätä varten räätälöidyn Python-kirjaston toteutuksen. Tämä määrittely-, siirto ja tallennusvaihe veivät hankkeen resursseista n. 50% ennakoidun n. 10%:n sijaan. Tämän vuoksi alkuperäistä tavoitetta tarkennettiin hankkeen aikana vähentämällä aineistolähteiden lukumäärää ja jättämällä kahdesta testialueesta toinen pois. Keskeisin muutos oli taimikon pituuden estimointi SMK:n ilmakuva- ja latvuston korkeusmalliaineiston avulla.

2 Aineisto

2.1 SMK:n toimittama aineisto

Hankkeessa käytetty aineisto perustui SMK:n jo valmisteluvaiheessa tekemään suunnitelmaan [1]. SMK:n tekemässä aineistolistauksessa oli kullekin aineistolle määritelty yksilöllinen tunnusnumero (Taulukko 1.). Hanketta varten koottu aineisto kattoi koko Suomen.

Taulukko 1. Projektissa käytetty aineisto

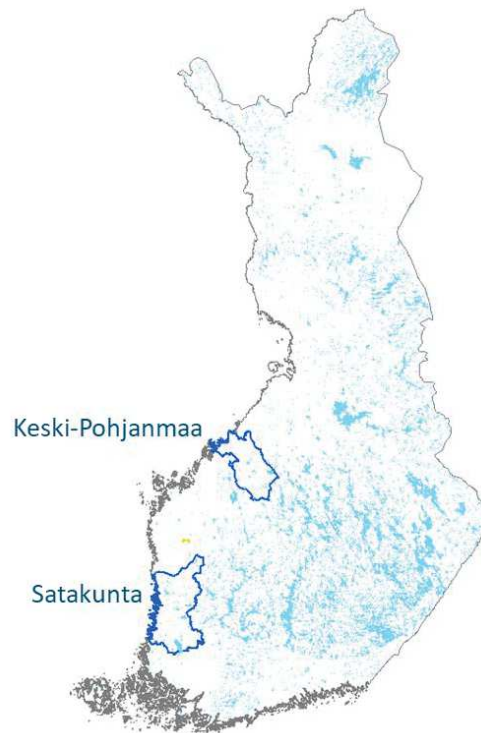
Aineiston tunnus-nro (1)	Aineiston tunnus	Aineisto	Kuvaus
1	D1	Metsävarakuviot	Aineisto ladattiin avoimesta metsään.fi-rajapinnasta Geopackage-formaatissa.
2, 10	D2_10	Kemera-aineisto	Aineisto ladattiin avoimesta metsään.fi-rajapinnasta Geopackage-formaatissa.
6, 8	D6_8	Metsänkäyttöilmoitukset	Aineisto ladattiin avoimesta metsään.fi-rajapinnasta Geopackage-formaatissa.
3	D3	Taimikon perustamisilmoitukset	Aineisto saatiin SMK:lta ArcGISin File Geodatabase -formaatissa (.gdb).
4, 11	D4_11	Toimenpidekuviot	Aineisto saatiin SMK:lta ArcGISin File Geodatabase -formaatissa (.gdb).
18	D18	Maastossa inventoidut taimikkokuviot	Aineisto saatiin SMK:lta ArcGISin File Geodatabase -formaatissa (.gdb).
19	D19_1, D19_2 2)	Maastoinventoiduille kuvioille tuotetut puustotulkintatiedot	Aineisto saatiin SMK:lta ArcGISin File Geodatabase -formaatissa (.gdb).
16	D16	Ilmakuvat	Ilmakuva-aineisto saatiin SMK:lta ECW-muotoisia kuvina, jotka oli rajattu karttalehtijaon mukaisiksi. Kolmekanavaisten kuvien resoluutio oli puoli metriä ja ne olivat projektiossa ETRS-TM35FIN. Kuvia oli vuosilta 2010-2017.
13	D13	Kasvillisuuden pintamalli (CHM)	CHM-aineisto saatiin SMK:lta GeoTiff-muotoisina kuvina, jotka oli rajattu karttalehtijaon mukaisiksi. Yksikanavaisten kuvien resoluutio oli yksi metri ja ne olivat projektiossa ETRS-TM35FIN. Kuvia oli vuosilta 2010-2017.

1) SMK:n määrittelemä tunnusnumero

2) VTT:n lisäämää alanimeroitua D19_1 ja D19_2 käytettiin erottelemaan inventointipuusto- ja laskentapuustotiedot toisistaan

2.2 Testialueet ja aineiston raja

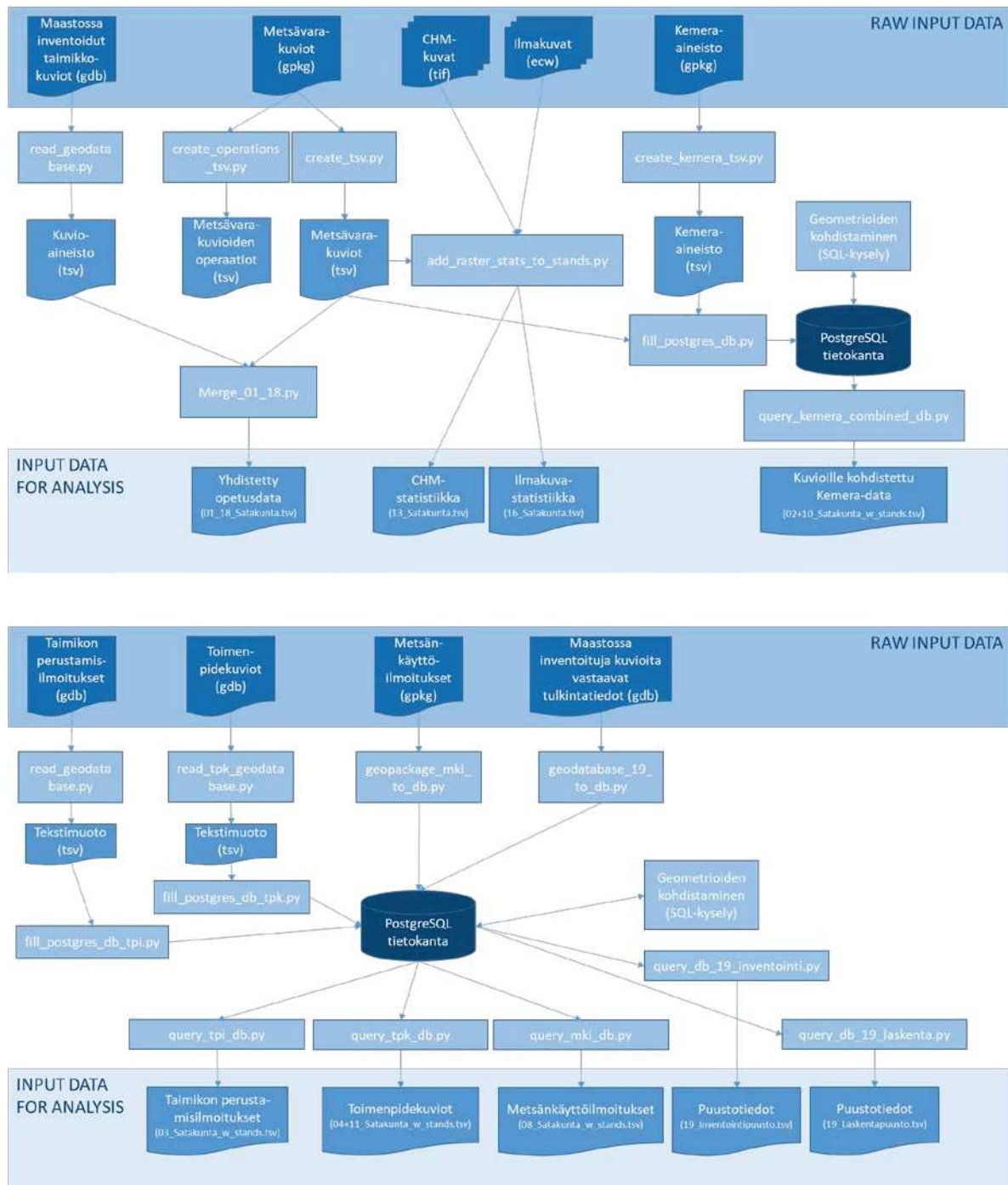
Projektin käynnistyskokouksessa määriteltiin testialueiksi Keski-Pohjanmaa ja Satakunta. Koska aineiston esikäsittely oli työlästä, sovittiin SMK:n kanssa hankkeen aikana, että testialueista Keski-Pohjanmaa jätetään pois ja keskityttiin Satakuntaan.



Kuva 1. Hankkeelle määritellyt testimaakunnat. Hankkeen aikana Keski-Pohjanmaa päätettiin jättää pois analyysistä.

2.3 Esikäsittely

Lähtöaineisto vaati runsaasti esikäsittelyä, jotta se saatiin muokattua mallinnuksen vaatimaan formaattiin. Esikäsittelyä varten tuotettiin joukko Python-ohjelmia pääosin Windows-ympäristössä ajettaviksi. Kuvien esikäsittelyssä tukeuduttiin Linux-ympäristöön sen tarjoamien kirjastojen vuoksi. Periaatekuva aineiston yhdistämisestä on esitetty Kuva 2.



Kuva 2. Lähtöaineiston esikäsittely ja analyysiaineiston koostaminen

Metsävarakuviot (D1) muokattiin tekstimuotoiseksi tiedostoksi, johon poimittiin tiedot kehitysluokkien A0, T1 ja T2 kuvioista (taimikkokuviot). Kuvioiden puulajikohtaiset tiedot aggregoitiin jakoon kuusi, mänty ja lehtipuut. Metsämuuttujista otettiin mukaan runkotilavuus, pohjapinta-ala, läpimitta, korkeus, runkoluku ja ikä. Aineistosta kerättiin myös poimittuihin kuvioihin liittyvä operaatiotieto.

Maastossa inventoidut taimikkokuviot (D18), sekä näille kuviolle laskettu, puustotulkintatiedoista inventointihetkeen kasvatetut tiedot (D19) kohdistettiin taimikkokuvioiden SMK:n näihin aineistoihin lisäämään kuvio-id:n avulla. Kemera-aineiston, Metsänkäyttöilmoitusten (MKI), toimenpidekuvioiden (TPK_MHTYO) ja taimikon perustamisilmoitusten (TPI) tiedot kohdistettiin kuviolle geometriatiedon perusteella. Näistä aineistoista otettiin mukaan kuviot, joiden yhteinen pinta-ala metsävaratiedon kuvioiden kanssa oli vähintään 80% kuvion pinta-alasta.

Ilmakuva- ja CHM-aineistoista laskettiin kehitysluokkien A0, T1 ja T2 metsikkökuvioille kaikille kuvien kanaville tilastotiedot minimi, maksimi, keskiarvo, mediaani ja keskihajonta (Taulukko 2). Metsikkökuvioiden reunasta poistettiin kolmen metrin puskuri, ja laskentaan otettiin ne kuvapikselit, joiden keskipiste osui jäljelle jääneen kuvion sisään.

Aineiston esikäsittelyn tuloksena piirvektorimatriisi (feature matrix), joka sisälsi Satakunnan alueelta puustotulkintatiedot metsävarakuvioista (aineisto 1.), Kemera-aineistot (aineistot 2. ja 10.), taimikon perustamisilmoitukset (aineisto 3.), toimenpidekuviot (aineistot 4. ja 11.), latvusmalliaineisto (aineisto 13.), ilmakuva-aineisto (aineisto 16.), maastossa inventoidut taimikkokuviot (aineisto 18.) ja aineiston 18. kuvioille puustotulkinnasta saadut tiedot (aineisto 19.1) ja niistä inventointihetken kasvatetut tiedot (aineisto 19.2). Metsänkäyttöilmoitukset (aineistot 6. ja 8.) jätettiin piirvektorimatriisista pois.

Tehtävässä 2. edellytettyä erillistä testausympäristöä varten VTT testasi Azure-ympäristöä pystyttämällä sinne virtuaaliserverin (Linux). Tämän ei kuitenkaan todettu hyödyttävän hankkeen kokonaistavoitetta, joten erillisestä testausympäristöstä luovuttiin. Osa aineistosta tallennettiin esikäsittelyvaiheessa PostgreSQL-tietokantaan käsittelyn yksinkertaistamiseksi.

Taulukko 2. Kuvioille lasketut Ilmakuva- ja latvusmallipiirteet

Ilmakuvapiirteet (kullekin kanavalle)		Latvusmallipiirteet	
Tunnus	Kuvaus	Tunnus	Kuvaus
mean	intensiteetin keskiarvo	mean	CHM keskiarvo
median	intensiteetin mediaani	median	CHM mediaani
min	intensiteetin minimi	max	maksimi
max	intensiteetin maksimi	std	keskihajonta
std	intensiteetin keskihajonta	MDM	keskiarvo/keskihajonta (Mean Div Median)
xxxRel	suhteellinen intensiteetti		
NDVI	Kasvillisuusindeksi (Normalized Difference Vegetation Index)		
MDM	intensiteetin keskiarvo		
contr	intensiteetin kontrasti		

3 Menetelmät

3.1 Analyysin rajaus

Analyysissä käytettiin maastossa inventoituja kuvioita ja niille mitattuja puustotietoja. Näistä kuvioista rajattiin osajoukko, jolle löytyi sekä ilmakuva-, että latvusmalli. Mikäli aineisto D19_2 oli oleellisena analyysissä, rajattiin kyseisestä aineistosta pois kuviot, jotka sisälsivät ajoitusristiriidan (n. 5000 kuviota).

Hankkeelle asetettuja tavoitteita lähestyttiin muodostamalla kuva kuvioon kohdistuneista mahdollisista toimenpiteistä (tietolähteet) ja niiden ajallisista riippuvuuksista. Kuva 3. esittää hypoteettisen kuvion tapahtumia aikajanaalla 2010 - 2018. Kuviolle tuotetut inventointipuusto- ja laskentapuustotiedot saadaan vaihtoehtoisista lähteistä (vihreät katkoviivanuolet). Kaikista Satakunnan taimikoista noin puolet oli inventoitu maastossa (aineisto D18). Niiden puustotietoja käytettiin menetelmäkehityksen referenssiaineistona (kuvion tavoitetiedot). Samoillem kuviolle oli SMK:n toimesta laskettu puustotulkinnasta tulleet tiedot, jotka oli kasvatettu kuvion maastoinventointihetkeen (aineisto D19_2). Tämä aineisto vastasi sitä tietoa, mikä kuviolla olisi, jos maastoinventointia ei olisi tehty.

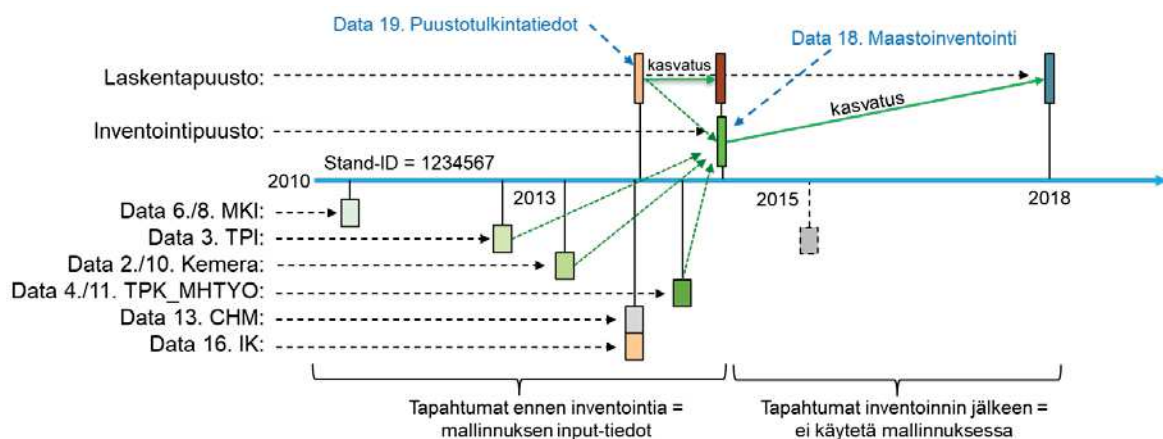
Kuvioiden puustotietojen korjausprosessissa erotettiin menetelmäkehitystä varten tapaukset a) ja b):

a) Kuviolla inventointipuustotiedot olemassa

- Inventointipuusto tuotettu kaukokartoitusaineistosta tai muita aineistoja yhdistämällä
- Tavoite: löytää kuviot, joiden puustotiedot väärää
- Tavoite: tuottaa kuviolle korjatut puustotiedot (jos väärää)

b) Inventointipuusto puuttuu

- Tavoite: tuottaa kuviolle puustotiedot



Kuva 3. Esimerkkikaavio aineistojen ajallisista riippuvuuksista. Kuvan aineisto Data 18. maastoinventointitieto sisältyy osaan (n. ½) kuvioista. (MKI = metsänkäyttöilmoitus, TPI = taimikon perustamisilmoitus, Kemera = Kemera-ilmoitus ('Kemera' = Kestävän metsätalouden rahoituslaki), TPK_MHTYO = toimenpidekuvio/metsänhoitotyö, CHM = latvusmalli, IK = ilmakuva)

SMK on käyttänyt puustotietojen oikeellisuuden arviointiin ja tulosten korjaukseen manuaalista lähestymistapaa, jossa eri lähteiden eriaikaisia aineistoja yhdistetään ja tulkitaan visuaalisesti ilmakuvia ja mahdollisesti CHM-aineistoa.

Alustavan analyysin tuloksena selvisi, että aineistojen D2_10 (Kemera), D3 (TPI) ja D4_11 (TPK_MHTYO) taimikoita oli vähän (ks. 3.2.1), ja että näissä lähteissä on puustotiedot vain pienessä osassa Kemera-ilmoituksia. Näiden aineistojen muu tietosisältö on toimenpide- tai työlajikoodeja, sekä niihin liittyviä aikaleimoja. Tämän aineiston tulkinta koneoppimisen menetelmin on huomattavasti vaativampaa kuin jatkuva-arvoisten numeeristen tietojen. Näiden tietojen käyttö taimikkotiedon parantamisessa vaatii todennäköisesti sääntöpohjaisen ja automaattisen koneoppimismenetelmän yhdistelmän. Koska kuvioille lasketun inventointipuuston eli aineisto D19_2:n puustotiedot olivat myös oletettavasti hyvin epävarmoja, päädyttiin esittämään SMK:lle puustotunnusten mallintamista ilma- ja latvusmallikuvien avulla. Alustava oletus oli, että näin pystyttäisiin estimoimaan kuvioiden puustotiedot aineistoa D19_2 paremmin. Aineistojen D_2_10, D3, ja D4_11 käyttö sovittiin SMK:n kanssa siirrettäväksi mahdollisiin jatkohankkeisiin, koska havaintojen pienen määrän vuoksi ja niiden käsittelyä ei katsottu kokonaisuuden kannalta hyödylliseksi.

Analyysin toteuttamisvaiheen suunnitelma oli seuraava (sovittiin SMK:n kanssa, 15.10.2018):

- Analyysin kohdeaineistoksi valittiin Satakunta. Maastossa inventoitujen kuvioiden aineistoa käytettiin vertailuaineistona. Nämä kuviot jaettiin kolmeen osaan: opetus-, validointi- ja testijoukkoon.
- Aineiston alustava analyysi, joka sisälsi muuttujien väliset Pearson-korrelaatiotaulukoiden ja tärkeimpien muuttujien jakaumien laskennan.
- Puustotunnusten estimointi käyttäen tähän soveltuvaa jatkuvien muuttujien mallinnusmenetelmää. Kohdemuuttujiksi valittiin keskiläpimitta (D), pituus (H), runkoluku (N) ja ikä (T), joille aineistossa oli puulajikohtaiset ositetiedot. Estimointimallit tuotettiin männylle (pine) ja kuuselle (spruce). Alustassa analyysissä tuloksia laskettiin myös lehtipuille (bl).
- Estimaattimallien suorituskky arvioitiin laskemalla testiaineiston suhteellinen RMS-virhe (RMSE%), suhteellinen harha (BIAS%) ja selitysaste (R^2).
- Myös syvien neuroverkkojen mahdollinen lisäarvo puustotunnusten estimointitarkkuuteen testattiin.
- Puustotunnusten estimaatit tuotettiin myös Satakunnan kuvioille, joille ei ole maastokäyntiä. Tulosten verifiointi jätetään SMK:n toteutettavaksi esim. nykytietoon verraten ja maastokäynnein (2019).

3.2 Aineiston alustava analyysi

3.2.1 Eri tietolähteiden määrä Satakunnan aineistossa

Alustavassa analyysissä pyrittiin varmistumaan kootun analyysiaineiston oikeellisuudesta ja eri lähteiden keskinäisestä konsistenssista. Löydetyt epä johdonmukaisuudet ja puutteet, kuten ajoitusristiriidat ja eri aineistoja linkittävän kuvio-ID:n puuttuminen korjattiin ennen jatkokäsittelyä. Tässä vaiheessa pyrittiin myös saamaan käsitys eri tietolähteiden merkityksestä puustotietojen estimoinnissa.

Satakunnan alueelta kootulle aineistolle laskettiin kutakin eri lähtöaineistotyyppiä sisältävien kuvioiden lukumäärät. Taulukko 3. sisältää tiedot kaikilta satakunnan alueen kuvioista, joiden kehitysluokka on A0, A1 tai A2. Yhteensä kuvioita oli 63620, joista suurin osa, 42812 kuviota, kuului kehitysluokkaan T2, 13295 kehitysluokkaan A0 ja loput 7513 kuviota kehitysluokkaan T1. Ilmakuva- ja CHM-tieto oli lähes kaikille kuvioille (n. 94 % kuvioista). Taulukon yläosassa on koostettu eri kehitysluokkiin kuuluvien kuvioiden määrä, alaosassa lukumäärät eri aineistoja sisältävistä kuvioista. Taulukkoon laskettu kuvioiden lukumäärä tarkoittaa, että ainakin kyseinen tieto (esim. Kemera-ilmoitus) on näin monelle kuviolle.

Menetelmäkehityksen kannalta oleellinen tieto on, kuinka moni kuvioista sisältää maastossa inventoidut tiedot. Tiedot näistä kuvioista on koottu Taulukkoon 4. Kuten edellisessä taulukossa, ilmoitetut lukumäärät kertovat, että ainakin kyseinen tieto löytyy näin monelta kuviolta.

Kun laskettiin kuviot, joilta löytyy useita tietolähteitä yhtä aikaisesti, päädyttiin huomattavasti pienempiin lukumääriin (Taulukko 5.). Kun mukaan otettiin T2 kuviot, joista löytyy maastoinventointitiedon lisäksi ilmakuva- ja CHM-tieto, verifioitu kasvatettu puustotulkintatieto (D19_2; vertailua varten), sekä Kemera-ilmoitus, on kuvioiden lukumäärä enää 445. Näistä kuvioista, joilla Kemera-tiedoissa on puuston runkoluku ja keskimääräinen läpimitta ilmoitettu, on vain 20.

Muista lähteistä ei suoraan ole saatavilla puustotulkintatietoja, vaan päivämäärä- ja kategoriatyypistä tietoa (kuten puulaji-, metsänhoitotyö- tai uudistamistyyppikoodi).

Taulukko 3. Yhteenvedo Satakunnan alueen taimikkokuvioiden (A0, T1 ja T2) tietolähteistä

Satakunta, kaikki kuviot (2010 - 2018), kehitysluokat A0, T1, T2						
Metsikkökuviot <> tietolähde	Tietolähde#	Kuvioiden lkm	%-lkm	Kuvioiden yhteis-PA [ha]	%-PA	Kuvioiden keskim. PA [ha]
Kuvioita, kehitysluokka A0	01	13295	20.9	15392.6	18.6	1.2
Kuvioita, kehitysluokka T1	01	7513	11.8	9951.1	12.0	1.3
Kuvioita, kehitysluokka T2	01	42812	67.3	57241.1	69.3	1.3
Kuvioita yhteensä	01	63620	100.0	82584.9	100.0	
Kuvioita, joille ilmakuvatieto	16	59951	94.2	77780.5	94.2	1.3
Kuvioita, joille puuston latvusmalli (CHM)	13	59484	93.5	77361.6	93.7	1.3
Toimenpidekuvioita	4 / 11	192	0.3	294.7	0.4	1.5
Kuvioita, joille Kemera-ilmoitus	2 / 10	1418	2.2	2231.8	2.7	1.6
Kuvioita, joille taimikon perustamisilmoitus	3	1546	2.4	2141.8	2.6	1.4
Kuvioita, joille tehty maastokäynti	18	28070	44.1	38321.8	46.4	1.4

Taulukko 4. Yhteenvedo Satakunnan alueen maastossa inventoitujen taimikkokuvioiden (A0, T1 ja T2) tietolähteistä

Satakunta, maastoinventoidut kuviot (2010 - 2018), kehitysluokat A0, T1, T2						
Metsikkökuviot <> tietolähde	Tietolähde#	Kuvioiden lkm	%-lkm	Kuvioiden yhteis-PA [ha]	%-PA	Kuvioiden keskim. PA [ha]
Kuvioita, kehitysluokka A0	01/18	232	0.8	203.4	0.5	0.9
Kuvioita, kehitysluokka T1	01/18	1577	5.6	2169.9	5.7	1.4
Kuvioita, kehitysluokka T2	01/18	26261	93.6	35948.6	93.8	1.4
Kuvioita yhteensä	01/18	28070	100.0	38321.8	100.0	
Kuvioita, joille ilmakuvatieto	16	27552	98.2	37664.9	98.3	1.4
Kuvioita, joille puuston latvusmalli (CHM)	13	27547	98.1	37654.4	98.3	1.4
Toimenpidekuvioita	4 / 11	101	0.4	153.8	0.4	1.5
Kuvioita, joille Kemera-ilmoitus	2 / 10	862	3.1	1408.7	3.7	1.6
Kuvioita, joille taimikon perustamisilmoitus	3	1150	4.1	1654.7	4.3	1.4
Kuvioita, joilla data 19. (1)	19.1 / 19.2	15353	54.7	21502.4	56.1	1.4

1) Aineistosta poistettu datasetin 19. kuviot, joilla aikaleimastirijaa

Taulukko 5. Yhteenvedo Satakunnan alueen maastossa inventoiduista, kehitysluokkaan T2 kuuluvista taimikkokuvioista, joille löytyy ilmakehän- ja latvusmallitieto, sekä puustotulkinnasta saatu inventointihetkeen kasvatettu tieto (aineisto D19_2), Kemera-ilmoitus ja Kemera-ilmoitus, joissa ilmoitettu runkoluku (N) ja pituus (H)

Satakunta, maastoinventoidut kuviot (2010 - 2018), kehitysluokka T2						
Metsikkökuviot <> tietolähde	Tietolähde#	Kuvioiden lkm	%-lkm	Kuvioiden yhteis-PA [ha]	%-PA	Kuvioiden keskim. PA [ha]
Kuvioita, joille ilmakuvatieto latvusmalli ja data 19	18/13/16/19	14470	51.5	20335.0	53.1	1.4
... kuviot, joille myös Kemera-ilmoitus	18/13/16/19/2/10	445	1.6	777.7	2.0	1.7
... kuviot, joille myös Kemera-tiedoista N ja H	18/13/16/19/2/10	20	0.1	39.0	0.1	1.9

3.2.2 Jakaumat ja korrelaatiotaulukot

Kootun Satakunnan aineiston rakenteen selvittämiseksi tulostettiin puustotunnusten jakaumat (histogrammit) eri osajoukkoja (esim. kehitysluokka) tarkastelemalla. Pyrkimyksenä oli varmistua aineiston yleisestä virheettömyydestä ja löytää mahdolliset epäjohtamismuutokset. Alustavassa analyysissä laskettiin myös maastossa inventoitujen puustotietojen ja muiden, eri lähteistä saatujen tietojen väliset Pearson-korrelaatiotaulukot.

Estimoitavien puustotunnusten ja lähtöaineiston muuttujien välisten riippuvuuksien selvittämiseksi laskettiin maastossa mitatun aineiston (D18) puustotunnusten pituus (H), rungon läpimitta (D) ja runkoluku (N) ja ilmakehänpiirteiden (D16) sekä latvusmallipiirteiden (D13) väliset korrelaatiotaulukot (Pearson-korrelaatio). Myös aineiston D19_2 puustotunnusten (D, H, N) ja aineistojen D16 ja D13 väliset korrelaatiotaulukot, sekä aineistojen D18 ja D19_2 keskinäiset korrelaatiotaulukot laskettiin. Käytetyt ilmakehän- ja latvusmallipiirteet on esitetty Taulukko 2. Korrelaatiotaulukot laskettiin neljälle eri pituusluokalle: 0 - 2 m, 2 - 4 m, 4 - 6 m ja > 6 m kehitysluokissa T1 ja T2. Kunkin pituusluokan laskennassa huomioitiin vain tapaukset (kuviot), joilla kaikkien puulajien pituudet olivat ko. luokalle määritellyllä välillä. Korrelaatiotaulukoiden laskennassa käytetyt aineistot on koottu Taulukko 6.

Taulukko 6. Korrelaatioaineistot

#	Puustotunnukset	x	Piirteet/Puustotunnukset
1	D18 (D, H, N)		D19_2 (D, H, N)
2	D18 (D, H, N)		Ilmakuvapiirteet
3	D18 (D, H, N)		Latvusmallipiirteet
4	D19_2 (D, H, N)		Ilmakuvapiirteet
5	D19_2 (D, H, N)		Latvusmallipiirteet

3.3 Puustotunnusten estimointi

3.3.1 Estimointiaineiston jako

Estimointia varten aineisto jaettiin opetus-, validointi ja testijoukkoihin. Tässä käytettiin satunnaistettua jakoa, joka perustui standID:n viimeiseen numeroon. Eri joukkojen edustavuus tarkistettiin vertaamalla eri joukkojen jakaumia valituilla metsämuuttujilla (H-pine, H-spruce). Mallinnuksessa kokeiltiin useita satunnaisesti valittuja jakoja ko. joukkoihin. Yleisimmät jakosuhteet eri joukkojen välillä olivat (train/validation/test): 60/20/20, 50/20/30 ja 50/30/20 (%). Tuloksissa ei havaittu merkittäviä eroja eri jaoilla tai jakosuhteilla.

3.3.2 Estimointimenetelmät

Mallinnuksessa käytettiin yhden tai useamman piilokerroksen täysin kytkettyä neuroverkkoa (fully connected MLP) lähtien yksinkertaisista verkoista (Työkalu: Matlab). Estimointimenetelmänä käytettiin pääosin yhden piilokerroksen neuroverkkoa. Näiden mallien opetusta ohjattiin kehysohjelmalla (wrapper), jossa toteutettiin kunkin mallinnettavan puustotunnuksen kannalta parhaiden piirteiden haku ns. brute force forward selection – menetelmällä. Menetelmässä käydään läpi kaikki tulomuuttujakombinaatiot, ja etsitään parhaiden piirteiden valikoima ja lukumäärä kyseiselle mallinnettavalle suureelle. Parhaaksi malliksi katsotaan se, joka tuottaa pienimmän testijoukon suhteellisen RMS-virheen. Johtuen Estimointimenetelmien stokastisesta alustuksesta ja eri piirteiden samankaltaisesta 'hyvyydestä' mallin tulomuuttujina, saadaan kehysohjelman toistoilla käytännössä eri hieman tuloksia. Toistamalla kehysohjelmaa useita kertoja on mahdollista saada tilastollista varmuutta mallin virhetunnuksiin ja parhaiden piirteiden valikoimaan.

Mahdollinen syvien neuroverkkojen (=deep neural network, DNN) mukanaan tuoma etu estimaattien tarkkuuteen tarkistettiin toteuttamalla usean piilokerroksen neuroverkkomallit männyn ja kuusen pituudelle. Nämä ns. syvät neuroverkkomallit toteutettiin Keras-ympäristössä, joka on neuroverkkojen käyttöön tarkoitettu korkean tason avoimen lähdekoodin ohjelmointiympäristö/kirjasto (framework), joka on toteutettu Python-ohjelmointikielellä (<https://keras.io/>). Termi 'Syvä neuroverkko' on tässä yhteydessä hieman harhaanjohtava - yleisesti tällä tarkoitetaan neuroverkkoja, joissa on kymmeniä tai satoja piilokerroksia.

Mallinnuksen yksikkönä oli kuvio, jolloin kohdemuuttujina olivat kuviokohtaiset puustotunnukset.

Yhteenveto valituista Estimointimenetelmistä:

1) Matala Neuroverkkomalli

- Yhden piilokerroksen multi-layer perceptron (MLP) neuroverkkomalli
- 10 - 26 ilmakuva- ja CHM-piirrettä mallin tulomuuttujina

- Piirteiden valinta: Brute force forward selection –menetelmä (wrapper)
- Toteutus: Matlab-ohjelma

2) 'Syvä' neuroverkkomalli (Deep Neural Network)

- 3 / 4 / 7 kerroksen täysin kytketty (MLP) neuroverkko
- Kaikki 26 IK- ja CHM-piirrettä tulomuuttujina
- Toteutus: Keras + Tensorflow (Python-ohjelma)

3.4 Menetelmä virheellisten kuvioiden löytämiseksi

Hankkeessa kehitettiin menetelmä, jolla voidaan kullekin kuviolle tuottaa eräänlainen laatuindeksi, joka kuvaa keskimääräisellä tasolla kuviolle tuotetun puustotunnuksen laatua tai odotettavissa olevaa virhettä. Menetelmässä kaikki kuviot, joille oli sekä ilmakehän-, että latvusmalliaineisto, koottiin lähtöaineistoksi. Tälle kuviojoukolle tehtiin ohjaamaton klusterointi (K-means), jonka tuloksena saatiin N kpl kuvioryhmiä. Nämä kuvioryhmät erosivat toisistaan klusteroinnin input-muuttujien (ilmakehän- ja latvusmallipiirteet) suhteen. Maastossa inventoiduille kuviolle tuotettiin estimointimallilla männyn pituuden estimaatti ja kuviokohtainen virhe, jonka avulla laskettiin kullekin klusterille virheen keskiarvo. Oletuksena oli, että klusterien keskimääräiset virheet poikkeavat toisistaan, jolloin kuviot (tai kuvioryhmät) voidaan erottaa toisistaan virheen suuruuden perusteella.

Menetelmällä saadaan jatkuva-arvoinen laatuindeksi kaikille kuviolle, jotka ovat olleet klusteroinnin input-aineistossa. Indeksiksi voi olla esim. edellä kuvattu virheen klusterikohtainen keskiarvo, tai asetetun kynnyksen ylittäneiden kuvioiden osuus klustereissa. Kuviolle voidaan myös tuottaa virhekategorioiden käyttäen vapaasti määriteltäviä virhekyynnyksiä.

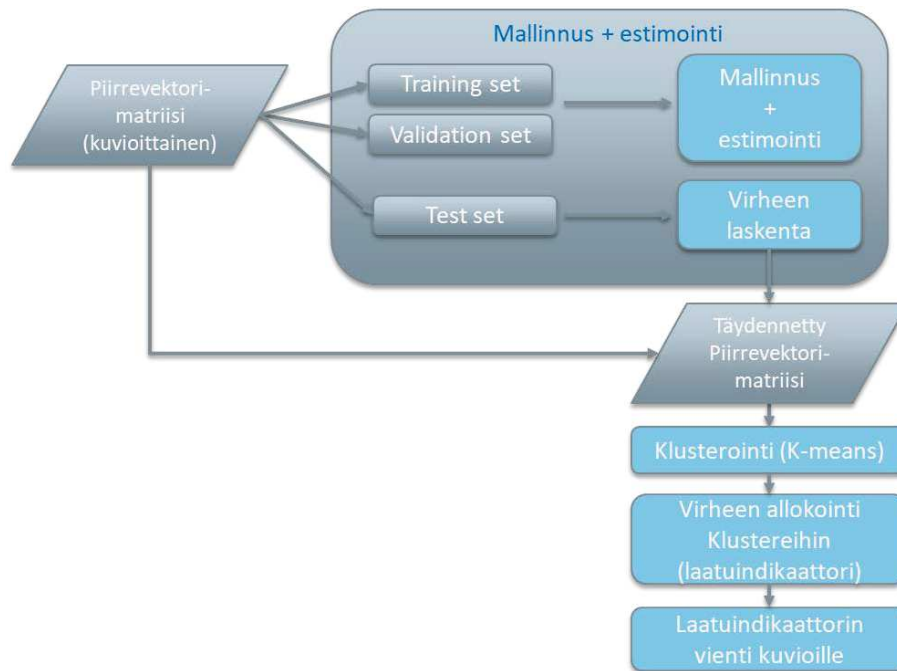
Menetelmän pääkohdat ovat (Kuva 4.):

1. Estimoidaan puustotunnukset ilmakehän- ja CHM-piirteiden ja maastoinventointitietojen avulla
2. Lasketaan mallin estimointivirhe opetus-, validointi- ja testikuviolle
3. Klusteroidaan koko aineisto (kaikki kuviot)
 - N klusteria ($N = 30, 40, \dots$)
 - Klusteroinnin tulomuuttujat: esim. mallin tulomuuttujat (+ mahdollisesti muita)
 - Kullekin kuviolle saadaan klusterinumero (label)
4. Allokoidaan kohdassa 2. laskettu virhe saatuihin klustereihin erikseen opetus- (validointi-) ja testijoukolle
 - Lasketaan laatuindeksi (= virhetunnus) opetus- ja testijoukolle
 - Oletus: virhe jakautuu epätasaisesti klustereihin. Tämä on toivottava tilanne.
5. Lasketaan laatuindeksi kaikille kuviolle, jotka olivat mukana klusteroinnissa

Kohdan 1. tarkoitus on tuottaa kuviolle puustotunnukset. Jos nämä saadaan jostain muusta lähteestä (esim. SMK:n aineisto), voidaan kyseinen kohta ohittaa.

Tavoitteena oli toteuttaa menetelmä, jonka avulla voidaan maastotyötä (tarkastukset) kohdentaa kuvioiden, joissa on odotettavissa suurimmat virheet. Menetelmästä toteutettiin prototyyppiohjelma ja tuotettiin virhetunnus (laatuindeksi) VTT:n estimointimallista saadulle

männyn pituudelle. Menetelmäkehityksessä opetus- ja validointikuviot yhdistettiin yhdeksi opetusjoukoksi. Menetelmän toimivuutta testattiin testikuvioille lasketun laatuindeksin avulla.



Kuva 4. Puustotunnusten laatuindeksin laskeminen

4 Tulokset

Esimerkkejä hankkeen päätuloksista on esitetty alla. Täydellisemmät tulostukset on esitetty liitteissä.

4.1 Puustotunnusten jakaumat

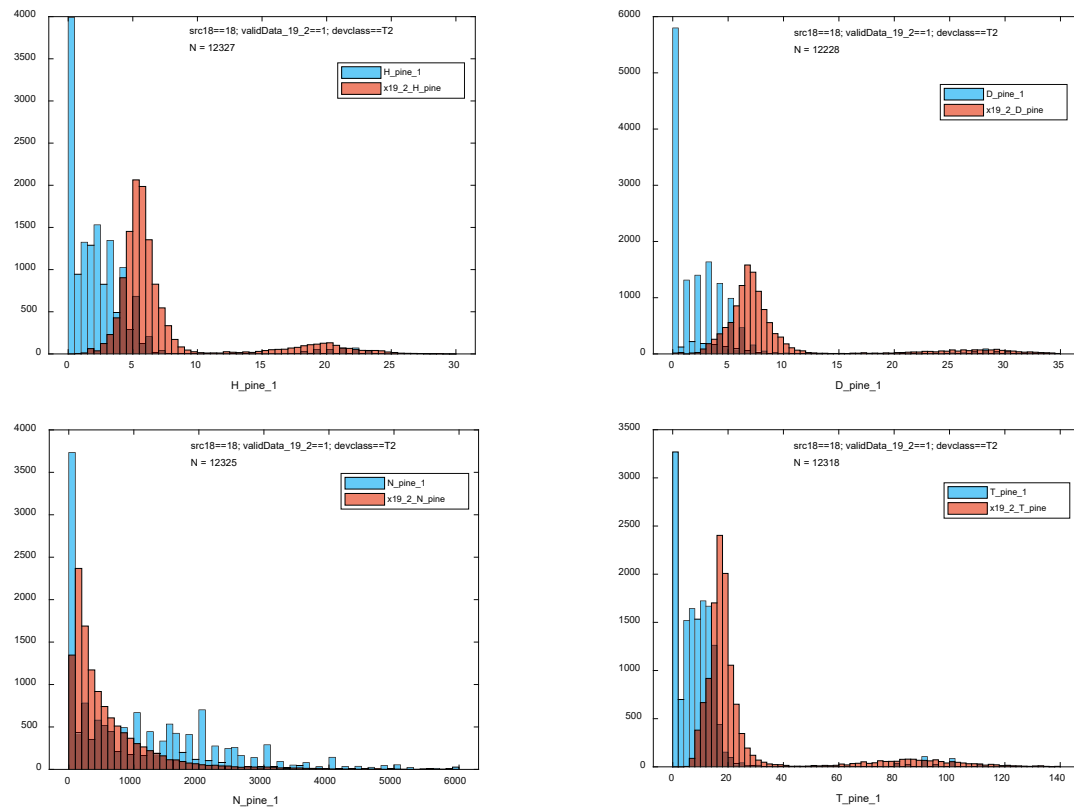
Maastossa mitattujen puustotietojen (aineisto D18) ja maastoinventointihetken puustotulkinnasta kasvatettujen puustotietojen (aineisto D19_2) runkoluku-, pituus- ja keskiläpimittajakaumat laskettiin puulajeille mänty, kuusi ja lehtipuu (LIITE A). Kehitysluokat A0, T1 ja T2 esitetty omilla kuvissaan.

Aineiston jakaumien tarkastelussa löydettiin kehitysluokkien A0 ja T1 kuvioaineistosta epäjohdonmukaisuuksia, jotka todennäköisesti johtuivat siitä, että kuvion kehitysluokka oli muutettu päätehakkuun tai muun toimenpiteen jälkeen, mutta puustotunnukset oli jätetty ennalleen. Tämä ei välttämättä ole ongelma SMK:n prosessien kannalta, mutta oli otettava huomioon aineiston analyysissä. Nämä kuviot oli kuitenkin helppo löytää kehitysluokittaisia puustotunnuksia tarkastelemalla. Koska kehitysluokkien A0 ja T1 osuudet maastossa inventoiduista kuvioista olivat varsin pieniä ja osassa kuvioita puustotiedot olivat vanhentuneita, keskityttiin mallinnuksessa kehitysluokan T2 taimikoihin.

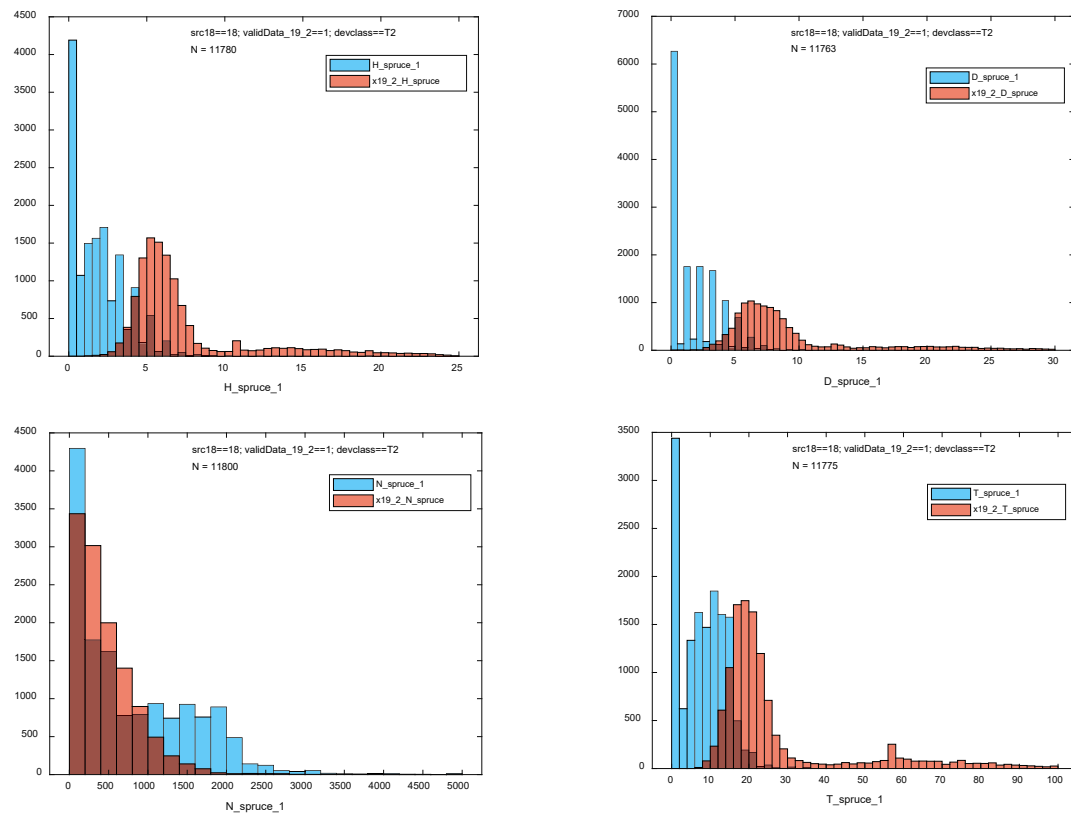
4.1.1 Aineiston D19_2 vertailu maastossa inventoituun tietoon

Puustotulkinnasta maastoinventointihetken kasvatettua puustotietoa (D19_2) verrattiin maastossa inventoituun tietoon tulostamalla muuttujien H, D, N, ja T jakaumat kehitysluokan T2 kuvioille (Kuva 5. mänty; Kuva 6. kuusi). Vertailu koski maastossa inventoituja kuvioita, joille oli myös aineiston D19_2 tieto.

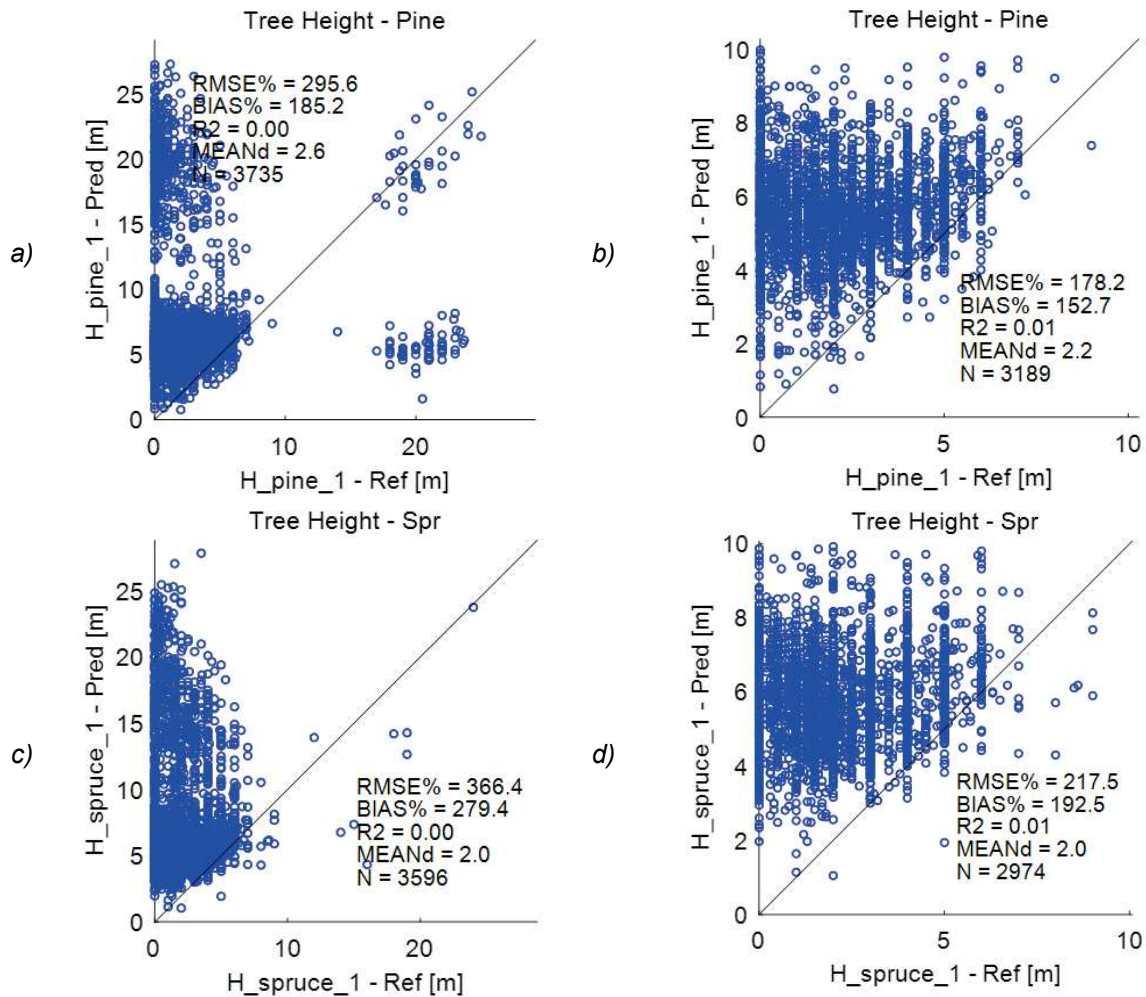
Yksittäisten kuvioiden männyn ja kuusen pituuden vastaavuutta tutkittiin vielä tarkemmin tulostamalla ko. aineistojen väliset hajontakuviot (Kuva 7.). Muiden muuttujien (D, N ja T) vastaavat hajontakuviot ks. LIITE C.



Kuva 5. Maastossa mitattujen (sininen) ja aineiston D19_2 (oranssi) puustotunnusten H-pine, D-pine, N-pine, ja T-pine jakaumat.



Kuva 6. Maastossa mitattujen (sininen) ja aineiston D19_2 (oranssi) puustotunnusten H-spruce, D-spruce, N-spruce, ja T-spruce jakaumat.



Kuva 7. Männyn (a ja b) ja kuusen (c ja d) pituuden hajontakuviot maastossa inventoidun tiedon ($H_{\text{pine}_1} - \text{Ref}$ ja $H_{\text{spruce}_1} - \text{Ref}$) ja aineiston D19_2 ($H_{\text{pine}_1} - \text{Pred}$ ja $H_{\text{spruce}_1} - \text{Pred}$) välillä. Kuvissa b) ja d) aineistot rajattu pituuksiin $0 \leq H < 10$ m. Tulostuksessa mukana kuviot, joiden kuvionumeron (standID) viimeinen numero 2, 4 tai 6 (mallinnuksen testijoukko).

Muuttujien H, D, ja T eri aineistojen jakaumissa oli huomattavia eroja. Sekä histogrammeista, että pituuden hajontakuvioista nähtiin, että aineistossa D19_2 oli varsin suuri yliarvio (myös keskiläpimitalla ja iällä, ks. LIITE C), kun taas runkoluvussa oli aliarvio.

Myös maastossa inventoiduissa kehitysluokan T2 kuvioiden puustotunnuksissa oli arvoja, jotka viittaavat varttuneeseen metsään, sillä männyn pituus saattoi olla yli 19 m. Nämä kuviot päätettiin jättää jatkossa pois mallinnuksesta. Estimoinnissa opetusaineistoon hyväksyttiin kuviot, joille $0 < H_{\text{species}} < 10$ m (species $\in$ [pine, spruce]).

4.2 Korrelaatiotaulukot

Taulukoissa Taulukko 7. - Taulukko 10. on esimerkit lasketuista korrelaatioista. Esimerkkien aineistot ovat kehitysluokan T2 maastoinventoitujen kuvioiden inventointipuusto- ja vastaavien kuvioiden puustotulkinnasta maastoinventointihetkeen kasvatetut laskentapuustotiedot. Puustomuuttajat olivat läpimitta (D), pituus (H) ja runkoluku (N) puulajeille mänty, kuusi koivu. Kunkin pituusluokan korrelaatiot ovat omassa taulukossaan.

Alle kuuden metrin taimikoissa korrelaatiot olivat hyvin alhaiset, mutta taimikon koon yltäessä yli kuuden metrin monet korrelaatiot olivat korkeita. Näitä taimikoita oli kuitenkin vain 43 kpl. Taulukot paljastavat myös, että runkoluvun korrelaatiot olivat aina alhaiset. Taulukko 10:ssä eri puulajien muuttajat olivat vahvasti korreloituneita.

Loput korrelaatiotaulukot esitetään LIITE B.

Taulukko 7. Korrelaatiotaulukko D18 vs. D19_2 (D, H, N); Kehitysluokka = T2, Pituusluokka 0 - 2 m

T2 0 - 2 m	D19_2 piirteet						N = 1044
D18 Piirteet	x19_2_N_pine	x19_2_D_pine	x19_2_H_pine	x19_2_N_spruce	x19_2_D_spruce	x19_2_H_spruce	
D_pine_1	0.127	-0.115	-0.121	0.206	-0.115	-0.123	
H_pine_1	0.103	-0.133	-0.140	0.241	-0.180	-0.183	
N_pine_1	-0.016	0.098	0.089	-0.122	-0.033	-0.024	
D_spruce_1	0.144	-0.196	-0.190	0.204	-0.116	-0.116	
H_spruce_1	0.147	-0.178	-0.180	0.238	-0.150	-0.152	
N_spruce_1	-0.078	-0.025	-0.016	0.074	0.037	0.021	
D_bl_1	0.077	-0.107	-0.113	0.152	-0.120	-0.118	
H_bl_1	0.043	-0.082	-0.087	0.138	-0.109	-0.113	
N_bl_1	-0.038	-0.034	-0.034	0.039	-0.004	-0.020	

Taulukko 8. Korrelaatiotaulukko D18 vs. D19_2 (D, H, N); Kehitysluokka = T2, Pituusluokka 2 - 4 m

T2 2 - 4 m	D19_2 piirteet						N = 795
D18 Piirteet	x19_2_N_pine	x19_2_D_pine	x19_2_H_pine	x19_2_N_spruce	x19_2_D_spruce	x19_2_H_spruce	
D_pine_1	0.206	-0.010	-0.012	0.058	-0.039	-0.034	
H_pine_1	0.218	-0.013	-0.018	0.060	-0.051	-0.053	
N_pine_1	0.164	-0.007	0.002	-0.031	-0.048	-0.047	
D_spruce_1	0.007	0.018	0.014	0.088	-0.027	-0.010	
H_spruce_1	-0.006	-0.027	-0.038	0.094	-0.007	-0.003	
N_spruce_1	-0.144	-0.066	-0.071	0.043	0.049	0.044	
D_bl_1	-0.018	-0.077	-0.069	0.156	-0.098	-0.096	
H_bl_1	0.144	-0.063	-0.060	0.064	-0.032	-0.033	
N_bl_1	0.118	-0.034	-0.042	0.125	-0.066	-0.062	

Taulukko 9. Korrelaatiotaulukko D18 vs. D19_2 (D, H, N); Kehitysluokka = T2, Pituusluokka 4 - 6 m

T2 4 - 6 m	D19_2 piirteet						N = 404
D18 Piirteet	x19_2_N_pine	x19_2_D_pine	x19_2_H_pine	x19_2_N_spruce	x19_2_D_spruce	x19_2_H_spruce	
D_pine_1	0.122	0.015	0.011	0.071	-0.017	-0.008	
H_pine_1	0.087	0.047	0.038	0.140	0.028	0.017	
N_pine_1	0.282	-0.015	-0.020	-0.054	-0.103	-0.099	
D_spruce_1	-0.016	0.043	0.050	0.037	0.030	0.039	
H_spruce_1	-0.038	0.090	0.083	0.120	0.036	0.026	
N_spruce_1	-0.181	-0.055	-0.049	0.045	0.023	0.005	
D_bl_1	-0.232	0.094	0.091	-0.044	0.071	0.070	
H_bl_1	0.121	0.038	0.037	-0.045	0.027	0.029	
N_bl_1	0.043	-0.111	-0.100	0.134	-0.074	-0.076	

Taulukko 10. Korrelaatiotaulukko D18 vs. D19_2 (D, H, N); Kehitysluokka = T2, Pituusluokka > 6 m

T2 6 m plus	D19_2 piirteet						N = 43
D18 Piirteet	x19_2_N_pine	x19_2_D_pine	x19_2_H_pine	x19_2_N_spruce	x19_2_D_spruce	x19_2_H_spruce	
D_pine_1	-0.198	0.781	0.816	-0.273	0.690	0.789	
H_pine_1	-0.195	0.814	0.845	-0.309	0.719	0.818	
N_pine_1	0.080	-0.329	-0.360	-0.008	-0.203	-0.223	
D_spruce_1	-0.225	0.782	0.815	-0.273	0.705	0.804	
H_spruce_1	-0.219	0.787	0.819	-0.304	0.724	0.819	
N_spruce_1	0.250	-0.095	-0.107	-0.192	-0.204	-0.217	
D_bl_1	-0.241	0.743	0.774	-0.254	0.717	0.809	
H_bl_1	-0.190	0.791	0.827	-0.341	0.730	0.824	
N_bl_1	-0.051	-0.124	-0.105	-0.155	0.054	0.003	

4.3 Puustotietojen estimointi

Kappaleessa 3.2.2 kuvattujen menetelmillä estimointiin kuvioiden puulajikohtaisia puustotunnuksia maastossa tulkittujen inventointitietojen avulla. Estimoidut puustotunnukset olivat taimikon pituus (H), keskiläpimitta (D), runkoluku (N) ja ikä (T) puulajeille mänty ja kuusi. Osassa testeistä muuttujina on vain männyn pituus ja kuusen pituus.

4.3.1 Piirrevalintatestit

Relevanttien ilmakeu- ja latvusmallipiirteiden määrän vähentämiseksi, ja vastaavasti toteutettavien estimointiajojen nopeuttamiseksi tehtiin männyn ja kuusen pituudelle piirrevalintatestit. Näille muuttujille etsittiin parhaat selittävät piirteet kappaleessa 3.2.2 esitetyllä brute force forward selection -menetelmällä, yhden piilokerroksen MLP-neuroverkolla. Koska testin tuloksena ei saatu suppeampaa yhteistä joukkoa piirteitä, joiden olisi voitu tulkita ennustavan hankkeessa käytettyjä metsämuuttujia yleisesti, käytettiin muiden muuttujien estimointimallien opetuksessa kaikkia 26 ilmakeu- ja latvusmallipiirrettä ja kuvattua piirrevalintamenetelmää. Piirrevalintatestin tulokset esitetään LIITE D.

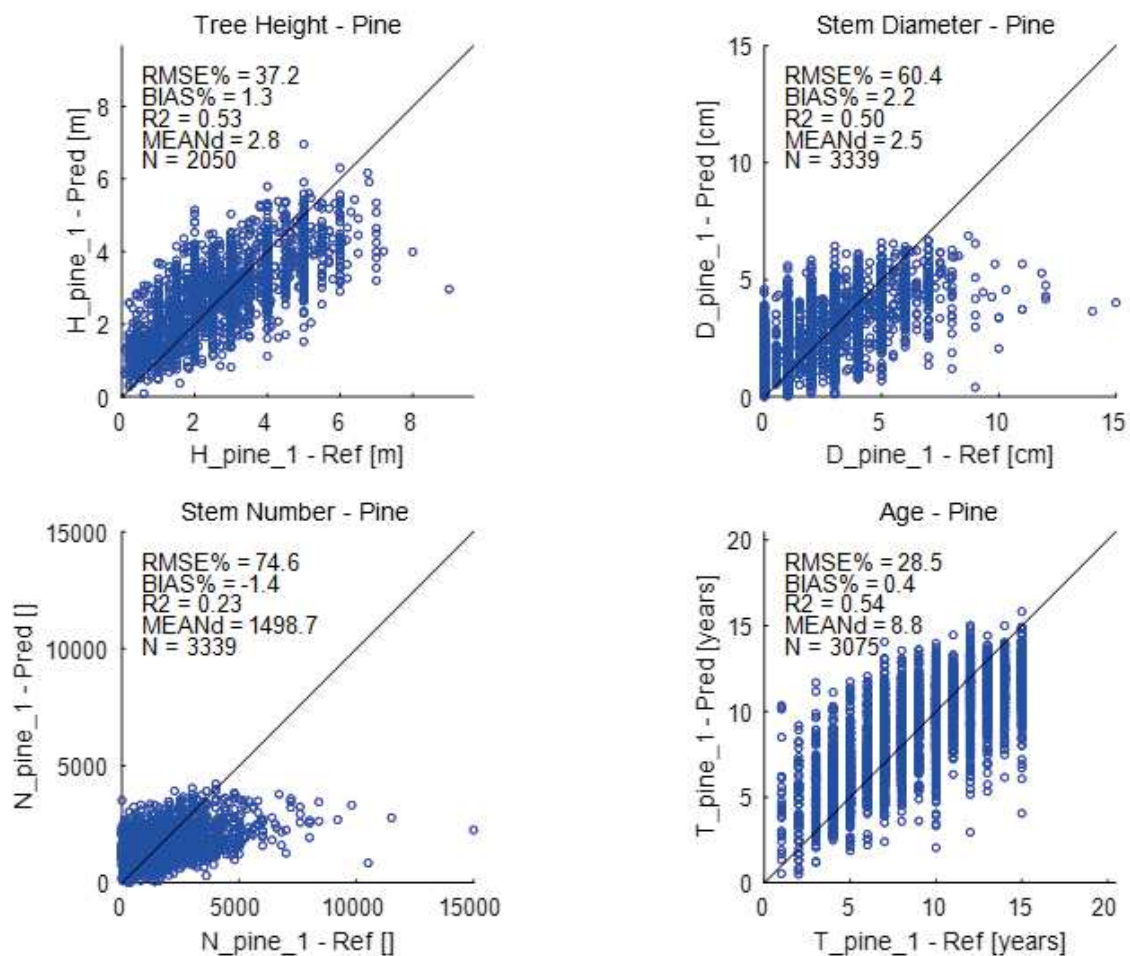
4.3.2 Puustotietojen estimointi yhden piilokerroksen neuroverkolla

Yhden piilokerroksen neuroverkkomallit tuotettiin kappaleessa 3.2.2 kuvatulla menetelmällä (yhden piilokerroksen MLP + piirrevalinta) muuttujille H, D, N ja T, ja puulajeille mänty (pine)

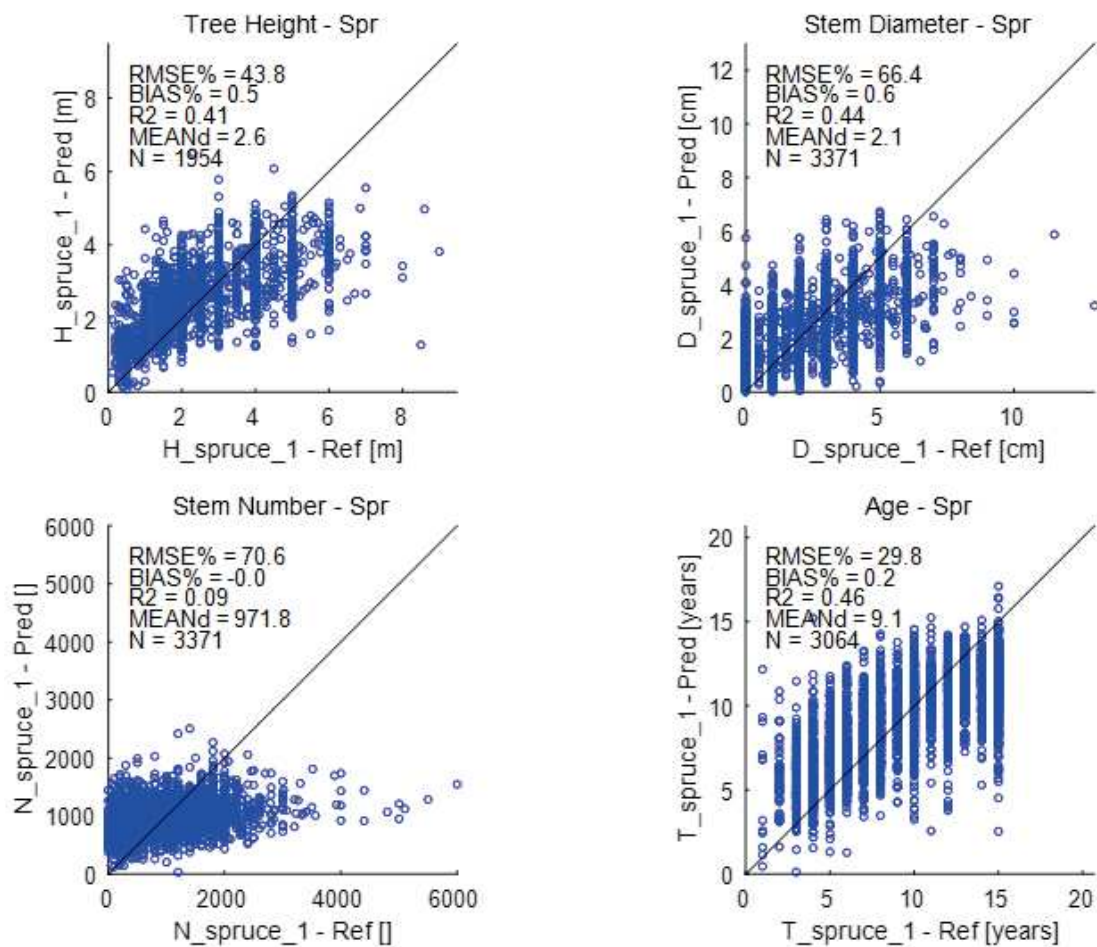
ja kuusi (spruce). Mallit tuotettiin käyttäen vain kehitysluokkaan T2 kuuluvia taimikkokuvioita. Yleisesti aineisto rajattiin myös vain kuvioihin joille mallinnettavan puulajin pituus oli välillä 0 m < H_species < 10 m (species∈[pine, spruce]), paitsi testeissä joissa käytettiin vain yli 3 m taimikoita (H_species > 3 m). Muuttujan ikä mallinnukseen otettiin vain kuviot joille T < 16 v.

Parhaat tulokset molemmille puulajeille saatiin muuttujille ikä (T) ja pituus (H), joiden suhteelliset RMS-virheet olivat 28.9% ja 37.9% ja selitysasteet 0.53 ja 0.52 vastaavasti. Mallien suhteellinen harha oli kaikille muuttujille erittäin pieni (|RMSE%| < 1% kaikilla muuttujilla, ks. Taulukko 12.).

Kuvissa Kuva 8. ja Kuva 9. on esimerkit tyypillisistä testiaineiston hajontakuvioida männyn ja kuusen estimaateille (0 m < H_species < 10 m (species∈[pine, spruce])). Yhteenveto 25 iteraation tuloksista on luvussa 4.3.5.



Kuva 8. Testiaineiston hajontakuviot männyn pituus- (H), keskiläpimitta- (D), runkoluku- (N) ja ikäestimaateille (T). (Ref = inventointitieto, Pred = estimaatti)



Kuva 9. Testiaineiston hajontakuviot kuusen pituus- (H), keskiläpimitta- (D), runkoluku- (N) ja ikäestimaateille (T). (Ref = inventointitieto, Pred = estimaatti)

4.3.3 Puustotietojen estimointi syvällä neuroverkolla

Männyn ja kuusen pituuden mallit toteutettiin myös käyttäen 3 - 7 piilokerrosta sisältävää syvää neuroverkkoa. Saavutetuissa tuloksissa ei pystytty näkemään oleellista eroa yhden piilokerroksen neuroverkkomalleihin verrattuna. Tästä syystä syvien neuroverkkojen käyttö päätettiin lopettaa näiden alustavien testien jälkeen. Tulokset syvien neuroverkkojen käytöstä on esitetty LIITTEESSÄ E.

4.3.4 Puustotietojen estimointi käyttäen aineistoa D19_2 selittävinä muuttujina (ilmakuvien ja CHM-aineiston lisäksi)

Maastoinventointihetkeen kasvatetun puustotulkinta-aineiston (D19_2) merkitystä taimikkotunnusten mallinnuksessa tutkittiin opettamalla männyn ja kuusen pituudelle neuroverkkomallit (yhden piilokerroksen MLP), joissa mallien tulomuuttujina oli ilmake- ja latvusmallipiirteiden lisäksi ko. aineistosta muuttujat pituus (D19_2_H_species), keskiläpimitta (D19_2_D_species), runkoluku (D19_2_N_species) ($species \in [pine, spruce]$). Aineisto rajattiin koskemaan vain taimikoita, joiden pituus $H < 10$ m, mikä raja- oli voimassa yhtäaikaan kuusen ja männyn pituudelle molemmissa aineistoissa (D18 ja D19_2). Mallit opetettiin kahdelle eri pituusluokalle, kohdemuuttujan H_species ($species \in [pine, spruce]$) mukaan.

Estimointi: yhden piilokerroksen MLP-neuroverkkomalli, 25 estimointikierrosta (iteraatiota).

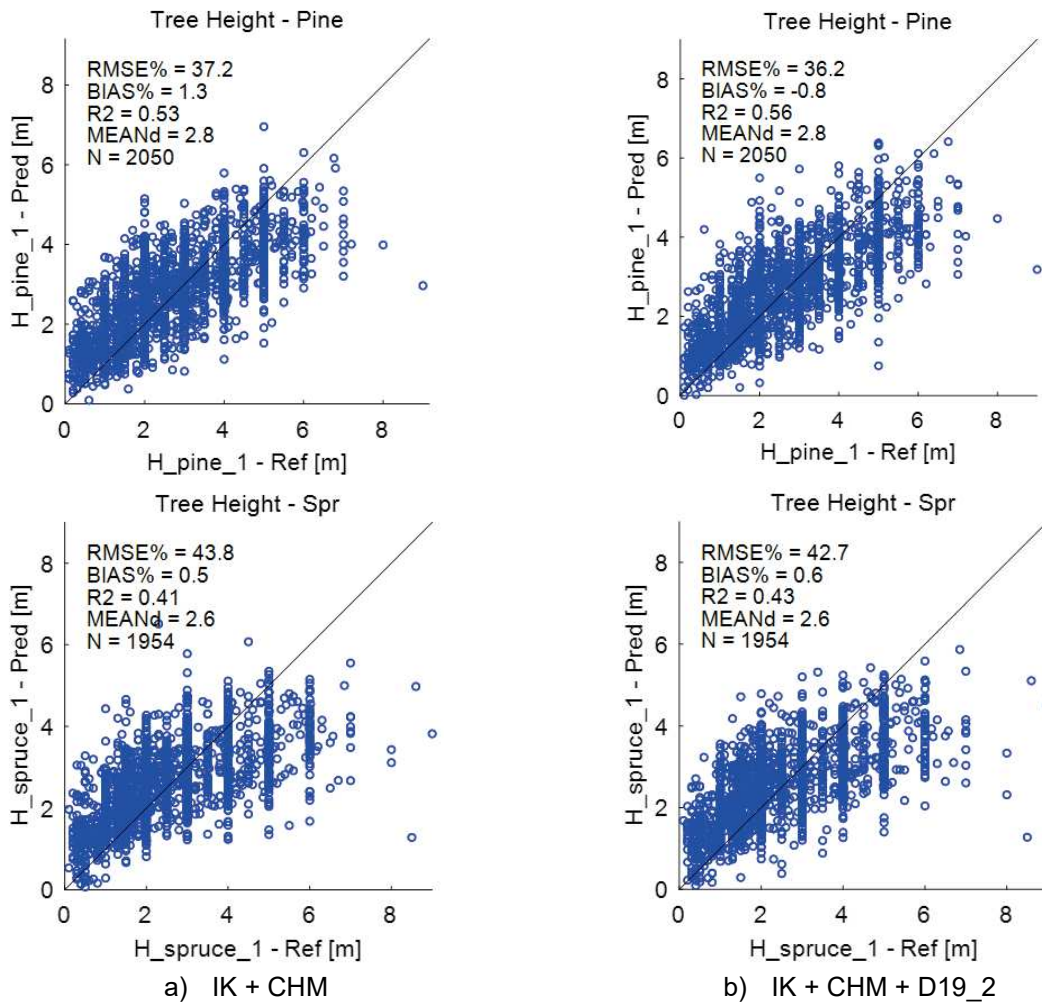
Tulomuuttujat:

- IK + CHM
- IK + CHM + Aineisto D19_2

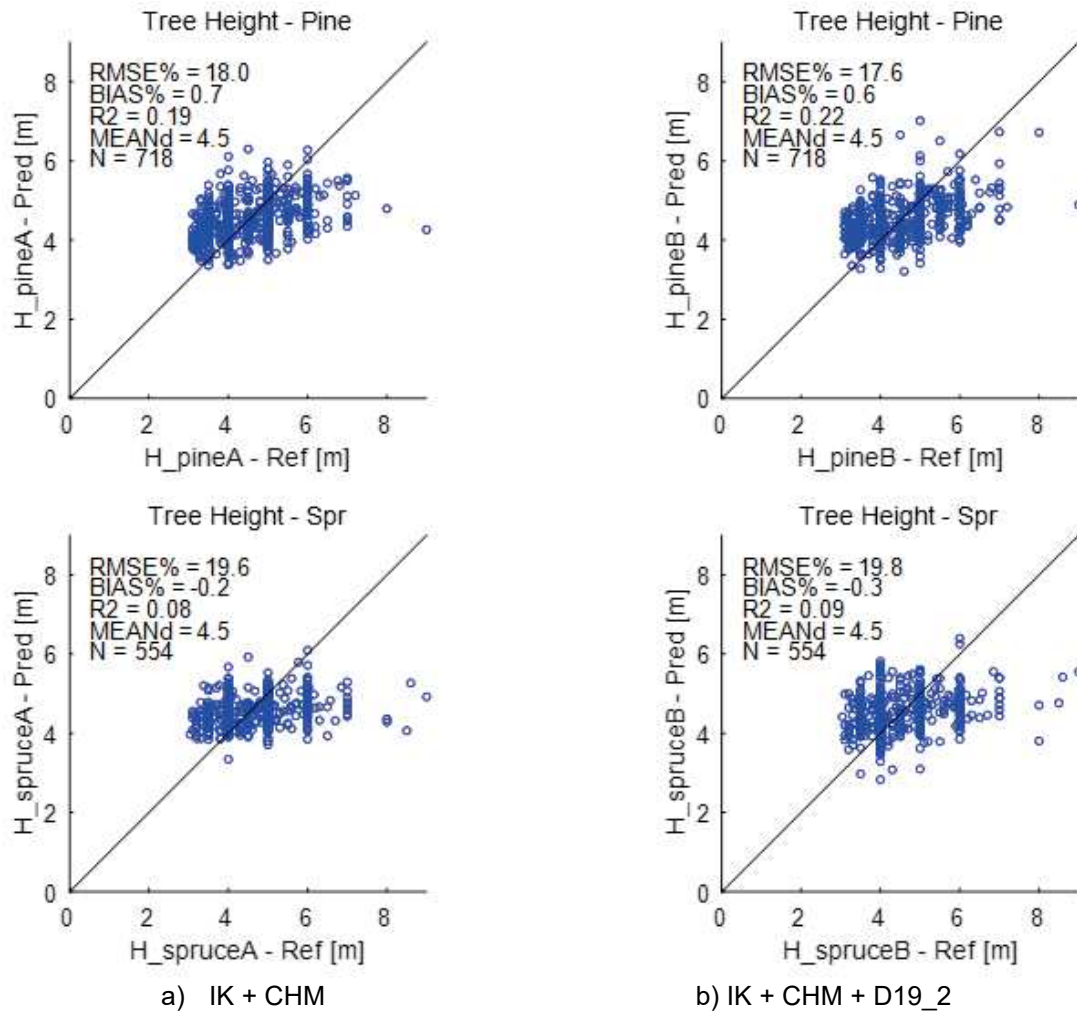
Pituusluokat:

- $0 < H_{\text{species}} < 10 \text{ m}$, $\text{species} \in [\text{pine}, \text{spruce}]$
- $3 < H_{\text{species}} < 10 \text{ m}$, $\text{species} \in [\text{pine}, \text{spruce}]$

Aineiston D19_2 käyttö mallien tulomuuttujina paransi männyn pituuden estimaattien tarkkuutta lievästi verrattuna malleihin, joissa oli mukana vain ilmakuva- ja latvusmallimuuttujat. Kuusella ei vastaavaa parannusta havaittu. Mallien tyypilliset hajontakuviot testiaineistolle on esitetty kuvissa Kuva 10. ja Kuva 11. Yhteenveto 25 iteraation tuloksista on luvussa 4.3.5.



Kuva 10. Estimoitujen puustotunnusten a) H_{pine} ja b) H_{spruce} hajontakuviot testiaineistoa vastaan tulostettuna. Mallin tulomuuttujat: a) IK+CHM = ilmakuva- ja latvusmalliaineisto mallin tulomuuttujina (kuvien 5. ja 6. hajontakuviot toistettu tässä vertailun helpottamiseksi), b) IK+CHM+D19_2 = aineisto D19_2 ilmakuva- ja latvusmalliaineiston piirteiden lisäksi (Ref = inventointitieto, Pred = estimaatti). Pituusluokka 1: $0 < H_{\text{species}} < 10 \text{ m}$.



Kuva 11. Estimoitujen puustotunnusten H_{pine} ja H_{spruce} hajontakuviot testiaineistoa vastaan tulostettuna. a) IK+CHM = ilmakuva- ja latvusmalliaineisto mallin tulomuuttujina, b) IK+CHM+D19_2 = aineisto D19_2 ilmakuva- ja latvusmalliaineiston piirteiden lisäksi (Ref = inventointitieto, Pred = estimaatti). Pituusluokka 2: $3 < H_{\text{species}} < 10$ m.

4.3.5 VTT:n estimaattien vertailu SMK:n puustotietoihin

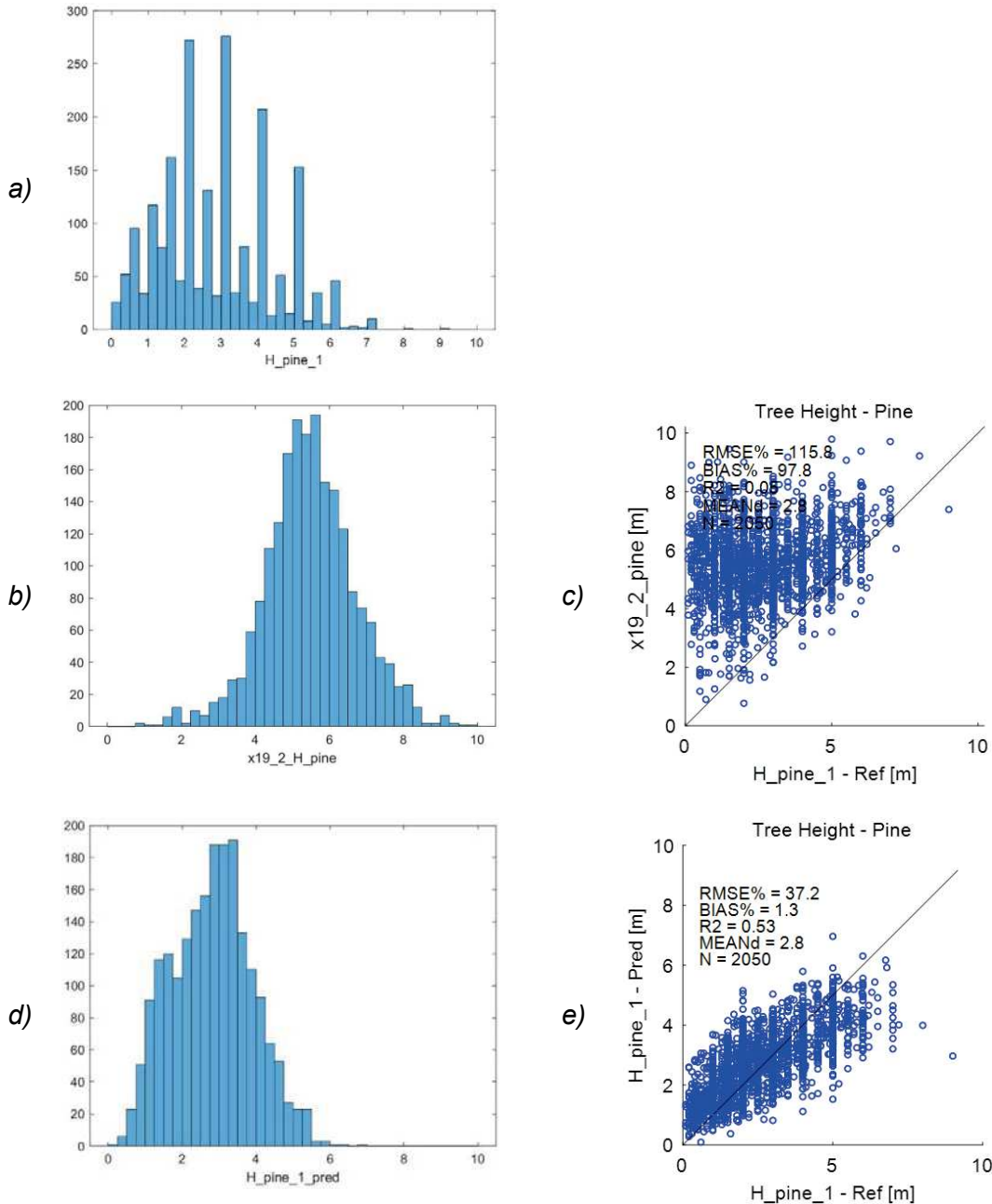
Koska Satakunnan aineistossa oli vain 20 kuviota, joille oli sekä maastotieto, että Kamera-ilmoituksesta saadut puustotunnukset (ks. Taulukko 5.), tehtiin aineistovertilu vain VTT:n estimaattien ja maastoinventointihetken kasvatettujen SMK:n puustotulkintatietojen (D19_2) välillä. Ylläolevien tulostaulukoiden lisäksi tulostettiin männyn ja kuusen pituuden jakaumat ja hajontakuviot testiaineistolle (Kuva 12. ja Kuva 13.). Testijoukon koko: mänty 2050, kuusi 1954 kuviota.

Kuvista on nähtävissä SMK:n puustotunnusten suuri yliarvio nuorilla taimikoilla (0 - 5 m). Suhteellinen RMS-virhe männyn pituudelle oli 116% ja kuusen pituudelle 135%, ja suhteelliset harhat olivat 98% ja 118% vastaavasti. Selitysasteet jäivät myös vaatimattomiksi (mänty: 0.05, kuusi: 0.03). Aineistossa D19_2 havaittu verrattain suuri harha johtunee osittain käytetystä laskentamenetelmästä, jossa tietyn pituuskynnyksen alittavat hilaruudut jätetään kuviokeskiarvon laskennassa huomiotta.

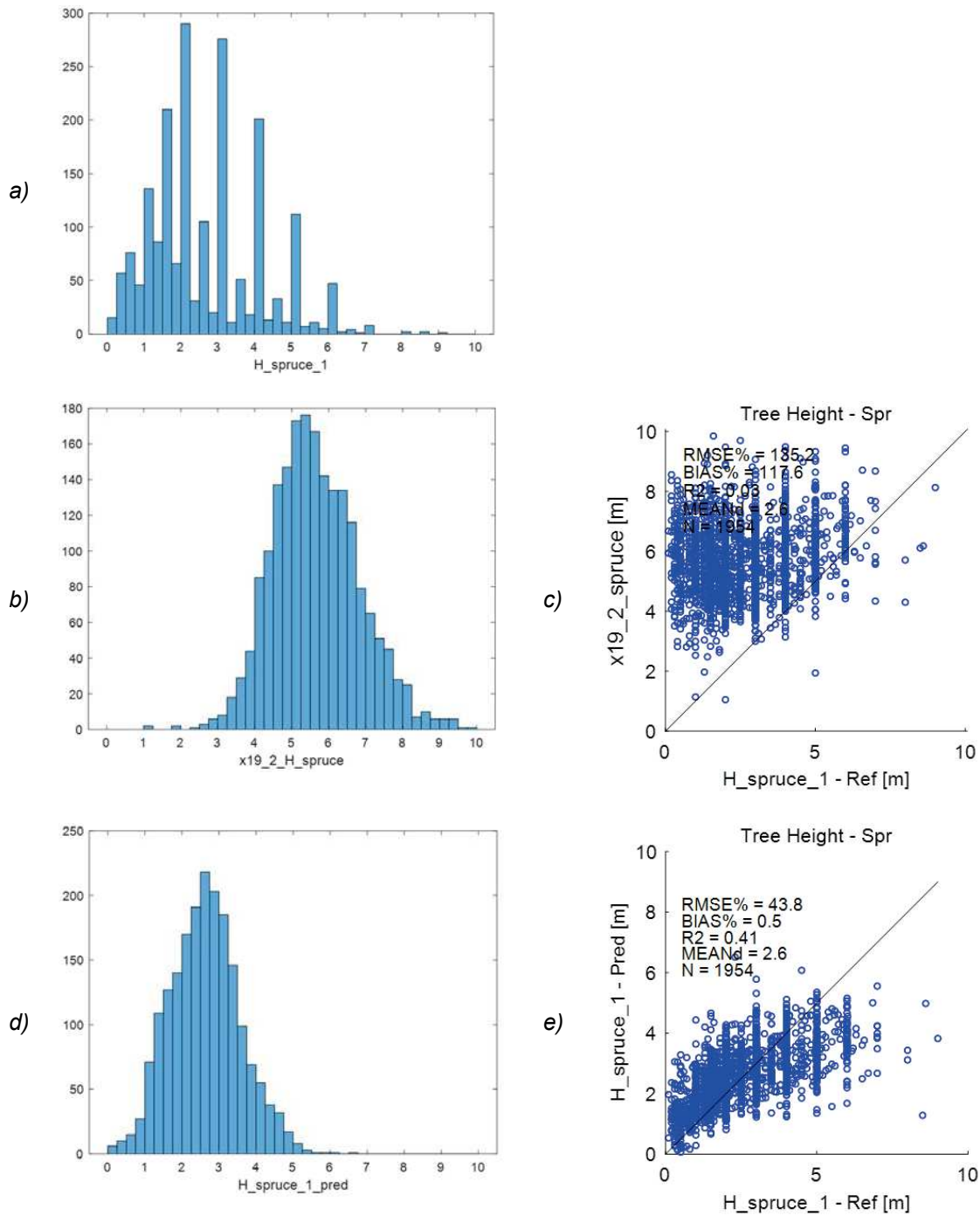
VTT:n estimaateissa harha oli erittäin pieni (sekä mänty että kuusi: 0.6%, .), joskin tyypillistä oli lievä yliarvio nuorissa ja aliarvio varttuneemmissa taimikoissa (Kuva 8., Kuva 9., Kuva 12., Kuva 13.). Estimaattien suhteelliset RMS-virheet ja selitysasteet olivat kohtalaisen hyviä.

Käytetyn testijoukon yksittäisten kuvioiden keskinäisessä vertailussa VTT:n MLP-malli tuotti paremman männyn pituusestimaatin 1777 (86.7%) testijoukon kuviolle kuin aineisto D19_2. Aineisto D19_2 antoi VTT:n estimaattia tarkemman pituuden 273 (13.3%) testikuviolle. Vastaavat luvut kuusen pituusestimaatille olivat MLP-estimaatti parempi: 1695 (86.8%), aineisto D19_2-estimaatti parempi: 259 (13.2%).

Muuttujien D, N ja T suhteen vertailu tuotti samankaltaiset tulokset. Jakaumat ja hajontakuviot näille muuttujille on esitetty LIITE C. Yhteenvedo ko. aineistojen vertailusta on esitetty kappaleessa 4.3.6.



Kuva 12. Eri lähteistä saadut männyn pituuden (H-pine) jakaumat ja hajontakuviot maastoinventointitietoa vastaan: Maastoinventointitieto (a), aineiston D19_2 jakauma (b) ja hajontakuvio (c), MLP-estimaatin jakauma (d) ja hajontakuvio (e). Tulokset testijoukon kuvioille.



Kuva 13. Eri lähteistä saadut kuusen pituuden (H -spruce) jakaumat ja hajontakuviot maastoinventointitietoa vastaan: Maastoinventointitieto (a), aineiston D19_2 jakauma (b) ja hajontakuvio (c), MLP-estimaatin jakauma (d) ja hajontakuvio (e). Tulokset testijoukon kuviolle.

4.3.6 Yhteenveto estimointimalliennuksen tuloksista ja tulosten vertailu

Männyn ja kuusen puustotunnuksille pituus (H), keskiläpimitta (D), runkoluku (N) ja ikä (T) tuotettujen estimointimallien tulokset (suhteellinen RMS-virhe, suhteellinen harha ja selitysaste) testijoukolle laskettuna on koottu taulukoihin Taulukko 11. - Taulukko 13. Taulukoihin on sisällytetty estimointi sekä yhden piilokerroksen (MLP-1), että syvällä (DNN) neuroverkkomallilla (vain pituus). Yhden piilokerroksen neuroverkkomallien (Malli = MLP-1) tulokset on laskettu 25 estimointikierron keskiarvoina ja myös 99% luottamusvälit on laskettu taulukkoon. Pituudelle tuotettiin mallit käyttämällä sekä ilmakehän- ja latvusmallipiirteitä (Mallin tulomuuttujat = IK+CHM), että näiden piirteiden lisäksi myös kuvioille puustotulkinnasta laskettuja ja maastoinventointihetkeen kasvatettuja tunnuksia (Mallin tulomuuttujat = IK+CHM+D19_2), joista tulokset on kirjattu omiin sarakkeisiinsa. Keskiläpimitalle ja puuston iälle tuotettiin mallit vain ilmakehän- ja latvusmallipiirteillä. Yhden piilokerroksen neuroverkoilla tuotettiin myös pituusmallit kahdessa eri pituusluokassa: H-species > 0 ja H-species > 3 m (species ∈ [pine, spruce]). Nämä tulokset on esitetty omilla riveillään.

Syvällä neuroverkolla toteutettujen mallien tulokset ovat yhden estimointikierron tuloksia, eikä luottamusvälejä täten ole laskettu. Syvien neuroverkkomallien tulomuuttujina ei myöskään käytetty muita, kuin ilmakehän- ja latvusmallipiirteitä.

Vertailun vuoksi on puustotulkinnasta maastoinventointihetkeen kasvatetun aineiston (D19_2) virhetunnukset laskettu taulukoihin Taulukko 14. - Taulukko 16. Nämä virhetunnukset on laskettu samalla testiaineistolla (= kuvioilla) kuin VTT:n estimaattien. Puustotunnuksilla D ja T (sekä mänty, että kuusi) oli aineistossa D19_2 ilmeisesti vanhentuneita tietoja maastoinventoiduissa kuvioissa. Tästä syystä virhetunnusten laskennassa näiden muuttujien lukualueet rajattiin $D19_2_D_pine \leq 15$, $D19_2_D_spruce \leq 15$, $D19_2_T_pine \leq 50$ ja $D19_2_T_spruce \leq 40$. Näiden laskennasta pois jätettyjen kuvioiden lukumääräinen osuus koko testiaineistosta olivat 9.8%, 9.8%, 15% ja 16% vastaavasti.

VTT:n estimaattien tulokset kaikilla testatuilla puustotunnuksilla paremmat kuin aineistolla D19_2 (ks. Taulukko 11. - Taulukko 16).

Taulukko 11. Testijoukolle laskettu suhteellinen RMS-virhe (RMSE%) MLP- ja DNN-neuroverkoestimaateille

		Mallin tulomuuttujat		99% luottamusväli				
		IK+CH M	IK+CHM+D 19_2	IK+CHM		IK+CHM+ D19_2		
Puusto- tunnus	Malli			low	high	low	high	Pituusluokka
H-pine	MLP-1	37.9	36.5	37.7	38.1	36.3	36.6	H-pine > 0
H-pine	MLP-1	18.0	17.7	17.9	18.0	17.5	17.8	H-pine > 3
H-pine	DNN	42.1	-	-	-	-	-	H-pine > 0
H-spruce	MLP-1	43.5	43.5	43.3	43.7	43.2	43.9	H-spruce > 0
H-spruce	MLP-1	20.3	19.9	19.9	20.6	19.7	20.2	H-spruce > 3
H-spruce	DNN	49.7	-	-	-	-	-	H-spruce > 0
N-pine	MLP-1	74.6	-	74.3	75.0	-	-	H-pine > 0
N-spruce	MLP-1	71.1	-	70.8	71.3	-	-	H-spruce > 0
D-pine	MLP-1	60.8	-	60.6	61.0	-	-	H-pine > 0
D-spruce	MLP-1	66.5	-	66.4	66.7	-	-	H-spruce > 0
T-pine	MLP-1	28.9	-	28.7	29.0	-	-	H-pine > 0
T-spruce	MLP-1	30.1	-	29.9	30.2	-	-	H-spruce > 0

Taulukko 12. Testijoukolle laskettu suhteellinen harha (BIAS%) MLP- ja DNN-neuroverkkoestimaateille

		Mallin tulomuuttujat		99% luottamusväli				
		IK+CH M	IK+CHM+D 19_2	IK+CHM		IK+CHM+ D19_2		
Puusto- tunnus	Malli			low	high	low	high	Pituusluokka
H-pine	MLP-1	0.6	-0.1	0.2	1.1	-0.5	0.3	H-pine > 0
H-pine	MLP-1	0.1	0.3	-0.1	0.2	0.0	0.5	H-pine > 3
H-pine	DNN	0.4	-	-	-	-	-	H-pine > 0
H-spruce	MLP-1	0.6	0.5	0.2	1.0	0.0	1.0	H-spruce > 0
H-spruce	MLP-1	-0.1	0.6	-0.6	0.5	-0.1	1.3	H-spruce > 3
H-spruce	DNN	-7.2	-	-	-	-	-	H-spruce > 0
N-pine	MLP-1	-0.9	-	-1.42	-0.28	-	-	H-pine > 0
N-spruce	MLP-1	0.7	-	-0.16	1.48	-	-	H-spruce > 0
D-pine	MLP-1	-0.0	-	-0.55	0.53	-	-	H-pine > 0
D-spruce	MLP-1	-0.4	-	-1.43	0.69	-	-	H-spruce > 0
T-pine	MLP-1	0.5	-	0.33	0.67	-	-	H-pine > 0
T-spruce	MLP-1	0.6	-	0.31	0.91	-	-	H-spruce > 0

Taulukko 13. Testijoukolle laskettu selitysaste (R^2) MLP- ja DNN-neuroverkkoestimaateille

Puusto- tunnus	Malli	Mallin tulomuuttujat		99% luottamusväli				Pituusluokka
		IK+CHM	IK+CHM+ D19_2	IK+CHM		IK+CHM+ D19_2		
				low	high	low	high	
H-pine	MLP-1	0.52	0.55	0.51	0.52	0.55	0.56	H-pine > 0
H-pine	MLP-1	0.18	0.21	0.17	0.18	0.19	0.22	H-pine > 3
H-pine	DNN	0.57	-	-	-	-	-	H-pine > 0
H-spruce	MLP-1	0.42	0.42	0.41	0.42	0.41	0.42	H-spruce > 0
H-spruce	MLP-1	0.06	0.08	0.05	0.07	0.07	0.08	H-spruce > 3
H-spruce	DNN	0.48	-	-	-	-	-	H-spruce > 0
N-pine	MLP-1	0.23	-	0.22	0.24	-	-	H-pine > 0
N-spruce	MLP-1	0.09	-	0.08	0.09	-	-	H-spruce > 0
D-pine	MLP-1	0.49	-	0.49	0.50	-	-	H-pine > 0
D-spruce	MLP-1	0.44	-	0.44	0.45	-	-	H-spruce > 0
T-pine	MLP-1	0.53	-	0.52	0.53	-	-	H-pine > 0
T-spruce	MLP-1	0.45	-	0.44	0.45	-	-	H-spruce > 0

Taulukko 14. Testijoukolle laskettu aineiston D19_2 suhteellinen RMS-virhe (RMSE%)

Puustotunnus	RMSE%	Pituusluokka
19_2_H_pine	115.8	H-pine > 0
19_2_N_pine	105.2	H-pine > 0
19_2_D_pine	175.2	H-pine > 0
19_2_T_pine	102.6	H-pine > 0
19_2_H_spr	135.2	H-pine > 0
19_2_N_spr	100.5	H-pine > 0
19_2_D_spr	237.6	H-pine > 0
19_2_T_spr	115.8	H-pine > 0

Taulukko 15. Testijoukolle laskettu aineiston D19_2 suhteellinen harha (BIAS%)

Puustotunnus	BIAS%	Pituusluokka
19_2_H_pine	97.8	H-pine > 0
19_2_N_pine	-57.0	H-pine > 0
19_2_D_pine	148.2	H-pine > 0
19_2_T_pine	79.9	H-pine > 0
19_2_H_spr	117.6	H-spruce > 0
19_2_N_spr	-50.9	H-spruce > 0
19_2_D_spr	207.2	H-spruce > 0
19_2_T_spr	97.3	H-spruce > 0

Taulukko 16. Testijoukolle laskettu aineiston D19_2 selitysaste (R^2)

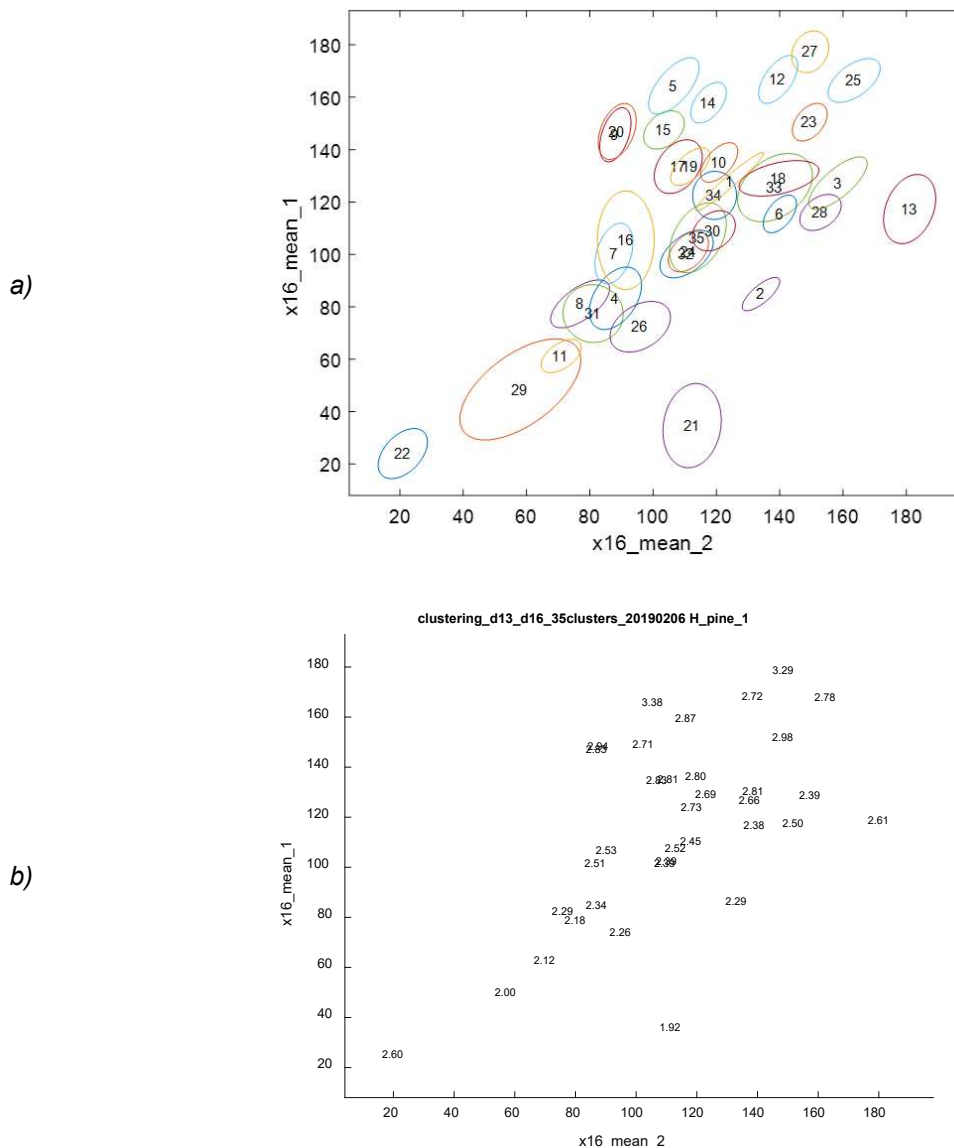
Puustotunnus	R^2	Pituusluokka
19_2_H_pine	0.05	H-pine > 0
19_2_N_pine	0.03	H-pine > 0
19_2_D_pine	0.01	H-pine > 0
19_2_T_pine	0.00	H-pine > 0
19_2_H_spr	0.03	H-spruce > 0
19_2_N_spr	0.01	H-spruce > 0
19_2_D_spr	0.01	H-spruce > 0
19_2_T_spr	0.00	H-spruce > 0

Puustotunnusten estimointi valitulla neuroverkkoon perustuvalla menetelmällä tuotti yllättävän hyviä tuloksia. Parhaat tulokset saatiin pituuden (H) ja iän (T) estimaateilla, ja myös rungon läpimitan (D) tarkkuus oli kohtalainen. Huonoimmat tulokset saatiin runkoluvulle (N). Männyn estimaatit olivat kuusen estimaatteja parempia kautta linjan. Kun mallinnuksessa rajoitettiin aineisto pituuden suhteen välille $3 < H < 10$ m, saatiin parempi keskimääräinen neliövirhe (aineiston suurempi keskipituus), mutta selitysasteet putosivat lähes nollaan.

Puustotulkinnasta maastoinventointihetkeen kasvatetun puustotiedon (D19_2) käyttö mallien tulomuuttujina paransi vain männyn pituuden estimaattia lievästi, kuusella samaa ei havaittu. Syvien neuroverkkojen käyttö mallinnuksessa ei lisännyt estimaattien tarkkuutta.

4.4 Virheellisten kuvioiden löytäminen (laatuindeksi)

Kappaleessa 3.3 esitettyä menetelmää testattiin männyn pituuden MLP-mallilla. Alla esitetyt esimerkkitulokset tuotettiin käyttämällä klusteroinnin tulomuuttujina ilmakehän ja latvusmallipiirteitä (yhteensä 19. tulomuuttujaa). Koko Satakunnan kuvioaineisto klusteroitiin 35 klusteriin ottamalla mukaan kaikki kehitysluokan T2 kuviot, joille oli ilmakehän ja latvusmallitieto (yhteensä 39610 kuviota). Tuloksena saatiin klusterien tilastotiedot (keskipisteet ja kovarianssimatriisit) ja kullekin kuviolle klusterinumero. Kukin klusteri edustaa kuviojoukkoa, jolle on ilmakehän ja latvusmallipiirteissä jotain yhteistä (esim. latvusmallin mediaani samankaltainen → kuvioiden taimikot samaa pituusluokkaa). Kuva 14.a) on klustereille annetut tunnusnumerot (klusterinumero) klusterien keskipisteissä ilmakehän punaisen (x16_mean_2) ja lähi-infra-kanavan (x16_mean_1) suhteen tulostettuna. Kunkin klusterin ympärillä on tulostettuna klusterin edustamien kuvioiden tulomuuttujien hajontaa kuvaava ellipsoidi. Kuva 14.b) on tulostettuna kunkin klusterin edustamien kuvioiden keskimääräinen männyn pituus laskettuna maastointitiedosta (H_pine_1).



Kuva 14. a) Satakunnan aineiston (kuvioiden) klusterinumero tulostettuna ilmakehän punaisen kanavan (x16_mean_2) ja lähi-infra-kanavan intensiteettien suhteen. b) Klustereihin opetusaineistosta allokoitu männyn keskimääräinen pituus (H_pine_1).

Kuvioiden virheen allokointi klustereihin

Kullekin maastossa inventoidulle kuvioille (jolle oli ilmakuva- ja latvusmalliaineisto, sekä validi tieto aineistosta D19_2) laskettiin männylle pituuden mallilla estimaatti ja sen kuviokohtainen virhe (Err_1). Näille kuvioille laskettiin myös suhteellinen virhe (RERR_1) ja sen itseisarvo (RAERR_1):

$Err_1 = \text{maastotieto} - \text{estimaatti}$

$RERR_1 = (\text{maastotieto} - \text{estimaatti}) / \text{maastotieto}$ (suhteellinen virhe)

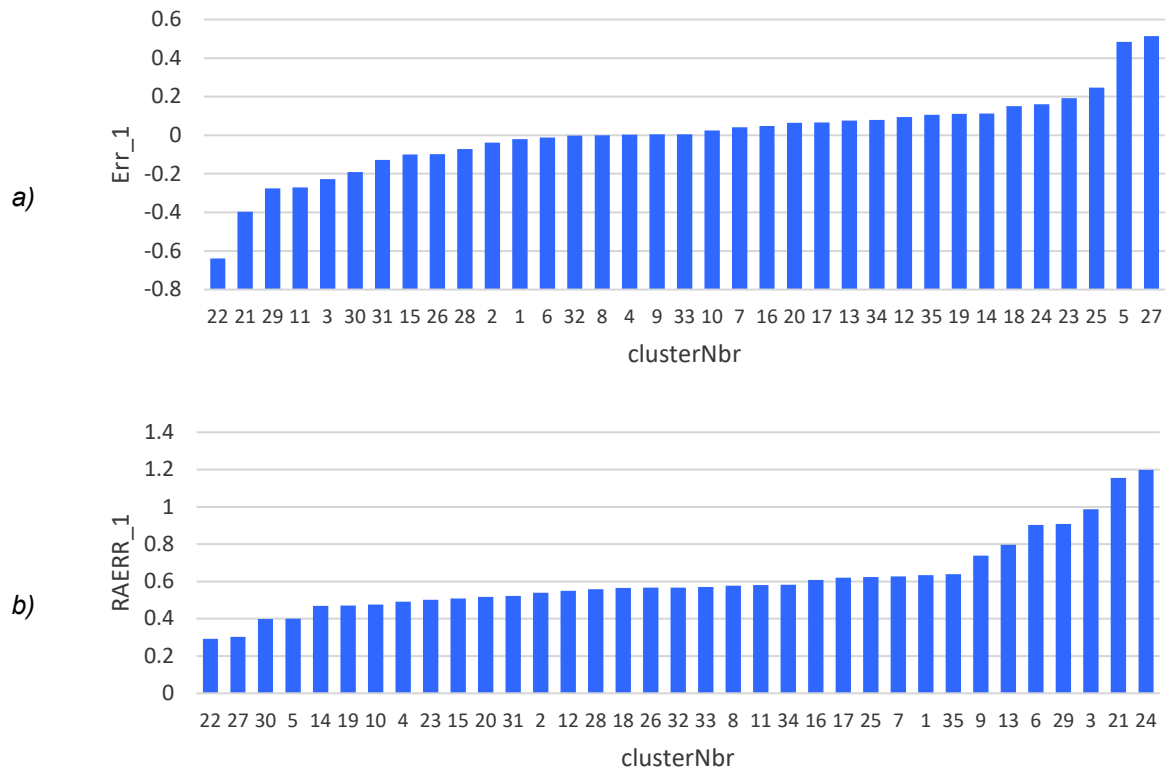
$RAERR_1 = \text{abs}(RERR_1)$, (suhteellisen virheen itseisarvo; $\text{abs}(\cdot) = \text{itseisarvo}$)

Menetelmän testausta varten maastossa inventoitu kuvioaineisto jaettiin opetus- ja testijoukoiksi. Tässä noudatettiin samaa kuviokohtaista jakoa, kuin MLP-estimointimallin opetusvaiheessa:

- laatuindeksin opetusjoukko = MLP estimointimallin opetus- + validointijoukko
- laatuindeksin testijoukko = MLP estimointimallin testijoukko

Näin varmistettiin, että käytetty opetusaineisto ei missään vaiheessa osallistunut estimaattien, tai niistä johdettujen testijoukon virheiden laskentaan.

Saaduille klustereille laskettiin edellä kuvattujen kuviokohtaisten virhetunnusten keskiarvot erikseen opetus- ja testiaineistolla. Kunkin klusterin keskiarvo on laskettu keskimäärin 189:n tai 83:n yksittäisen kuvion (jolla maastotieto) virheen keskiarvona opetus- ja testijoukolla vastaavasti. Tehdyssä testissä testijoukko simuloi kuvioita, joille ei ole maastossa mitattua tietoa. Kuva 15. on klustereihin allokoitu opetusjoukon absoluuttinen virhe (Err_1) ja suhteellisen virheen itseisarvo (RAERR_1) esitetty järjestettynä pylväsdiagrammina.



Kuva 15. Klustereihin allokoitujen männyn pituuden MLP estimaatin virheet opetusjoukolle: a) absoluuttinen virhe (Err_1), b) suhteellisen virheen itseisarvo (RAERR_1). Klusterit järjestettynä kasvavan virheen mukaisesti.

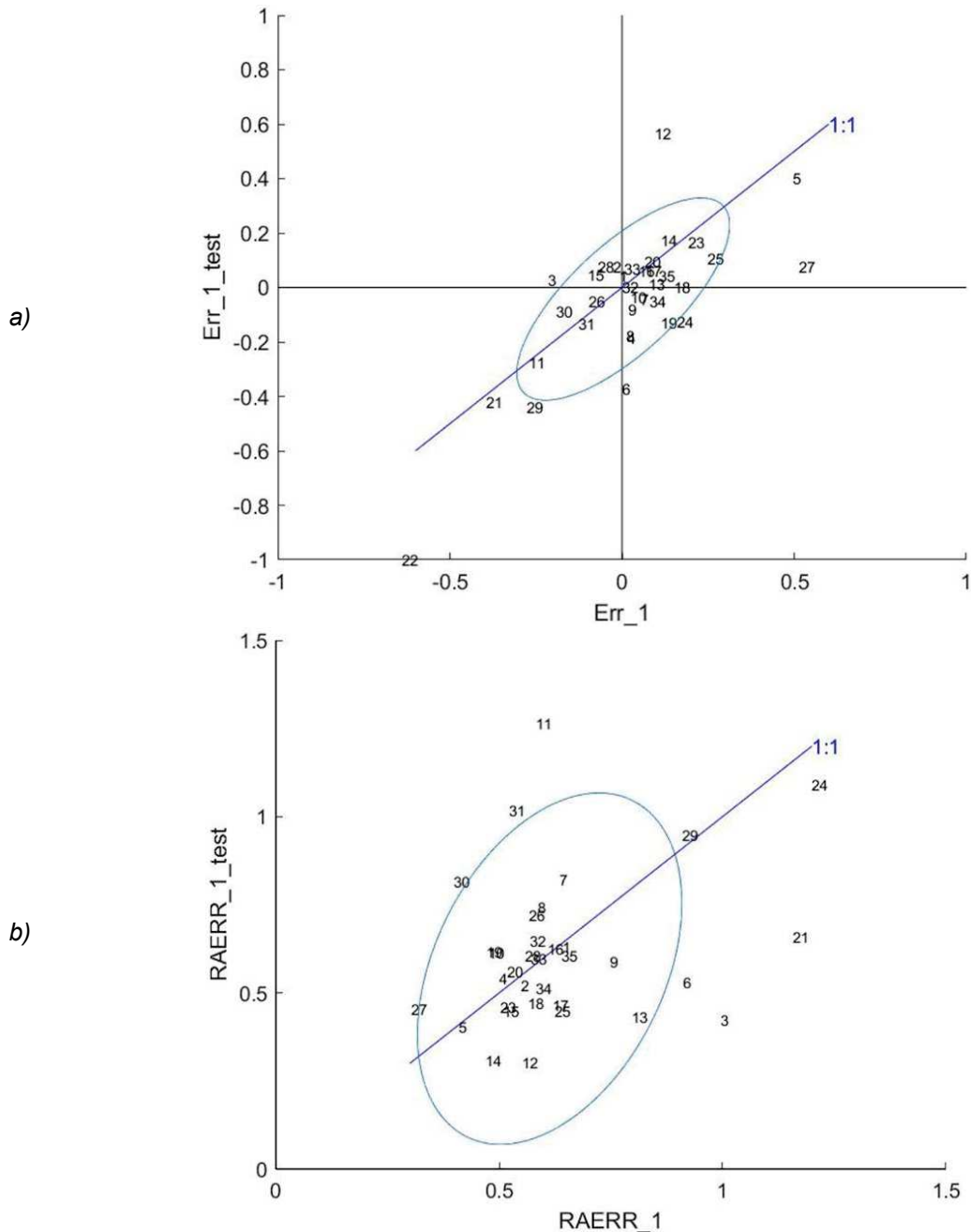
Sekä opetus- että testijoukon virheiden jakautumisessa klustereihin oli odotettua hajontaa. Opetus- ja testijoukon virheitä (Err_1 ja RAERR_1) vertailtiin tulostamalla klusterit (klusterien tunnusnumerot) hajontakuviona opetusjoukon virheen ja testijoukon virheen suhteen (Kuva 16.). Koska klustereille allokoitujen virhetunnukset ovat kyseiseen klusteriin kuuluvien kuvioiden virheiden keskiarvoja, on etumerkillisten virhetunnusten (Err_1, RERR_1) tulkinnessa huomattava, että positiivisten ja negatiivisten virheiden summautuminen nollaan (i.e. keskiarvo = 0) ei tarkoita, että kyseinen klusteri edustaisi virheettömiä kuvioita (positiiviset ja negatiiviset virheet osittain kumoavat toisensa). Tästä syystä on tarpeen tarkastella myös kuvioiden virheen itseisarvoa kuvaavia tunnuksia (RAERR_1).

Kuva 16.a) klusterit, joilla on suurin etäisyys virheakselien origoon, edustavat harhaisia kuvioita (positiivinen virhe: pituuden aliarvio, negatiivinen virhe: pituuden yliarvio). Ideaalisessa tapauksessa klusterit asettuisivat kuvassa 1:1 -suoralle, jolloin kuvioryhmille (klustereille) allokoitu opetusjoukon virhe ennustaisi täydellisesti testijoukon kuvioiden keskimääräistä virhettä. Käytännössä virhe-ennusteessa on kuvan mukaista hajontaa.

Tässä esimerkissä rajattiin suurimman virheen (joko opetus- tai testijoukon) sisältävät klusterit käyttämällä klusterien hajontaa ko. virheakselien suhteen. Kuvaan piirretty ellipsi on virheiden keskinäisen jakauman hajontaan suhteutettu kynnys (1.5 x kovarianssi). Ellipsin ulkopuolelle rajattujen klusterien oletettiin edustavan kuvioita, joille on tarpeen tehdä maastotarkastus. Näitä klustereita oli yhdeksän ja niiden tunnusnumerot olivat: 3, 5, 6, 12, 21, (22), 24, 27 ja 29. Näistä klusterin no: 22 kuvioita ei sisältynyt testijoukkoon yhtään ja opetusjoukossakin vain 4 klusteriin 22 kuuluvaa kuviota, joten tämä klusteri oli erikoistapaus.

Etumerkillisen virhetunnuksen Err_1 lisäksi vertailtiin klusterien jakautumista suhteellisen virheen itseisarvon (RAERR_1) suhteen. Vertailussa nähtiin (Kuva 16.b)), että klusterit 5, 12

Klusteroinnin tuloksena kuviolle saatiin klusterinumero, jonka avulla esimerkin vertailussa löydettyihin kuvioryhmiin (klustereihin) kuuluvat kuviot voidaan kohdentaa. Esimerkin yhdeksään klusteriin liittyvien kuvioden lukumäärät on esitetty Taulukko 17. Yhteensä ko. klustereissa oli 2148 kuviota, joista huomattavan suuri osuus 1854 (85%) oli maastossa inventoituja. Esimerkissä valitulla kynnysarvolla suurin maastotarkastustarve oli 324 kuviolla.



Kuva 16.a) Klusterien tunnusnumerot tulostettuna opetus- ja testijoukon absoluuttisen virheen suhteen (Err_1/Err_1_test). b) Klusterien tunnusnumerot tulostettuna opetus- ja testijoukon suhteellisen virheen suhteen ($RAERR_1/RAERR_1_test$). Kuvien ellipsi = 1.5 x virhetunnusten hajonta (kovarianssi).

Klusteroinnin tuloksena kuviolle saatiin klusterinumero, jonka avulla esimerkin vertailussa löydettyihin kuvioryhmiin (klustereihin) kuuluvat kuviot voidaan kohdentaa. Esimerkin yhdeksään klusteriin liittyvien kuvioden lukumäärät on esitetty Taulukko 17. Yhteensä ko. klustereissa oli 2148 kuviota, joista huomattavan suuri osuus 1854 (85%) oli maastossa inventoituja. Esimerkissä valitulla kynnysarvolla suurin maastotarkastustarve oli 324 kuviolla.

Tietyn virhekynnyksen ylittävien kuvioryhmiä rajaamisen sijaan klusterit voidaan laittaa myös valitun virhetunnuksen/tunnusten mukaiseen ohjeelliseen priorisointijärjestykseen (tästä esimerkki Kuva 17.). Käytettävä virhetunnus voidaan myös laskea useamman puustotiedon pohjalta.

Edellä esitetyt laatuindeksiin liittyvät tulokset ovat alustavia, ja tuotettu ensimmäisten kokeilujen perusteella. Valitulla testiaineistolla saavutetut tulokset olivat lupaavia: menetelmällä voitiin tunnistaa testiaineistosta suurimman keskimääräisen virheen sisältävät kuvioryhmiä. Menetelmä vaatii vielä lisää kehitystyötä ja laajempia testejä maastossa inventoiduilla kuvioilla.

Taulukko 17. Kuvioiden lukumäärä esimerkin virhekynnyksen ylittävissä klustereissa

Klusteri#	Kuvioita/klusteri (tot)	Maastoinventoidut kuviot	Muut kuviot
3	166	121	45
6	54	39	15
11	360	322	38
21	442	380	62
22	9	4	5
24	120	96	24
29	404	354	50
30	218	176	42
31	405	362	43
Yhteensä (kpl):	2178	1854	324



Kuva 17. Esimerkki MLP-estimaatin virheestä (RAERR_1) johdettuna satakunnan taimikkokuvioille. Kuvan koko: 1500 x 1500 m. Kuvassa virhe on jaoteltu seitsemään luokkaan (pieni virhe = vihreä, suuri virhe = punainen).

5 Johtopäätökset ja yhteenveto

5.1 Johtopäätökset

Alustava analyysi

Aineiston alustava analyysi paljasti, että Satakunnan alueen taimikoilla oli varsin vähän Kemera-ilmoituksia, taimikon perustamisilmoituksia ja toimenpideilmoituksia (Taulukko 3.), joten käsiteltävä aineisto oli siis hyvin harvaa. Tämä osaltaan johti hankkeen tavoitteiden tarkennukseen ja aineistolähteiden karsimiseen. Korrelaatiotaulukoista ja aineistoille lasketuista jakaumista voitiin päätellä, että SMK:n toimittaman, maastossa inventoiduille kuvioille puustotulkinnasta maastoinventointihetkeen kasvatetun puustotiedon (D19_2) pituus, keskiläpimitta ja ikä oli yliarvioitu. Kyseinen harha johtunee osittain SMK:n tulkinnan laskentamenetelmästä, jossa hila-aineistosta johdetaan kuviokohtainen puustotieto sääntöpohjaisilla kaavoilla.

Puustotunnusten estimointi

Testit Satakunnan alueen aineistolla osoittivat, että ilmakuvien ja latvusmallin avulla oli mahdollista saada SMK:n nykytietoa (D19_2) tarkempi taimikon pituustieto 87%:lle kuvioista. Myös läpimitan, runkoluvun ja iän suhteen tilanne on samankaltainen (LIITE C). Tässä hankkeessa lasketut mallit perustuivat kuvioille keskiarvoina laskettuihin ilmake- ja latvusmallipiirteisiin. Estimointimenetelmänä käytettiin täysin kytkettyä yhden piilokerroksen neuroverkkoa. Taimikkotietojen estimointi alkuperäisen ilmake- ja laseraineiston avulla voisi nostaa estimaattien tarkkuutta edelleen. Tämä edellyttää, että myös mallinnuksessa käytettävä referenssiaineisto on tarkkaa ja sen resoluutio vastaa kaukokartoitusaineistoa.

Aineiston D19_2 puustotunnusten harhasta huolimatta niiden käyttö mallinnuksen tulomuuttujina paransi männyn pituuden estimaattia, joskin marginaalisesti (Taulukko 11., Taulukko 13). Kuusen pituuden arviointi ei tarkentunut. Syvien neuroverkkojen käyttö ei parantanut tehdyissä testeissä taimikoiden puustotietojen estimointitarkkuutta.

Virheellisten kuvioiden löytäminen

Hankkeessa kehitettiin menetelmä virheellisten kuvioiden löytämiseksi. Menetelmä perustui koko kuvioaineiston klusterointiin, jolla pyrittiin löytämään toistensa kaltaiset kuvioryhvät (klusterit). Menetelmällä voitiin tuottaa kuvioille klusterikohtainen laatuindeksi, tai virhetunnus, joka kuvaa kyseisen kuvioryhmän keskimääräistä arviointivirhettä. Satakunnan taimikkokuvioille (kehitysluokka T2) laskettiin menetelmän avulla virhetunnukset, jotka perustuivat VTT:n neuroverkkomallilla tuotetun männyn pituuteen. Virhetunnuksen toimivuutta testattiin jakamalla aineisto opetus- ja testijoukkoihin. Esimerkkitestissä pystyttiin tunnistamaan testiaineiston kuviot, joilla oli keskimäärin suuri virhe. Asetetun virhekyynnyksen ylitti yhteensä 2178 kuviota, joista 324 kuviolle ei ollut tehty maastoinventointia.

Kehitettyyn menetelmään liittyen tehtiin seuraavat johtopäätökset:

- Kuvattu menetelmä on yleiskäyttöinen, sillä voidaan tuottaa hyvyysindeksi mille tahansa muuttujalle (sekä jatkuvat, että kategoriatyypiset muuttujat). Käytännön toimivuuden edellytyksenä on, että kohdemuuttujan (puustotunnuksen) virhe kohdentuu klustereihin epätasaisesti. Tehdyissä alustavissa testeissä näin myös tapahtui.
- Tehdyissä testeissä käytettiin klusteroinnin tulomuuttujina pääasiassa ilmake- ja latvusmallipiirteitä. Myös kategoriatyypisten piirteiden (kehitysluokka, kasvillisuustyyppi jne.) käyttö klusteroinnissa on mahdollista. Näiden tuoman lisäarvon tutkiminen jää mahdollisiin jatkohankkeisiin.

- Toteutetussa esimerkkianalyysissä keskityttiin vain yhteen puustotunnukseen. Menetelmän toimivuuden varmistaminen vaatii lisätestejä ja eri puustotunnusten vertailua. Eri puustotunnuksille saaduista virheistä voidaan laskea yhdistetty laatuindeksi (esim. painotettu keskiarvo).

5.2 Yhteenvedo ja ehdotus jatkotoimenpiteistä

Hankkeen tavoitteena oli tuottaa automaattinen menetelmä taimikkotiedon parantamista varten. Menetelmän avulla virheellistä taimikkotietoa voidaan korjata ja paikantaa taimikot, joihin maastotarkastukset tulee kohdentaa. Analyysiaineiston esikäsittelyvaihe vei ennakoitua huomattavasti suuremman osan hankkeen resursseista, minkä seurauksena alkuperäisiä tavoitteita tarkennettiin hankkeen kuluessa. Yhdessä SMK:n kanssa sovittiin, että menetelmäkehityksessä ei käytetty kaikkia alun perin suunniteltuja aineistoja ja että käytetään vain Satakunnan testialuetta. Toisaalta hankkeen tehtäviin lisättiin taimikon pituuden estimointi SMK:n ilmakeu- ja latvuston korkeusmalliaineiston avulla.

Hankkeessa analysoitiin ensin SMK:n eri lähteistä koottua Satakunnan alueen kuviotietoa aineiston rakenteen, korrelaatioiden ja puustotietojen jakaumien selvittämiseksi. Yksi tämän vaiheen tuloksia oli, että SMK:n toimittamissa taimikkotiedoissa oli huomattavan paljon harhaa. Alustavan analyysin tulosten perusteella estimointiin taimikoiden puustotiedot ilmakeu- ja latvusmallipiirteiden ja maastossa inventoitujen taimikkokuvioiden avulla. VTT:n neuroverkkoihin perustuvan estimoinnin tuloksia vertailtiin SMK:n toimittamiin puustotietoihin. Kaikkien testattujen puustotietojen suhteellinen keskineliövirhe (RMSE%), suhteellinen harha (BIAS%) ja selitysaste (R^2) olivat VTT:n estimaateilla paremmat kuin SMK:n toimittamilla puustotunnuksilla.

Hankkeessa kehitettiin menetelmä niiden kuvioiden löytämiseksi, joissa on odotettavissa suurin estimointivirhe (taimikkotiedon laatuindeksi). Menetelmä perustuu ohjatun ja ohjaamattoman oppimisen menetelmien yhdistämiseen. Menetelmää testattiin alustavasti Satakunnan taimikkokuviolla ja tulokset olivat lupaavia. Menetelmän käyttökelpoisuuden arviointi vaatii jatkokehittelyä ja lisätestausta yhteistyössä SMK:n kanssa.

Ehdotus jatkotoimenpiteistä

- 1) Ilmakeuihin ja latvusmalliaineistoon perustuvan puustotunnusten tulkintamenetelmän jatkokehitys

Tässä hankkeessa mallinnettiin puustotunnuksia lähtien kuviotasolle lasketuista ilmakeu- ja latvusmallipiirteistä. Koska kuvausaineisto on resoluutioltaan huomattavasti kuviokokoa hienojakoisempaa, on mahdollista, että taimikoista saadaan tarkempia estimaatteja, kuin tässä hankkeessa lasketut. Menetelmäkehitys tulisi jakaa esioperatiiviseen ja operatiiviseen vaiheeseen. Esioperatiivisessa vaiheessa selvitetään, mikä lisäarvo kehityksestä on mahdollista saada, sekä käytettävät menetelmät, tietoratkaisut ja niiden tarvitsemat prosessointi- ja tiedonhallintaresurssit. Yksi tämän vaiheen tehtävistä on myös selvittää, mitkä ovat vaatimukset maastossa tehtävän inventointityön määrälle ja laadulle. Operatiivisessa vaiheessa menetelmä tulee toteuttaa SMK:n laskentajärjestelmille asettamien vaatimusten mukaisesti.

- 2) Automaattisen taimikkotiedon tuotantomenetelmän jatkokehitys

Jatketaan tässä hankkeessa tehtyä työtä laajentamalla aineistopohjaa usean maakunnan alueelle ja käyttämällä aineistolähteitä D2_10, D3, D4_11 ja mahdollisesti myös D6_8 taimikkotiedon parantamisessa.

- Laajennetaan analyysin aineistoa ottamalla mukaan Satakunnan lisäksi 3 - 5 maakuntaa Etelä-Suomesta. Tavoitteena on varmistaa Kemera-ilmoitusten, taimikon perustamisilmoitusten ja toimenpidekuvioiden edustus analyysissä. Jatkokehityksessä hyödynnetään tämän hankkeen tuloksina tuotettuja esikäsittelyohjelmia.
- Analysoidaan Kemera-ilmoitusten, taimikon perustamisilmoitusten, toimenpidekuvioiden ja metsänkäyttöilmoitusten merkitys taimikkotiedon parantamisessa.
- Suunnitellaan edellisessä kohdassa mainittujen aineistojen tulkintaketju niiden hyödyntämiseksi taimikkotiedon parantamisessa. Tässä lähtöoletuksena on, että tarvittava tulkintasysteemi on yhdistelmä perinteisiä sääntöpohjaista tulkintaa (asiantuntijajärjestelmä) ja koneoppimismenetelmiä.
- Yhdistetään tulkintaketjun tulos/tulokset ilmake- ja latvusmallipohjaiseen estimaattiin niin, että kuviolle saadaan todennäköisimmin oikea puustotieto.
- Tehdään suunnitelma operatiivisen menetelmän toteuttamisesta.

3) Taimikkotiedon laatuindeksin jatkokehitys

Kehitetään edellä kuvattu laatuindeksi taimikoiden maastotarkastusten priorisointiin.

- Arvioidaan menetelmän käyttökelpoisuus maastossa inventoitua kuvioaineistoa käyttäen.
- Kehitetään menetelmä esioperatiiviseen vaiheeseen.
- Tehdään suunnitelma operatiivisen menetelmän toteuttamisesta.

4) Taimikoiden puustotunnusten estimointi Suomen alueelle VTT:n tässä hankkeessa kehittämällä menetelmillä.

VTT:n tässä hankkeessa estimoidut puustotunnukset osoittautuivat SMK:n toimittamia puustotietoja paremmiksi. Jatkohankkeessa estimoidaan koko Suomen taimikkokuviot käyttäen tässä hankkeessa toteutettuja neuroverkkoihin perustuvia malleja.

- Kuvioille laskettuja ilmake- ja latvusmallipiirteitä käytetään mallien selittävinä muuttujina.
- Estimoidaan puulajikohtaiset puustotunnukset pituus (H), runkoluku (N), keskiläpimitta (D) ja ikä (T) puulajeina mänty, kuusi ja lehtipuut.
- Estimoinnin alueellinen yksikkö on kuvio. Käytettävät kuviorajaukset SMK:n avoimen rajapinnan kuvioaineistosta.
- Kokonaispuustotunnusten estimoinnista sovitaan erikseen.

Lähdeviitteet

[1] Taimikkotiedon tuottamisen automatisointi, tavoitteen ja aineistojen kuvaus 20180427.pdf, SMK, 27.4.2018

[2] Taimikkotiedon tuottamisen automatisointi.pdf (kalvosarja), SMK, 20.4.2018

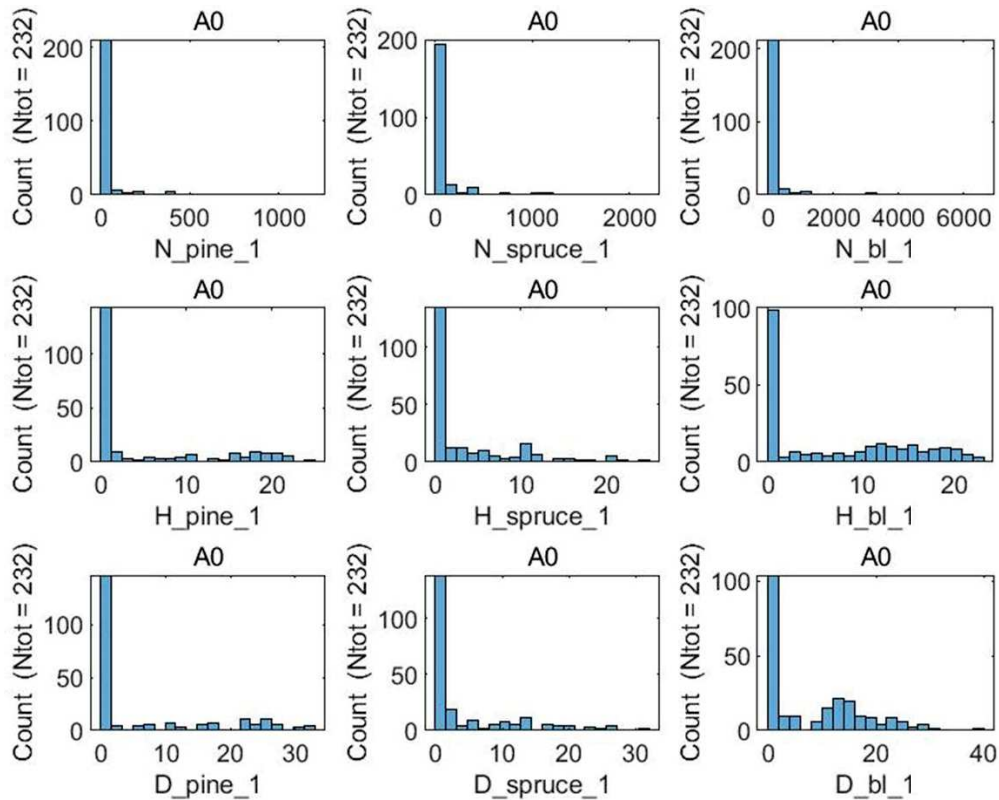
[3] Projektisuunnitelma, Taimikkotiedon Tuottamisen Automatisointi, VTT, 15.5.2018

Liitteet /

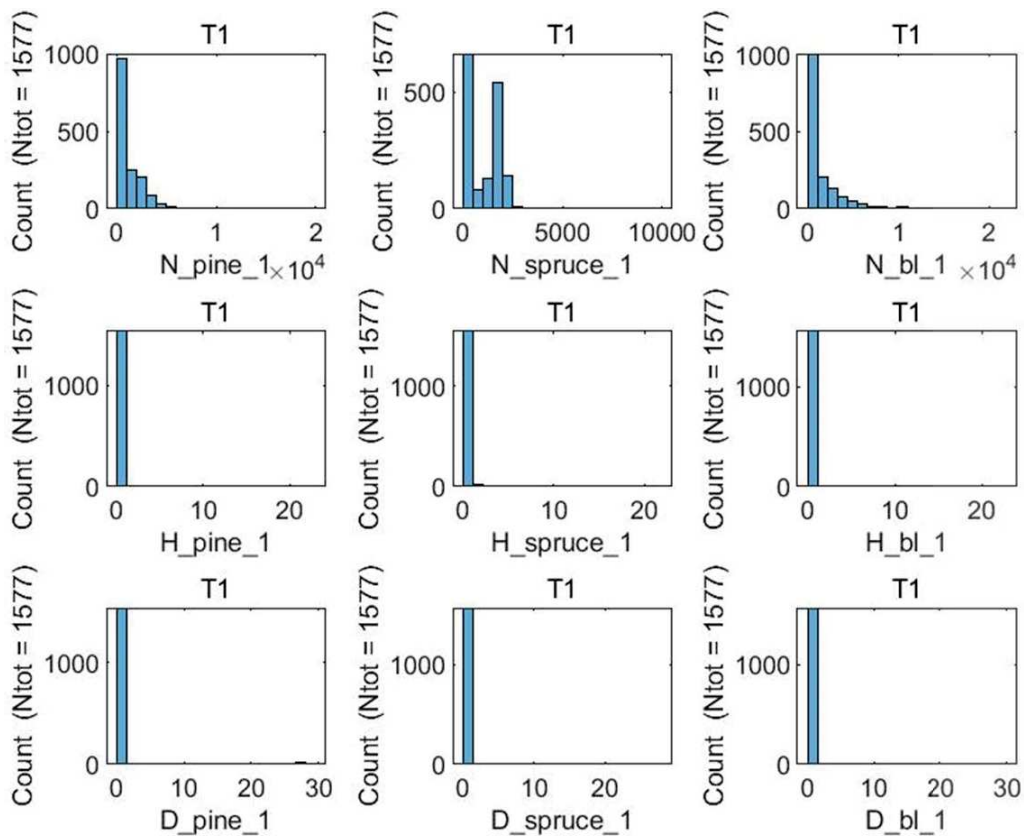
LIITE A	AINEISTON JAKAUMAT
LIITE B	KORRELAATIOTAULUKOT
LIITE C	AINEISTON D19_2 VERTAILU MAASTOSSA INVENTOITUUN TIETOOON
LIITE D	PIIRREVALINTATESTIN TULOKSET
LIITE E	TESTIT SYVILLÄ NEUROVERKOILLA
LIITE F	KOKONAISPUUSTOTUNNUSTEN LASKENTA JA ESTIMOINTI
LIITE G	KÄYTETYT TERMIT, LYHENTEET JA MERKINNÄT

LIITE A AINEISTON JAKAUMAT

MAASTOSSA MITATUT KUVIOT (D18)

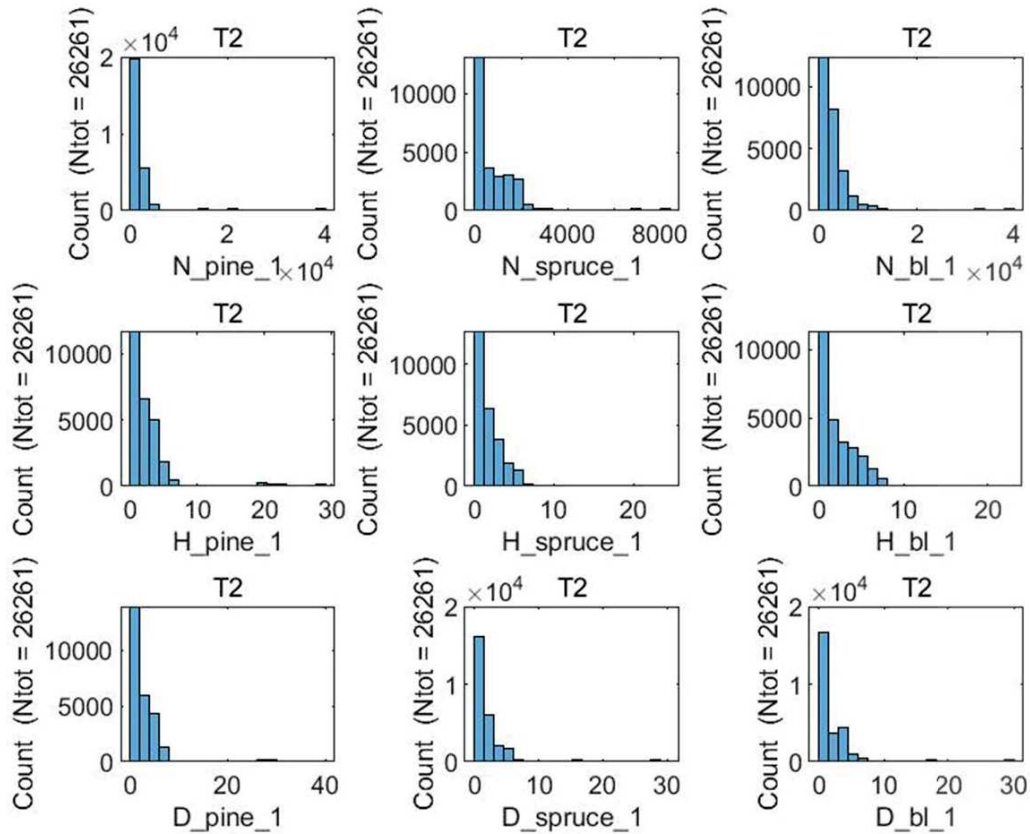


Kuva 18. Runkoluvun (N), pituuden (H) ja keskiläpimitan (D) jakaumat maastossa inventoiduille Satakunnan kuvioille. Kehitysluokka A0.



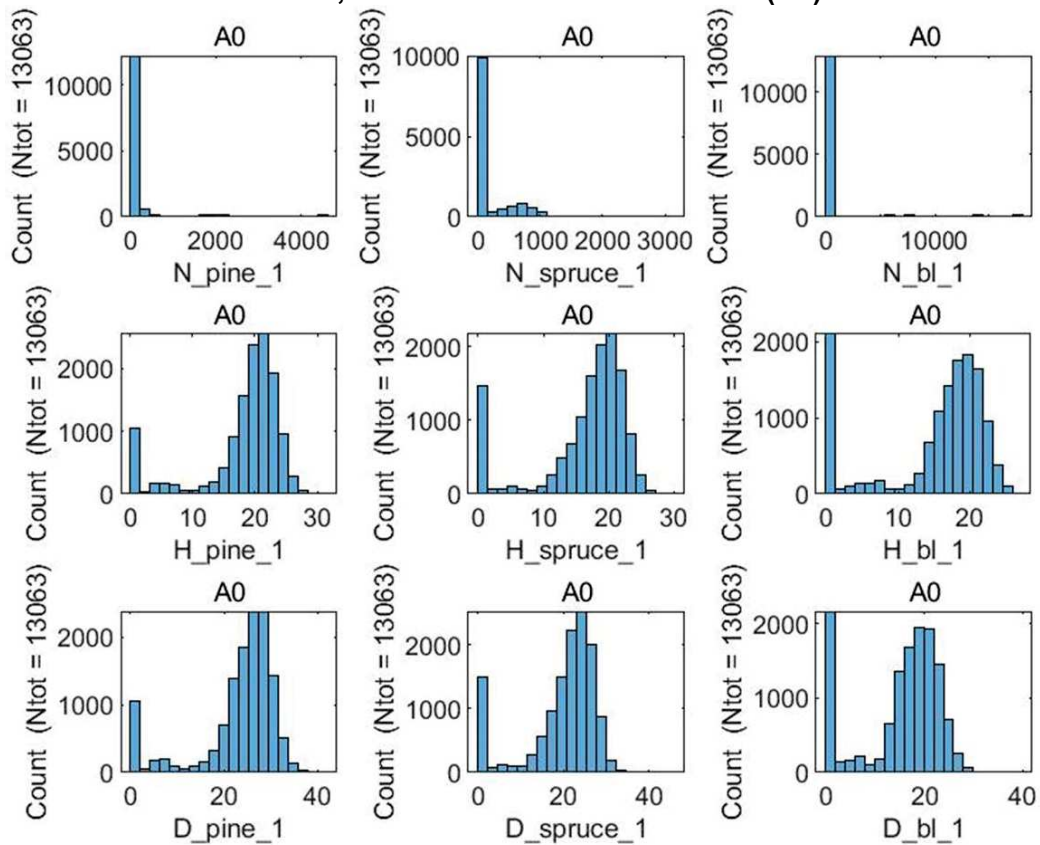
Kuva 19. Runkoluvun (N), pituuden (H) ja keskiläpimitan (D) jakaumat maastossa inventoiduille Satakunnan kuvioille. Kehitysluokka T1.

MAASTOSSA MITATUT KUVIOT (D18)

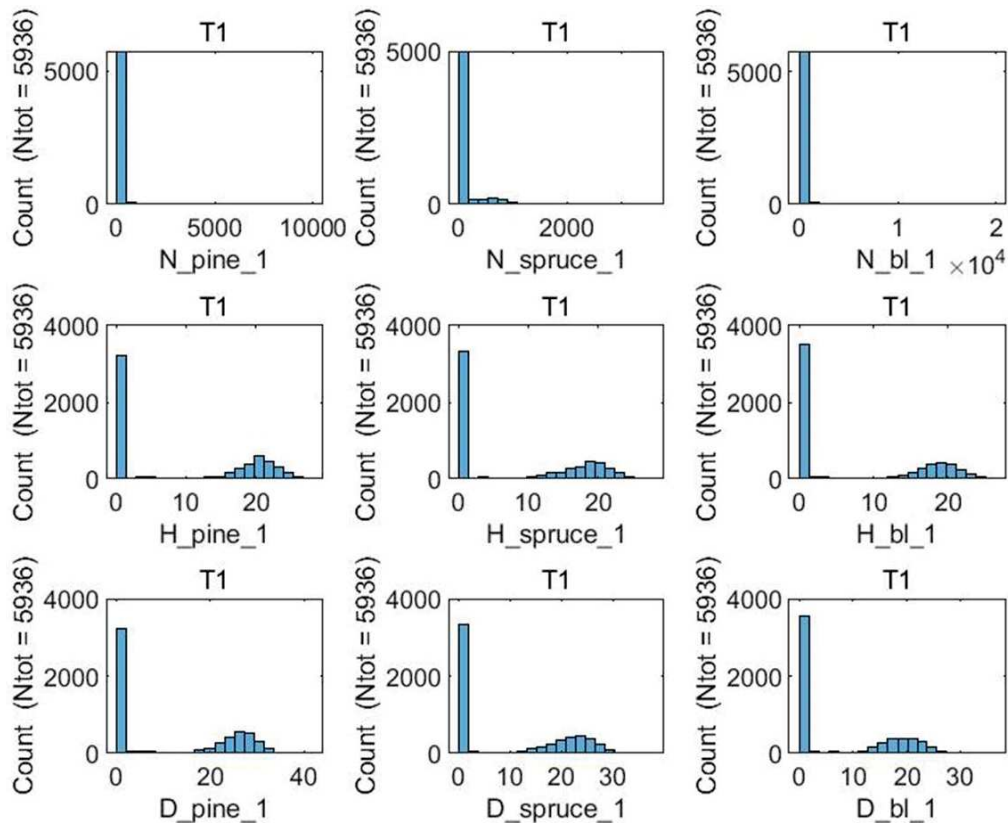


Kuva 20. Runkoluvun (N), pituuden (H) ja keskiläpimitan (D) jakaumat maastossa inventoiduille Satakunnan kuviolle. Kehitysluokka T2.

KUVIOT, JOILLE EI MAASTOKÄYNTIÄ (D1)

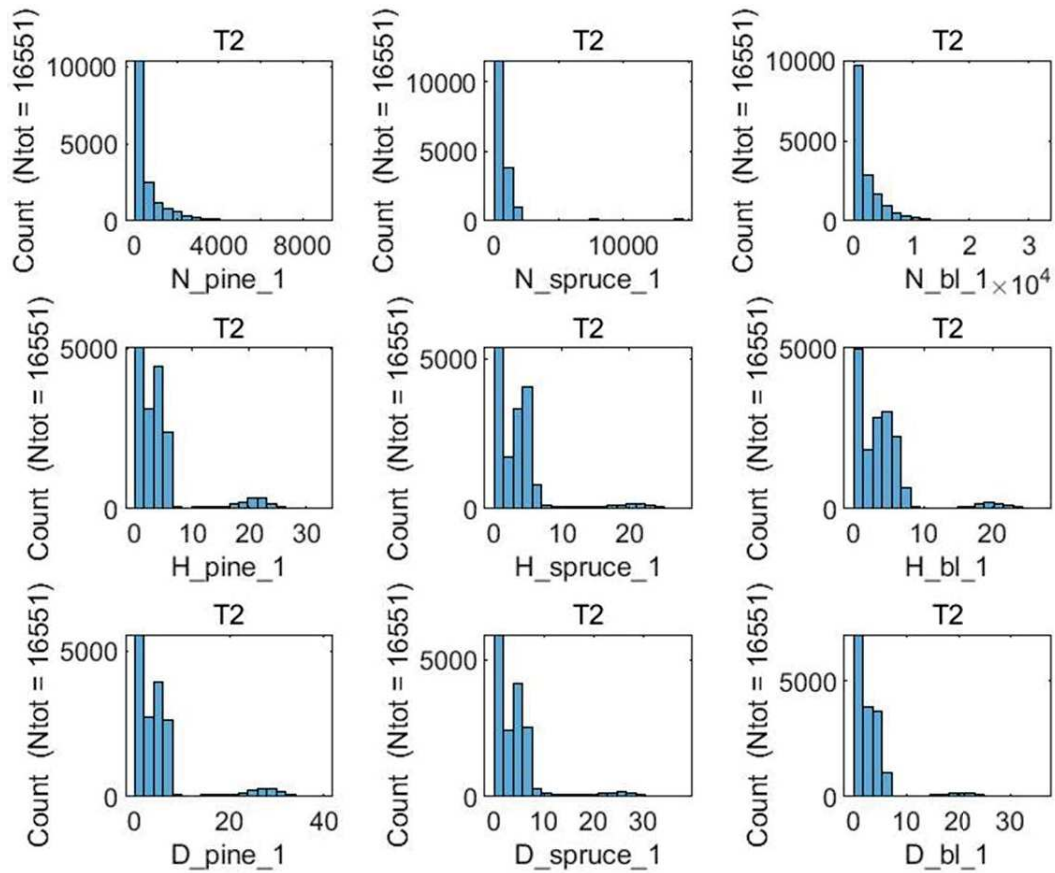


Runkoluvun (N), pituuden (H) ja keskiläpimitan (D) jakaumat Satakunnan kuvioille, joille ei maastokäyntiä. Kehitysluokka A0.



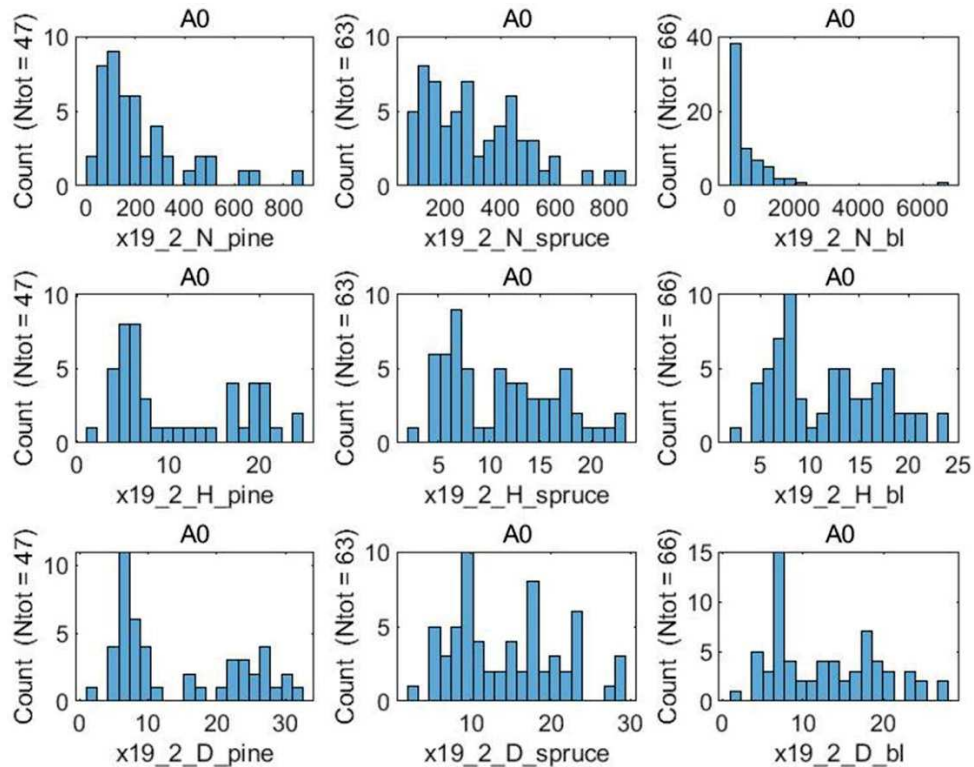
Runkoluvun (N), pituuden (H) ja keskiläpimitan (D) jakaumat Satakunnan kuvioille, joille ei maastokäyntiä. Kehitysluokka T1.

KUVIOT, JOILLE EI MAASTOKÄYNTIÄ (D1)

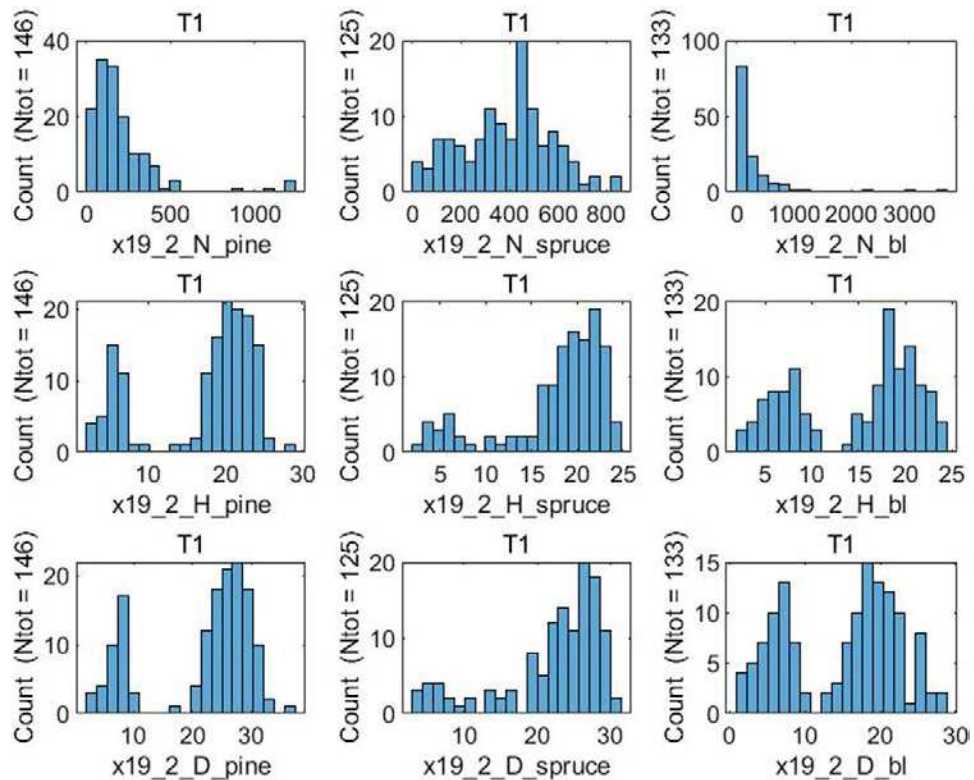


Runkoluvun (N), pituuden (H) ja keskiläpimitan (D) jakaumat Satakunnan kuvioille, joille ei maastokäyntiä. Kehitysluokka T1.

PUUSTOTULKINNASTA TUOTETUT PUUSTOTUNNUKSET (D19_2)

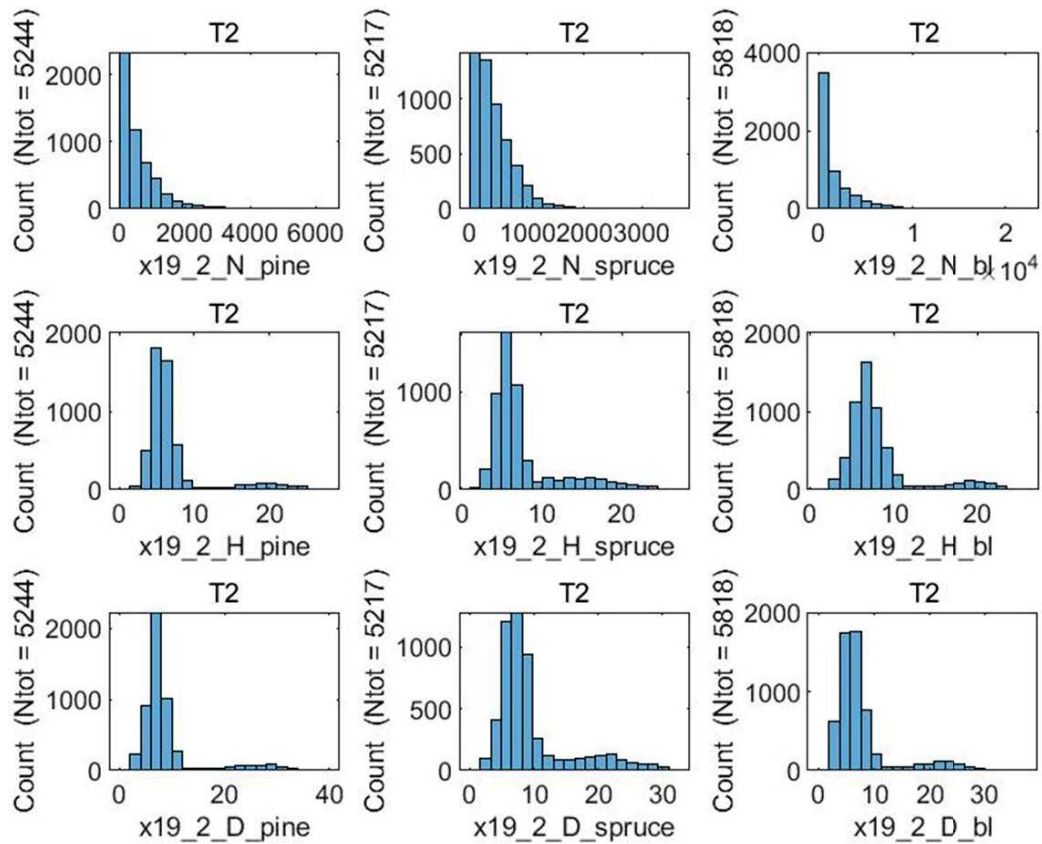


Puustotulkinnasta maastossa inventoiduille Satakunnan kuvioille tuotettujen puustotunnusten N, H ja D jakaumat. Kehitysluokka A0.



Puustotulkinnasta maastossa inventoiduille Satakunnan kuvioille tuotettujen puustotunnusten N, H ja D jakaumat. Kehitysluokka T1.

PUUSTOTULKINNASTA TUOTETUT PUUSTOTUNNUKSET (D19_2)



Puustotulkinnasta maastossa inventoiduille Satakunnan kuvioille tuotettujen puustotunnusten N H ja D jakaumat. Kehitysluokka T2.

LIITE B KORRELAATIOTAULUKOT

Taulukko 18. Maastossa inventoidun puustotiedon ja latvusmallipiirteiden välinen korrelaatiotaulukko (D18 vs. D13), Kehitysluokka = T1, Pituusluokka 0 - 2 m

T1 0 - 2 m	CHM-piirteet					N = 67
D18 piirteet	x13_max_1	x13_mean_1	x13_median_1	x13_std_1	x13_MDM	
D_pine_1						
H_pine_1	-0.057	-0.222	-0.193	-0.106	-0.014	
N_pine_1	-0.127	-0.159	-0.191	0.013	0.123	
D_spruce_1						
H_spruce_1	-0.013	-0.170	-0.167	0.016	0.017	
N_spruce_1	0.155	0.121	0.147	0.017	-0.171	
D_bl_1						
H_bl_1	-0.045	-0.115	-0.079	-0.097	-0.073	
N_bl_1	-0.152	-0.153	-0.131	-0.119	-0.008	

Taulukko 19. Maastossa inventoidun puustotiedon ja latvusmallipiirteiden välinen korrelaatiotaulukko (D18 vs. D13), Kehitysluokka = T2, Pituusluokka 0 - 2 m

T2 0 - 2 m	CHM-piirteet					N = 1044
D18 piirteet	x13_max_1	x13_mean_1	x13_median_1	x13_std_1	x13_MDM	
D_pine_1	-0.054	-0.048	-0.027	-0.081	-0.045	
H_pine_1	-0.077	-0.117	-0.101	-0.102	-0.062	
N_pine_1	-0.033	-0.017	-0.020	-0.005	-0.032	
D_spruce_1	-0.046	-0.033	-0.020	-0.055	-0.070	
H_spruce_1	-0.088	-0.088	-0.078	-0.080	-0.082	
N_spruce_1	-0.023	-0.030	-0.010	-0.039	0.042	
D_bl_1	-0.041	-0.097	-0.076	-0.074	-0.122	
H_bl_1	-0.056	-0.110	-0.111	-0.047	-0.067	
N_bl_1	-0.094	-0.067	-0.047	-0.069	-0.057	

Taulukko 20. Maastossa inventoidun puustotiedon ja latvusmallipiirteiden välinen korrelaatiotaulukko (D18 vs. D13), Kehitysluokka = T2, Pituusluokka 2 - 4 m

T2 2 - 4 m	CHM-piirteet					N = 795
D18 piirteet	x13_max_1	x13_mean_1	x13_median_1	x13_std_1	x13_MDM	
D_pine_1	-0.024	-0.035	0.022	-0.097	0.007	
H_pine_1	0.029	0.060	0.127	-0.045	-0.040	
N_pine_1	-0.013	-0.025	0.022	-0.070	0.044	
D_spruce_1	-0.024	0.017	0.005	0.023	-0.029	
H_spruce_1	0.005	0.016	0.009	0.037	-0.039	
N_spruce_1	0.006	0.018	-0.004	0.036	-0.023	
D_bl_1	-0.072	0.040	0.078	-0.029	0.005	
H_bl_1	-0.026	0.102	0.153	-0.022	-0.059	
N_bl_1	-0.015	-0.022	-0.007	-0.055	-0.029	

Taulukko 21. Maastossa inventoidun puustotiedon ja latvusmallipiirteiden välinen korrelaatiotaulukko (D18 vs. D13), Kehitysluokka = T2, Pituusluokka 4 - 6 m

T2 4 - 6 m	CHM-piirteet				N = 404
D18 piirteet	x13_max_1	x13_mean_1	x13_median_1	x13_std_1	x13_MDM
D_pine_1	-0.040	-0.007	0.048	-0.079	-0.091
H_pine_1	-0.013	0.143	0.232	-0.025	-0.151
N_pine_1	-0.050	-0.093	-0.069	-0.109	-0.037
D_spruce_1	0.010	-0.020	0.013	-0.021	-0.048
H_spruce_1	-0.013	0.067	0.121	-0.005	-0.107
N_spruce_1	-0.107	-0.008	0.015	-0.081	-0.135
D_bl_1	0.014	-0.008	0.001	0.067	-0.023
H_bl_1	-0.029	0.126	0.154	-0.007	-0.080
N_bl_1	-0.029	0.100	0.107	-0.021	-0.084

Taulukko 22. Maastossa inventoidun puustotiedon ja latvusmallipiirteiden välinen korrelaatiotaulukko (D18 vs. D13), Kehitysluokka = T2, Pituusluokka > 6 m

T2 6 m plus	CHM-piirteet				N = 43
D18 piirteet	x13_max_1	x13_mean_1	x13_median_1	x13_std_1	x13_MDM
D_pine_1	0.423	0.826	0.860	0.492	-0.053
H_pine_1	0.447	0.848	0.877	0.532	-0.063
N_pine_1	-0.369	-0.452	-0.428	-0.351	-0.323
D_spruce_1	0.411	0.838	0.873	0.489	-0.099
H_spruce_1	0.452	0.840	0.871	0.526	-0.085
N_spruce_1	-0.187	-0.220	-0.255	-0.128	0.073
D_bl_1	0.433	0.822	0.858	0.509	-0.061
H_bl_1	0.460	0.849	0.880	0.537	-0.073
N_bl_1	-0.040	0.003	-0.005	-0.123	-0.041

Taulukko 23. Maastossa inventoidun puustotiedon ja ilmakuvapiirteiden välinen korrelaatiotaulukko (D18 vs. D16), Kehitysluokka = T1, Pituusluokka 0 - 2 m

T10 - 2 m	Ilmakuvapiirteet										N = 67
D18 piirteet	x16_mean_1	x16_mean_2	x16_mean_3	x16_nirRel	x16_redRel	x16_greenRel	x16_std_1	x16_std_2	x16_std_3	x16_NDVI	x16_MDM
D_pine_1											
H_pine_1	0.247	0.135	0.149	0.174	-0.174	-0.101	0.014	-0.054	-0.072	0.177	-0.109
N_pine_1	0.009	-0.073	-0.144	0.101	0.032	-0.232	0.121	0.207	0.159	0.055	0.069
D_spruce_1											
H_spruce_1	0.286	0.111	0.073	0.261	-0.174	-0.266	0.021	0.000	-0.026	0.231	-0.233
N_spruce_1	0.005	0.107	0.134	-0.125	0.065	0.151	-0.341	-0.287	-0.234	-0.104	0.030
D_bl_1											
H_bl_1	0.121	0.089	0.084	0.070	-0.059	-0.055	-0.163	-0.061	-0.051	0.059	0.058
N_bl_1	-0.212	-0.196	-0.211	-0.065	0.127	-0.043	0.095	0.202	0.169	-0.088	0.217
D18 piirteet	x16_min_1	x16_min_2	x16_min_3	x16_max_1	x16_max_2	x16_max_3	x16_median_1	x16_median_2	x16_median_3	x16_contr	
D_pine_1											
H_pine_1	0.179	0.082	0.149	0.116	0.087	0.092	0.251	0.146	0.164	-0.133	
N_pine_1	0.176	0.071	0.036	0.017	0.107	-0.003	0.020	-0.071	-0.135	-0.038	
D_spruce_1											
H_spruce_1	0.327	0.149	0.136	0.083	0.070	0.038	0.326	0.151	0.123	-0.140	
N_spruce_1	0.010	-0.003	-0.100	-0.074	-0.077	0.028	-0.035	0.078	0.102	0.113	
D_bl_1											
H_bl_1	0.145	0.053	-0.006	0.087	0.103	0.129	0.077	0.071	0.064	0.031	
N_bl_1	-0.041	-0.172	-0.205	-0.196	-0.058	-0.135	-0.223	-0.203	-0.218	0.199	

Taulukko 24. Maastossa inventoidun puustotiedon ja ilmakuvapiirteiden välinen korrelaatiotaulukko (D18 vs. D16), Kehitysluokka = T2, Pituusluokka 0 - 2 m

T20 - 2 m	Ilmakuvapiirteet										N = 1044
D18 piirteet	x16_mean_1	x16_mean_2	x16_mean_3	x16_nirRel	x16_redRel	x16_greenRel	x16_std_1	x16_std_2	x16_std_3	x16_NDVI	x16_MDM
D_pine_1	0.161	-0.059	0.016	0.208	-0.231	-0.074	0.032	-0.053	-0.031	0.221	-0.034
H_pine_1	0.237	-0.087	0.026	0.305	-0.342	-0.102	0.084	-0.021	0.009	0.325	-0.062
N_pine_1	0.005	-0.002	-0.007	0.007	0.003	-0.019	-0.052	-0.036	-0.018	0.008	-0.035
D_spruce_1	0.189	-0.074	0.014	0.249	-0.267	-0.104	0.074	-0.039	-0.016	0.262	-0.027
H_spruce_1	0.242	-0.110	-0.001	0.341	-0.356	-0.157	0.113	-0.030	0.000	0.355	-0.054
N_spruce_1	-0.007	-0.107	-0.072	0.086	-0.085	-0.046	0.096	0.138	0.120	0.085	0.096
D_bl_1	0.193	-0.099	-0.003	0.285	-0.298	-0.130	0.060	-0.047	-0.029	0.298	0.010
H_bl_1	0.246	-0.118	0.000	0.355	-0.378	-0.152	0.145	0.033	0.055	0.374	-0.019
N_bl_1	0.044	-0.119	-0.069	0.162	-0.155	-0.096	0.112	0.096	0.105	0.165	0.036
D18 piirteet	x16_min_1	x16_min_2	x16_min_3	x16_max_1	x16_max_2	x16_max_3	x16_median_1	x16_median_2	x16_median_3	x16_contr	
D_pine_1	0.098	0.001	0.055	0.068	-0.054	-0.003	0.155	-0.054	0.018	-0.056	
H_pine_1	0.122	-0.038	0.052	0.133	-0.013	0.047	0.236	-0.073	0.034	-0.041	
N_pine_1	-0.050	-0.019	-0.022	-0.041	0.023	0.022	0.015	-0.003	-0.003	0.018	
D_spruce_1	0.107	0.016	0.091	0.118	-0.062	-0.002	0.186	-0.066	0.017	-0.092	
H_spruce_1	0.148	-0.011	0.064	0.154	-0.046	0.003	0.244	-0.094	0.010	-0.062	
N_spruce_1	-0.065	-0.180	-0.137	-0.028	-0.011	-0.041	-0.005	-0.094	-0.067	0.149	
D_bl_1	0.075	-0.046	0.048	0.121	-0.024	0.003	0.190	-0.097	-0.005	-0.040	
H_bl_1	0.081	-0.113	-0.021	0.150	0.004	0.031	0.252	-0.102	0.011	0.037	
N_bl_1	-0.071	-0.201	-0.136	0.071	-0.017	-0.024	0.053	-0.102	-0.058	0.146	

Taulukko 25. Maastossa inventoidun puustotiedon ja ilmakuvapiirteiden välinen korrelaatiotaulukko (D18 vs. D16), Kehitysluokka = T2, Pituusluokka 2 - 4 m

T2 2 - 4 m	Ilmakuvapiirteet										N = 795
D18 piirteet	x16_mean_1	x16_mean_2	x16_mean_3	x16_nirRel	x16_redRel	x16_greenRel	x16_std_1	x16_std_2	x16_std_3	x16_NDVI	x16_MDM
D_pine_1	-0.068	-0.072	-0.091	0.012	0.035	-0.059	-0.078	-0.066	-0.048	-0.008	-0.007
H_pine_1	-0.096	-0.115	-0.112	0.024	0.014	-0.053	-0.073	-0.085	-0.053	0.008	0.046
N_pine_1	-0.223	0.047	-0.091	-0.209	0.295	-0.024	-0.176	-0.054	-0.067	-0.261	-0.017
D_spruce_1	0.036	-0.033	0.054	0.020	-0.104	0.091	0.097	-0.011	0.027	0.061	0.016
H_spruce_1	0.037	-0.045	0.029	0.039	-0.097	0.053	0.045	-0.066	-0.025	0.069	0.018
N_spruce_1	0.138	-0.031	0.062	0.116	-0.182	0.035	0.160	0.145	0.157	0.156	0.023
D_bl_1	0.081	0.073	0.233	-0.088	-0.138	0.297	0.289	0.148	0.207	0.015	0.034
H_bl_1	-0.006	-0.169	-0.160	0.183	-0.095	-0.171	-0.044	-0.016	-0.002	0.150	0.053
N_bl_1	0.160	-0.091	-0.040	0.237	-0.208	-0.122	-0.001	-0.031	-0.006	0.240	-0.042
Ilmakuvapiirteet											
D18 piirteet	x16_min_1	x16_min_2	x16_min_3	x16_max_1	x16_max_2	x16_max_3	x16_median_1	x16_median_2	x16_median_3	x16_contr	
D_pine_1	0.032	0.024	-0.013	-0.027	-0.044	-0.046	-0.068	-0.074	-0.090	0.001	
H_pine_1	0.044	0.039	0.007	-0.021	-0.035	-0.037	-0.106	-0.125	-0.121	-0.012	
N_pine_1	-0.031	0.005	-0.080	-0.134	0.093	0.016	-0.224	0.041	-0.086	0.081	
D_spruce_1	0.049	0.130	0.163	0.044	-0.031	0.035	0.038	-0.040	0.040	-0.157	
H_spruce_1	0.093	0.170	0.175	0.038	-0.060	0.003	0.036	-0.053	0.015	-0.178	
N_spruce_1	-0.037	-0.126	-0.053	0.057	-0.018	0.011	0.142	-0.023	0.058	0.064	
D_bl_1	-0.053	0.076	0.248	0.193	0.175	0.284	0.080	0.056	0.207	-0.180	
H_bl_1	0.077	-0.072	-0.080	0.003	-0.060	-0.073	-0.018	-0.162	-0.153	0.072	
N_bl_1	0.093	-0.083	-0.075	0.041	-0.101	-0.080	0.161	-0.074	-0.027	0.052	

Taulukko 26. Maastossa inventoidun puustotiedon ja ilmakuvapiirteiden välinen korrelaatiotaulukko (D18 vs. D16), Kehitysluokka = T2, Pituusluokka 4 - 6 m

T2 4 - 6 m	Ilmakuvapiirteet										N = 404
D18 piirteet	x16_mean_1	x16_mean_2	x16_mean_3	x16_nirRel	x16_redRel	x16_greenRel	x16_std_1	x16_std_2	x16_std_3	x16_NDVI	x16_MDM
D_pine_1	-0.071	-0.058	-0.065	-0.018	0.031	-0.004	-0.102	-0.137	-0.105	-0.024	0.119
H_pine_1	-0.093	-0.076	0.064	-0.097	-0.054	0.193	0.110	0.025	0.083	-0.035	0.151
N_pine_1	-0.201	0.026	-0.115	-0.149	0.259	-0.034	-0.229	-0.049	-0.078	-0.204	0.059
D_spruce_1	0.009	0.029	0.035	-0.028	0.004	0.037	0.007	-0.122	-0.076	-0.014	0.060
H_spruce_1	0.019	0.037	0.128	-0.074	-0.054	0.160	0.144	0.025	0.088	-0.019	0.071
N_spruce_1	0.132	-0.065	0.074	0.095	-0.200	0.054	0.225	0.115	0.150	0.149	-0.037
D_bl_1	0.043	0.123	0.304	-0.177	-0.082	0.336	0.324	0.141	0.193	-0.071	-0.008
H_bl_1	0.007	-0.145	-0.176	0.145	-0.050	-0.163	-0.109	-0.087	-0.098	0.111	0.198
N_bl_1	0.050	-0.196	-0.171	0.205	-0.133	-0.170	-0.013	0.004	-0.006	0.186	0.099
Ilmakuvapiirteet											
D18 piirteet	x16_min_1	x16_min_2	x16_min_3	x16_max_1	x16_max_2	x16_max_3	x16_median_1	x16_median_2	x16_median_3	x16_contr	
D_pine_1	0.019	0.115	0.058	-0.106	-0.040	-0.006	-0.083	-0.074	-0.079	-0.068	
H_pine_1	-0.067	0.031	0.119	0.020	-0.023	0.079	-0.109	-0.096	0.032	-0.102	
N_pine_1	-0.034	-0.061	-0.169	-0.107	0.038	-0.057	-0.206	0.014	-0.115	0.163	
D_spruce_1	0.070	0.250	0.197	-0.011	-0.042	0.036	-0.002	0.010	0.014	-0.211	
H_spruce_1	0.024	0.164	0.207	0.008	-0.006	0.105	0.007	0.019	0.101	-0.196	
N_spruce_1	-0.071	-0.120	-0.025	0.184	0.007	0.035	0.146	-0.046	0.082	0.045	
D_bl_1	-0.035	0.163	0.379	0.180	0.171	0.298	0.029	0.105	0.267	-0.328	
H_bl_1	0.114	-0.029	-0.091	0.026	-0.106	-0.096	-0.032	-0.167	-0.193	0.058	
N_bl_1	-0.062	-0.236	-0.251	0.037	-0.102	-0.118	0.043	-0.186	-0.162	0.244	

Taulukko 27. Maastossa inventoidun puustotiedon ja ilmakuvapiirteiden välinen korrelaatiotaulukko (D18 vs. D16), Kehitysluokka = T2, Pituusluokka > 6 m

T2 6 m plus	Ilmakuvapiirteet										N = 43
D18 piirteet	x16_mean_1	x16_mean_2	x16_mean_3	x16_nirRel	x16_redRel	x16_greenRel	x16_std_1	x16_std_2	x16_std_3	x16_NDVI	x16_MDM
D_pine_1	-0.568	-0.443	-0.529	-0.222	0.457	-0.147	-0.329	0.078	0.015	-0.326	0.830
H_pine_1	-0.569	-0.462	-0.549	-0.210	0.455	-0.163	-0.338	0.082	0.007	-0.319	0.866
N_pine_1	0.119	0.265	0.314	-0.110	-0.077	0.254	0.144	-0.013	0.036	-0.031	-0.180
D_spruce_1	-0.547	-0.451	-0.508	-0.211	0.421	-0.125	-0.300	0.072	0.014	-0.304	0.853
H_spruce_1	-0.529	-0.460	-0.545	-0.177	0.425	-0.183	-0.308	0.086	0.012	-0.286	0.857
N_spruce_1	-0.079	-0.054	-0.149	0.057	0.031	-0.122	-0.133	-0.094	-0.134	0.019	-0.087
D_bl_1	-0.559	-0.401	-0.402	-0.291	0.418	0.003	-0.241	0.098	0.059	-0.350	0.814
H_bl_1	-0.526	-0.478	-0.590	-0.149	0.433	-0.234	-0.362	0.045	-0.027	-0.273	0.832
N_bl_1	0.409	0.157	-0.110	0.368	-0.129	-0.434	-0.202	-0.126	-0.093	0.272	-0.081
	Ilmakuvapiirteet										
D18 piirteet	x16_min_1	x16_min_2	x16_min_3	x16_max_1	x16_max_2	x16_max_3	x16_median_1	x16_median_2	x16_median_3	x16_contr	
D_pine_1	-0.155	-0.447	-0.365	-0.230	0.015	-0.056	-0.543	-0.504	-0.567	0.409	
H_pine_1	-0.153	-0.458	-0.367	-0.212	0.024	-0.069	-0.547	-0.529	-0.592	0.410	
N_pine_1	-0.028	0.349	0.356	0.201	0.322	0.371	0.072	0.239	0.280	-0.285	
D_spruce_1	-0.120	-0.419	-0.311	-0.179	0.031	-0.040	-0.528	-0.518	-0.553	0.352	
H_spruce_1	-0.131	-0.452	-0.365	-0.175	0.024	-0.072	-0.505	-0.526	-0.586	0.400	
N_spruce_1	0.025	0.105	-0.067	-0.431	-0.245	-0.258	-0.079	-0.050	-0.139	0.033	
D_bl_1	-0.160	-0.356	-0.190	-0.171	0.082	0.041	-0.540	-0.471	-0.456	0.243	
H_bl_1	-0.117	-0.468	-0.417	-0.262	-0.023	-0.126	-0.497	-0.538	-0.623	0.446	
N_bl_1	0.655	0.239	-0.151	-0.034	-0.076	-0.062	0.378	0.138	-0.113	0.137	

Taulukko 28. Puustotulkinnasta maastoinventointihetkeen kasvatetut laskentapuustotietojen ja latvusmallipiirteiden välinen korrelaatiotaulukko (D19_2 vs. D13), Kehitysluokka = T1, Pituusluokka 0 - 2 m

T1 0 - 2 m	Latusmallipiirteet				N = 67
D19_2 piirteet	x13_max_1	x13_mean_1	x13_median_1	x13_std_1	x13_MDM
x19_2_N_pine	0.039	0.006	0.000	-0.051	-0.026
x19_2_D_pine	-0.202	0.224	0.254	0.017	-0.274
x19_2_H_pine	-0.254	0.207	0.251	-0.051	-0.355
x19_2_N_spruce	-0.406	0.068	0.178	-0.351	-0.706
x19_2_D_spruce	-0.238	0.213	0.253	-0.013	-0.345
x19_2_H_spruce	-0.255	0.222	0.262	-0.024	-0.353

Taulukko 29. Puustotulkinnasta maastoinventointihetkeen kasvatetut laskentapuustotietojen ja latvusmallipiirteiden välinen korrelaatiotaulukko (D19_2 vs. D13), Kehitysluokka = T2, Pituusluokka 0 - 2 m

T2 0 - 2 m	CHM-piirteet				N = 1044
D19.2 piirteet	x13_max_1	x13_mean_1	x13_median_1	x13_std_1	x13_MDM
x19_2_N_pine	-0.139	-0.071	-0.037	-0.162	-0.075
x19_2_D_pine	0.015	0.400	0.336	0.305	0.012
x19_2_H_pine	0.008	0.432	0.370	0.314	0.017
x19_2_N_spruce	-0.290	0.058	0.132	-0.253	-0.077
x19_2_D_spruce	-0.039	0.453	0.440	0.185	-0.109
x19_2_H_spruce	-0.041	0.466	0.447	0.206	-0.100

Taulukko 30. Puustotulkinnasta maastoinventointihetkeen kasvatetut laskentapuustotietojen ja latvusmallipiirteiden välinen korrelaatiotaulukko (D19_2 vs. D13), Kehitysluokka = T2, Pituusluokka 0 - 2 m

T2 2 - 4 m	CHM-piirteet				N = 795
D19.2 piirteet	x13_max_1	x13_mean_1	x13_median_1	x13_std_1	x13_MDM
x19_2_N_pine	-0.110	0.041	0.220	-0.274	-0.035
x19_2_D_pine	0.116	0.326	0.189	0.339	0.091
x19_2_H_pine	0.124	0.367	0.226	0.361	0.125
x19_2_N_spruce	-0.185	-0.121	0.034	-0.305	-0.050
x19_2_D_spruce	0.147	0.375	0.206	0.386	0.177
x19_2_H_spruce	0.146	0.419	0.246	0.413	0.185

Taulukko 31. Puustotulkinnasta maastoinventointihetkeen kasvatetut laskentapuustotietojen ja latvusmallipiirteiden välinen korrelaatiotaulukko (D19_2 vs. D13), Kehitysluokka = T2, Pituusluokka 0 - 2 m

T2 4 - 6 m	CHM-piirteet				N = 404
D19.2 piirteet	x13_max_1	x13_mean_1	x13_median_1	x13_std_1	x13_MDM
x19_2_N_pine	-0.051	-0.055	0.041	-0.287	-0.047
x19_2_D_pine	0.037	0.167	0.121	0.216	0.015
x19_2_H_pine	0.050	0.212	0.170	0.234	0.021
x19_2_N_spruce	-0.144	-0.193	-0.106	-0.301	-0.123
x19_2_D_spruce	0.145	0.311	0.245	0.325	-0.047
x19_2_H_spruce	0.155	0.344	0.271	0.354	-0.040

Taulukko 32. Puustotulkinnasta maastoinventointihetkeen kasvatetut laskentapuustotietojen ja latvusmallipiirteiden välinen korrelaatiotaulukko (D19_2 vs. D13), Kehitysluokka = T2, Pituusluokka > 6 m

T2 6 m plus	CHM-piirteet				N = 43
D19.2 piirteet	x13_max_1	x13_mean_1	x13_median_1	x13_std_1	x13_MDM
x19_2_N_pine	-0.206	-0.245	-0.209	-0.377	-0.001
x19_2_D_pine	0.436	0.815	0.791	0.627	0.034
x19_2_H_pine	0.456	0.847	0.827	0.630	0.030
x19_2_N_spruce	-0.254	-0.372	-0.355	-0.331	-0.037
x19_2_D_spruce	0.498	0.730	0.737	0.534	-0.056
x19_2_H_spruce	0.506	0.797	0.809	0.560	-0.068

Taulukko 33. Puustotulkinnasta maastoinventointihetken kasvatetut laskentapuustotietojen ja ilmakuvapiirteiden välinen korrelaatiotaulukko (D19_2 vs. D16), Kehitysluokka = T1, Pituusluokka 0 - 2 m

T10 - 2 m	Ilmakuvapiirteet										N = 67
D19_2 piirteet	x16_mean_1	x16_mean_2	x16_mean_3	x16_nirRel	x16_redRel	x16_greenRel	x16_std_1	x16_std_2	x16_std_3	x16_NDVI	x16_MDM
x19_2_N_pine	0.099	0.122	0.169	0.030	-0.100	0.063	0.204	-0.009	-0.048	0.047	-0.236
x19_2_D_pine	-0.345	-0.460	-0.456	0.149	-0.102	-0.119	-0.060	0.041	0.197	0.158	0.420
x19_2_H_pine	-0.349	-0.490	-0.477	0.186	-0.132	-0.143	-0.080	0.027	0.174	0.194	0.459
x19_2_N_spruce	-0.173	-0.415	-0.330	0.307	-0.325	-0.122	-0.086	0.091	0.146	0.339	0.463
x19_2_D_spruce	-0.353	-0.470	-0.481	0.171	-0.073	-0.185	-0.146	-0.020	0.118	0.161	0.420
x19_2_H_spruce	-0.366	-0.489	-0.495	0.177	-0.088	-0.177	-0.125	-0.003	0.133	0.170	0.435
	Ilmakuvapiirteet										
D19_2 piirteet	x16_min_1	x16_min_2	x16_min_3	x16_max_1	x16_max_2	x16_max_3	x16_median_1	x16_median_2	x16_median_3	x16_contr	
x19_2_N_pine	-0.123	0.116	0.240	0.093	-0.032	-0.074	0.124	0.138	0.175	-0.297	
x19_2_D_pine	-0.071	-0.404	-0.524	-0.213	-0.285	-0.291	-0.349	-0.459	-0.448	0.523	
x19_2_H_pine	-0.058	-0.446	-0.551	-0.249	-0.316	-0.335	-0.354	-0.488	-0.470	0.550	
x19_2_N_spruce	0.066	-0.601	-0.532	-0.247	-0.283	-0.355	-0.193	-0.415	-0.338	0.529	
x19_2_D_spruce	-0.022	-0.402	-0.549	-0.314	-0.353	-0.350	-0.357	-0.467	-0.469	0.541	
x19_2_H_spruce	-0.040	-0.428	-0.565	-0.300	-0.359	-0.353	-0.370	-0.486	-0.484	0.558	

Taulukko 34. Puustotulkinnasta maastoinventointihetken kasvatetut laskentapuustotietojen ja ilmakuvapiirteiden välinen korrelaatiotaulukko (D19_2 vs. D16), Kehitysluokka = T2, Pituusluokka 0 - 2 m

T20 - 2 m	Ilmakuvapiirteet										N = 1044
D19_2 piirteet	x16_mean_1	x16_mean_2	x16_mean_3	x16_nirRel	x16_redRel	x16_greenRel	x16_std_1	x16_std_2	x16_std_3	x16_NDVI	x16_MDM
x19_2_N_pine	0.011	-0.062	-0.100	0.118	-0.029	-0.198	-0.161	-0.186	-0.186	0.079	-0.037
x19_2_D_pine	-0.416	-0.335	-0.341	-0.123	0.169	-0.008	0.089	0.071	0.098	-0.141	0.346
x19_2_H_pine	-0.436	-0.369	-0.372	-0.110	0.165	-0.029	0.081	0.064	0.091	-0.132	0.382
x19_2_N_spruce	0.147	-0.214	-0.027	0.316	-0.412	-0.016	0.072	-0.102	-0.048	0.363	0.177
x19_2_D_spruce	-0.455	-0.478	-0.490	0.010	0.107	-0.185	0.122	0.042	0.062	-0.038	0.410
x19_2_H_spruce	-0.464	-0.492	-0.498	0.011	0.104	-0.181	0.132	0.041	0.061	-0.036	0.414
	Ilmakuvapiirteet										
D19_2 piirteet	x16_min_1	x16_min_2	x16_min_3	x16_max_1	x16_max_2	x16_max_3	x16_median_1	x16_median_2	x16_median_3	x16_contr	
x19_2_N_pine	0.264	0.200	0.096	-0.039	-0.107	-0.176	-0.007	-0.067	-0.097	-0.142	
x19_2_D_pine	-0.099	-0.173	-0.127	-0.225	-0.238	-0.196	-0.412	-0.342	-0.348	0.115	
x19_2_H_pine	-0.095	-0.185	-0.132	-0.236	-0.266	-0.221	-0.435	-0.377	-0.381	0.121	
x19_2_N_spruce	0.147	0.051	0.215	0.060	-0.215	-0.109	0.110	-0.225	-0.056	-0.221	
x19_2_D_spruce	-0.041	-0.154	-0.142	-0.224	-0.364	-0.338	-0.456	-0.474	-0.489	0.121	
x19_2_H_spruce	-0.033	-0.144	-0.123	-0.226	-0.375	-0.339	-0.462	-0.487	-0.497	0.101	

Taulukko 35. Puustotulkinnasta maastoinventointihetken kasvatetut laskentapuustotietojen ja ilmakuvapiirteiden välinen korrelaatiotaulukko (D19_2 vs. D16), Kehitysluokka = T2, Pituusluokka 2 - 4 m

T22 - 4 m	Ilmakuvapiirteet										N = 795
D19_2 piirteet	x16_mean_1	x16_mean_2	x16_mean_3	x16_nirRel	x16_redRel	x16_greenRel	x16_std_1	x16_std_2	x16_std_3	x16_NDVI	x16_MDM
x19_2_N_pine	-0.090	-0.277	-0.317	0.282	-0.078	-0.339	-0.370	-0.387	-0.386	0.193	0.158
x19_2_D_pine	-0.238	-0.106	-0.091	-0.144	0.095	0.109	0.195	0.151	0.152	-0.134	0.088
x19_2_H_pine	-0.253	-0.114	-0.109	-0.146	0.112	0.094	0.186	0.159	0.153	-0.143	0.085
x19_2_N_spruce	0.292	-0.124	0.184	0.256	-0.488	0.169	0.138	-0.138	-0.038	0.388	0.038
x19_2_D_spruce	-0.298	-0.108	-0.196	-0.151	0.221	-0.024	0.173	0.275	0.230	-0.197	0.055
x19_2_H_spruce	-0.303	-0.138	-0.211	-0.134	0.201	-0.026	0.180	0.250	0.204	-0.178	0.077
	Ilmakuvapiirteet										
D19_2 piirteet	x16_min_1	x16_min_2	x16_min_3	x16_max_1	x16_max_2	x16_max_3	x16_median_1	x16_median_2	x16_median_3	x16_contr	
x19_2_N_pine	0.217	0.106	-0.037	-0.116	-0.265	-0.301	-0.136	-0.293	-0.326	-0.034	
x19_2_D_pine	-0.088	-0.060	-0.033	-0.110	-0.026	-0.031	-0.223	-0.104	-0.095	0.028	
x19_2_H_pine	-0.090	-0.075	-0.053	-0.131	-0.019	-0.039	-0.236	-0.110	-0.111	0.050	
x19_2_N_spruce	0.083	0.154	0.327	0.172	-0.107	0.073	0.280	-0.131	0.161	-0.320	
x19_2_D_spruce	-0.122	-0.203	-0.229	-0.172	0.019	-0.067	-0.273	-0.094	-0.188	0.225	
x19_2_H_spruce	-0.103	-0.171	-0.199	-0.180	-0.003	-0.084	-0.279	-0.127	-0.206	0.191	

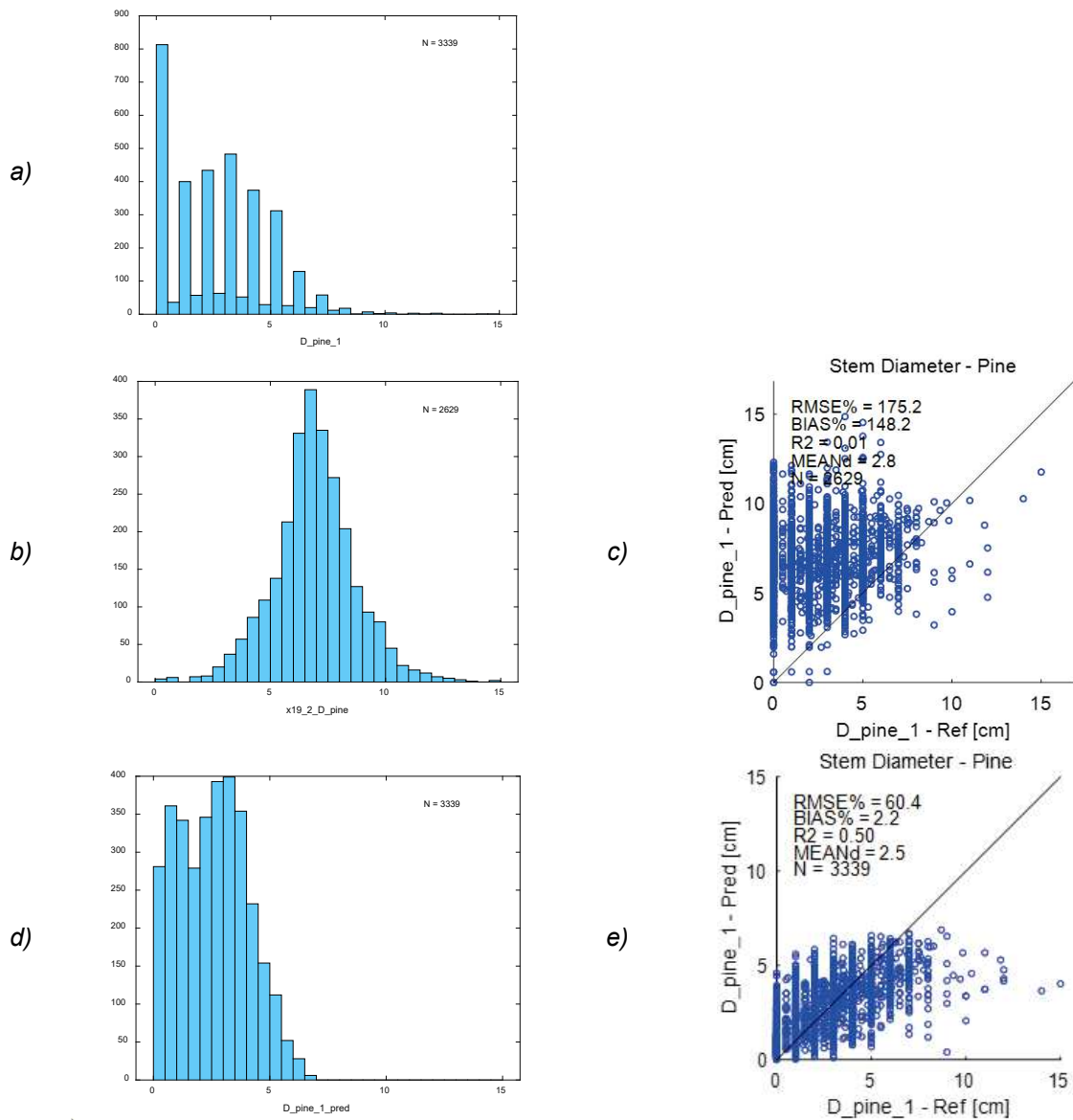
Taulukko 36. Puustotulkinnasta maastoinventointihetken kasvatetut laskentapuustotietojen ja ilmakuva-aihteiden välinen korrelaatiotaulukko (D19_2 vs. D16), Kehitysluokka = T2, Pituusluokka 4 - 6 m

T2 4 - 6 m	Ilmakuvapiirteet										N = 404
D19_2 piirteet	x16_mean_1	x16_mean_2	x16_mean_3	x16_nirRel	x16_redRel	x16_greenRel	x16_std_1	x16_std_2	x16_std_3	x16_NDVI	x16_MDM
x19_2_N_pine	-0.104	-0.316	-0.470	0.295	0.019	-0.447	-0.603	-0.588	-0.616	0.162	0.168
x19_2_D_pine	-0.092	0.023	0.054	-0.127	0.057	0.130	0.164	0.149	0.160	-0.098	0.043
x19_2_H_pine	-0.107	-0.005	0.017	-0.113	0.063	0.104	0.157	0.146	0.155	-0.092	0.070
x19_2_N_spruce	0.231	-0.120	0.066	0.239	-0.343	-0.019	-0.004	-0.167	-0.100	0.303	-0.008
x19_2_D_spruce	-0.141	-0.026	-0.001	-0.125	0.069	0.116	0.271	0.334	0.323	-0.102	0.031
x19_2_H_spruce	-0.140	-0.045	-0.027	-0.104	0.064	0.091	0.256	0.311	0.298	-0.087	0.063
	Ilmakuvapiirteet										
D19_2 piirteet	x16_min_1	x16_min_2	x16_min_3	x16_max_1	x16_max_2	x16_max_3	x16_median_1	x16_median_2	x16_median_3	x16_contr	
x19_2_N_pine	0.203	0.121	-0.205	-0.200	-0.333	-0.453	-0.139	-0.316	-0.451	0.103	
x19_2_D_pine	-0.065	-0.036	0.024	0.010	0.002	0.059	-0.089	0.015	0.039	-0.011	
x19_2_H_pine	-0.061	-0.048	-0.003	-0.010	-0.015	0.037	-0.104	-0.014	0.000	0.015	
x19_2_N_spruce	0.072	0.057	0.106	0.090	-0.088	-0.028	0.220	-0.108	0.073	-0.113	
x19_2_D_spruce	-0.153	-0.213	-0.127	-0.015	0.047	0.062	-0.119	-0.026	-0.009	0.152	
x19_2_H_spruce	-0.129	-0.198	-0.129	-0.017	0.028	0.041	-0.121	-0.049	-0.037	0.149	

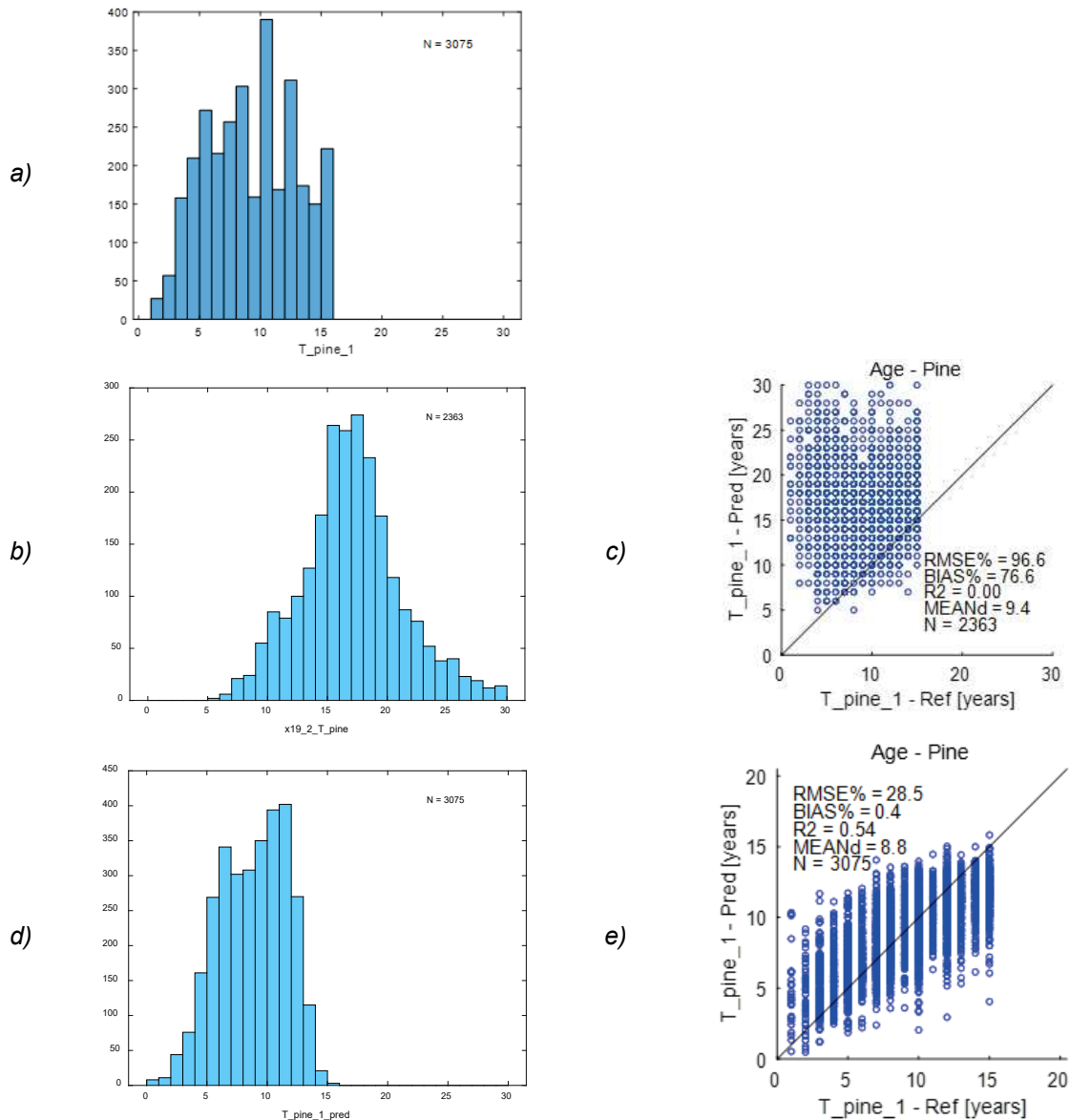
Taulukko 37. Puustotulkinnasta maastoinventointihetken kasvatetut laskentapuustotietojen ja ilmakuva-aihteiden välinen korrelaatiotaulukko (D19_2 vs. D16), Kehitysluokka = T2, Pituusluokka > 6 m

T2 6 m plus	Ilmakuvapiirteet										N = 43
D19_2 piirteet	x16_mean_1	x16_mean_2	x16_mean_3	x16_nirRel	x16_redRel	x16_greenRel	x16_std_1	x16_std_2	x16_std_3	x16_NDVI	x16_MDM
x19_2_N_pine	-0.032	-0.208	-0.283	0.225	-0.063	-0.280	-0.511	-0.657	-0.666	0.153	-0.110
x19_2_D_pine	-0.445	-0.322	-0.439	-0.164	0.402	-0.153	-0.268	0.242	0.163	-0.262	0.700
x19_2_H_pine	-0.444	-0.368	-0.496	-0.121	0.394	-0.211	-0.313	0.185	0.107	-0.234	0.731
x19_2_N_spruce	0.332	0.080	0.293	0.188	-0.398	0.113	0.337	-0.151	-0.076	0.299	-0.301
x19_2_D_spruce	-0.450	-0.260	-0.342	-0.252	0.418	-0.035	-0.224	0.252	0.207	-0.339	0.635
x19_2_H_spruce	-0.523	-0.330	-0.414	-0.265	0.460	-0.058	-0.280	0.214	0.162	-0.363	0.732
	Ilmakuvapiirteet										
D19_2 piirteet	x16_min_1	x16_min_2	x16_min_3	x16_max_1	x16_max_2	x16_max_3	x16_median_1	x16_median_2	x16_median_3	x16_contr	
x19_2_N_pine	0.066	0.209	0.014	-0.190	-0.429	-0.490	-0.072	-0.179	-0.259	-0.129	
x19_2_D_pine	-0.053	-0.526	-0.426	-0.117	0.005	-0.041	-0.431	-0.375	-0.474	0.465	
x19_2_H_pine	-0.032	-0.528	-0.455	-0.139	-0.023	-0.082	-0.430	-0.421	-0.529	0.487	
x19_2_N_spruce	-0.022	0.249	0.345	0.445	-0.034	0.030	0.324	0.115	0.325	-0.383	
x19_2_D_spruce	-0.027	-0.340	-0.297	-0.192	0.142	0.074	-0.431	-0.320	-0.391	0.357	
x19_2_H_spruce	-0.049	-0.377	-0.314	-0.221	0.113	0.041	-0.506	-0.397	-0.468	0.371	

**LIITE C AINEISTON D19_2 JA VTT:n MLP-ESTIMAATTIEN
VERTAILU MAASTOSSA INVENTOITUUN TIETOO**

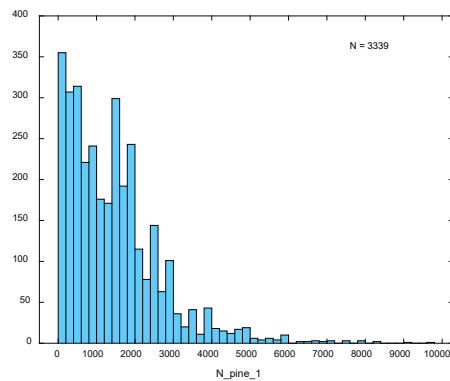


Kuva 21. Eri lähteistä saadut männyn keskiläpimitan (D_{pine}) jakaumat ja hajontakuviot maastoinventointitietoa vastaan: Maastoinventointitieto (a), aineiston D_{19_2} jakauma (b) ja hajontakuvio (c), MLP-estimaatin jakauma (d) ja hajontakuvio (e). Kuvissa b) ja c) on mukana vain testikuvio, joille $D_{19_2_D_pine} < 15$ cm (ks. 4.3.6).

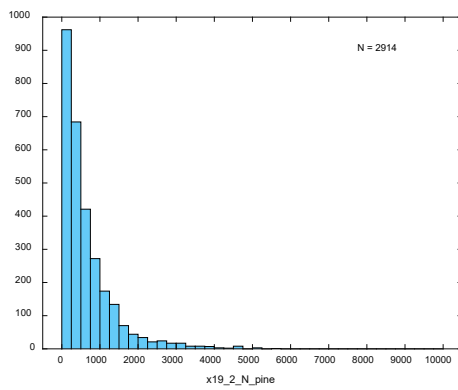


Kuva 22. Eri lähteistä saadut männyn iän (T_{pine}) jakaumat ja hajontakuviot maastoinventointitietoa vastaan: Maastoinventointitieto (a), aineiston D19_2 jakauma (b) ja hajontakuviot (c), MLP-estimaatin jakauma (d) ja hajontakuviot (e). MLP-estimointi rajattiin koskemaan kuvioita, joille $T_{\text{pine}_1} < 16$. Kuvassa b) ja c) ovat mukana vain testikuvio, joille D19_2_ $T_{\text{pine}} < 30$ cm (ks. 4.3.6).

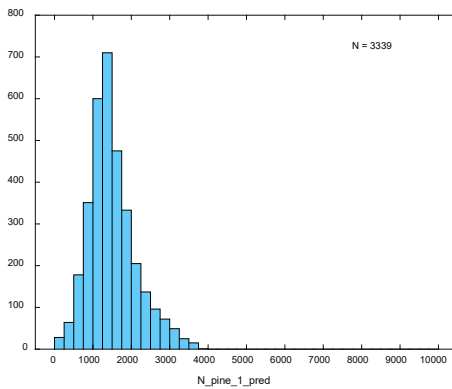
a)



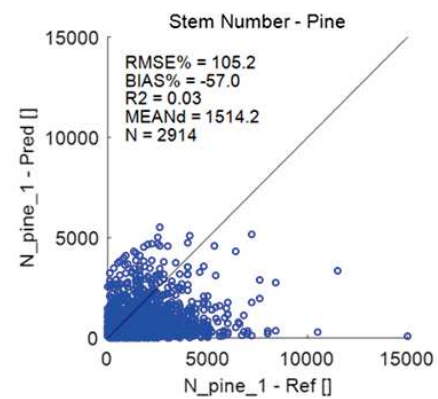
b)



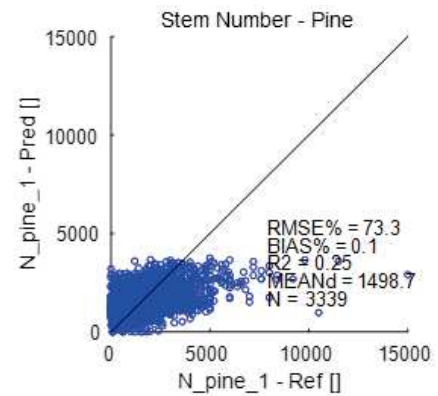
d)



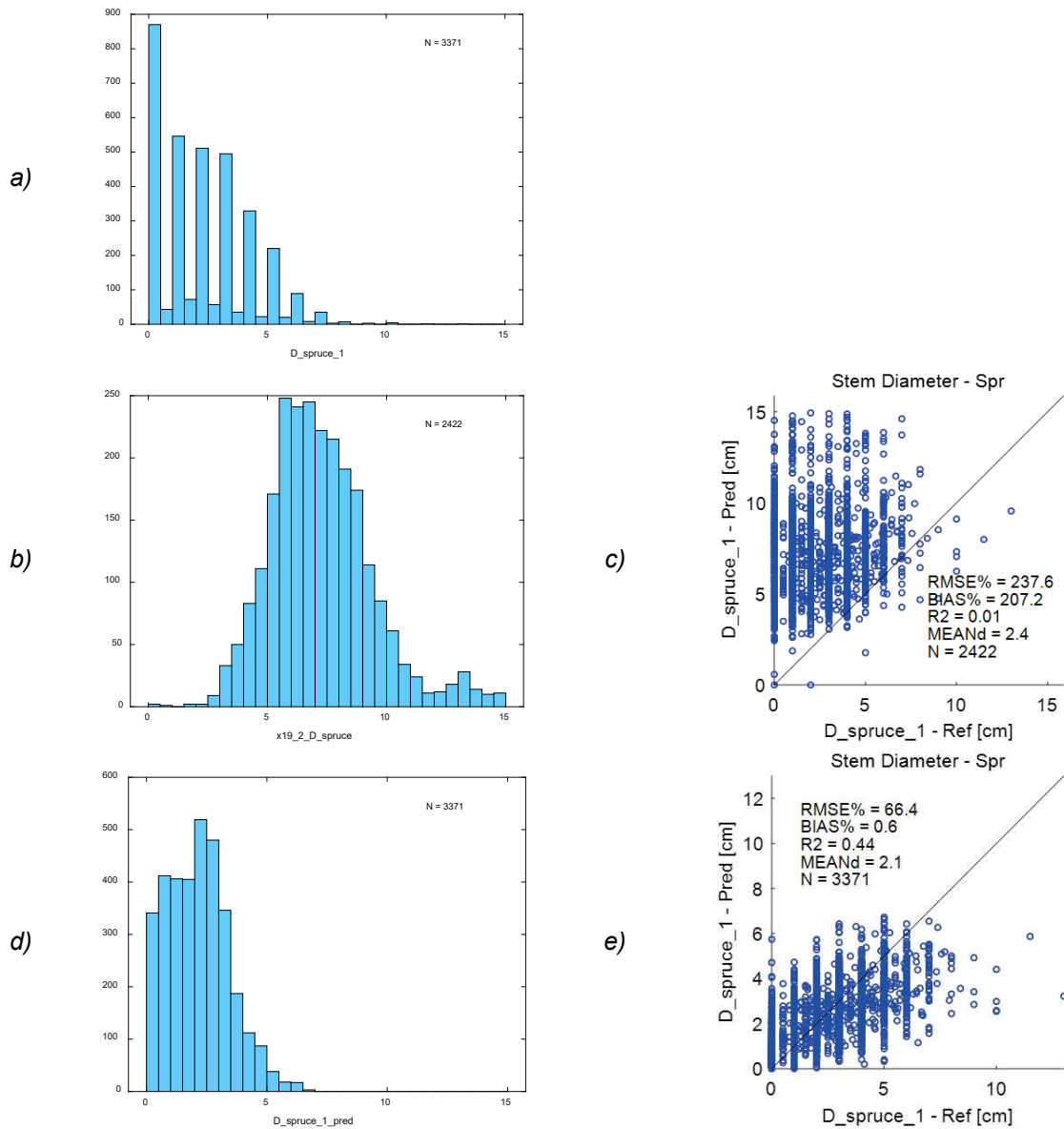
c)



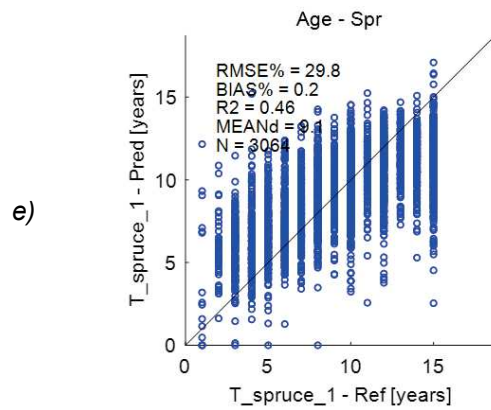
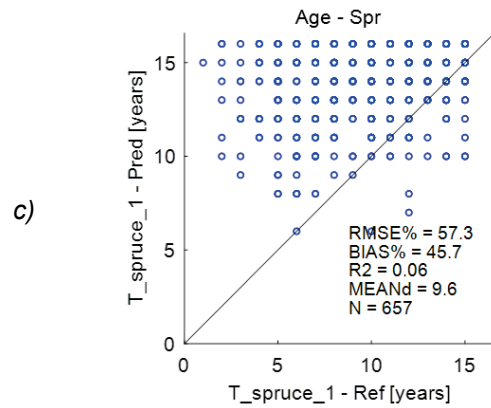
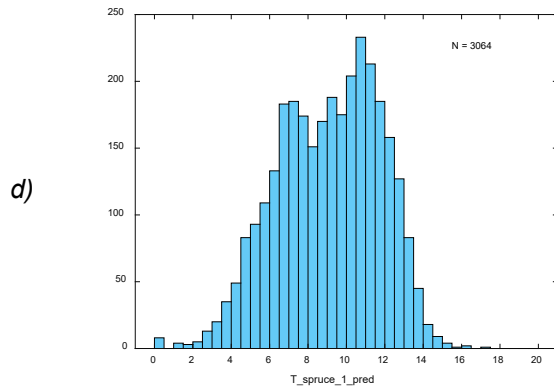
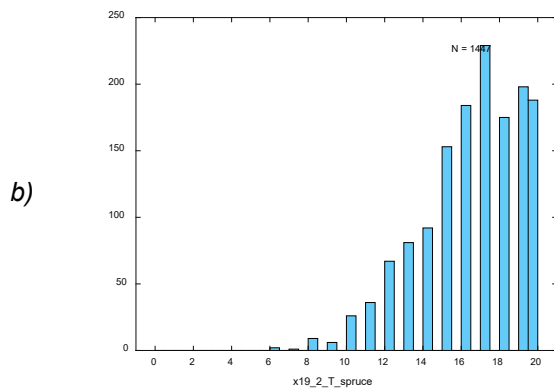
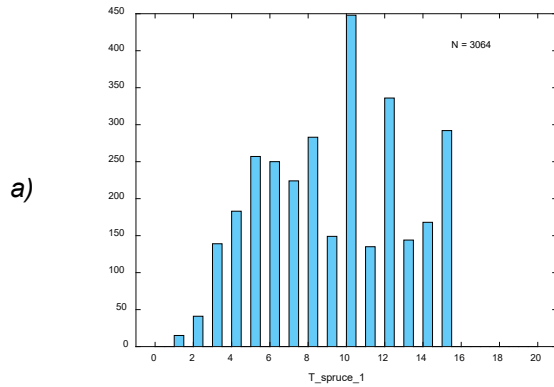
e)



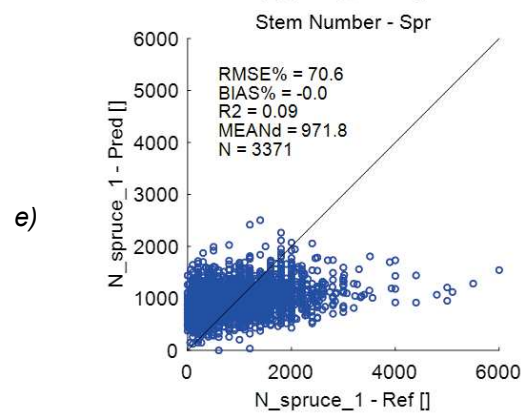
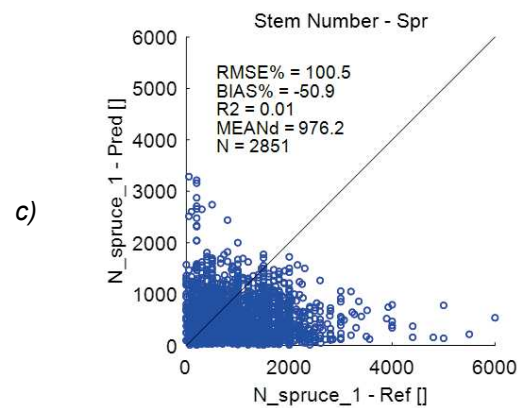
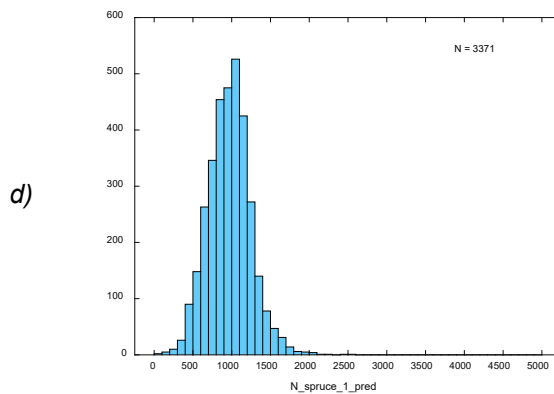
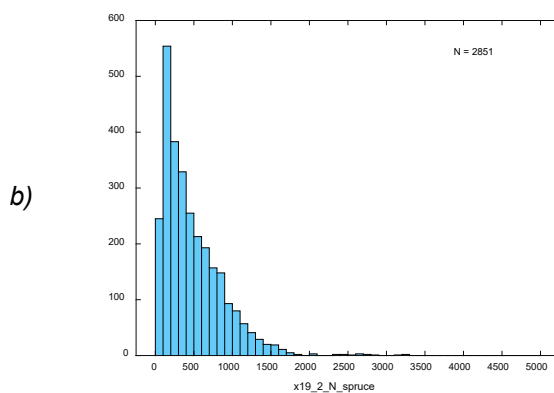
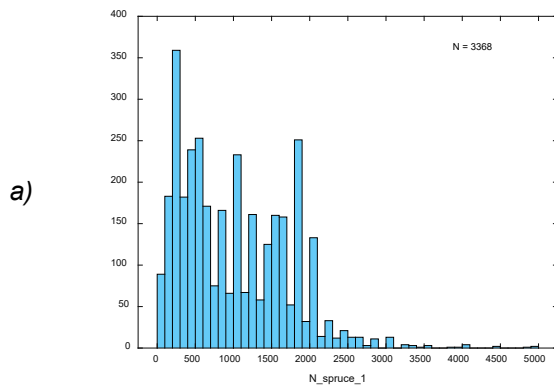
Eri lähteistä saadut männyn runkoluvun (N-pine) jakaumat ja hajontakuviot maastoinventointitietoa vastaan: Maastoinventointitieto (a), aineiston D19_2 jakauma (b) ja hajontakuviot (c), MLP-estimaatin jakauma (d) ja hajontakuviot (e).



Kuva 23. Eri lähteistä saadut kuusen keskiläpimitan (D -spruce) jakaumat ja hajontakuviot maastoinventointitietoa vastaan: Maastoinventointitieto (a), aineiston $D19_2$ jakauma (b) ja hajontakuviot (c), MLP-estimaatin jakauma (d) ja hajontakuviot (e). Kuvissa b) ja c) on mukana vain testikuvio, joille $D19_2_D_spruce < 15$ cm (ks. 4.3.6).



Eri lähteistä saadut kuusen iän (T_{spruce}) jakaumat ja hajontakuviot maastoinventointitietoa vastaan: Maastoinventointitieto (a), aineiston D19_2 jakauma (b) ja hajontakuvio (c), MLP-estimaatin jakauma (d) ja hajontakuvio (e). MLP-estimointi rajattiin koskemaan kuvioita, joille $T_{\text{spruce_1}} < 16$. Kuvassa b) ja c) ovat mukana vain testikuviot, joille $D19_2_T_{\text{spruce}} < 30$ cm (ks. 4.3.6).



Eri lähteistä saadut kuusen runkoluvun (N_{spruce}) jakaumat ja hajontakuviot maastoinventointitietoa vastaan: Maastoinventointitieto (a), aineiston D19_2 jakauma (b) ja hajontakuvio (c), MLP-estimaatin jakauma (d) ja hajontakuvio (e).

LIITE D PIIRREVALINTATESTIN TULOKSET

Piirrevalintatestit

Männyn ja kuusen pituudelle tehtiin piirrevalintatestit. Näille muuttujille etsittiin parhaat selittävät piirteet kappaleessa 3.2.2 esitetyllä brute force forward selection -menetelmällä, yhden piilokerroksen MLP-neuroverkolla. Kehysohjelmalle syötettiin kaikki 26 laskettua ilmakuva- ja latvusmallipiirrettä input-muuttujiksi. Testin tuloksena saatiin taulukot parhaista piirteistä kyseisten muuttujien estimoinnissa valitulla menetelmällä (Taulukko 38.). Piirteille laskettiin niiden estimointikierrosten osuus (%) kaikista 25 kierroksesta, joissa kyseinen piirre tuli valituksi kierroksen parhaan tuloksen saavuttaneeseen malliin (sarake P(pine) ja P(spruce)).

Paras selittävä piirre molemmille kohdemuuttujille on latvusmallin kuviolle lasketun mediaaniarvo (x13_median). Myös aineiston hajontaa keskiarvoa kuvaavat piirteet olivat parhaiden piirteiden joukossa. Yllättävä havainto on myös relevanttien piirteiden kohtalaisen suuri määrä = 18 (25 iteraation mediaani). Myöhemmissä estimointitesteissä käytettiin tässä testatulle kahdelle muuttujalle löydettyjä 18 parhaan tulomuuttujan valikoimia.

Koska testin tuloksena ei saatu suppeampaa yhteistä joukkoa piirteitä, joiden olisi voitu tulkita ennustavan hankkeessa käytettyjä metsämuuttujia yleisesti, käytettiin muiden muuttujien mallinnuksessa kaikkia 26 ilmakuva- ja latvusmallipiirrettä ja esitettyä piirrevalintamenetelmää.

- Relevantit piirteet = piirteet, jotka valittu parhaaseen malliin piirrevalintakierroksella (iteraatiolla)
- Paras malli = malli, joka tuotti testijoukolle pienimmän RMSE%:n piirrevalintakierroksella

Taulukko 38. Piirrevalintatestin tulokset puustotunnuksille H_pine_1 ja H_spruce_1. Taulukossa esitetty niiden estimointikierrosten osuus (%) kaikista 25 kierroksesta, joissa kyseinen piirre tuli valituksi kierroksen parhaan tuloksen saavuttaneeseen malliin (P(pine) ja P(spruce)). Harmaalla korostetut rivit indikoivat parhaiden piirteiden lukumäärän mediaania 25 kierrokselta.

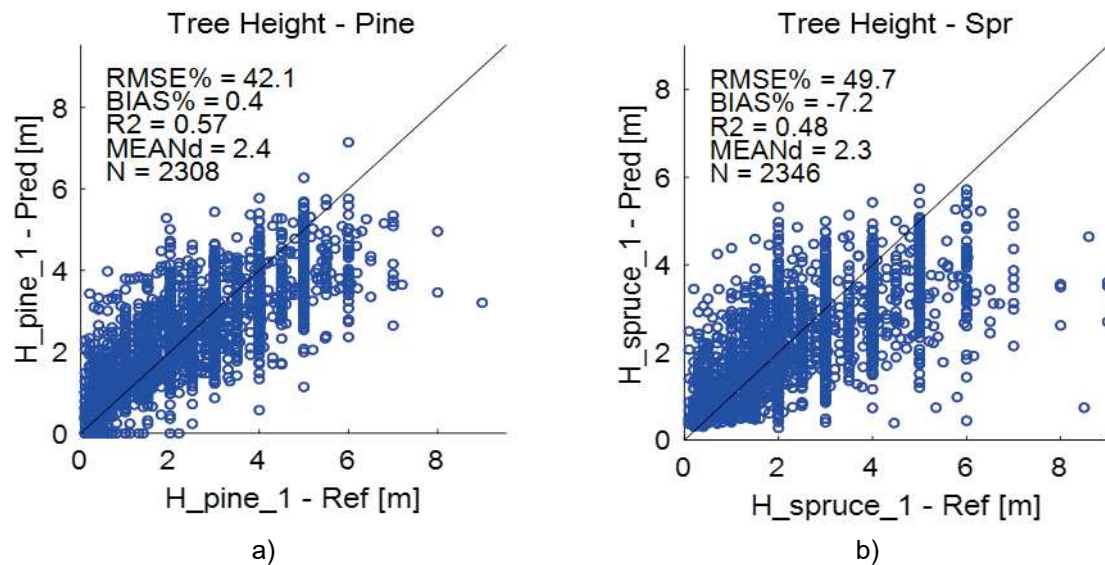
No	H_pine_1	P(pine)	H_spruce_1	P(spruce)
1	x13_median	100	x13_median	100
2	x13_std	96	x16_std_3	96
3	x16_std_3	88	x16_std_2	80
4	x13_mean	88	x16_min_3	76
5	x16_std_2	84	x16_median_1	76
6	x16_min_3	84	x16_median_2	76
7	x16_median_3	80	x16_median_3	72
8	x16_median_2	76	x16_std_1	68
9	x16_mean_3	72	x16_contr	68
10	x16_NDVI	72	x16_mean_2	64
11	x16_mean_2	68	x16_redRel	64
12	x16_redRel	68	x16_NDVI	64
13	x16_median_1	68	x16_max_2	64
14	x16_mean_1	64	x13_mean	64
15	x16_nirRel	64	x16_mean_3	60
16	x16_greenRel	60	x16_greenRel	60
17	x16_max_1	60	x16_MDM	60
18	x13_max	60	x16_min_1	60
19	x16_std_1	56	x16_max_1	60
20	x16_MDM	56	x16_max_3	60
21	x16_contr	56	x13_max	60
22	x16_max_2	56	x16_mean_1	56
23	x16_min_2	52	x16_nirRel	56
24	x16_max_3	52	x13_MDM	56
25	x13_MDM	52	x16_min_2	52
26	x16_min_1	44	x13_std	40

LIITE E TESTIT SYVILLÄ NEUROVERKOILLA

Syvien neuroverkkojen käyttö puustotunnusten estimoinnissa

Puustotunnusten mallinnuksessa testattiin myös ns. syviä neuroverkkoja (Deep Neural Network, DNN). Tavoite oli selvittää, onko käytetyssä aineistossa jotain sellaista informaatiota, jota yhden piilokerroksen neuroverkko ei pysty mallintamaan. Tässä toteutettiin estimointimallit männyn ja kuusen pituudelle käyttäen muutamaa 3 - 7 piilokerrosta sisältävää täysin kytkettyä (fully connected) syvää neuroverkkoa. Kaikkien mallien viimeisessä kerroksessa käytettiin ns. ReLU-yksikköä (ReLU = Rectified Linear Unit), joka on tyypillinen jatkuva-arvoisten muuttujien mallinnuksessa. Testit toteutettiin Keras/Tensorflow ohjelmointiympäristössä. Mallien opetusvaiheen optimointimenetelminä käytettiin Adam- ja RmsProp -algoritmeja (<https://keras.io/optimizers/#adam>, <https://keras.io/optimizers/#rmsprop>).

Tyypilliset syvillä neuroverkoilla toteutettujen männyn ja kuusen estimointimallien hajontakuviot testiaineistolle on esitetty Kuva 24. Saavutetuissa tuloksissa ei pystytty näkemään oleellista eroa yhden piilokerroksen neuroverkkomalleihin verrattuna. Tästä syystä syvien neuroverkkojen käyttö päätettiin lopettaa näiden alustavien testien jälkeen.



Kuva 24. a) Syvällä neuroverkkomallilla tuotettujen männyn (H-pine) ja b) kuusen (H-spruce) pituusestimaattien hajontakuviot testijoukolle (Ref = inventointitieto, Pred = estimaatti)

LIITE F Kokonaispuustotunnusten laskenta ja estimointi

Satakunnan taimikkokuvioille laskettiin pituuden ja rungon läpimitan kokonaispuustotunnukset (H ja D) puulajikohtaisista tunnuksista (H_species ja D_species, species = [pine, spruce, BL]) runkoluvulla (N_species) painotettuna keskiarvona. Mikäli puulajikohtainen tunnus puuttui, oli nolla tai mikäli ko. puulajin runkoluku puuttui tai oli nolla, jätettiin tämä pois kokonaistunnuksen laskennassa. Kokonaisrunkoluku (N) laskettiin kaikkien puulajien runkolukujen summana. Kuviot, joille pituus ja rungon läpimitta saatiin laskettua, muodostivat siis osajoukot kaikista satakunnan taimikkokuvioista. Runkoluku laskettiin kaikille kuvioille, mutta osalla kuvioista runkoluku N oli nolla. Taulukkoon 39 on koottu eri kokonaispuustotunnuksille toteutuneet kuviolukumäärät.

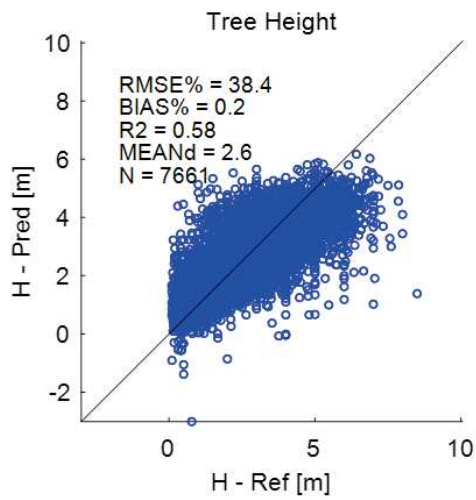
Taulukko 39. Niiden satakunnan kuvioiden lukumäärät, joille kokonaispuustotunnukset laskettiin

Puustotunnus	Kaikkien kuvioiden lukumäärä (SaKu)	T2 Kuviot, joille ilma- ja CHM-aineisto (yht. 39636 kuviota)
Pituus (H)	46214	36627
Rungon läpimitta (D)	38431	32135
Runkoluku (N)	63620 1)	39636 2)

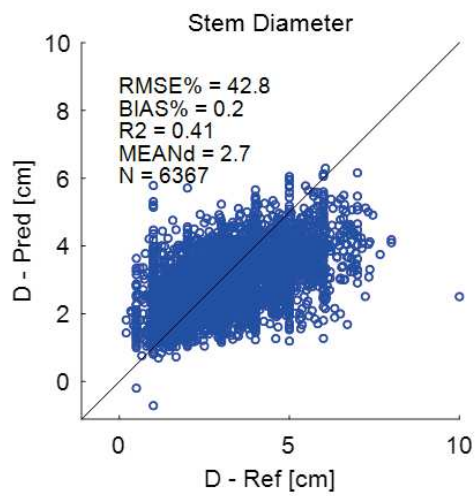
1) Näistä 16952 kuviolla N = 0

2) Näistä 2608 kuviolla N = 0

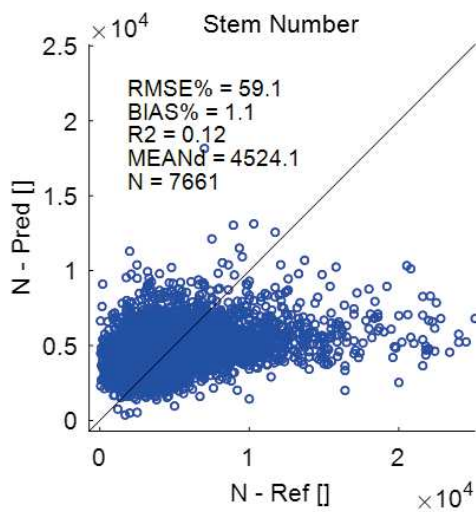
Pituuden (H), runkoläpimitan (D) ja runkoluvun (N) estimointimallit toteutettiin yhden piilokerroksen MLP neuroverkolla, kuten puulajikohtaisille tunnuksille. Tuotettujen estimaattimallien tyypilliset hajontakuviot testiaineistolle on esitetty Kuva 25. Aineisto rajattiin kuvioihin, joille pituus oli välillä 0 m < H < 10 m. Puun kokonaistunnuksen virheet ja selitysaste ovat samaa luokkaa kuin männyn ja kuusen puulajikohtaisille tunnuksille saadut. Parhaat tulokset saatiin puun pituudelle ja runkoläpimitalle. Runkoluvun testitulokset olivat odotetusti vaatimattomammat.



a)



b)



c)

Kuva 25. Testiaineiston hajontakuviot pituus- (H), keskiläpimitta- (D) ja runkolukuestimaateille (N). (Ref = inventointitieto, Pred = estimaatti)

LIITE G KÄYTETYT TERMIT, LYHENTEET JA MERKINNÄT

BIAS	Harha
BIAS%	Suhteellinen harha
CHM	Puuston latvusmalli (Canopy Height Model)
DNN	Deep neural network
H_pine_1	Puustotunnus - männyn pituus. Merkintä '_1' viittaa SMK:n kuvioaineistossa inventointipuustotunnukseen
yleisesti: X_species_k	yleisesti: $X \in [H, D, N, T]$ $k \in [1, 2]$ species $\in$ [pine, spruce] (tai [pine, spr]) Vastaavat muuttujat aineistolle D19_2: esim: D19_2_H_pine tai x19_2_H_pine (jakaumakuviissa)
IK	Ilmakuvat
Kemera	Kestävän metsätalouden rahoituslaki
K-means	Data-analyysissä ja koneoppimisessa laajalti käytetty klusterointi-algoritmi (aka migrating means). https://en.wikipedia.org/wiki/K-means_clustering
Inventointipuusto	Kuviolle tuotettu puustotieto. Tämä voi olla peräisin maastoinventoinnista tai muista lähteistä (muuttujan tunnuksessa merkintä '*_1' viittaa tähän)
Laskentapuusto	Kuviolle tuotettu puustotieto, joka on kasvatettu inventointipuustosta (muuttujan tunnuksessa merkintä '*_2' viittaa tähän)
MEANd	Referenssiaineiston keskiarvo
MKI	Metsänkäyttöilmoitus
MLP	Multi layer perceptron
N	Runkoluku (puustomuuttuja) tai näytteiden lukumäärä (hajontakuviot)
Pred	Prediction; ennuste tai estimaatti
R ²	Selitysaste
Ref	Reference; maastotieto (maastossa inventoidut kuviot)
RMSE	Keskimääräinen neliövirhe (Root Mean Squared Error)
RMSE%	Suhteellinen keskimääräinen neliövirhe (Relative Root Mean Squared Error)
TPI	Taimikon perustamisilmoitus
TPK_MHTYO	Toimenpidekuvio/metsänhoitotyö