

**Lahopuumäärän ennustaminen ja
kartoitus lentokonelaserkeilauksen
avulla**

Metsävaratieteen ja -teknologian pro gradu -
tutkielma maatalous- ja metsätieteiden
maisterin tutkintoa varten

Helsingin yliopisto, metsätieteiden laitos
huhtikuu 2011

Topi Tanhuanpää

Tiedekunta/Osasto — Fakultet/Sektion — Faculty Maatalous-metsätieteellinen tiedekunta		Laitos — Institution — Department Metsätieteiden laitos	
Tekijä — Författare — Author Topi Tanhuanpää			
Työn nimi — Arbetets titel — Title Lahopuumäärän ennustaminen ja kartoitus lentokonelaserkeilauksen avulla			
Oppiaine — Läroämne — Subject Metsävaratiede ja -teknologia			
Työn laji — Arbetets art — Level Maisterin tutkielma	Aika — Datum — Month and year Huhtikuu 2011	Sivumäärä — Sidoantal — Number of pages 47s. + liitteet 2 s.	
Tiivistelmä — Referat — Abstract Lahopuun määrästä ja sijoittumisesta ollaan kiinnostuneita paitsi elinympäristöjen monimuotoisuuden, myös ilmakehän hiilen varastoinnin kannalta. Tutkimuksen tavoitteena oli kehittää aluepohjainen laserkeilausdataa hyödyntävä malli lahoppukohteiden paikantamiseksi ja lahoppuun määrän estimoimiseksi. Samalla tutkittiin mallin selityskyvyn muuttumista mallinnettavan ruudun kokoa suurennettaessa. Tutkimusalue sijaitsi Itä-Suomessa Sonkajärvellä ja koostui pääasiassa nuorista hoidetuista talousmetsistä. Tutkimuksessa käytettiin harvapulssista laserkeilausdataa sekä kaistoittain mitattua maastodataa kuolleesta puuaineksesta. Aineisto jaettiin siten, että neljäsosa datasta oli käytössä mallinnusta varten ja loput varattiin valmiiden mallien testaamiseen. Lahoppuun mallintamisessa käytettiin sekä parametrissa että ei-parametrissa mallinnusmenetelmää. Logistisen regressioon avulla erikokoisille (0,04, 0,20, 0,32, 0,52 ja 1,00 ha) ruuduille ennustettiin todennäköisyys lahoppuun esiintymiselle. Muodostettujen mallien selittävät muuttujat valittiin 80 laserpiirteiden ja näiden muunnoksien joukosta. Mallien selittävät muuttujat valittiin kolmessa vaiheessa. Aluksi muuttujia tarkasteltiin visuaalisesti kuvaamalla ne lahoppuun määrän suhteen. Ensimmäisessä vaiheessa sopivimmiksi arvioitujen muuttujien selityskykyä testattiin mallinnuksen toisessa vaiheessa yhden muuttujan mallien avulla. Lopullisessa usean muuttujan mallissa selittävien muuttujien kriteerinä oli tilastollinen merkitsevyys 5 % riskitasolla. 0,20 hehtaarin ruutukoolle luotu malli parametrisoitiin muun kokoisille ruuduille. Logistisella regressiolla toteutetun parametrissa mallintamisen lisäksi, 0,04 ja 1,0 hehtaarin ruutukokojen aineistot luokiteltiin ei-parametrissa CART-mallinnuksen (Classification and Regression Trees) avulla. CART-menetelmällä etsittiin aineistosta vaikeasti havaittavia epälineaarisia riippuvuuksia laserpiirteiden ja lahoppuun määrän välillä. CART-luokittelu tehtiin sekä lahoppuustoisuuden että lahoppuutilavuuden suhteen. CART-luokituksella päästiin logistista regressiota parempiin tuloksiin ruutujen luokituksessa lahoppuustoisuuden suhteen. Logistisella mallilla tehty luokitus parani ruutukoon suurentuessa 0,04 ha:sta (kappa 0,19) 0,32 ha:iin asti (kappa 0,38). 0,52 ha:n ruutukoolla luokituksen kappa-arvo kääntyi laskuun (kappa 0,32) ja laski edelleen hehtaarin ruutukokoon saakka (kappa 0,26). CART-luokitus parani ruutukoon kasvaessa. Luokitustulokset olivat logistista mallinnusta parempia sekä 0,04 ha:n (kappa 0,24) että 1,0 ha:n (kappa 0,52) ruutukoolla. CART-malleilla määritettyjen ruutukohtaisten lahoppuutilavuuksien suhteellinen RMSE pieneni ruutukoon kasvaessa. 0,04 hehtaarin ruutukoolla koko aineiston lahoppuun määrän suhteellinen RMSE oli 197,1 %, kun hehtaarin ruutukoolla vastaava luku oli 120,3 %. Tämän tutkimuksen tulosten perusteella voidaan todeta, että maastossa mitatun lahoppuun määrän ja tutkimuksessa käytettyjen laserpiirteiden yhteys on pienellä ruutukoolla hyvin heikko, mutta vahvistuu hieman ruutukoon kasvaessa. Kun mallinnuksessa käytetty ruutukoko kasvaa, pienialaisten lahoppuukeskittymien havaitseminen kuitenkin vaikeutuu. Tutkimuksessa kohteen lahoppuustoisuus pystyttiin kartoittamaan kohtuullisesti suurella ruutukoolla, mutta pienialaisten kohteiden kartoittaminen ei onnistunut käytetyillä menetelmillä. Pienialaisten kohteiden paikantaminen laserkeilauksen avulla edellyttää jatkotutkimusta erityisesti tiheäpulsin laserdatan käytöstä lahoppuinventoinneissa.			
Avainsanat — Nyckelord — Keywords lahoppu, laserkeilaus, mallinnus, harvinaisten ilmiöiden kartoitus, ruutukoon vaikutus			
Säilytyspaikka — Förvaringsställe — Where deposited Viikin tiedekirjasto ja metsätieteiden laitos			
Muita tietoja — Övriga uppgifter — Further information			

Tiedekunta/Osasto — Fakultet/Sektion — Faculty Faculty of Agriculture and Forestry		Laitos — Institution — Department Department of Forest Sciences	
Tekijä — Författare — Author Topi Tanhuanpää			
Työn nimi — Arbetets titel — Title Prediction and mapping of coarse woody debris with airborne laser scanning			
Oppiaine — Läroämne — Subject Forest Resource Science and Technology			
Työn laji — Arbetets art — Level Master's thesis		Aika — Datum — Month and year April 2011	Sivumäärä — Sidoantal — Number of pages 47s. + appendices 2 s.
Tiivistelmä — Referat — Abstract <p>There is an ever growing interest in coarse woody debris (CWD). This is because of its role in maintaining biodiversity and storing atmospheric carbon. The aim of this study was to create an ALS-data utilizing model for mapping CWD and estimating its volume. The effect of grid cell size change to the model's performance was also considered.</p> <p>The study area is located in Sonkajärvi in eastern Finland and it consisted mostly of young commercially managed forests. The study utilized low-frequency ALS-data and precise strip-wise field inventory of CWD. The data was divided into two parts: one fourth of the data was used for modeling and the remaining three fourths for validating the models that were constructed.</p> <p>Both parametric and non-parametric modelling practices were used for modelling the area's CWD. Logistic regression was used to predict the probability of encountering CWD in grid cells of different sizes (0.04, 0.20, 0.32, 0.52 and 1.00 ha). The explanatory variables were chosen among 80 ALS-based variables and their conversions in three stages. Firstly, the variables were plotted against CWD volumes. Secondly, the best variables plotted in the first stage were examined in single-variable models. Thirdly, variables to the final multivariable model were chosen using 95 % level of significance. The 0.20 ha model was parametrized to other grid cell sizes. In addition to parametric model constructed with logistic regression, 0.04 ha and 1.0 ha grid cells were also classified with CART-modelling (Classification and Regression Trees). With CART-modelling, non-linear dependencies were sought between ALS-variables and CWD. CART-models were constructed for both CWD existence and volume.</p> <p>When the existence of CWD in the study grid cells was considered, CART-modelling resulted in better classification than logistic regression. With logistic model the goodness of classification was improved as the grid cell size grew from 0.04 ha (kappa 0.19) to 0.32 ha (kappa 0.38). On 0.52 ha cell size, kappa value of the classification started to diminish (kappa 0.32) and was furthermore diminished to 1.0 ha cell size (kappa 0.26). The CART classification improved as the cell size grew larger. The results of CART-modelling were better than those of the logistic model in both 0.04 ha (kappa 0.24) and 1.0 ha (kappa 0.52) cell sizes. The relative RMSE of the cellwise CWD volume predicted with CART-models diminished as the cell size was enlarged. On 0.04 ha grid cell size the RMSE of the total CWD volume of the study area was 197.1 % and it diminished to 120.3 % as the grid cell size was enlarged to 1.0 ha.</p> <p>On the grounds of the results of this study it can be stated that the link between CWD and ALS-variables is weak but becomes slightly stronger when cell size increases. However, when cell size increases, small-scale variation of CWD becomes more difficult to spot. In this study, the existence of CWD could be estimated somewhat accurately, but the mapping of small-scale patterns was not successful with the methods that were used. Accurate locating of small-scale CWD variation requires further research, particularly on the use of high density ALS-data in CWD inventories.</p>			
Avainsanat — Nyckelord — Keywords coarse woody debris, laser scannig, modelling, inventory of rare phenomena, effect of cell size			
Säilytyspaikka — Förvaringsställe — Where deposited Viikki Science Library and Department of Forest Sciences			
Muita tietoja — Övriga uppgifter — Further information			

ALKUSANAT

Tämä tutkielma on tehty Helsingin yliopiston metsätieteiden laitokselle osana Annika Kankaan johtamaa *Mapping rare forest characteristics using remote sensing material and field data*-tutkimushanketta. Haluan kiittää Annika Kangasta pitkäjänteisestä ja kannustavasta ohjauksesta. Lisäksi kumarran syvään Mikko Vastarannalle ja Ilona Pietilälle, joiden neuvot olivat suurena apuna työn valmiiksi saattamisessa.

Helsingissä huhtikuussa 2011

Topi Tanhuanpää

Sisällys

1. Johdanto	1
1.1 Lahopuu tutkimuskohteena.....	1
1.2 Lahopuumäärän mittaaminen	3
1.3 Kuolleen puuston mallintaminen	5
1.4 Laserkeilaus metsäsuunnittelun apuvälineenä	7
1.5 Ruutukoon vaikutus harvinaisten ilmiöiden tutkimuksessa	10
1.6 Tutkimuksen tavoitteet	11
2. Aineisto	13
2.1 Maastodata.....	13
2.2 ALS-data.....	14
3. Menetelmät	16
3.1 Datan käsittely	16
3.2 Logistinen regressio	17
3.3 Luokituksen hyvyyden arviointi	21
3.4 CART-mallinnus	22
3.5 Lahopuuestimaattien luotettavuuden arviointi.....	24
4. Tulokset.....	25
4.1 Logistiset lahopuumallit	25
4.2. Aineiston luokittelu logistisilla malleilla.....	27
4.3 CART: luokitus	28
4.4 CART: tilavuusestimaatit	29
5. Tulosten tarkastelu	31
5.1 Logistisen mallin parametrit	31
5.2 Luokitus logistisilla malleilla.....	32
5.3 Luokitus CART-malleilla	35
5.4 CART-tilavuusestimaatit	37
5.5 Ruutukoon vaikutus lahopuun inventoinnissa	38
5.6 Laserkeilauksen mahdollisuudet lahopuun inventoinnissa	40
6. Johtopäätökset.....	42
Lähteet.....	43
Liitteet	48

1. Johdanto

1.1 Lahopuu tutkimuskohteena

Lahopuun määrästä ollaan kiinnostuneita sekä luonnon monimuotoisuuden (esim. Harmon 1986, Siitonen 2001) että nykyään yhä enenevässä määrin myös siihen varastoituneen hiilen kannalta (esim. Gough ym. 2007). Pystyyn kuolleet ja maahan kaatuneet puut ovat välttämättömiä monien lajien elinolosuhteiden kannalta, pesäpuina tai ravintona. Samaan aikaan lahopuu toimii myös hiilivarastona, jossa ilmasta orgaaniseen ainekseen sidottu hiilidioksidi säilyy kymmeniä vuosia kiinteässä muodossa, vapautuen hiljalleen takaisin ilmakehään hajotustoiminnan seurauksena. Sekä monimuotoisuuden ylläpitäminen että hiilen varastointi koetaan tärkeinä ja tavoiteltavina toimintoina. Metsäluonnon lajirunsaus on arvo sinänsä, mutta ääritapauksissa myös metsän elinvoimaisuuden ja terveyden kulmakivi. Ilmakehän hiilen varastointi ja kierto taas vaikuttavat suoraan ilmastoon ja sitä kautta ihmisen elinolosuhteisiin.

Lahopuun määrän vaikutus Suomen metsäluontoon on huomattava, sillä noin neljäsosa maan eliölajeista on suoraan riippuvaisia lahopuun määrästä (Siitonen 2001). Pelkkä lahopuun määrä ei kuitenkaan ole kattava indikaattori kaikkien lahopuuta tarvitsevien lajien runsaudelle. Pelkkää lajirunsausta mitattaessa kuolleiden runkojen lukumäärä on merkittävin yksittäinen tekijä, kun taas vaateliaimpien lajien määrään vaikuttaa eniten eri lahoasteiden esiintyminen (Stockland 2001).

Luonnontilaisten metsien ja talousmetsien ero lahopuun määrän suhteen on merkittävä. Luonnontilaisissa tai luonnontilaisen kaltaisissa pohjoisissa havumetsissä lahopuun määrä on tutkimusten mukaan keskimäärin noin 40–70 kuutiometriä hehtaarilla (Linder ym. 1997, Karjalainen ja Kuuluvainen 2002). Valtakunnan metsien kymmenennen inventoinnin (VMI10) tulosten perusteella (Ihalainen ja Mäkelä 2009) koko Suomen metsien keskimääräinen lahopuun määrä oli 5,4 kuutiometriä hehtaarilla. Puuntuotannossa olevissa metsissä lahopuun määrä oli keskimäärin 4,7 ja suojelluissa metsissä 12,7 kuutiometriä hehtaarilla. Suuret erot luonnontilaisten, suojeltujen ja talouskäytössä olevien metsien lahopuun määrissä selittyvät metsien käsittelyn ja käsittelyhistorian kautta. Luonnontilaiseksi tai luonnontilaisen kaltaiseksi voidaan

luokitella metsikkö, jossa ei ole merkkejä ihmisen toiminnasta. Tällaisilla kohteilla kaikki kuollut puuaines on jäänyt metsään jo vähintään satojen vuosien ajan, jolloin metsän lahoppuvaranto on kasvanut ”luonnolliselle” tasolle. Suojeltujen metsien lähihistoria voi olla hyvin moninainen, koska suojelun perusteet vaihtelevat. Suojeltu alue voi olla vanhaa metsää, jonka piirteet ovat lähellä luonnontilaista tai sen kaltaista metsää. Toisaalta suojelun perusteena saattaa olla esimerkiksi maisemalliset syyt, jolloin suojeltu metsikkö voi olla myös talouskäytössä. Tällöin lahoppumäärä on todennäköisesti huomattavasti luonnontilaisia metsiköitä alempi. Pienimmät keskimääräiset lahoppumäärät tavataan ymmärrettävästi talousmetsissä, joissa lähes kaikki puuaines kerätään talteen. Näissä metsissä kuollut puusto koostuu lähinnä yksittäisistä pystyyn kuolleista puista, tuulenskaadoista tai hakkuun yhteydessä metsään jääneistä rungoista.

Lahoppuun määrän lisäksi myös sen spatiaalinen jakaantuminen eroaa voimakkaasti talousmetsien ja luonnontilaisten metsien välillä. Molemmissa tapauksissa lahoppuun esiintyminen on satunnaista, mutta talousmetsissä lahoppumäärän hajonta kasvaa huomattavasti (Gibb ym. 2005). Luonnontilaisista metsistä poiketen kuollutta puustoa havaitaan talousmetsissä lähinnä pistemäisinä esiintyminä. Tämä johtuu pääosin siitä, että luonnonpoistumaa vastaava määrä puuta kerätään metsästä harvennushakkuilla. Toisaalta myös lainsäädäntö vaikuttaa lahoppuun esiintymiseen. Laki metsän hyönteis- ja sienituhojen torjunnasta (1991) velvoittaa metsänomistajan huolehtimaan myrskyn metsäpalon tai lumen vaurioittamien puiden poistosta, mikäli vahingoittuneita havupuita on enemmän kuin 10 % hehtaarikohtaisesta runkoluvusta. Lain mukaan myös kaikki yli kahdenkymmenen vahingoittuneen havupuun ryhmät tulee korjata metsästä. Tuulelle alttiilla alueilla pistemäiset esiintymät voivat toisaalta talousmetsissäkin olla tilavuudeltaan hyvin suuria. Mikäli esimerkiksi tuulituho tapahtuu asteittain monen vuoden aikana ja kerralla kuolleen puuston määrä ei ole merkittävä, edellä mainittu laki ei velvoita kuolleen puuston korjaamiseen pois metsästä. Koska lahoppuuta esiintyy talousmetsissä niukasti luonnontilaisiin metsiin verrattuna, yksittäisten runsalahoppuustoisten kohteiden arvo luonnon monimuotoisuuden kannalta on suuri. Lahoppuusta riippuvaisten vaatelioiden lajien elinolosuhteita kehitettäessä tulisikin keskittyä näiden lahoppukeskittymien säilyttämiseen ja kehittämiseen sen sijaan, että kaikkialle pyrittäisiin saamaan hieman enemmän lahoppuuta (Punttila 2000). Lahoppuun

tarkka inventointi muiden metsikkötunnusten tapaan edesauttaisi tärkeimpien kohteiden säilymistä ja mahdollista kehittämistä. Paikannetut kohteet voitaisiin ottaa paremmin huomioon esimerkiksi suojelualueiden suunnittelussa tai operatiivisessa hakkuiden suunnittelussa.

1.2 Lahopuumäärän mittaaminen

Lahopuun määrää koskevaa tietoa, kuten muitakin metsäaluetta kuvaavia tunnuksia, voidaan kerätä kahdella tavalla. Mittaukset voidaan joko tehdä kattavasti koko alueelle tai otokselle, joka edustaa mitattavaa aluetta. Kattavan eli niin sanotun wall-to-wall-mittaamisen etuna on, että tutkittavalle alueelle ei jää tuntemattomia kohteita, vaan koko alueelta saadaan suoria havaintoja. Yksinkertaisimmassa muodossaan kattava mittaaminen tarkoittaa metsäalueen jokaisen neliömetrin mittaamista haluttujen tunnusten tai ominaisuuksien suhteen. Korkeiden kustannustensa johdosta tällainen mittaaminen on kuitenkin mahdotonta muussa kuin tutkimustarkoituksessa. Myös kuvioittainen arviointi voidaan kuitenkin luokitella kattavaksi mittaamiseksi. Arviointi perustuu homogeenisten metsikkökuvioiden silmänvaraiseen määrittämiseen ja riittävien mittausten tekemiseen näiden kuvioiden sisällä (Koivuniemi ja Korhonen 2006). Menetelmän subjektiivisuudesta johtuen suuri osa mittauksen tarkkuudesta riippuukin mittaajan ammattitaidosta (Haara ja Korhonen 2004). Kuvioittaisella arvioinnilla saadaan kuitenkin kattavaa tietoa kuviotasolla, sillä arvioitsija havainnoi silmämääräisesti koko kuvion ominaisuudet ja tekee tarpeelliseksi kokiessaan täydentäviä mittauksia, jotta kuviotiedot kuvaisivat todellisuutta mahdollisimman hyvin. Kuviotietoa ei kuitenkaan voi soveltaa määritettyä kuviota pienemmälle alueelle, joten mitattu tieto on käyttökelpoista vain kuvion koon asettamaan alarajaan saakka.

Edustavan otoksen keräämiseen perustuva mittaaminen on yleensä huomattavasti kattavaa mittaamista tehokkaampaa: samoilla kustannuksilla voidaan mitata huomattavasti laajempia alueita. Otantaan perustuvia mittausmenetelmiä käytetäänkin erityisesti suurien metsäalueiden mittaamisessa. Esimerkiksi Valtakunnan metsien inventoinnin (VMI) maastotieto kerätään koko maan kattavalla koealaverkolla (Tomppo 1998a). Otantapohjaiset menetelmät soveltuvat suuralueille erityisesti niiden joustavuuden takia: Koealaverkon tiheyttä voidaan muuttaa tarkkuus- ja kustannusvaatimusten mukaan. Tällöin suuralueille voidaan määritellä koealaverkko, josta kerätty maastotieto on mitattavien ominaisuuksien osalta luotettavuudeltaan

halutulla tasolla. Koealoilta mitattu tieto on kuitenkin luonteeltaan pistemäistä, eikä mittausyksikön ulkopuolelta saada mitään tietoa. Tällöin mitatulta alueelta ei saada wall-to-wall-tietoa vaan otoksen ulkopuoliset kohteet oletetaan keskimäärin otoksen kaltaisiksi. Alueiden keskitunnuksia estimoitaessa tällä ei juuri ole merkitystä, mutta tilanne on ongelmallinen, mikäli tutkittava ilmiö halutaan paikantaa mitatun datan perusteella esimerkiksi teemakarttojen muodossa. Suurimmasta osasta tutkittua aluetta on tällöin käytettävissä ainoastaan otoksen perusteella estimoitua tietoa, jota ei sinänsä ole mielekästä paikantaa, koska se ei kuvaa todellista tilannetta millään yksittäisellä kohteella. Ainoastaan otokseen kuuluneiden koealojen todelliset tiedot tunnetaan ja niinpä ainoastaan koealatietojen paikantaminen on mielekästä. VMI:n tapauksessa koealojen väliin jäävän tuntemattoman alueen ongelma on ratkaistu satelliittikuvien hyödyntämiseen perustuvalla mallinnuksella, jolla koealoilta mitatut metsää kuvaavat tunnuksat yleistetään maasto-otoksen ulkopuolelle jääneille alueille. Tällöin tutkittavalle alueelle ei jää lainkaan täysin tuntemattomia alueita, koska satelliittikuvat kattavat koko tutkittavan alueen.

Pienialaisuus ja harvinaisuus asettavat omat rajoitteensa menetelmille, joilla lahopuun määrää mitataan. Puuntuotannollisia tunnuksia mittaavat menetelmät on kehitetty verrattain yleisten ilmiöiden määrittämiseen. Tästä syystä perinteiset metsän mittauksen menetelmät eivät toimi tehokkaasti, kun tarkasteltavan ilmiön esiintyminen on lahopuun tapaan harvaa ja epäsäännöllistä (Ståhl 1998, Gove ym. 2002). Periaatteessa elävän puuston tilavuuden mittausmenetelmät toimivat myös lahopuulle. Ongelman aiheuttaa lahopuusta saatujen havaintojen pieni määrä, joka johtuu ilmiön harvinaisuudesta. Koska havaintoja saadaan vähemmän kuin elävästä puustosta, myös estimaattien tarkkuus on tiettyä pinta-alaa tarkasteltaessa elävän puuston estimaatteja huonompi (Kangas 2006). Tarkasteltavan pinta-alan kasvaessa estimaattien tarkkuus paranee ja suuraluetasolla tuloksia voidaankin pitää luotettavina.

Tämänhetkiset koko Suomea koskevat lahopuumäärän estimaatit lasketaan VMI:n yhteydessä, joten ne perustuvat koko maan kattavaan koealaverkostoon. Lahopuuestimaatit esitetään huomattavasti elävän puuston estimaatteja suuremmille alueille. Elävän puuston tunnuksat voidaan VMI:n maastokoealojen perusteella laskea luotettavasti esimerkiksi metsäkeskustasolla (Tomppo 1998a) kun taas lahopuustoa

kuvaavat estimaatit on jaettu alueellisesti varsin karkeasti Etelä- ja Pohjois-Suomeen (esim. Ihalainen ja Mäkelä 2009). Suuresta pinta-alasta johtuen näitä estimaatteja voidaan pitää lahopuun kokonaismäärän osalta luotettavina vaikka mittausten menetelmä onkin suunniteltu elävän puuston tehokkaaseen mittaamiseen. Periaatteessa lahopuusto voitaisiin yleistää koealaverkon väliin jäävälle alueelle elävän puuston tunnusten tapaan, jolloin saataisiin koko maan kattavaa wall-to-wall-tietoa lahopuun esiintymisestä. Oletettavaa kuitenkin on, että lahopuun harva ja epäsäännöllinen esiintyminen tekee yleistämisestä huomattavasti elävän puuston yleistämistä epävarmempaa. Jotta otantapohjaisella mittausten menetelmällä päästäisiin samalle tarkkuustasolle harvinaisen ja yleisemmän ilmiön mittaamisessa, harvinaisen ilmiön mittaamiseen jouduttaisiin osoittamaan enemmän mittauksia.

1.3 Kuolleen puuston mallintaminen

Metsäsuunnittelun kannalta tulevaisuuden lahopuun määrän ennustaminen on yhtäläillä tärkeää kuin metsässä jo olevan lahopuun määrän estimointi. Mikäli pyritään mallintamaan metsikköön syntyvää lahopuuta, on olennaista kuinka paljon metsässä kuolee puuta ja kuinka kauan kuollut puu säilyy metsässä. Eid ja Øyen (2003) mallinsivat puuston kuolleisuutta tasaikäisissä metsiköissä. Tutkimuksessa luotiin mallit puuston selviytymistodennäköisyydelle sekä runkoluvun pienenemiselle. Selviytymistä selitettiin hehtaariohtaisella runkoluvulla, kasvupaikalla, metsikön iällä sekä havupuiden osuudella pohjapinta-alasta. Runkoluvun pienenemistä kuvattiin erikseen vallitsevan puulajin mukaan männiköille, kuusikoille ja lehtipuun metsille. Mallissa selittäviä muuttujia olivat runkoluku, kasvupaikka ja puulajin osuus pohjapinta-alasta. Ainoastaan mäntyvaltaisen metsikön mallissa käytettiin kaikkia kolmea muuttujaa. Kuusikossa kuusen osuus pohjapinta-alasta ja koivikossa kasvupaikka eivät olleet tilastollisesti merkitseviä selittäjiä. Næsset (1999) puolestaan mallinsi toista keskeistä lahopuun määrään vaikuttavaa tekijää eli kuolleen puun hajoamisnopeutta. Tutkimuksessa muodostettiin lineaarinen malli, joka kuvasi kuolleen puun tiheyttä ajan funktiona. Tiheyden muutosta selitettiin kuolleen rungon ja sen hajoamisympäristön ominaisuuksilla (ks. Næsset 1999). Estimaatit kuolevan puuston määrästä ja hajoamisajasta mahdollistavat lahopuun määrän kehityksen mallintamisen. Ranius ym. (2004) mallinsivat tutkimuksessaan kuolleen puun keskimääräistä tilavuutta pitkällä aikavälillä pohjoisissa vanhoissa kuusivaltaisissa havumetsissä. Malli perustui oletuksille lahopuun hajoamiseen kuluva ajasta ja kuolevien puun koosta suhteessa

muuhun puustoon eli tutkimuksessa oli mukana sekä hajoamis- että kilpailuelementti. Metsikkötunnuksista mallissa olivat mukana elävän puuston tilavuus sekä metsikön vuotuinen kasvu. Mallin avulla estimoitiin keskimääräistä maisematason lahopuumäärää luonnontilaisessa metsässä.

Mallintamisen päämääränä voi olla myös metsässä jo olevan kuolleen puuston kuvaaminen. Tällaisia malleja voidaan käyttää esimerkiksi aluetason lahopuumäärien (Ihalainen ja Mäkelä 2009) tai muiden monimuotoisuusindikaattorien kuten kolopuiden lukumäärän estimoinnissa (Fan ym. 2003). Metsikön lahopuumäärän ja kuolleen puuston nykytilaa on pyritty mallintamaan sekä maastossa mitattujen metsikkötunnusten että kaukokartoituksen avulla saatujen tunnusten avulla. Fan ym. (2003) käyttivät metsikkötunnuksia mallintaessaan kolopuiden esiintymistä 0,4 hehtaarin koealoille. Metsikön iän ja pohjapinta-alan avulla mallinnettiin todennäköisyyttä kolopuiden hehtaarikohtaiselle esiintymiselle. Mallit soveltuivat erityisesti suuraluetason estimaattien laadintaan metsänhoitotoimenpiteiden vaikutuksesta kolopuiden määrään.

Monissa tutkimuksissa on myös selvitetty mahdollisuutta mallintaa kuolleen puuston nykytilaa eri kaukokartoitusmenetelmien avulla. Bater ym. (2009) ennustivat tutkimuksessaan pystypuiden laatua Lidar-datan (Light detection and ranging) avulla. Puut jaoteltiin yhdeksään luokkaan (1-9), joista seitsemään viimeiseen luokkaan kuuluneet puut olivat kuollutta puustoa. Luokan järjestysnumeron kasvaessa runkojen karsiintuneisuus ja lahoaste kasvoivat. Ennustaminen tapahtui regressiomallien avulla siten, että tietyn puuluokan malli käsitti myös kaikki edellisiin puuluokkiin kuuluneet rungot. Tutkimuksessa käytetyistä selittäjistä kasvillisuudesta heijastuneiden laserpulssien varianssin luonnollinen logaritmi selitti kuolleiden puiden luokkaa parhaiten. Kun kuolleiden runkojen osuutta runkojen kokonaismäärästä ennustettiin 25 x 25 metrin ruuduille, osuuden keskineliövirheen neliöjuuri (root mean square error, RMSE) oli 16,8 %. Pasher ja King (2009) ennustivat kuollutta puustoa kaukokartoituksen avulla. Tutkimuksessa oli käytössä korkean resoluution ilmakuvaineisto, jolla lahopuuta havainnoitiin suoraan ja regressiomallin kautta. Eri aallonpituusalueita hyödyntävällä suoralla havainnoinnilla löydettiin ja paikannettiin 84 % testiaineiston latvuston läpi erottuviksi luokitelluista kuolleista puista.

Regressiomalleihin perustuvalla lähestymistavalla pyrittiin kuolleiden puiden paikantamisen sijaan mallintamaan alueen kokonaislahopuumäärää. Testiaineistossa lahopuun kokonaismäärään sisällytettiin myös ne latvuston peittoon jääneet puut, jotka oli luokiteltu latvuston läpi erottumattomiksi ja joita ei näin ollen yritettykään paikantaa. Mallinnuksen tarkkuus jäi kuitenkin huomattavasti kuolleiden runkojen määrän estimaattia heikommaksi. Kuolleen puuston tilavuusestimaatin suhteellinen RMSE oli 62,1 %. Myös edellä mainitut VMI:n kuollutta puustoa koskevat estimaatit, samoin kuin muut VMI:n tunnuksot, voidaan yleistää koealaverkon väliin jäävälle alueelle mallintamisen avulla. Tässä monilähdeinventoinniksi kutsutussa menetelmässä maastokoealoilta mitatut tunnuksot yleistetään tarkasteltavalle alueelle k-lähimmän naapurin menetelmällä (k-*nn*). Menetelmä hyödyntää satelliittikuvia, joiden avulla mitatut tunnuksot yleistetään suuremmalle alueelle (Tomppo 1998b). Jokaiselle satelliittikuvan pikselille, jolla ei ole maastossa mitattuja tunnuksia, johdetaan tunnuksot muutaman sävyarvoiltaan samankaltaisimman mitatun pikselin avulla. Tunnuksot lasketaan mitattujen koealojen keskiarvoina.

1.4 Laserkeilaus metsäsuunnittelun apuvälineenä

Lentokoneesta tehtävä laserkeilaus (ALS, Airborne laser scanning) on aktiivinen kaukokartoitusmenetelmä, joka hyödyntää aallonpituudeltaan noin yhden mikrometrin laservaloa. Menetelmä perustuu tarkkaan välimatkan mittaamiseen laserin lähtöpisteen ja heijastumiskohteen välillä (Wehr ja Lohr 1999). Laserkeilaimesta lähetetään laserpulssi havainnoitavaan kohteeseen, josta pulssi heijastuu yhtenä tai useampana laserkaikuna. Laserpulssin kulkema matka mitataan sen ajan avulla, joka laservalolta on kulunut edestakaiseen matkaan laserkeilaimen ja heijastuksen aiheuttaneen kohteen välillä. Kun etäisyyksiä mittaaviin laserpulsseihin lisätään tarkka GPS- (Global positioning system) paikannus sekä laserkeilaimen liikettä ja asentoa kuvaava IMU- (Inertial measurement unit) tieto, voidaan jokaiselle heijastuneelle laserpulssille määrittää tarkka kolmiulotteinen sijainti (Lim 2003).

Metsäalueesta mitatun lasertiedon tulkinta perustuu laserkaikujen heijastumiskorkeuksiin. Osa yksittäisestä laserpulssista heijastuu latvuksen ylimmistä osista kun taas osa voi kulkea latvuksen aukoista aina maanpinnalle saakka. Yhdestä lähetetystä laserpulssista voidaan tallentaa useita paluukaikujen heijastumiskorkeuksia (Næsset 2004). Keilattavien kohteiden korkeuden määrittämisen osalta on olennaista, että

maanpinnan korkeus, johon ylempää heijastuneiden pulssien korkeuksia verrataan, on määriteltä tarkasti. Maanpinnan korkeusmalli (DTM, Digital Terrain Model) muodostetaan alimpien heijastuneiden kaikujen avulla, joiden oletetaan heijastuneen maanpinnasta (Axelsson 2000). Heijastuneiden kaikujen maanpintaan suhteutetut heijastumiskorkeudet lasketaan vähentämällä DTM:n tuottama maanpinnan korkeus havaintokorkeudesta.

Laserkeilauksella saavutetaan monia etuja perinteisiin mittausmenetelmiin verrattuna. Suurien alueiden mittauksessa käytettäviin, otantaan perustuviin menetelmiin verrattuna ALS-data on kattavampaa. Lentokoneesta tapahtuvan mittauksen avulla saadaan yksittäisten koalojen sijaan havaintoja koko tutkittavalta metsäalueelta (wall-to-wall), jolloin esimerkiksi mitattavien ominaisuuksien pienialaisen vaihtelun havaitseminen on mahdollista. Keilauksen spatiaalinen tarkkuus määräytyy keilauksen pulssitiheyden (pulssia/m²) perusteella. Pulssitiheyteen vaikuttavat keilaimen pulssitaajuus ja avauskulma sekä ilma-aluksen nopeus ja lentokorkeus. Käytännössä pulssitiheys tarkoittaa otantatiheyttä mitattavasta kohteesta.

Metsikkötason tunnusten estimointiin on perinteisesti käytetty kuvioittaista arviointia. Laserkeilauksen avulla on kuitenkin saavutettu kuvioittaista arviointia tarkempia estimaatteja esimerkiksi elävää puustoa kuvaavien puustotunnusten osalta (Næsset 1997, Næsset 2002, Maltamo ym. 2006, Uuttera ym. 2006, Packalen ym. 2007, Holopainen ym. 2010). Laserkeilaus on metsänmittausmenetelmänä kattavan tiedon ja luotettavien estimaattien lisäksi kilpailukykyinen myös kustannustensa suhteen (Holopainen ja Talvitie 2006, Uuttera ym. 2006). Kuvioittaiseen arviointiin verrattuna laserkeilausmenetelmässä tarvitaan vähemmän kalliita maastomittauksia.

Laserkeilauksen hyödyntämistä metsävarojen inventoinnin tarkoituksiin on tutkittu paljon (esim. Næsset 2002, Packalen ym. 2007, Vehmas ym. 2009, Holopainen ym. 2010, Yu ym. 2010, Korpela ym. 2010, Vastaranta ym. 2011). Laserdatan tulkinnessa on käytössä kaksi erillistä päämenetelmää: yleensä tiheäpulsiselle (yli 5 pulssia/m²) datalle tehtävä yksinpuin-tulkinta (ITD, Individual tree detection) (Hyypä ja Inkinen 1999), sekä tavallisesti harvapulsiselle datalle (alle 1 pulssi/m²) tehtävä aluepohjainen tulkinta (ABA, Area-based approach) (Næsset 2002). ITD:ssä jokaisesta metsäalueen

puusta pyritään saamaan riittävästi havaintoja puutunnusten estimoimiseksi. Menetelmä perustuu yksittäisten puiden pituuksien, latvusten koon ja sijaintien määrittämiseen laserdatan avulla (Hyypä ja Inkinen 1999). ABA-menetelmässä heijastuneista pulsseista lasketaan tilastollisia tunnuksia valitun kokoisille hilaruuduille tai segmenteille, jotka toimivat otosyksikköinä. Lasketut tunnuksat voivat kuvata esimerkiksi korkeutta, jolla heijastuneiden pulssien kumulatiivinen kertymä on halutun suuruinen, pulssien kumulatiivista kertymää halutulla korkeudella tai laserpulssien intensiteettiä tietyllä korkeudella (Næsset 2002). Menetelmän toisessa vaiheessa tarkka maastokoealoilta mitattu puustotieto yleistetään koko tutkittavalle alueelle laskettujen lasertunnusten avulla. Yleistämisessä käytetään joko ei-parametrisia menetelmiä (Maltamo ym. 2006) tai puustotunnuksia estimoivia regressiomalleja (Næsset 2002, 2004). ABA-menetelmän avulla on päästy elävää puustoa kuvaavien tunnusten estimoinnissa kuvioittaista arviointia tarkempiin tuloksiin jo harvapulssisen ALS-datan avulla (Næsset 2004, Suvanto ym. 2005).

Lahopuumäärän ennustamiseksi yksittäisten lahopuiden tarkkuudella lahopuista pitäisi saada suoria havaintoja latvuston yläpuolelta tehtävän laserkeilauksen avulla. Palaneita metsiä lukuun ottamatta suurin osa kuolleesta puustosta on kuitenkin maassa (Uotila ym. 2001, Ihalainen ja Mäkelä 2009), jolloin vallitseva latvuskerros ja aluskasvillisuus vaikeuttavat kuolleiden puiden havaitsemista huomattavasti. Vehmaan ym. (2009) tutkimuksessa määritettiin lahopuukeskittymien esiintymistä elävän puuston latvustossa havaittavissa aukoissa ALS-datan avulla. Tutkimuksessa käytetyn datan pulssitiheys oli 3,9 pulssia/m², joka riitti latvuston korkeuden pienialaisen vaihtelun mallintamiseen. Lahopuukeskittymien löytämiseksi tutkimuksessa oletettiin, että suuren puun kuoleminen aiheuttaa aukon elävään latvukseen. Oletuksen perusteella lahopuuston etsiminen keskitettiin korkeusmallissa havaittuihin latvusaukkoihin. Aukkojen ulkopuolelle jäävään lahopuun määrää ei tutkittu. Tutkimuksessa tarkasteltiin erityyppisistä kaiuista (first of many, last of many ja only) laskettujen ALS-piirteiden kykyä selittää latvusaukkojen jakautumista kolmeen luokkaan: lahopuustoinen (DDW), kasvillisuuden peittämä (VEGE) ja tyhjä (CLEAR), jolla ei havaittu kumpaakaan edellisistä. CLEAR-luokkaan kuuluvat aukot pystyttiin erottamaan kahdesta muusta luokasta kaikkien muiden paitsi first of many-tyyppisten kaikujen perusteella. Luokkien

DDW ja VEGE väillä havaittiin tilastollisesti merkitseviä eroja kaikilla kaikutyypeillä, kun kaikujen kumulatiivinen kertymä oli yli 70 %.

Lahopuun määrää voidaan ennustaa ALS-datasta myös ABA-menetelmällä elävän puuston tavoin. Pesonen ym. (2008) mallinsivat suojelualan lahopuumäärää regressiomallien avulla. Tutkimuksessa pysty- ja maalahopuulle sovitettiin omat ALS-datasta laskettuja piirteitä tai niiden muunnoksia hyödyntävät tilavuusmallit, joilla ennustettiin lahopuun määrää koelatasolla. ALS-piirteet laskettiin erikseen ensimmäisille ja viimeisille heijastuneille kiauille. Lopullisissa malleissa käytettiin sekä ensimmäisistä että viimeisistä kiauista laskettuja ALS-piirteitä. Tutkimuksessa maalahopuun määrä pystyttiin estimoimaan pystylahopuun määrää tarkemmin (tilavuuden RMSE maalahopuulla 51,6 % ja pystylahopuulla 78,8 %).

1.5 Ruutukoon vaikutus harvinaisten ilmiöiden tutkimuksessa

Hilamuotoisessa aineistossa hilaruutujen koko määrittää spatiaalisen tarkkuustason, johon esimerkiksi aineistosta luoduilla malleilla on mahdollista päästä. Mikäli esimerkiksi lahopuuinto on tehty hehtaarin ruuduille, voidaan ruudulla todeta olevan keskimäärin tietty määrä lahopuuta. Lahopuumäärän jakautumisesta ruudun sisällä ei saada mitään tietoa. Tällöin tutkittavan ilmiön pienialaista, hehtaarin ruutukoon tapauksessa alle hehtaarin vaihtelua ei pystytä havaitsemaan. Ruutukoon suurentuessa tämä ilmiöiden pienialaisen vaihtelun paikannustarkkuus luonnollisesti heikkenee entisestään.

Ruutukokoa suurennettaessa suurempien ruutujen tunnusten arvot lasketaan niiden pienempien ruutujen keskiarvona, joista suurempi ruutu koostuu. Tällöin suurempiin ruutukokoihin siirryttäessä erittäin paljon ja erittäin vähän lahopuuta sisältävien ruutujen osuus pienenee ja yksittäisten ruutujen saamat arvot lähestyvät alueen keskiarvoa. Yksittäinen 0,04 hehtaarin ruutu saattaa sisältää lahopuuta vaikkapa 20 kuutiometriä hehtaarilla. Kyseinen ruutu voi kuitenkin olla hehtaarin alueella ainut lahopuustoinen ruutu. Tällöin lahopuustoltaan runsas kohde häviää suurempaan hehtaarin ruutuun, jonka keskiarvoksi muodostuisi tällöin vain 0,8 kuutiota hehtaarilla.

Toisaalta spatiaalisesti hyvin tarkkakaan aineisto ei välttämättä takaa hyviä mallinnustuloksia. Vaikka mallinnettavasta ilmiöstä kerätty tieto on spatiaalisesti

tarkkaa, ilmiön pienialainen vaihtelu saattaa olla hyvin satunnaista. Tällöin ruutukoon pienentyessä tarkasteltavan ilmiön hajonta kasvaa, mikä johtaa otannassa isompiin keskivirheisiin aluetasolla, jos koalojen määrä pysyy vakiona. (Loetsch ym. 1973, Koivuniemi 2003). Ilmiö on havaittavissa kaikessa otannan avulla tehtävässä inventoinnissa. Harvinaisten ilmiöiden, kuten lahopuumäärän hajonta on hyvin suurta esimerkiksi elävään puustoon verrattuna. Tästä johtuen harvinaisista ilmiöistä saadaan elävää puustoa heikompia estimaatteja kun mittaustiheys on sama. Lahopuun mallinnuksessa käytettävän aineiston ruutukoko tulisikin valita siten, että mallin luotettavuus suhteessa ruutukokoa kasvatettaessa menetettyyn informaatioon on käyttötarkoitukseen sopiva. Tämä tarkoittaa, että esimerkiksi koko maata koskevat estimaatit lahopuun kokonaismäärästä voidaan tuottaa hyvinkin suurella ruutukoolla kun taas monimuotoisuutta indikoivien lahopuukeskittymien paikantamisessa vaaditaan spatiaalisesti huomattavasti tarkempaa tietoa.

Hyyppä ja Hyyppä (2001) tutkivat metsikön koon vaikutusta kaukokartoitusmenetelmien tarkkuuteen. Tutkimuksessa verrattiin optisten satelliittikuvien (SPOT ja Landsat TM), SAR-tutkakuvien (synthetic aperture radar) ja AISA-spektrometrikuvien avulla yleistetyn puuston tilavuuden virhettä metsikön pinta-alan funktiona. Pienin tarkasteltu pinta-ala oli yksi hehtaari ja suurin 20 hehtaaria. Tilavuuden keskivirhe oli kaikilla menetelmillä suurimmillaan yhden hehtaarin tarkastelualueella. Tällöin keskivirheeksi muodostui 47–58 %. Kaikilla menetelmillä tulosten tarkkuus parani tarkastellun pinta-alan kasvaessa. Muutos tulosten tarkkuudessa oli voimakkainta siirryttäessä yhden hehtaarin tarkastelutasolta 1,4 hehtaarin tasolle. Tarkasteltavan pinta-alan kasvaessa keskivirheen pieneneminen hidastui.

1.6 Tutkimuksen tavoitteet

Tämän tutkimuksen tavoitteet voidaan jakaa kahteen osaan. Ensimmäisen osan tavoitteena oli rakentaa ALS-dataa hyödyntävät mallit alueen lahopuuston kuvaamista varten. Mallinnuksen yhteydessä haluttiin myös verrata parametrisella ja ei-parametrisella eli aineistopohjaisella menetelmällä saatavia tuloksia. Näin ollen lahopuun esiintymistä mallinnettiin sekä logistisen regression että CART-analyysin avulla. Menetelmistä ensimmäinen edustaa parametrista ja toinen ei-parametrista mallinnusta. Tutkimuksen ensimmäisessä osassa rakennettuja malleja käytettiin pohjana toisen osan tutkimuskysymyksen ratkaisemisessa.

Tutkimuksen toisessa osassa keskeinen tutkimuskysymys kuului: Miten mallinnuksessa käytetty ruutukoko vaikuttaa mallinnettujen lahoppuuestimaattien tarkkuuteen? Kysymykseen pyrittiin vastaamaan kahdesta eri näkökulmasta. Ensimmäisenä tarkasteltiin mallien kykyä luokitaa tutkimusalueella sijaitsevia ruutuja lahoppuustoisuuden suhteen. Toisessa lähestymistavassa tarkasteltiin mallin avulla tehdyn luokituksen aiheuttamaa lahoppuutilavuuden virhettä. Ennako-oletus ruutukoon vaikutuksesta oli, että ruutujen oikeinluokitus paranee ja luokituksen aiheuttama lahoppuutilavuuden virhe pienenee mallinnuksessa käytettävän ruutukoon kasvaessa.

2. Aineisto

2.1 Maastodata

Tutkimuksessa käytetty maastodata mitattiin kesän 2007 aikana Keski-Suomessa Sonkajärvellä. UPM Kymmenen ja Metsähallituksen omistamista metsistä koostuva tutkimusalue on pinta-alaltaan noin 300 hehtaaria ja käsittää pääasiassa nuorta ja hoidettua mäntyvaltaista talousmetsää. Koealueella on kuitenkin myös yksittäisiä vanhempia metsiköitä. Aineisto kerättiin mittaamalla sadan metrin levyisiä kaistoja, jotka oli sijoitettu koealueelle satunnaisotannan avulla. Kaistojen pituudet vaihtelivat 188 metristä 2824 metriin. Kaikki kaistoilla havaitut tyviläpimitaltaan yli 10 senttimetriset lahopuut mitattiin. Lisäksi niistä määritettiin puulaji ja lahoaste. Yli 1,3 metrisille vielä pystyssä oleville ja kaatuneille kuolleille puille määritettiin tilavuus Laasasenahon (1982) runkokäyrämallien avulla. Alle 1,3 metristen mitattujen puiden tilavuudet määritettiin Huberin kaavalla (ks. Schreuder ym. 1993). Kaatuneiden puiden molempien päiden koordinaatit määritettiin kaatumissuunnan ja puun pituuden avulla. (Pesonen ym. 2009.)

Maastossa mitatut kaistat yhdistettiin keinotekoisesti yhtenäiseksi koealueeksi siten, että kaistojen suunta säilyi samana mittaussuunnan kanssa (pohjois-etelä) (Pesonen ym. 2008). Tällöin koealueeksi muodostui 1,1 km x 2,6 km suorakaide, joka koostui kokonaisuudessaan 7160:sta 20 m x 20 m ruudusta (kuva 1).



Kuva 1. Lahopuun alueellinen jakautuminen kaistamosaiikin sisällä 20 x 20 metrin ruutukoolla.

2.2 ALS-data

Koalue laserkeilattiin heinäkuussa 2006. Keilaus tehtiin Optech ALTM 3100 -keilaimella keskimäärin 2500 metrin lentokorkeudelta 50 kHz:n pulssitaajuudella. Avauskulma oli ± 15 astetta, jolloin yhden keilauslinjan leveydeksi muodostui 1350 metriä ja keskimääräiseksi pulssitiheydeksi noin 0,5 laserhavaintoa neliömetriltä. Sekä ensimmäiset että viimeiset kaiut tallennettiin. Maanpinnan korkeusmalli muodostettiin Axelsson ym. (2000) kehittämällä menetelmällä neliömetrin ruutukoolle. Menetelmässä jokaisen ruudun maanpinnan korkeus määritetään niiden viimeisten kaikujen avulla, joiden oletetaan heijastuneen maanpinnasta. Ruuduille, joilla ALS-data ei sisältänyt maaosumia, maanpinnan taso määritettiin Delaunayn kolmioinnin ja bilineaarisen

interpoloinnin avulla. ALS-aineistosta laskettiin 80 laserpulssien heijastumista kuvaavaa piirrettä:

- Kaikujen kumulatiiviset osuudet 0, 1, 5, 10, 20,...90, 95, 99 ja 100 % korkeuksille
- Korkeudet, joilla saavutettiin kaikujen 0, 1, 5, 10, 20,...90, 95, 99 ja 100 %:n kumulatiivinen kertymä
- Kaikujen intensiteetit 10, 30,...90 % korkeuksilla
- Kaikujen heijastumiskorkeuksien keskihajonta

Piirteet laskettiin 0,04 hehtaarin (20 m x 20 m) ruutukoolle, sekä erikseen ensimmäisille ja viimeisille heijastuneille kaiuille. Lasketut piirteet noudattelevat useissa laserkeilaustutkimuksessa käytettyjä piirteitä (esim. Naasset 2002, 2004, Suvanto 2005). Koska datan avulla haluttiin mallintaa myös maalahopuun esiintymistä, piirteiden laskennassa käytettiin yleisesti käytettyä kahden metrin leikkausrajaa (esim. Naasset 2004, Suvanto ym. 2005) alemmaa heijastuneita kaikuja. Maanpinnan yläpuolisiksi heijastuksiksi luokiteltiin kaikki 0,5 metriä korkeusmallin maanpinnan tason yläpuolelta heijastuneet pulssit. Tätä alemmaa heijastuneet pulssit jätettiin ALS-piirteiden laskennassa huomiotta.

3. Menetelmät

3.1 Datan käsittely

Tutkimuksen yhtenä päätavoitteena oli tarkastella ruutukoon vaikutusta lahopuun kartoituksen tarkkuuteen. Alkuperäisestä 7150:stä 0,04 hehtaarin ruudusta muodostettiin suurempia ruutuja ArcGis-ohjelmiston *block statistics* -työkalun avulla. Uutta ruudukkoa luotaessa määritettiin, montako alkuperäisen ruudukon solua eli perusruutua haluttiin liittää yhteen korkeus- ja leveyssuunnassa. Näin ollen yksi muokattu ruutu koostui aina tasamäärästä alkuperäisiä 0,04 hehtaarin perusruutuja. Suurennettujen ruutujen ALS-piirteet ja lahopuutilavuudet laskettiin perusruutujen aritmeettisena keskiarvona. Tilanteissa, joissa uusi ruutu muodostui osittain tyhjille ruuduille (no data), ALS-piirteet ja lahopuutilavuus laskettiin niiden perusruutujen aritmeettisena keskiarvona, joista tieto löytyi. Aineiston tyhjät ruudut olivat todellisuudessa kaistoihin kuuluneita suo- ja vesialueita jotka rajattiin tarkastelun ulkopuolelle.

Metsien lahopuusta suurin osa on maalahopuuta (esim. Ihalainen ja Mäkelä 2009), ja koska alle 0,5 metrin korkeudelta heijastuneet laserpulssit oli piirteitä irrotettaessa jätetty huomiotta, ALS-dataa ei voitu käyttää lahopuun suoraan mittaamiseen. Vaikka lähempää maan pintaa heijastuneet pulssit olisikin sisällytetty käytettyyn dataan, harvapulssisen aineiston avulla yksittäisten kohteiden löytäminen olisi joka tapauksessa ollut hyvin epävarmaa. Lahopuun suoran havainnoinnin sijaan aluepohjainen data kuvasi metsikön rakennetta. Mallinnuksessa lahopuun oletettiin siis korreloivan ALS-piirteiden kanssa epäsuorasti metsikön rakenteen kautta.

Alkuperäinen ruudukko koostui sata metriä leveistä kaistoista. Jotta suuremmille ruuduille lasketut laserpiirteet kuvaisivat metsiköiden rakennetta, perusruutujen tuli sijaita vierekkäin myös maastossa. Tämän takia suurennettujen ruutujen leveyden haluttiin olevan jaollinen kahdellakymmenellä ja lisäksi sadan piti olla jaollinen uusien ruutujen leveydellä. Tällöin sata metriä leveä mittauskaista koostui aina tasamäärästä kaksikymmentä metriä leveitä perusruutuja. Rajoituksella pyrittiin minimoimaan niiden tilanteiden määrä, joissa samaan laskentaruutuun päätyisi metsiköitä monesta eri paikasta. Muutoin kaistojen leveyssuuntaisissa saumakohdissa syntyisi systemaattisesti ruutuja, jotka eivät kuvaisi todellista metsikköä. Puhtaaseen metsikköpohjaiseen tarkasteluun ei kuitenkaan päästy kahdesta syystä. Kaistojen sisällä kulkevia

metsikkörajoja ei pystytty ottamaan huomioon ruutupohjaisessa tarkastelussa. Myös kaistojen pitkittäisiä saumakohtia oli ruutuja suurennettaessa käytännössä mahdotonta välttää, ja nekin jätettiin huomiotta ALS-piirteitä yleistettäessä. Ruutujen koosta riippuen aineisto sisälsi myös vaihtelevan määrän tyhjiä soluja, joille ei ollut määritetty lahopuun tilavuutta tai laskettu ALS-piirteitä. Tämän takia esimerkiksi 0,04 hehtaarin aineiston 7150 solun joukosta vain 6460 oli käytettävissä mallinnukseen ja testaukseen. Suurempia ruutuja muodostettaessa kaikki 7150 solua olivat kuitenkin käytössä. Aineiston keskeiset tunnuksot erikokoisille ruuduille laskettuna on koottu taulukkoon 1.

Taulukko 1. Tutkimuksessa käytettyjen ruutujen mitat sekä ruuduilla havaitun lahopuun määrän vaihtelu.

ruudun mitat, m		pinta-ala, ha	ruutujen lkm.	lahopuun määrä, m ³ /ha			
leveys	korkeus			minimi	maksimi	keskiarvo	σ
20	20	0.04	6460	0	69.2	2.7	5.9
20	100	0.20	1365	0	37.9	2.9	4.3
20	160	0.32	849	0	30.5	2.9	3.8
20	260	0.52	543	0	29.7	3.0	3.9
100	100	1.00	286	0	25.7	2.8	3.4

Tyhjien ruutujen poistamisen jälkeen jokaisen ruutukoon aineisto jaettiin mallinnus- ja testausosaan siten, että $\frac{1}{4}$ aineistosta oli varattu mallintamista varten ja loppu $\frac{3}{4}$ mallien testausta varten. Jakaminen tapahtui valitsemalla systemaattisesti joka neljäs ruutu mallinnusdataan, jolloin molemmat aineiston osat kattoivat tasaisesti koko tutkimusalueen.

3.2 Logistinen regressio

Koska aiemmissa tutkimuksissa lahopuun tilavuuden suoralla mallintamisella ei ole saavutettu kovinkaan tarkkoja tilavuusestimaatteja (esim. Kotamaa 2007), tutkimuksen päähuomio keskitettiin lahopuustoisten ruutujen luokituksen onnistumiseen. Logistinen regressio on lineaarisen regression erityistapaus ja se soveltuu kuvaamaan kategorisia muuttujia, kuten tässä tapauksessa lahopuuluokkia. Yksinkertaisimmillaan logistinen malli kertoo saavutetaanko jokin valittu taso vai ei. Tällöin kyse on binäärisestä ongelmasta.

$$Z = \begin{cases} 1, & \text{jos tapahtuma on tosi} \\ 0, & \text{jos tapahtuma on epätosi} \end{cases} \quad (1)$$

Logistista regressiota on käytetty esimerkiksi kuvaamaan koe-eläinten vastetta erilaisille myrkyille. Tällöin koe-eläimen kuolemaa voidaan kuvata arvolla 1 ja selviämistä arvolla 0 (Dobson 1990). Metsätieteissä logistisilla malleilla on kuvattu esimerkiksi metsäalueella esiintyviä tuhoja (Yli-Kojola 2002).

Logistinen malli on muotoa (Hosmer ja Lemeshow 2000)

$$\pi(\mathbf{x}) = \frac{e^{g(\mathbf{x})}}{1+e^{g(\mathbf{x})}}, \quad (2)$$

ja se voidaan muuntaa parametreiltaan lineaariseen logit-muotoon

$$g(\mathbf{x}) = \ln \left[\frac{\pi(\mathbf{x})}{1-\pi(\mathbf{x})} \right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p, \quad (3)$$

missä $\beta_0, \beta_1, \dots, \beta_p$ ovat mallin parametreja ja x_1, x_2, \dots, x_p mallin selittäviä muuttujia.

Logistinen regressiomalli sovitetaan aineistoon mallin logit-muunnoksen avulla (kaava 3). Lineaarinen mallin sovittamisessa käytetään usein pienimmän neliösumman menetelmällä. Menetelmä ei kuitenkaan ole käyttökelpoinen logit-funktion tapauksessa, jossa sekä yhtälön oikea että vasen puoli riippuvat estimoitavista parametreista (Kangas 2001). Lisäksi logistisesta mallista estimoiduilla pienimmän neliösumman estimaateilla ei ole samoja tilastollisia ominaisuuksia kuin lineaarisen mallin vastaavilla estimaateilla (Hosmer ja Lemeshow 2000). Logit-funktio sovitetaan aineistoon iteratiivisella menetelmällä. Tässä tutkimuksessa logit-funktioiden sovittamiseen käytettiin R-ohjelmiston glm-funktiota (R Development Core Team 2007).

Aineistoon sovitetun logit-muunnoksen (kaava 3) avulla voidaan logistisella mallilla (kaava 2) estimoida kullekin havaintoyksikölle todennäköisyys ennustettavan ilmiön toteutumiselle. Malli tuottaa lukuarvon välillä $[0,1]$, jota tulkitaan valitun kynnsarvon avulla. Tilanteessa, jossa mahdollisia lopputuloksia on kaksi, eli jokin asia on tosi tai epätosi, on luonnollista olettaa, että molempien tapausten todennäköisyys on yhtä suuri.

Tällöin raja-arvoksi valitaan 0,5. Havaintoyksiköt, joissa logistisen mallin arvo on yli valitun raja-arvon, tapahtuma on tosi. Jos raja-arvo ei ylity, tapahtuma on epätosi.

Mallinnusaineisto jaettiin siten, että ruudut joista oli mitattu maastossa enemmän kuin yksi kuutiometri lahpuuta hehtaarilla luokiteltiin lahpuustoisiksi. Näissä ruuduissa logistisen mallin logit-muunnos sai arvon 1. Muissa ruuduissa muunnos sai arvon 0. Logit-muunnos sovitettiin mallinnusaineistoon, minkä jälkeen logistisella mallilla voitiin estimoida jokaiselle testiaineiston ruudulle todennäköisyys yli yhden kuutiometrin hehtaarikohtaisen lahpuutilavuuden esiintymiselle. Testiaineiston ruutu luokiteltiin lahpuustoiseksi, kun logistisella mallilla estimoitu todennäköisyys oli yli 0,5.

Mallinnuksessa lahpuustoisuuden raja-arvona käytettiin yhtä kuutiometriä hehtaarilla. Ruudut, joilla oli maastomittausten mukaan vähemmän lahpuuta, luokiteltiin lahpuuttomiksi. Raja-arvo valittiin ruutujen varmempaa tunnistamista silmälläpitäen. Oletuksena oli, että hyvin pienet lahpuumäärät (alle $1 \text{ m}^3/\text{ha}$) heikentävät mallin selityskykyä merkittävästi ja siksi nämä hyvin vähän lahpuuta sisältävät ruudut luokiteltiin lahpuuttomiksi. Oletusta testattiin mallintamalla myös täysin lahpuuttomien ruutujen esiintymistä 0,04 hehtaarin ruuduille. Täysin lahpuuttomien ruutujen luokitus osoittautui testissä heikommaksi kuin hyvin vähän lahpuuta sisältävien ruutujen luokitus. Täysin lahpuuttomien ruutujen luokitus sai kappa-arvon (ks. luku 3.3) 0,175, kun hyvin vähän lahpuuta sisältävien ruutujen luokituksen kappa oli 0,194. Myös mallinnus osoittautui hyvin hankalaksi täysin lahpuuttomilla ruuduilla, sillä ruutukoon suurentuessa lahpuuttomien ruutujen määrä pieneni voimakkaasti, jolloin toimivien mallien tekeminen muuttui mahdottomaksi. Yhden hehtaarin ruutukoolla mallinnusaineistossa oli vain kaksi täysin lahpuutonta ruutua. Lahpuustoisuuden rajaaminen alle yhteen kuutiometriin hehtaarilla paransi mallintamisen edellytyksiä huomattavasti (taulukko 2). Raja vastaa noin kolmasosaa tutkimusalueen keskimääräisestä lahpuumäärästä ($2,7\text{--}3,0 \text{ m}^3/\text{ha}$).

Taulukko 2. Hyvin vähälahopuustoisten ja täysin lahopuuttomien ruutujen määrä viidellä eri ruutukoolla (suluissa % -osuus kaikista ko. ruutukoon ruuduista).

		ruutukoko, ha				
		0,04	0,20	0,32	0,52	1,00
$V_{\text{lahopuu}} = 0$		2860 (44 %)	167 (12 %)	60 (7 %)	23 (4 %)	9 (3 %)
$0 < V_{\text{lahopuu}} < 1$		721 (11 %)	366 (27 %)	246 (29 %)	147 (27 %)	67 (23 %)

Lahopuun esiintymistä selittävät muuttujat valittiin 0,2 hehtaarin ruutuihin jaetun aineiston avulla. Muiden ruutukokojen malleille ei valittu erikseen parhaita mahdollisia muuttujia, koska yhtenä tutkimuksen päätarkoituksena oli ruutukoon vaikutuksen tutkiminen. Muuttujien valinta tehtiin kolmessa vaiheessa siten, että parhaimmat muuttujat jatkoivat seuraavaan vaiheeseen huonompien pudotessa pois vertailusta. Ensimmäisen vaiheen karsinta tehtiin silmämääräisesti hajontakuvien avulla. Jokainen yksittäinen ALS-piirre kuvattiin solun lahopuumäärän suhteen ja muodostuneesta hajontakuviosta arvioitiin, selittikö kyseinen piirre lahopuun määrää. Valinnassa kiinnitettiin huomiota myös siihen, että samantyyppisten ALS-piirteiden ryhmästä valittiin mukaan pääsääntöisesti vain kaksi parasta.

Toisessa vaiheessa lahopuumäärää mallinnettiin jokaisella yksittäisellä mukaan valitulla ALS-piirteellä erikseen. Koska logistinen regressio on lineaarisen regression erityistapaus, ALS-piirteen tai jonkin sen muunnoksista tuli olla lahopuumäärän suhteen lineaarinen. Ensimmäisen vaiheen silmämääräisellä arvioinnilla ei pystytty tarkasti määrittämään parasta mahdollista piirteen linearisoivaa muunnosta. Siksi kaikille ensimmäisessä vaiheessa valituille piirteille tehtiin kolme muunnosta. Käytetyt muunnokset olivat tarkasteltavan piirteen neliöjuuri, luonnollinen logaritmi ja käänteisluku. Lasketun ALS-piirteen ja sen kolmen muunnoksen kykyä selittää lahopuun määrää arvioitiin yksitellen lineaarisen regression avulla. Selityskyvyn mittareina käytettiin selityssastetta ja jäännöshajontaa.

Kolmanteen vaiheeseen ja varsinaisen mallin muodostamiseen valittiin toisessa vaiheessa parhaiten lahopuumäärää selittäneet muuttujat. Mallin viimeistelyssä seurattiin muuttujien tilastollisen merkitsevyyden ja mallin jäännöshajonnan muutoksia erilaisilla selittävien muuttujien yhdistelmillä. 0,2 hehtaarin ruutukoolle tehty malli parametrisoitiin muun kokoisille ruuduille. Koska selittävät muuttujat pysyivät samoina

kaikissa malleissa, ruutukoon vaikutusta pystyttiin tarkastelemaan ilman mallien erilaisuudesta johtuvaa epätarkkuutta. Kun selittävät muuttujat olivat samat, ainoa mallien välillä muuttuva tekijä oli mallien parametreihin vaikuttava ruutukoko.

3.3 Luokituksen hyvyden arviointi

Testiaineiston luokituksen onnistumista arvioitiin vertaamalla logistisella mallilla ennustettua lahpuustoisuutta maastossa mitattuun lahpuumäärään. Koska mallinnuksessa vain yli yhden kuutiometrin hehtaariohtaiset lahpuutilavuudet otettiin huomioon, myös luokitusta arvioitaessa tätä pienemmät lahpuumäärät jätettiin huomiotta. Vertailtavina tunnuksina käytettiin oikeinluokitusprosenttia (kaava 4) ja kappa-arvoa (kaava 5). Kappa-arvo otettiin mukaan arviointiin, koska se ottaa huomioon paitsi oikeinluokitettujen ruutujen määrän, myös luokkien määrälliset suhteet. Oikeinluokitusprosentti lasketaan

$$\text{oikeinluokitus} - \% = \frac{N_{\text{oikein}}}{N} * 100, \quad (4)$$

missä N_{oikein} on oikeinluokitettujen ruutujen määrä ja N on ruutujen kokonaismäärä. Kappa-arvo lasketaan (ks. Campbell 2007)

$$K = \frac{p_o - p_e}{1 - p_e}, \quad (5)$$

missä p_o on havaittujen luokkien, eli tässä tapauksessa lahpuustoisten ja lahpuuttomien ruutujen osuus ja p_e on oikein estimoitujen luokkien osuus satunnaisluokituksessa. Kappa-arvo vaihtelee välillä 0-1 ja sitä voidaan tulkita esimerkiksi taulukossa 3 esitellyllä tavalla:

Taulukko 3. Kappa-arvon tulkinta Landis ja Koch (1977) mukaan.

kappa-arvo	luokituksen onnistuminen
< 0,00	huono
0,00 - 0,20	heikko
0,21 - 0,40	keskinkertainen
0,41 - 0,60	kohtalainen
0,61 - 0,80	hyvä
0,81 - 1,00	lähes täydellinen

3.4 CART-mallinnus

Toinen tutkimuksessa käytetty mallinnusmenetelmä oli CART (Breiman ym. 1984). Logistisesta regressiosta poiketen CART on ei-parametrinen data-mining menetelmä, jolla voidaan automaattisesti seuloa suuria aineistoja monien muuttujien suhteen. Käytännössä tämä tarkoittaa, että mallinnusprosessissa voidaan käydä läpi suurempi määrä mahdollisia selittäjiä ja niiden yhdistelmiä kuin parametrisilla menetelmillä. CART:n avulla aineisto voidaan jakaa halutun ominaisuuden suhteen mahdollisimman samankaltaisiin luokkiin ja selvittää mitkä aineiston piirteet määrittävät kunkin luokan. Lähtökohtana on yksi luokka, johon kuuluu koko aineisto. Seuraavana aineisto jaetaan kahteen haaraan, jotka ovat sisäisesti mahdollisimman homogeenisiä halutun aineiston osan suhteen. Aineiston luokittelussa käytettiin R-ohjelmiston RPART-pakettia (Therneau ja Atkinson 1997), joka toimii CART-mallin tavoin. RPART määrittää jokaisen haaran homogeenisuuden eräänlaisten puhtausfunktioiden (impurity function) avulla (kaava 6)

$$I(A) = \sum_{i=1}^C f(piA), \quad (6)$$

missä C on luokkien lukumäärä, piA on niiden havaintojen osuus luokassa A , jotka kuuluvat luokkaan i . Puhtausfunktiona voidaan käyttää informaatioindeksiä (kaava 7) tai Gini-indeksiä (kaava 8).

$$f(p) = p(1 - p) \quad (7)$$

$$f(p) = -p \log(p) \quad (8)$$

Luokitusta voidaan ohjata kompleksisuusparametrin (complexity parameter, CP) avulla. Parametrilla määritetään kuinka paljon viimeisen jaon tulee vähintään parantaa luokituksen hyvyttä, jotta algoritmin suorittaminen jatkuisi. Tutkimuksessa CP asetettiin tasolle 0,0005. Tällöin ohjelma jatkoi aineiston jakamista, kunnes paras jako nosti luokituksen hyvyttä kuvaavan vakion selitystasetta vähintään 0,0005:llä.

Automaattinen luokitus voidaan jakaa kahteen vaiheeseen. Ensimmäisessä vaiheessa algoritmi käy läpi koko aineiston ja muodostaa eräänlaisen puun, jonka haarojen päistä löytyvät osajoukot eli ”lehdet” ovat mahdollisimman samankaltaisia valitun ominaisuuden suhteen. Aineiston jokaista haaroittumista kuvataan aineiston jonkin muuttujan kynnyksarvolla. Ensimmäisen vaiheen luokittelusta on usein tuloksena rakenteeltaan hyvin monimutkainen puu eli aineisto jaetaan suureen määrään luokkia. Toisessa vaiheessa puun jokainen haara käydään läpi ja määritetään ne haarat, jotka otetaan mukaan lopulliseen malliin. Haarot valitaan määrittämällä ristiinvalidoinnin avulla CP, jolla muodostetun puun haarojen yhteenlaskettu riski on pienin (Therneau ja Atkinson 1997). Tässä tapauksessa riskillä tarkoitetaan väärästä luokituksesta aiheutunutta puhtaustuksen arvon menetystä kerrottua luokituksen todennäköisyydellä.

CART:n avulla pyrittiin löytämään lahoppumäärän ja ALS-piirteiden välillä mahdollisesti olevia riippuvuuksia, joita on vaikea havaita tavanomaisen parametrin mallintamisen avulla. Ruuduista muodostettiin menetelmän avulla lahoppuuston suhteen mahdollisimman samankaltaisia ryhmiä. Ryhmät määriteltiin ALS-piirteiden avulla. Luokitus tehtiin suurimman ja pienimmän resoluution aineistoille (0,04 ha ja 1,0 ha). Molemmissa tapauksissa luokitus tehtiin ¼ aineistolle, samaan tapaan kuin logistinen mallinnus. Tällöin luokituksen yleistettävyyttä voitiin tarkastella luokittelemalla mallinnuksesta erillään pidetty testiaineisto merkitseviksi havaittujen ALS-piirteiden avulla.

CART-menetelmällä tuotettiin kaksi ennustetta sekä 0,04 että 1,0 hehtaarin ruuduille. Ensimmäisen ennusteen tarkoituksena oli saada vertailukohta logistisilla malleilla

tehdyllä luokituksella. Ruudut siis jaettiin lahoppuustoisiin ja lahoppuuttomiin. Lahoppuustoisuuden raja-arvona käytettiin logistisen mallin tapaan yhtä kuutiometriä lahoppuuta hehtaarilla. Tällöin mallinnusaineisto luokiteltiin kahteen luokkaan luokkamuuttujan avulla: 0 = ei lahoppuuta, 1 = lahoppuustoinen. Luokituksen onnistumista arvioitiin oikeinluokitusprosentin ja kappa-arvon avulla. Lisäksi CART:n avulla muodostettiin ruutukohtainen lahoppuuston tilavuusennuste. Todellista tilavuutta ennustettaessa mallinnusaineisto luokiteltiin maastossa mitatun lahoppuutilavuuden avulla, joka on jatkuva muuttuja. Tuloksena saaduille luokkakohtaisille lahoppuukeskiarvoille laskettiin luokituksen onnistumista kuvaavien tunnusten sijaan tilavuuden keskivirhe.

3.5 Lahoppuuestimaattien luotettavuuden arviointi

CART-mallinnuksen yhteydessä estimoitujen ruutukohtaiset lahoppuutilavuuksien luotettavuutta arvioitiin keskineliövirheen neliöjuuren (RMSE) avulla. RMSE laskettiin kaavalla

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{n-1}}, \quad (9)$$

jossa x_i = lahoppuun todellinen määrä ruudussa i , \hat{x}_i = ennustettu lahoppuumäärä ruudussa i ja n = ruutujen lukumäärä. Tunnusta käytettiin kuvaamaan niin yksittäisten luokkien kuin myös koko aineiston lahoppuutilavuusennusteiden luotettavuutta.

Lahoppuuluokkien ja -mallien välistä vertailua varten laskettiin myös suhteellinen keskineliövirheen neliöjuuri (RMSE %). Laskenta tapahtui kaavalla

$$RMSE \% = \frac{RMSE_j}{\bar{x}_j}, \quad (10)$$

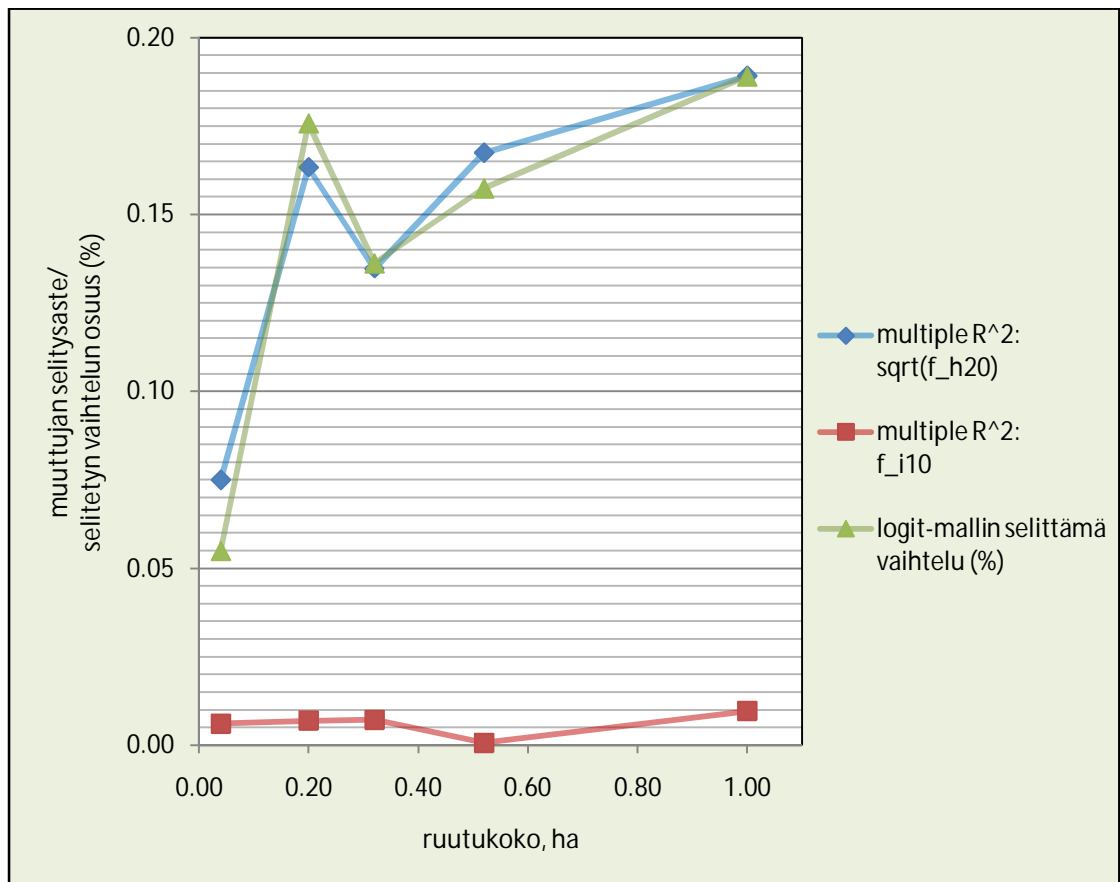
jossa $RMSE_j$ = lahoppuutilavuuden RMSE luokassa j ja \bar{x}_j = luokan j todellinen lahoppuukeskiarvo.

4. Tulokset

4.1 Logistiset lahoppuumallit

Ruutukoon vaikutusta lahoppuennusteen luotettavuuteen tarkasteltiin mallintamalla lahoppuun esiintymistä kaikille viidelle ruutukoolle. Yksittäisten muuttujien paremmuutta verrattiin yhden selittäjän mallien avulla. Lopullisen, useamman muuttujan mallin selittäviksi muuttujiksi valikoituivat neliöjuuri korkeudesta, jolla ensimmäisten heijastuneiden kaikujen kumulatiivinen kertymä on 20 prosenttia ($\sqrt{f_{h20}}$) sekä 10 %:n korkeudelta heijastuneiden ensimmäisten kaikujen intensiteetti (f_{i10}). Molemmat muuttujat olivat lopullisessa, 0,2 hehtaarin ruudukolle tehdyssä mallissa tilastollisesti merkitseviä 5 %:n riskitasolla.

Yhden muuttujan mallien avulla tarkastellut selitysasteet vaihtelivat muuttujan ($\sqrt{f_{h20}}$) osalta välillä 0,075-0,189 ja muuttujan f_{i10} osalta välillä 0,0007-0,0096 (kuva 2).



Kuva 2. Muuttujien selitysasteet yhden muuttujan malleissa sekä lopullisen logit-mallin selittämän vaihtelun osuus ruutukoon funktiona.

Muuttujan ($\sqrt{f_{h20}}$) sisältämän yhden muuttujan mallin selitysaste parani ruutukoon kasvaessa. 0,20 hehtaarin ruutukoolla, jolla piirteet valittiin, selitysaste oli 0,163. Suurimmalla yhden hehtaarin ruutukoolla selitysaste saavutti arvon 0,189. Toisen selittävän muuttujan, f_{i10} :n sisältämän yhden muuttujan mallin selitysasteet olivat selvästi ($\sqrt{f_{h20}}$):n vastaavia pienempiä. 0,20 hehtaarin ruutukoolla selitysaste oli 0,0069 ja yhden hehtaarin ruutukoolla vastaavasti 0,0096.

Viidelle ruutukoolle parametrisoidut lahopuun esiintymistä kuvaavat logit-muotoiset mallit olivat seuraavat:

$$g(x)_{(0,04)} = -3.75022 + 1.13907 * (\sqrt{f_{h20}}) + 0.05322 * f_{i10} \quad (11)$$

$$g(x)_{(0,20)} = -7.8053 + 2.5696 * (\sqrt{f_{h20}}) + 0.2056 * f_{i10} \quad (12)$$

$$g(x)_{(0,32)} = -7.3084 + 2.42298 * (\sqrt{f_{h20}}) + 0.17131 * f_{i10} \quad (13)$$

$$g(x)_{(0,52)} = -9.7689 + 2.6937 * (\sqrt{f_{h20}}) + 0.4515 * f_{i10} \quad (14)$$

$$g(x)_{(1,00)} = -6.7162 + 2.5916 * (\sqrt{f_{h20}}) + 0.1783 * f_{i10} \quad (15)$$

Kun 0,20 hehtaarin ruuduille tehty malli sovitettiin lopuille neljälle ruutukoolle (0.04, 0.32, 0.52 ja 1.0 ha), piirteen f_{i10} tilastollinen merkitsevyys heikkeni (taulukko 4). Kahdessa tapauksessa muuttuja ei ollut tilastollisesti merkitsevä 5 %:n riskitasolla. 0,32 hehtaarin ruutukoolle parametrisoidussa mallissa muuttujan p-arvo ($p = 0,061$) oli vain hiukan halutun tason yläpuolella, mutta yhden hehtaarin ruutukoolla vastaava arvo nousi jo selvästi yli 10 %:n riskitason ($p = 0,188$). Piirre ($\sqrt{f_{h20}}$) sitä vastoin pysyi tilastollisesti merkitseväenä 5 %:n riskitasolla kaikkien viiden ruutukoon malleissa. Koko mallin selityskyky nousi selvästi ruutukoon kasvaessa. 0,04 hehtaarin ruutukoon mallin avulla pystyttiin selittämään 5,49 % lahopuustoisuuden kokonaisvaihtelusta, kun hehtaarin ruutukoon mallilla vastaava osuus oli 18,9 %.

Taulukko 4. Logististen mallien parametrien kertoimet, keskiarvot ja p-arvot sekä valmiiden mallien selittämän vaihtelun osuus kokonaisvaihtelusta.

		mallin parametrit			selitetyn vaihtelun osuus, %	
		vakio	sqrt(f_{h20})	f_{i10}		
mallinnettava ruutukoko	0,04 ha	kerroin	-3,750	1,139	0,053	5,49 %
		keskiarvo	0,455	0,117	0,025	
		p-arvo	<2E-16	<2E-16	0,035	
	0,20 ha	kerroin	-7,805	2,570	0,206	17,58 %
		keskiarvo	1,246	0,338	0,071	
		p-arvo	3,74E-10	2,93E-14	0,004	
	0,32 ha	kerroin	-7,308	2,423	0,171	13,62 %
		keskiarvo	1,713	0,455	0,091	
		p-arvo	1,98E-05	9,95E-08	0,061	
	0,52 ha	kerroin	-9,769	2,694	0,452	15,74 %
		keskiarvo	2,401	0,621	0,145	
		p-arvo	4,71E-05	1,44E-05	0,002	
	1,00 ha	kerroin	-6,716	2,592	0,178	18,90 %
		keskiarvo	2,719	0,795	0,135	
		p-arvo	0,014	0,001	0,188	

4.2. Aineiston luokittelu logistisilla malleilla

Mallinnusaineistosta erotetulle testiaineistolle tuotettiin ruutukohtaiset ennusteet lahoppuun esiintymiselle. Ennusteet tehtiin kaikille viidelle ruutukoolle sijoittamalla niitä varten parametrisoidut logit-mallit (kaavat 8-12) yksitellen logistisen mallin yhtälöön (kaava 2). Kun malli sai arvon 0,5 tai enemmän, ruutu määritettiin lahoppuustoiseksi. Yhteenveto mallien oikeinluokituskyvystä on esitetty taulukossa 5.

Taulukko 5. Lahoppuustoisuuden oikeinluokitusprosentti ja luokituksen kappa-arvo logististen mallien avulla 0,04; 0,2; 0,32; 0,52 ja 1,0 hehtaarin ruuduilla.

		ruutukoko, ha				
		0,04	0,20	0,32	0,52	1,00
oikeinluokitus %		60,8	70,3	73,4	74,9	76,2
kappa		0,19	0,34	0,38	0,32	0,26

Pienimmällä, 0,04 hehtaarin ruutukoolla oikeinluokitusprosentti oli 60,8 ja kappa-arvo 0,19. Siirryttäessä tarkastelussa seuraavaan 0,2 hehtaarin ruutukokoon oikeinluokitusprosentti nousee lähes kymmenen prosenttiyksikköä. Samalla kappa-arvo nousee 0,34:än. Huomionarvoista 0,2 hehtaarin ruutukoossa on, että malleissa käytetyt

selittävät muuttujat valittiin tämän ruutukoon aineistosta. Seuraavaksi suurin, 0,32 hehtaarin ruutukoko, tuottaa oikeinluokitusprosentiksi 73,4 ja kappa-arvoksi 0,38. Suuruusjärjestyksessä keskimmäisen ruutukoon jälkeen oikeinluokitusprosentti jatkaa hidastuvaa kasvuaan ollen 0,52 hehtaarin ruutukoolla 74,9 % ja yhden hehtaarin ruutukoolla 76,2 %. Samaan aikaan kappa-arvot kääntyvät kuitenkin laskuun. 0,52 hehtaarin ruudun kappa-arvoksi muodostui 0,32 ja yhden hehtaarin ruudulle vastaavasti 0,26.

4.3 CART: luokitus

Logistisista malleista poiketen CART-luokitus tehtiin vain aineiston 0,04 hehtaarin ja 1,0 hehtaarin ruutukoolle. 0,04 hehtaarin ruutukoolla pienimmän virheen luokitus jakoi mallinnusaineiston kahtia muuttujan f_{h20} avulla (liite 1). Lahopuustoisiksi luokiteltiin ruudut, joilla muuttujan f_{h20} arvo oli 6,525 tai enemmän. Lahopuuttomiksi luokiteltiin ruudut, joilla muuttujan f_{h20} arvo oli alle 6,525. 1,0 hehtaarin ruutukoolla ei pystytty tuottamaan virheen minimoivaa luokitusta lahopuuttomiin ja lahopuustoisiin ruutuihin. Sen sijaan tyydyttiin CART:n ensimmäisen vaiheen malliin, joka otti huomioon ainoastaan mallin selitysvoiman kasvun. Näin ollen suurimman ruutukoon aineisto jaettiin kahdesti (liite 2). Ensimmäinen jako tapahtui muuttujan f_{hstd} suhteen, joka kuvaa ensimmäisten heijastuneiden kaikujen heijastumiskorkeuksien hajontaa. Kun muuttuja sai arvon 2,553 tai enemmän, ruutu luokiteltiin lahopuustoiseksi. Mikäli taas muuttujan arvo oli pienempi kuin 2,553, suoritettiin toinen jako muuttujan l_{p1} suhteen, joka kuvaa korkeutta, jolla viimeisten heijastuneiden kaikujen kumulatiivinen kertymä oli 1 %. Mikäli muuttujan l_{p1} arvo oli alle 7,117, ruutu luokiteltiin lahopuustoiseksi. Lahopuuttomia olivat ruudut joissa l_{p1} sai arvon 7,117 tai enemmän.

Myös CART-luokituksen hyvyttä arvioitiin oikeinluokitusprosentin ja kappa-arvon avulla (taulukko 6). 0,04 hehtaarin ruutukoon malli luokitteli testiaineiston siten, että 3285 ruutua ennustettiin lahopuustoisiksi ja 1562 lahopuuttomiksi. Oikeinluokitusprosentti oli 60,5, mutta luokkien koon huomioon ottava kappa-arvo kuitenkin vain 0,24. Yhden hehtaarin ruutukoolla malli ennusti 169 ruutua lahopuustoisiksi ja 45 lahopuuttomiksi. Suurimmalla ruutukoolla Oikeinluokitusprosentti oli 82,2 ja luokituksen kappa-arvo 0,52. CART-mallin avulla tehtyjen luokitusten kappa-arvot olivat logistisen mallin vastaavia arvoja paremmat kummallakin testatulla ruutukoolla.

Taulukko 6. Lahopuustoisuuden oikeinluokitusprosentti ja luokituksen kappa-arvo CART-mallien avulla 0,04 ja 1,0 hehtaarin ruuduilla

		ruutukoko, ha	
		0,04	1,00
oikeinluokitus %		60,5 %	82,2 %
kappa		0,24	0,52

4.4 CART: tilavuusestimaatit

Pienimmän resoluution eli 0,04 hehtaarin ruutukoon automaattinen luokitus tuotti tulokseksi neljä lahopuuluokkaa (taulukko 7). Luokkien lahopuukeskiarvot olivat 2,345, 6,24, 10,04 ja 26,24 kuutiometriä hehtaarilla. Luokitteluparametreiksi valikoituivat seuraavat ALS-piirteet: korkeus, jolla ensimmäisten heijastuneiden kaikujen kumulatiivinen kertymä on 70 % (f_{h70}), viimeisten heijastuneiden kaikujen kumulatiivinen kertymä korkeudella 0 (l_{p0}) sekä korkeus, jolla viimeisten heijastuneiden kaikujen kumulatiivinen kertymä on 20 % (l_{h20}) (liite 3). Vähiten lahopuuta sisältävään ensimmäiseen luokkaan luokiteltiin kuuluvaksi 4532 ruutua, toiseen luokkaan 128 ruutua, kolmanteen 113 ruutua ja eniten lahopuuta sisältävään neljänteen luokkaan 72 ruutua.

Taulukko 7. CART-mallin tuottamat tilavuusluokat, luokkien koot ja tilavuuksien RMSE:t 0,04 hehtaarin ruutukoolla.

	luokat, lahopuuta m ³ /ha			
	2,354	6,24	10,04	26,24
Todellinen keskiarvo (m ³ /ha)	2,39	5,95	12,47	13,57
Ruutuja luokassa	4532	128	113	72
RMSE (m ³ /ha)	4,66	8,48	15,73	17,21
suht. RMSE	195,3 %	142,5 %	126,1 %	126,8 %
RMSE (m ³ /ha), koko aineisto	5,7			
suht. RMSE, koko aineisto	197,1 %			

Tilavuuden suhteellinen RMSE laskettiin havaitun RMSE:n ja testiaineistosta havaittujen luokakeskiarvojen avulla. Kolmessa ensimmäisessä luokassa suhteellinen RMSE laski mallinnettavan luokan keskitilavuuden noustessa. Keskitilavuudeltaan suurimman luokan kohdalla RMSE kääntyi aavistuksen nousuun. Vähiten lahopuuta sisältävään luokkaan suhteellinen RMSE oli ylivoimaisesti suurin ja samalla luokkaan

kuului suurin osa ruuduista (93,5 %). Koko aineiston RMSE oli lähes kaksinkertainen havaittuun keskitilavuuteen verrattuna.

Suurimman resoluution eli 1,0 hehtaarin ruutuja luokiteltaessa aineistosta erottui mallinnusaineiston perusteella kaksi luokkaa (taulukko 8). Luokkien lahoppumäärien keskiarvot olivat 1,617 ja 4,999 kuutiometriä hehtaarilla. Ainoana luokitteluperusteena oli ALS-piirre, joka kuvasi korkeutta, jolla viimeisten heijastuneiden kaikuksen kumulatiivinen kertymä oli 99 % (I_{h99}) (liite 4). Keskimääräiseltä lahoppuutilavuudeltaan pienempään luokkaan kuului tässä tapauksessa 128 ruutua ja suurempaan luokkaan loput 86 ruutua.

Taulukko 8. CART-mallin tuottamat tilavuusluokat, luokkien koot ja tilavuuksien RMSE:t 1,0 hehtaarin ruutukoolla.

	luokat, lahoppuuta m ³ /ha	
	1,617	4,999
Todellinen keskiarvo (m ³ /ha)	1,72	4,46
Ruutuja luokassa	128	86
RMSE (m ³ /ha)	1,62	5,00
suht. RMSE	94,0 %	112,0 %
RMSE (m ³ /ha), koko aineisto	3,40	
suht. RMSE, koko aineisto	120,3 %	

Yhden hehtaarin ruutukoolla RMSE nousi mallinnettavan luokan keskitilavuuden noustessa. Koko aineiston RMSE oli 3,4 kuutiometriä hehtaarilla ja suhteellinen RMSE jäi 120,3 %:in, eli huomattavasti pienemmäksi kuin 0,04 hehtaarin ruutukoolla. Tässäkin tapauksessa suhteellinen RMSE on laskettu testiaineistossa havaittujen luokkien todellisten keskitilavuuksien perusteella.

5. Tulosten tarkastelu

5.1 Logistisen mallin parametrit

Koska mallin haluttiin olevan sama kaiken kokoisilla ruuduilla, oli tehtävä valinta mallinnuksen ”nollatasosta” eli ruutukoosta, jonka avulla mallinnuksessa käytetyt ALS-muuttujat valittiin. Kokonaislahopuunmäärän ja ALS-piirteiden väliset korrelaatiot olivat hyvin heikkoja, mikä aiheutti vaikeuksia mallin muuttujien valinnassa. Mahdollisimman pienten ruutujen käyttäminen mallinnuksen perustasona oli heikoista korrelaatioista huolimatta mielekästä, koska pyrittiin löytämään spatiaalisesti mahdollisimman tarkka tapa kartoittaa lahopuun esiintymistä.

Mallien nollatason kiinnittäminen 0,2 hehtaariin osoittautui ongelmalliseksi, sillä valittujen muuttujien selityskyky vaihteli eri ruutukokojen välillä. Mallin selittäjien haluttiin kuitenkin olevan samat erikokoisilla ruuduilla, jotta ruutukoon vaikutusta voitaisiin arvioida luotettavasti. Näin ollen muuttujia ei pudotettu pois tai vaihdettu, vaikka kahdessa tapauksessa muuttujan f_{i10} tilastollinen merkitsevyys nousikin yli asetetun 5 %:n riskitason (taulukko 4). Muuttujan f_{i10} tilastollinen merkitsevyys jäi alle tavoitellun 5 %:n riskitason ainoastaan suurimmalla, yhden hehtaarin ruutukoolla. Tällöin muuttuja f_{i10} oli tilastollisesti merkitsevä vasta 20 %:n riskitasolla. Muuttujan ($\sqrt{f_{h20}}$) suhteen samaa ongelmaa ei ollut, vaan se oli tilastollisesti merkitsevä 5 %:n riskitasolla jokaisella ruutukoolla. Myös muuttujan ($\sqrt{f_{h20}}$) yhden muuttujan malleilla tutkittu selitysaste oli jokaisella ruutukoolla huomattavasti muuttujan f_{i10} selitysastetta korkeampi (kuva 2). Tulosten perusteella voidaan todeta, että muuttuja ($\sqrt{f_{h20}}$) oli mallin toiminnan kannalta keskeisessä osassa, kun taas muuttujan f_{i10} säilyttäminen mallissa saattoi jopa heikentää mallin toimintaa.

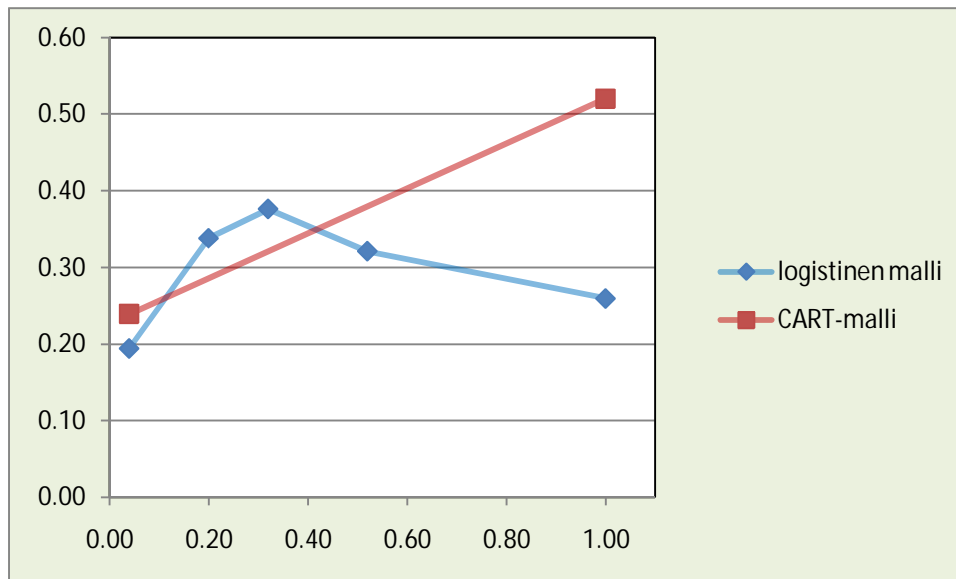
Kun tarkastellaan mallien selittämän vaihtelun osuutta aineiston kokonaisvaihtelusta, huomataan että selitetyn vaihtelun osuus 0,2 hehtaarin ruutukoolla on selvästi lähimpiä ruutukokoja parempi (kuva 2). Parempi selityskyky johtuu siitä, että 0,2 hehtaarin ruutukoko toimi mallinnuksen nollatasona eli muuttujina käytetyt laserpiirteet valittiin sen avulla. Vasta hehtaarin ruutukoolla parametrisoidun mallin selityskyky kohoaa nollatasona toimineen 0,2 hehtaarin mallin selityskyvyn yläpuolelle (kuva 2). Selitetyn vaihtelun osuuden ja muuttujan ($\sqrt{f_{h20}}$) selitysasteen kuvaajat ovat hyvin samankaltaiset.

Molemmissa mallinnuksen nollassa arvo poikkeaa voimakkaasti lähimpien ruutukokojen arvoista. Logistisen mallintamisen avulla olisi luultavasti saavutettu parempia tuloksia, mikäli lahoppuustoa kuvaavien mallien selittäjät olisi valittu jokaiselle ruutukoolle erikseen, mihin myös myöhemmin käsiteltävät CART-luokituksen tulokset viittaavat. Tällöin ruutukoon vaikutusta olisi kuitenkin ollut vaikeampi havainnoida.

5.2 Luokitus logistisilla malleilla

Logistisilla malleilla tehtyjen luokitteluiden tulokset olivat odotetun kaltaisia. Muuttujia valitessa havaitut korrelaatiot lahoppuun määrän ja ALS-piirteiden välillä eivät antaneet odottaa kovinkaan hyviä tuloksia luokituksen suhteen. Mallinnuksen toisessa vaiheessa yhden muuttujan malleilla lasketut selitysasteet olivat muuttujalla ($\sqrt{f_{h20}}$) 0,16 ja vastaavasti muuttujalla f_{i10} vain 0,007. Vaikka korrelaatiot ovat hyvin heikkoja, ne ovat samankaltaisia muun tutkimuksen kanssa. Esimerkiksi Pesonen ym. (2010a) havainnoivat lahoppuuta ALS-dataa hyödyntäen. Tutkimuksessa lahoppuuta parhaiten selittävien ALS-piirteiden korrelaatiot yhden muuttujan mallissa olivat 0,10 (f_{h60}), 0,09 (f_{hstd}) ja 0,01 (f_{p90}). Toisessa tutkimuksessaan Pesonen ym. (2010b) havaitsivat samansuuntaisia korrelaatioita ALS-piirteiden ja lahoppuun määrän välillä. Parhaan selittävän muuttujan (f_{h50}) korrelaatio vaihteli käytetystä ruutukoosta riippuen 0,051:stä 0,224:ään. Vaikka tässä tutkimuksessa käytetyt selittävät muuttujat poikkeavatkin edellisissä tutkimuksissa käytetyistä, ALS-piirteiden ja lahoppuun määrän heikko korrelaatio käy ilmi molemmista tutkimuksista. Tässä tutkimuksessa muuttujia valittiin osittain silmämääräisesti hajontakuvien perusteella. Visuaalisen tulkinnan vaiheessa hyvin samankaltaisten piirteiden joukosta jatkoon valittu piirre ei välttämättä ollut joukon paras.

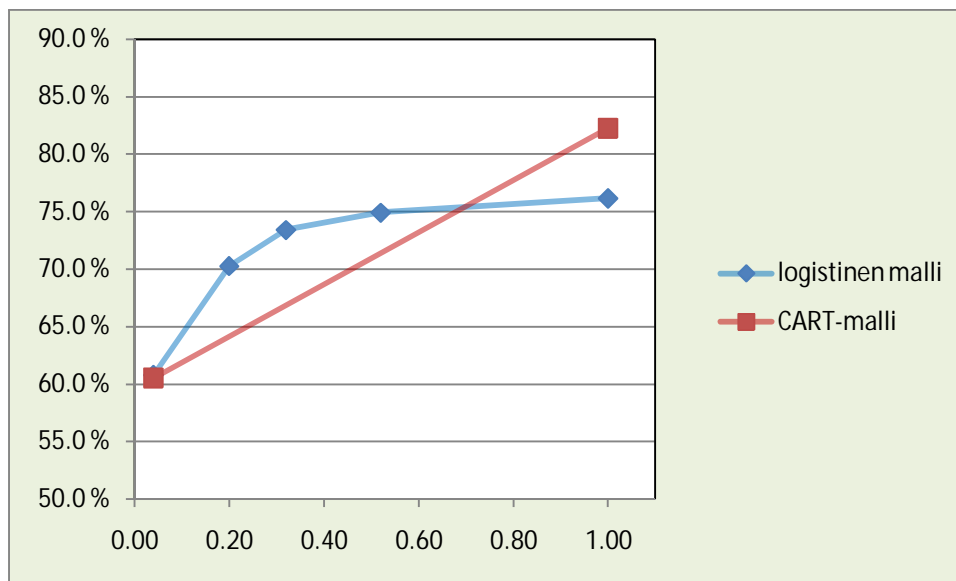
Logististen mallien avulla tehdyn luokituksen onnistumista kuvaava kappa-arvo osoittaa mallin luokittelukyvyyn parantuvan ruutukoon kasvaessa 0,32 hehtaarin ruutukokoon asti (kuva 3).



Kuva 3. Logistisen ja CART-mallin avulla luokiteltujen aineistojen kappa-arvot 0,04 – 1,0 hehtaarin ruuduilla.

Pienimmällä 0,04 hehtaarin ruutukoolla havaittu kappa-arvo oli vain 0,194, mikä viittaa heikkoon yhteyteen mallin tuottamien estimaattien ja todellisen lahopuumäärän välillä. Koska kappa-arvo oli kuitenkin selvästi nollan yläpuolella, voidaan todeta mallin tuovan lisäarvoa täysin satunnaiseen luokitukseen verrattuna jo kaikkein pienimmällä ruutukoolla. Kappa-arvolla mitattu luokituksen hyvyys saavuttaa maksiminsa 0,32 hehtaarin ruutukoolla. Kappa-arvo 0,38 kertoo luokituksen kohtuullisesta onnistumisesta (Landis ja Koch 1977). Tätä suuremmilla ruuduilla kappa-arvo alkaa pienentyä melko nopeasti. Osasyynä luokituksen huonontumiseen on luultavasti mallinnusaineiston ruutujen lukumäärä, joka pienenee ruutukoon suurentuessa.

Toinen luokituksen hyvyttä kuvaava tunnus eli oikeinluokitusprosentti kasvaa jokaisessa tapauksessa ruutukokoa suurennettaessa (kuva 4). Tunnus kuvaa kuitenkin huonosti mallin kykyä ennustaa lahopuun esiintymistä. Mikäli lahopuustoisten ja lahopuuttomien ruutujen määrät olisivat lähellä toisiaan, oikeinluokitusprosentti antaisi luotettavan arvion mallin hyvydestä. Tutkimusaineistossa kuitenkin lahopuuttomien ruutujen määrä pienenee ruutukoon kasvaessa ja lahopuumäärän keskiarvoistuessa. Tällöin malli voi tuottaa korkean oikeinluokitusprosentin vaikka yhtäkään lahopuutonta ruutua ei olisi paikannettu oikein. Oikeinluokitusprosentti kuvaakin luokituksen hyvyttä edes kohtuullisesti vain pienimmän kokoisilla ruuduilla, joilla lahopuuttomien ruutujen määrä on lähellä lahopuustoisten ruutujen määrää.



Kuva 4. Logistisen ja CART-mallin avulla luokiteltujen aineistojen oikeinluokitusprosentit 0,04 – 1,0 hehtaarin ruuduilla.

Luokituksen onnistuminen pienimmillä ruuduilla oli odotetun kaltaista. Metsiköiden puusto pystytään ALS-datan avulla ennustamaan melko tarkasti jo alle 0,04 hehtaarin pinta-aloille. Suvanto ym. (2005) saavuttivat tutkimuksessaan elävän puuston tilavuutta kuvaavalla prosenttipistemallilla alle 20 %:n suhteellisen RMSE:n koealatasolla. Koealan koko oli tutkimuksessa noin 0,025 hehtaaria. Lahopuumäärän ennustaminen ja kartoittaminen on kuitenkin elävää puustoa vaikeampaa. Tämä johtuu siitä, että laserkeilauksella saadaan elävästä puustosta suoria havaintoja, jolloin sen ominaisuuksia voidaan mallintaa suoraan, kun taas miltei kaikki lahoppuuta koskeva mallintaminen tapahtuu epäsuorasti elävän puuston kautta. Tutkimuksessa käytetyssä aineistossa kaikki alle 0,5 metrin korkeudelta heijastuneet kaiut luokiteltiin ALS-piirteitä laskettaessa maan pinnasta heijastuneiksi. Luoduilla malleilla ennustaminen perustuu käytännössä mallien kykyyn kuvata niitä metsikön piirteitä, jotka parhaiten selittävät metsikön lahoppuun määrää. Ainoan poikkeuksen tekevät pystyyn kuolleet puut, joista periaatteessa olisi mahdollista saada myös suoria havaintoja laserin avulla. Harvan pulssitiheyden ja pienen latvuskoon vuoksi pystyyn kuolleiden puiden erottuminen elävän puuston joukosta on hyvin epätodennäköistä. Epäsuoran mallintamisen ja lahoppuuston satunnaisen sijoittumisen yhteisvaikutuksen takia lahoppuuestimatit ovat elävää puustoa koskevia estimaatteja heikompia. 0,2 hehtaarin ruutukoolla luokitus onnistui huomattavasti pienintä ruutukokoa paremmin. Kappa-arvo kasvoi arvosta 0,19 arvoon

0,32. Tulosten paranemista selittää ainakin osittain lahopuumäärän hajonnan pieneneminen ruutukoon kasvaessa (taulukko 1).

Jos havaintoruutu koostuu useammasta kuin yhdestä metsiköstä, tuloksena saadaan keskiarvoinen ruutu, joka ei kuvaa todellista metsikköä. Ruutukoon suurentuessa todennäköisyys tällaiselle erilaisten metsiköiden sekoittumiselle kasvaa, mikä selittää luokitustulosten selvää heikkenemistä suuremmilla ruuduilla. Kaikkien ruutukokojen ruudut, suurinta 1,0 hehtaarin ruutukokoa lukuun ottamatta, koostettiin pienimmästä ruutukoosta 20 metriä leveinä kaistoina, joiden pituus muuttui ruutukoon kasvaessa. Ruutujen suurentuessa myös todennäköisyys kahden mittauskaistan saumaan osumiselle kasvaa. Tällöin perusruuduista koostettu suurempi ruutu koostuu väistämättä paitsi vähintään kahdesta eri metsiköstä, myös maantieteellisesti eri paikoista kerätystä aineistosta. Tällöin samaan ruutuun saattaa päätyä alueita esimerkiksi tuulelle alttiilta ja suojaisilta alueilta, joilla lahopuumäärät voivat olla hyvinkin erilaiset.

5.3 Luokitus CART-malleilla

Sekä kappa-arvoa että oikeinluokitusprosenttia vertailtaessa voidaan todeta, että testiaineiston luokittelu onnistui CART:n avulla selvästi logistista mallintamista paremmin kummallakin testatulla ruutukoolla (kuvat 3 ja 4). 0,04 hehtaarin ruuduilla CART-mallin avulla tehty luokitus tuotti kappa-arvon 0,24, mikä kertoo keskinkertaisesta luokituksen onnistumisesta. Hehtaarin ruutukoolla kappa-arvo oli 0,52, eli luokitusta voidaan tällöin pitää jo kohtuullisena. Logistisen ja CART-mallinnuksen välistä vertailua ei voida tehdä 0,2-0,5 hehtaarin ruuduille, koska kyseisiä ruutukokoja ei luokiteltu CART:n avulla.

Kun lahopuuston esiintymistä mallinnettiin hehtaarin ruutukoolle, jouduttiin CART-mallin luomista hieman yksinkertaistamaan. Lopulliseen luokitukseen valittiin mallinnuksen ensimmäisessä vaiheessa tuotettu malli. CART-mallinnuksen toinen vaihe eli riskitasoltaan korkeimpien puun haarojen poistaminen sivuutettiin, jolloin aineistoa jakavat laserpiirteet valittiin ainoastaan alkuperäisen CP:n avulla (Therneau ja Atkinson 1997). Ristiinvalidoinnilla saavutettavan pienimmän riskitason mallissa aineistoa ei olisi jaettu kertaakaan, jolloin kaikki ruudut olisi luokiteltu lahopuustoisiksi ja oikeinluokitusprosentiksi olisi tullut 72. Kappa-arvo olisi tällaisessa tapauksessa ollut 0, koska kaikki ruudut olisi ennustettu samaan luokkaan kuuluviksi. Erikoinen tulos johtui

luultavimmin liian pienestä mallinnusaineistosta, jossa ruutuja oli ainoastaan 72 kappaletta. Testiaineisto päätettiin luokitella karsimattoman mallin avulla, koska sen luokituskyky todettiin riippumattomassa aineistossa verrattain hyväksi. Ainoastaan yhtä lopputulosta ennustava malli olisi myös ollut lahopuun esiintymisen mallintamisen kannalta hyödytön.

CART-mallinnuksen ja logististen mallien avulla tehtyjä luokitteluja ei voida verrata suoraan, sillä logistisen mallin selittäjät olivat samat jokaisella ruutukoolla. CART taas jakoi aineiston jokaisella ruutukoolla parhaiten lahopuuston esiintymistä selittävien piirteiden suhteen. Jotta kahden menetelmän tuottamat luokitukset olisivat täysin vertailukelpoisia, myös logistisen mallin selittäjät olisi pitänyt valita jokaiselle ruutukoolle erikseen. Vaihtoehtoisesti CART-mallinnuksen yhteydessä olisi voitu valita logistista mallinnusta vastaava nollassa, jolla aineistoa jakavat piirteet olisi valittu. Mallin kiinnittäminen tiettyyn tasoon olisi kuitenkin ollut CART:n idean vastaista, sillä data-mining-menetelmänä se on parhaimmillaan juuri haluttua ominaisuutta parhaiten kuvaavien selittäjien etsinnässä.

Tutkimuskysymyksen kannalta jo logistisilla malleilla saadut tulokset riittivät ruutukoon vaikutuksen arviointiin. Näin ollen CART:n avulla voitiin syventää lahopuuston mallintamisen tarkastelua ilman, että ruutukoon vaikutuksen tutkiminen olisi asettanut rajoitteita mallinnuksen toteuttamiseen. Vaikka käytetyt mallinnusmenetelmät eivät olleetkaan tulostensa suhteen täysin vertailukelpoisia, niitä yhdistämällä pystyttiin tekemään päätelmiä lahopuun mallintamisesta. Sekä logistisessa että CART-mallinnuksessa ALS-piirre f_{h20} oli pienellä ruutukoolla selittävänä muuttujana, joskin logistisessa mallissa piirteestä käytettiin neliöjuurimuunnosta. Logistisen mallinnuksen tapauksessa piirre kiinnitettiin malliin, eikä suuremmilla ruuduilla mahdollisesti paremmin lahopuuta selittäviä muuttujia etsitty. CART:n avulla sen sijaan havaittiin, että ruutukoon suurentuessa toiset muuttujat (f_{hstd} ja I_{p1}) selittävät tutkittavaa ilmiötä pienellä ruutukoolla valittua muuttujaa (f_{h20}) paremmin. Tulos viittaa siihen, että myös logistisella mallinnuksella olisi todennäköisesti voitu saavuttaa parempia tuloksia nollassa suuremmilla ruuduilla, mikäli mallit olisi pelkän parametrisoinnin sijaan luotu kokonaan uudestaan jokaiselle ruutukoolle. Malleissa käytettyjen muuttujien vertailun lisäksi CART-mallinnuksen avulla saatiin estimaatit

parhaille mahdollisille luokituksille, joihin käytetyllä aineistolla voitiin päästä pienimmällä ja suurimmalla käytetyllä ruutukoolla.

5.4 CART-tilavuusestimaatit

Tutkimusalueen keskimääräinen lahopuutilavuus vaihteli laskennassa käytetystä ruutukoosta riippuen 2,7:stä 3,0 kuutiometriin hehtaarilla (taulukko 1). 0,04 hehtaarin ruuduilla tehdyssä luokituksessa 4845:sta luokitellusta solusta 4532 (93,5 %) kuului lahopuukeskiarvoiltaan pienimpään tilavuusluokkaan (2,354 m³/ha) (taulukko 7). Kun ennustettu lahopuukeskiarvo nousi, vastaaviin luokkiin kuuluvien ruutujen määrä pieneni odotetusti. Suuria lahopuutilavuuksia oli siis selvästi pieniä tilavuuksia vähemmän. Luokkien todelliset keskimääräiset lahopuutilavuudet kasvoivat mallinnettavan luokan odotetun lahopuutilavuuden kasvaessa. Toteutuneet keskiarvot eivät kuitenkaan yltäneet mallin ennustamiin keskiarvoihin kuin pienimmällä ruutukoolla. Suuremmilla ruuduilla odotetun ja toteutuneen lahopuukeskiarvon välinen ero kasvoi ruutukoon kasvaessa. Luokituksen mukainen testausaineiston keskimääräinen lahopuumäärä oli noin 3,0 kuutiometriä hehtaarilla, kun vastaava maastomittauksiin perustuva määrä oli noin 2,9 kuutiometriä hehtaarilla. CART-malliin perustuva ennuste siis yliarvioi alueen lahopuutilavuutta hieman (3,7 %). RMSE:tä tarkasteltaessa huomataan, että tilavuusennusteen virhe suurenee odotetusti mallin ennustaman keskimääräisen lahopuutilavuuden kasvaessa (taulukko 7). Mallinnettujen lahopuuluokkien suhteellinen RMSE kuitenkin pieneni kunkin luokan lahopuukeskiarvon kasvaessa. Lahopuukeskiarvoltaan pienimmän luokan RMSE oli miltei kaksinkertainen todelliseen lahopuukeskiarvoon verrattuna (suhteellinen RMSE 191,3 %) kun taas lahopuukeskiarvoltaan suurimman luokan RMSE oli noin neljäsosan havaittua lahopuukeskiarvoa suurempi (suhteellinen RMSE 126,8 %).

Yhden hehtaarin ruutukoolla 214:sta luokitellusta ruudusta 128 kuului kahdesta tilavuusluokasta pienempään (1,617 m³/ha) (taulukko 8). Vastaavasti suurempaan tilavuusluokkaan (4,999 m³/ha) kuului 86 ruutua. Koska hehtaarin ruutukoolla malli jakaa aineiston vain kerran, selittäjäksi valikoituneen ALS-piirteiden tulkitseminen on helppoa. Piirre I_{h99} kuvaa korkeutta, jolla viimeisten heijastuneiden laserpulssien kumulatiivinen kertymä on 99 %, eli arvot ovat todennäköisesti hyvin lähellä puuston valtapituutta tutkittavalla ruudulla. Tällöin malli siis jakaa aineiston kahteen lahopuuluokkaan elävän puuston pituuden perusteella siten, että lahopuustoltaan

niukempaan luokkaan kuuluvat valtapituudeltaan alle 15,74 metriset ruudut ja lahoppuustoltaan runsaampaan luokkaan valtapituudeltaan yli 15,74 metriset ruudut. Malli ennustaa odotetusti suuremman lahoppuutilavuuden ruuduille, jossa elävän puuston valtapituus on yli 15,74 metriä. Todelliset luokkakohtaiset lahoppuukeskiarvot olivat hyvin lähellä ennustettuja keskiarvoja. Pienemmässä luokassa ennustettu lahoppuukeskiarvo oli noin 7 % pienempi kuin todellinen. Vastaavasti suuremmassa luokassa ennustettu keskiarvo oli noin 10 % todellista suurempi. Luokituksen mukaiseksi testausaineiston lahoppuukeskiarvoksi muodostui noin 3,0 kuutiometriä hehtaarilla, kun vastaava maastomittauksiin perustuva määrä oli noin 2,8 kuutiometriä hehtaarilla. Aluetasolla hehtaarin ruutujen CART tilavuusestimaatit siis yliarvioivat lahoppuuston keskimääräistä tilavuutta hieman (5,3 %). Pienimmästä ruutukoosta poiketen hehtaarin ruutukoolle sekä lahoppuutilavuuden absoluuttinen että suhteellinen RMSE kasvoivat suurempaan tilavuusluokkaan siirryttäessä (taulukko 8).

CART-mallin tuottamat tilavuusestimaatit ovat kohtuullisia, kun otetaan huomioon tutkittavan ilmiön harvinaisuus ja satunnaisuus. Myös elävän puuston tilavuutta koskevat suhteelliset RMSE:t nousevat hyvin korkeiksi, mikäli tutkittavan puulajin esiintyminen on hyvin epäsäännöllistä. Esimerkiksi kuvioittaisella arvioinnilla estimoitujen puusto-ositteiden tilavuusestimaattien suhteellinen RMSE oli koivun osalta Haaran ja Korhosen (2004) tutkimuksessa 135,6 %. Koivun kokonaistilavuuden suhteellinen RMSE oli samassa tutkimuksessa 65 %. Kotamaa (2007) mallinsi maalahoppuun tilavuutta lineaarisella regressiolla Juuasta kerättyjen ALS-piirteiden avulla. Myös tässä tutkimuksessa lahoppuun suhteellinen RMSE nousi koealatasolla (0,025 ha) yli 140 prosenttiin.

5.5 Ruutukoon vaikutus lahoppuun inventoinnissa

Erikokoisille ruuduille täysin vertailukelpoiset ALS-piirteet tulisi laskea jokaiselle ruutukoolle perusruutukoon tapaan yksittäisten laserkaikujen avulla. Laskentojen yksinkertaistamiseksi tässä tutkimuksessa päädyttiin kuitenkin laskemaan suurempien ruutujen ALS-piirteet perusruutujen piirteiden keskiarvoina. Mallinnuksessa käytetyissä ALS-piirteissä ero laserkaikujen avulla laskettujen ja keskiarvoistettujen piirteiden välillä oletettiin hyvin pieneksi.

Lahopuun ennustaminen perustui ALS-datan kykyyn kuvata kohteena olevaa metsikköä. Ruutukoon muuttuessa irrotettujen ALS-piirteiden kyky kuvata kohteena olevaa metsikköä muuttui. Muutos johtui toisaalta siitä, että suuremmalta ruudulta saadaan enemmän havaintoja kohdemetsiköstä. Toisaalta ruutukoon suureneminen johti väistämättä siihen, että yhä suurempi määrä ruuduista koostui useammasta eri metsiköstä peräisin olevista peruseruuduista. Metsiköiden sekoittuminen suurilla ruuduilla sai aikaan tilanteen, jossa mallinnettava ruutu ei kuvannut enää todellista metsikköä, vaan kahden tai useamman metsikön sekoitusta.

Ruutukoon vaikutusta arvioitaessa on otettava huomioon erityisesti kaksi tuloksiin vaikuttavaa tekijää. Ensimmäinen koskee mallintamista logistisen regression avulla. Koska mallien avulla haluttiin saada tietoa ruutukoon vaikutuksesta, samoja muuttujia käytettiin jokaiselle ruutukoolle. Käytännössä tämä aiheutti kuitenkin ongelmia, sillä samat muuttujat eivät selittäneet lahopuun määrää pienillä ja suurilla ruuduilla. Toinen tulosten tulkinnassa huomioon otettava seikka on, että CART-mallinnuksessa kummallekin testatulle ruutukoolle valittiin omat selittävät muuttujat. Tällöin ruutukoon vaikutukseen ei päästä suoraan käsiksi vaan luokitustulokseen vaikuttaa myös selittäjien vaihtuminen. CART-mallien tulosten avulla voitiin joka tapauksessa arvioida käytetyn aineiston hyödyllisyyttä lahopuuston inventoinnissa ja määrittää vertailutaso logistisilla malleilla saaduille tuloksille.

Selittävien muuttujien selityskyvyn muutokset ruutukoon kasvaessa johtuvat mitä luultavimmin mallinnettavan lahopuuston hyvin erilaisesta esiintymisestä erikokoisilla ruuduilla. Pienillä ruuduilla lahopuutilavuuksien vaihtelu oli voimakasta: ruuduista täysin lahopuuttomia tai hyvin vähälahopuustoisia (alle 1 m³/ha) oli jopa 55 %. Samalla yksittäisten ruutujen lahopuutilavuudet nousivat lähes seitsemäänkymmeneen kuutiometriin hehtaarilla. Yhteys elävää puustoa kuvaavien piirteiden ja pienialaisten lahopuukeskittymien sijoittumisen välillä oli hyvin heikko. Mallinnettaessa lahopuuta suuremmilla ruuduilla lahopuumäärät lähestyivät koko alueen keskiarvoa. Suuret, yli neljäkymmenen kuutiometrin hehtaarikohtaiset lahopuutilavuudet hävisivät aineistosta kokonaan jo 0,2 hehtaarin ruutukokoon siirryttäessä ja hehtaarin ruutukoolla esiintyi maksimissaan enää noin 25 kuutiometrin hehtaarikohtaisia tilavuuksia. Samalla täysin lahopuuttomien ja vähälahopuustoisten ruutujen osuus väheni noin 25 %:iin. Edellä

kuvatut muutokset mallinnuksen kohteessa paransivat yksittäisten laserpiirteiden selityskykyä.

5.6 Laserkeilauksen mahdollisuudet lahopuun inventoinnissa

Saatujen tulosten perusteella voidaan todeta, että laserkeilausaineistolla voidaan löytää lahopuustoiset kohteet kohtuullisen tarkasti, mikäli tarkastelussa käytetty mittakaava on riittävän suuri. Sekä logistisen että CART-mallintamisen avulla saatujen tulosten perusteella havaittiin, että pienialaisen vaihtelun kuvaaminen elävän puuston rakennetta kuvaavan ALS-datan avulla on hyvin epävarmaa. Lahopuumäärän ja elävän puuston välinen riippuvuus on hyvin satunnainen pienillä pinta-aloilla. Sekä lahopuuluokituksen hyvyttä että tilavuustunnusten keskivirhettä kuvaavat tunnuksat paranevat selvästi ruutukoon kasvaessa. Pesonen ym. (2010) saivat tutkimuksessa samansuuntaisia tuloksia: parhaiden lahopuumäärää selittävien ALS-piirteiden selitysasteet paranivat selvästi ruutukoon kasvaessa. Luotettaviin tuloksiin vaadittava suuri ruutukoko tekee kuitenkin pienialaisten lahopuukeskittymien havaitsemisesta mahdotonta. Pienialaiset, mutta runsaat lahopuukeskittymät katoavat suuriin ruutuihin keskiarvoistumisen johdosta. Suurelle ruudulle ennustettu keskimääräistä korkeampi lahopuumäärä voi tällöin yhtä hyvin viitata yksittäisiin runsalahopuustoisiin kohteisiin tutkittavan ruudun sisällä tai tasaiseen, keskimääräistä runsaampaan lahopuumäärään.

Tutkimuksen perusteella harvapulssisen ALS-datan avulla ei pystytä paikantamaan pienialaisia lahopuukohteita. Tämä tarkoittaa, että esimerkiksi metsikkökuvion sisältä ei pystytä paikantamaan mahdollista monimuotoisuuden kannalta merkittävää lahopuukeskittymää. Lahopuun pienialainen vaihtelu on metsikkötunnuksiin nähden satunnaista ja sen paikantaminen vaatisi joko tarkempaa tietoa latvustossa mahdollisesti havaittavista aukoista (Vehmas ym. 2009) tai suoria havaintoja lahopuusta. Aluskasvillisuus ja vallitseva latvuserros tekevät suorasta havainnoinnista vaikeaa tai jopa mahdotonta, mutta luotettavampiin tuloksiin päästäisiin luultavasti tiheämmän ALS-datan mukanaan tuomilla mahdollisuuksilla. Tiheämmän datan avulla voitaisiin tarkastella elävässä latvuksessa havaittuja pienialaisiakin muutoksia ja käyttää niitä selittäjinä lahopuun esiintymistä mallinnettaessa. Latvusaukkojen lisäksi mahdollisia selittäjiä voisivat olla metsän terveyttä kuvaavat piirteet. Toisaalta eri ajankohtina tehtyjen laserkeilausten välisellä muutostulkinnalla voitaisiin löytää jopa yksittäiset tarkastelujakson aikana kaatuneet puut (Yu ym. 2004). Tämä voisi mahdollistaa paitsi

kuolleen puun tarkan paikantamisen, myös sen tilavuuden ja puulajin estimoinnin yksinpuintulkinnan keinoin.

Hieman suuremman mittakaavan lahopuuinventoinneissa, joissa vaadittu tarkkuustaso noudattelisi esimerkiksi kuviokokoa, myös harvapulssisen ALS-datan avulla voitaisiin saada kohtuullisia tuloksia. Koivuniemen ja Korhosen (2006) mukaan keskimääräinen metsikkökuvion koko on Etelä-Suomessa noin 1,5-2 hehtaaria. Tämän tutkimuksen tulosten perusteella jo hehtaarin pinta-aloilla voitaisiin erotella vähälahopuustoiset metsikkökuviot runsaampilahopuustoisista kohtuullisella tarkkuudella (taulukko 8). Tutkimuksen perusteella ei ruutupohjaisen lähestymistavan johdosta pystytä suoraan ottamaan kantaa siihen, miten lahopuun ennustaminen onnistuisi kuviotasolla. Todennäköistä kuitenkin on, että lahopuumäärän estimaatit tarkentuisivat, mikäli menetelmää pystyttäisiin soveltamaan homogeenisille metsikkökuvioille. Yksittäisten kuvioiden lahopuumäärän arviointi vaatisi käytännössä ALS-aineiston segmentointia siten, että muodostuneet segmentit kuvaisivat mahdollisimman hyvin todellisia metsikkökuvioita. Tällöin kuviokohtainen lahopuumäärä voitaisiin luultavasti ennustaa kohtuullisella tarkkuudella. Ennusteiden tarkkuuden kannalta kuviokokoo nousisi kuitenkin merkittäväksi tekijäksi: pienille kuvioille tehtyjen ennusteiden tarkkuus olisi suurten kuvioiden ennusteita heikompi. Kuviokohtaisessa tarkastelussa joidenkin jo olemassa olevien kuviotietojen, kuten esimerkiksi kasvupaikan liittäminen lahopuumalliin voisi parantaa ennusteiden luotettavuutta huomattavasti.

6. Johtopäätökset

Tämän tutkimuksen tavoitteena oli mallintaa metsäalueen lahopuun määrää harvapulssiseen ALS-datan avulla sekä tarkastella mallinnuksessa käytettävän ruutukoon vaikutusta aluepohjaisella menetelmällä saatuihin lahopuuestimaatteihin. Ennako-oletusten mukaisesti sekä lahopuustoisten ruutujen luokittelu että ruutukohtaiset estimaatit lahopuun tilavuudesta paranivat ruutukoon kasvaessa. Tulosten perusteella voidaan todeta, että talousmetsissä oleva lahopuu ja harvapulssisesta laseraineistosta irrotetut ALS-piirteet korreloivat keskenään hyvin heikosti pienillä ruuduilla. Ruutukoon suurentuessa ALS-piirteiden selityskyky paranee, mutta samalla menetetään tieto pienistä runsalahopuustoisista kohteista. Pienialaisten lahopuukeskittymien paikantaminen harvapulssisen ALS-datan avulla ei onnistunut tutkimuksessa käytetyillä menetelmillä. Jatkotutkimus tulisi suunnata tiheäpulssista ALS-dataa hyödyntäviin menetelmiin, joilla elävän puuston latvusrakenne pystytään kuvaamaan tarkemmin.

Lähteet

- Axelsson, P. 2000. DEM generation from laser scanner data using adaptive TIN models. *International Archives of Photogrammetry and Remote Sensing* 33: 110-117.
- Bater, C., Coops, N., Gergel, S., LeMay, V., Collins, D. 2009. Estimation of standing dead tree class distribution in northwest coastal forests using lidar remote sensing. *Canadian Journal of Forest Research* 39: 1080-1091.
- Breiman, L., Friedman, J. H., Olshen, R. D., Stone, C. J. 1984. *Classification and Regression Trees*. Taylor & Francis Group, Boca Raton, Fl. 358 s.
- Campbell, J. 2007 *Introduction to remote sensing*. Guildford Press, New York. 626 s.
- Dobson, A. 1990. *An Introduction to Generalized Linear Models*. Chapman and Hall, Cornwall. 174 s.
- Eid, T., Øyen, B.-H. 2003. Models for Prediction of Mortality in Even-aged Forests. *Scandinavian Journal of Forest Research* 18: 64-77.
- Fan, Z., Larsen, D. R., Shifley, S. R., Thompson, F. R. 2003. Estimating cavity tree abundance by stand age and basal area, Missouri, USA. *Forest Ecology and Management* 179: 231-242.
- Gibb, H., Ball, J., Johansson, T., Atlegrim, O., Hjätèn, J., Danell, K. 2005. Effects of management on coarse woody debris volume and composition in boreal forests of northern Sweden. *Scandinavian Journal of Forest Research* 20: 213-222.
- Gough, C., Vogel, C., Kazanski, C., Nagel, L., Flower, C., Curtis, P. 2007. Coarse woody debris and the carbon balance of a north temperate forest. *Forest Ecology & Management* 244: 60-67.
- Gove, J., Ducey, M., Valentine, H. 2002. Multistage point relascope and randomized branch sampling for downed coarse woody debris estimation. *Forest Ecology & Management* 155: 153-162.
- Haara, A., Korhonen, K. T. 2004. Kuvioittaisen arvioinnin luotettavuus. *Metsätieteen aikakauskirja* 4/2004: 489-508.
- Harmon, M. E., Franklin, J. F., Swanson, F. J., Sollins, P., Gregory, S. V., Lattin, J. D., Anderson, N. H., Cline, S. P., Aumen, N. G., Sedell, J. R., Lienkaemper, G. W., Cromack, K. JR., Cummins, K. W. 1986. Ecology of Coarse Woody Debris in Temperate Ecosystems. *Advances in Ecological Research* 15: 133-302.
- Holopainen, M., Vastaranta, M., Rasinmäki, J., Kalliovirta, J., Mäkinen, A., Haapanen, R., Melkas, T., Yu, X., Hyypä, J. 2010. Uncertainty in timber assortment estimates predicted from forest inventory data. *European Journal of Forest Research* 129: 1131-1142.

- , Talvitie, M., 2006. Effects of Data Acquisition Accuracy on Timing of Stand Harvests and Expected Net Present Value. *Silva Fennica* 40: 531-543.
- Hosmer, D. & Lemeshow, S. 2000. *Applied Logistic Regression*. John Wiley & Sons, New York. 375s.
- Hyyppä, H. J., Hyyppä, J. M. 2001. Effects of Stand Size on the Remote Sensing-Based Forest Inventory. *IEEE Transactions on Geoscience and Remote Sensing* 39: 2613-2621.
- Hyyppä, J., Inkinen, M. 1999. Detecting and estimating attributes for single trees using laser scanner. *The Photogrammetric Journal of Finland* 16(2): 27-42.
- Ihalainen, A., Mäkelä, H. 2009. Kuolleen puuston määrä Etelä- ja Pohjois-Suomessa 2004–2007. *Metsätieteen aikakauskirja* 1/2009: 35-56.
- Kangas, A. 2006. Sampling rare populations. Teoksessa: Kangas, A. & Maltamo, M. (toim.) *Managing Forest Ecosystems. Forest Inventory–Methodology and Applications*. Springer, Dordrecht. 362 s.
- 2001. Tietolaatikko 8.1. Logistinen regressio. Teoksessa: Maltamo, M., Laukkanen, S. (toim.) *Metsää kuvaavat mallit*. *Silva Carelica* 36. Joensuun yliopisto. 239 s.
- Karjalainen, L. & Kuuluvainen, T. 2002. Amount and Diversity of Coarse Woody Debris within a Boreal Landscape Dominated by *Pinus sylvestris* in Vienansalo Wilderness, Eastern Fennoscandia. *Silva Fennica* 36: 147-167.
- Koivuniemi, J., Korhonen, K. 2006 Inventory by compartments. Teoksessa: Kangas, A. & Maltamo, M. (toim.) *Managing Forest Ecosystems. Forest Inventory–Methodology and Applications*. Springer, Dordrecht. 362 s.
- 2003. Metsiköihin ja paikannettuihin koealoihin perustuvan kuvioittaisen arvioinnin tarkkuus. *Helsingin yliopiston metsävarojen käytön laitoksen julkaisu* 36. 160 s.
- Korpela, I., Ørka, H. O., Maltamo, M., Tokola, T., Hyyppä, J. 2010. Tree Species Classification Using Airborne LiDAR – Effects of Stand and Tree Parameters, Downsizing of Training Set, Intensity Normalization, and Sensor Type. *Silva Fennica* 44(2): 319-339.
- Kotamaa, E. 2007. Maalahopuun määrän ennustaminen laserkeilauksella talousmetsissä. Kandidaatin tutkielma. Joensuun yliopiston metsätieteellinen tiedekunta. 35 s.
- Laasasenaho, J. 1982. Taper Curve and Volume Functions for Pine, Spruce and Birch. *Communicationes Instituti Forestalis Fenniae* 108, 74 s.
- Laki metsän hyönteis- ja sienituhojen torjunnasta. 8.2.1991/263.
- Landis, R. J., Koch, G. G. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33: 159-174.

- Lim, K., Treitz, P., Wulder, M., St-Onge, B., Flood, M. 2003. LiDAR remote sensing of forest structure. *Progress in Physical Geography* 27(1): 88-106.
- Linder, P., Elfving, B., Zacrisson, O. 1997. Stand structure successional trends in virgin boreal forest reserves in Sweden. *Forest Ecology & Management* 98: 17-33.
- Loetsch, F., Zöhrer, F., Haller, K. E. 1973. *Forest Inventory*. BLV Verlagsgesellschaft mbH, München. 469 s.
- Næsset, E. 2004. Practical Large-scale Forest Stand Inventory Using a Small-footprint Airborne Scanning Laser. *Scandinavian Journal of Forest Research* 19: 164-179.
- 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote sensing of environment* 80: 88-99.
- 1999. Decomposition rate constants of *Picea abies* log in southeastern Norway. *Canadian Journal of Forest Research* 29: 372-381.
- 1997. Estimating Timber Volume of Forest Stands Using Airborne Laser Scanner Data. *Remote Sensing of Environment* 61: 246-253.
- Packalén, P., Maltamo, M. 2007. The k-MSN method for the prediction of species-specific stand attributes using airborne laser scanning and aerial photographs. *Remote Sensing of Environment* 109: 328-341.
- Pasher, J., King, D. 2009. Mapping dead wood distribution in a temperate hardwood forest using high resolution airborne imagery. *Forest Ecology & Management* 258: 1536–1548.
- Pesonen, A., Maltamo, M., Kangas, A. 2010a. The comparison of airborne laser scanning-based probability layers as auxiliary information for assessing coarse woody debris. *International Journal of Remote Sensing* 31: 1245-1259.
- , Kangas, A., Maltamo, M., Packalén, P., 2010b. Effects of auxiliary data source and inventory unit size on the efficiency of sample-based coarse woody debris inventory. *Forest Ecology and Management* 259: 1890-1899.
- , Leino, O., Maltamo, M., Kangas, A. 2009. Comparison of field sampling methods for assessing coarse woody debris and use of airborne laser scanning as auxiliary information. *Forest Ecology & Management* 257: 1532-1541.
- , Maltamo, M., Eerikäinen, K., Packalèn, P. 2008. Airborne laser scanning-based prediction of coarse woody debris volumes in a conservation area. *Forest Ecology & Management* 255: 3288-3296.
- Punttila, P. 2000. Metsien suojelualueverkon merkitys lahopuukovakuoriaisten elinkelpoisten populaatioiden säilymiselle Etelä-Suomessa. *Suomen ympäristö* 440: 49-96.

Ranius, T., Jonsson, B. G., Kruys, N. 2004. Modelling dead wood in Fennoscandian old-growth forests dominated by Norway spruce. *Canadian Journal of Forest Research* 34: 1025-1034.

R Development Core Team. 2007. R: A Language and Environment for Statistical Computing. www.r-project.org (viitattu 28.3.2011).

Schreuder, H. T., Gregoire, T. G., Wood, G. B. 1993. *Sampling Methods for Multiresource Forest Inventory*. Wiley & Sons, New York. 446 s.

Siitonen, J., Hottola, J., Immonen, A. 2009. Differences in stand characteristics between brook-side key habitats and managed forests in southern Finland. *Silva Fennica* 43: 21–37.

Siitonen, J., 2001. Forest management, coarse woody debris and saproxylic organisms: Fennoscandian boreal forests as an example. *Ecological bulletins* 49: 11-41.

Stokland, J. 2001. The coarse woody debris profile: an archive of recent forest history and an important biodiversity indicator. *Ecological Bulletins* 49: 71-83.

Ståhl, G. 1998. Transect Relascope Sampling – A Method For the Quantification of Coarse Woody Debris. *Forest Science* 44: 58-63.

Suvanto, M., Maltamo, M., Packalén, P., Kangas, J. 2005. Kuviokohtaisten puustotunnusten ennustaminen laserkeilauksella. *Metsätieteen aikakauskirja* 4/2005: 413-428.

Therneau, T. M., Atkinson, E. J. 1997. An Introduction to Recursive Partitioning Using the RPART Routines. www.mayo.edu/hsr/techrpt/61.pdf (viitattu 28.3.2011).

Tomppo, E., Henttonen, H., Korhonen, K. T., Aarnio, A., Ahola, A., Heikkinen, J., Ihalainen, A., Mikkilä, H., Tonteri, T., Tuomainen, T. 1998a. Etelä-Pohjanmaan metsäkeskuksen alueen metsävarat ja niiden kehitys 1968–97. *Metsätieteen aikakauskirja* 2B/1998: 293-374.

–, Katila, M., Moilanen, J., Mäkelä, H., Peräsaari, J. 1998b. Kunnittaiset metsävaratiedot 1990–94. *Metsätieteen aikakauskirja* 4B/1998: 619-839.

Uotila, A., Maltamo, M., Uuttera, J., Isomäki, A. 2001. Stand structure in semi-natural and managed forests in eastern Finland and Russian Karelia. *Ecological Bulletins* 49: 149-158.

Uuttera, J., Anttila, P., Suvanto, A., Maltamo, M. 2006. Yksityismetsien metsävaratiedon keruuseen soveltuvilla kaukokartoitusmenetelmillä estimoitujen puustotunnusten luotettavuus. *Metsätieteen aikakauskirja* 4/2006: 507-519.

Vastaranta, M., Holopainen, M., Yu, X., Hyyppä, J., Hyyppä, H., Viitala, R. 2011. Predicting stand-thinning maturity from airborne laser scanning data. *Scandinavian Journal of Forest Research* 26: 187-196.

Vehmas, M., Packalén, P., Maltamo, M. 2009. Assessing deadwood existence in canopy gaps by using ALS data. Teoksessa: (toim.), *SilviLaser 2009*, October 14-16, College Station, Texas USA: 137-144.

Wehr, A., Lohr, U. 1999. Airborne laser scanning—an introduction and overview. *ISPRS Journal of Photogrammetry & Remote Sensing* 54: 68-82.

Yli-Kojola, H. 2002. Tuulituhojen esiintyminen ja ennustaminen metsikkö- ja puutasolla Suomessa. *Metsänarvoimistieteen lisensiaatin työ*. Helsingin yliopisto, Metsävarojen käytön laitos. 102 s.

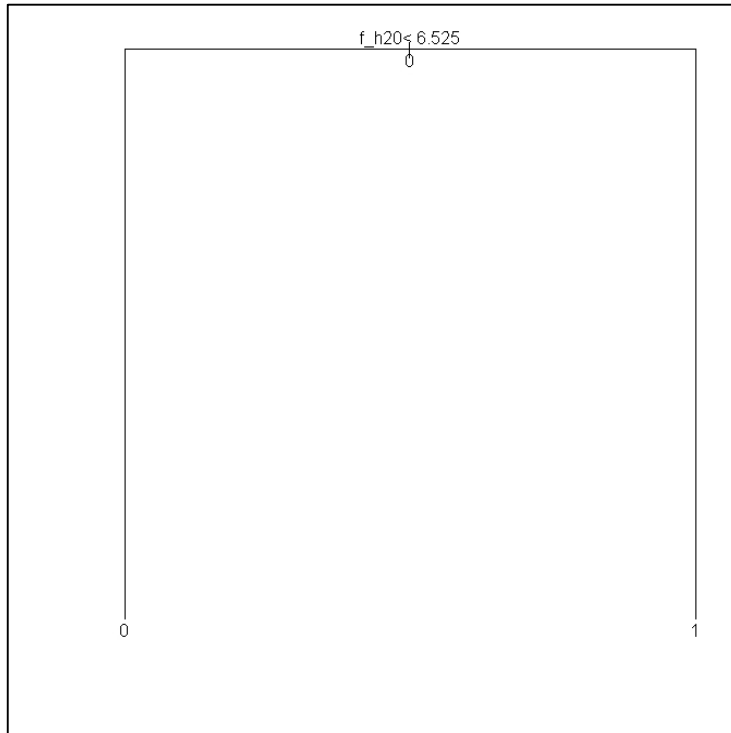
Yu, X., Hyypä, J., Vastaranta, M., Holopainen, M., Viitala, R. 2011. Predicting individual tree attributes from airborne laser point clouds based on random forests technique. *ISPRS Journal of Photogrammetry & Remote Sensing* 66: 28-37.

–, Hyypä, J., Kaartinen, H., Maltamo, M. 2004. Automatic detection of harvested trees and determination of forest growth using airborne laser scanning. *Remote Sensing of Environment* 90: 451-462.

Liitteet

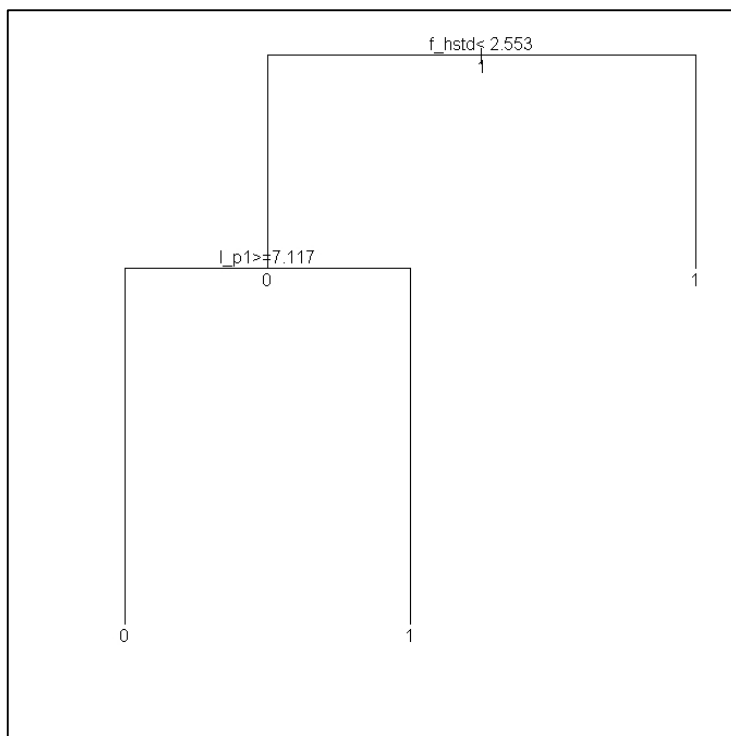
Liite 1

CART-malli 0,04 hehtaarin ruutukoolle: luokiteltavien ruutujen lahpuustoisuus



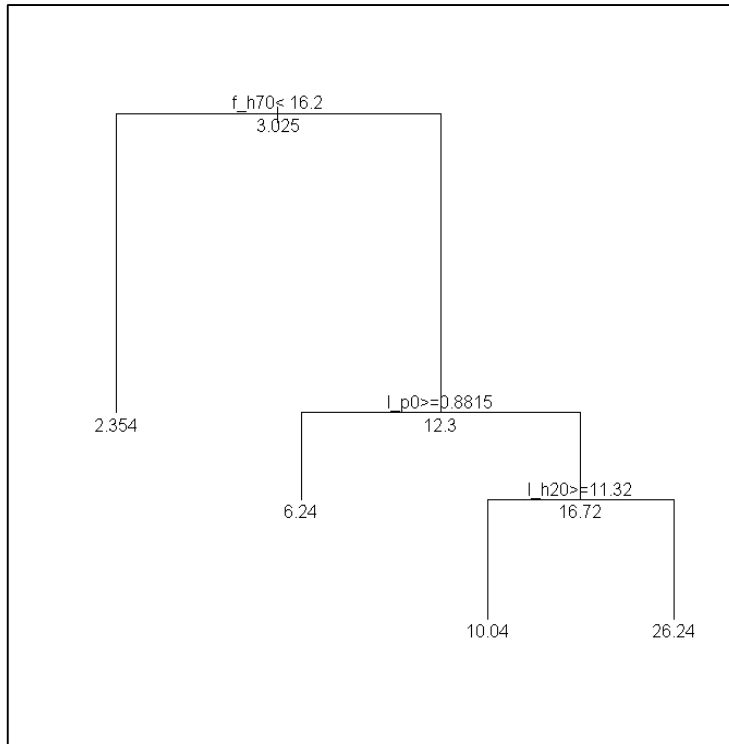
Liite 2

CART-malli 1,0 hehtaarin ruutukoolle: luokiteltavien ruutujen lahpuustoisuus



Liite 3

CART-malli 0,04 hehtaarin ruutukoolle: ruutukohtainen lahpuutilavuus



Liite 4

CART-malli 1,0 hehtaarin ruutukoolle: ruutukohtainen lahpuutilavuus

