



Arto Haara

Arto Haara

Metsävaratietojen laskennallinen ajantasaistus yleistyy metsäsuunnittelussa – onko luotettavuuden arviointiin mahdollisuuksia?

Haara, A. 2002. Metsävaratietojen laskennallinen ajantasaistus yleistyy metsäsuunnittelussa – onko luotettavuuden arviointiin mahdollisuuksia? *Metsätieteen aikakauskirja* 3/2002: 479–492.

Metsätalouden eri osa-alueilla pyritään jatkuvasti kustannussäästöihin. Kustannusten vähentämiseen tähtävien toimenpiteiden vaikutukset jäävät kuitenkin valitettavan usein selvittämättä. Yksityismetsien metsäsuunnittelun tietotarpeita tyydyttävään kuvioittaiseen arviointiin on haettu viime aikoina kustannussäästöjä erilaisia korvaavia ja täydentäviä arvioimismenetelmiä testaamalla. Metsäkeskusten suorittamassa yksityismetsien metsäsuunnittelussa metsälön kaikki kuviot käydään läpi maastoinventoinneilla noin kymmenen vuoden välein. Laskennallisessa ajantasaistuksessa kuvioaineistoa päivitetään tilastollisilla malleilla toimenpiteet huomioon ottaen. Tällöin rajoiltaan muuttumattomilla kuviolla ei tarvitse suorittaa puustomittauksia jokaisella inventointikerralla. Kasvuennusteiden luotettavuusarvioita ei kuitenkaan tehdä, vaan ajantasaistettua kuvioaineistoa käytetään kuten varsinaista uutta inventointiaineistoa.

Tässä katsauksessa tarkastellaan erilaisia menetelmiä, joiden avulla voidaan tuottaa ajantasaistetuille tunnuksille luotettavuusarvioita. Nämä luotettavuusarviot palvelevat tilanteissa, joissa olisi päätettävä, käytetäänkö kuviotiedoston ajantasaistusta kuvioittaisen inventoinnin tukena tai korvaamassa osittain uutta inventointitietoa. Suunnittelijalle luotettavuusarviot antavat tukea päätöstilanteessa, jossa valitaan kuviot, joille ajantasaistus voidaan tehdä luotettavasti.

Asiasanat: ennustaminen, epävarmuus, kuvioittainen arviointi, laskennallinen ajantasaistus

Yhteystiedot: Metla, Joensuun tutkimuskeskus, PL 68, 80101 Joensuu

Sähköposti arto.haara@metla.fi

Hyväksytty 18.9.2002

I Johdanto

Metsätalouden suunnittelu- ja tutkimustehtäviin käytettävät ohjelmistot sisältävät yleensä monimutkaisen simulointijärjestelmän, jolla kuvataan metsikön kehitystä ja simuloitujen toimenpiteiden vaikutusta tähän kehitykseen (mm. Eid 1990, Jons-son 1993, Siitonen 1993, 1996). Järjestelmät sisältävät mm. kuolleisuus-, uudistumis-, kasvu- ja tuot- tosmalleja, joiden tuottaman informaation ja erilais- ten metsänkäsittelyvaihtoehtojen simuloinnin avul- la saadaan selville metsän tuotantomahdollisuudet. Suunnittelussa ja päätöksenteossa tarvitaan arvioita myös mallien ennusteiden ja niiden perusteella las- kettujen tuotantomahdollisuuksien (esim. hakkuu- mahdollisuusarvioiden) luotettavuudesta. Niiden tuottaminen voi kuitenkin olla vaikeaa jo järjestel- mien monimutkaisuuden vuoksi.

Kehitysenusteiden ja suunnittelulaskelmien läh- tötiedoiksi tarvitaan paikkaansidottua metsävara- tietoa, jota Suomessa kerätään yleisesti kuvioit- taisella arvioinnilla. Alueellisten metsäkeskusten tekemä yksityismetsien alueellinen suunnittelu kattaa n. 72 % yksityismetsien pinta-alasta ja vuo- sittain suunnitellaan vajaa 1 000 000 hehtaaria (Ok- sanen-Peltola 1999). Vuonna 1999 ajantasaistet tila- kohtaiset suunnitelmat kattoivat 63 prosenttia yk- sityismetsien pinta-alasta (Karppinen ym. 2002). Alueellisessa suunnittelussa tiloja, joille ei tilata metsäsuunnitelmaa, kutsutaan välialueiksi. Alueel- lisessä ja tilakohtaisessa suunnittelussa käytetään samaa perustietosisältöä.

Laskennallisen ajantasaistuksen käyttö on yleisty- mässä kuvioittaisen arvioinnin yhteydessä (Anttila 2002). Laskennallisessa ajantasaistuksessa metsäva- ratiedot päivitetään malleilla simuloiden myös teh- tyjen toimenpiteiden vaikutukset. Ajantasaistuksen tuloksia voidaan käyttää suunnittelijan tukena uu- dessa inventoinnissa arvioitaessa metsikkötunnuk- sia (Anttila 2002). Osalla kuvioista ajantasaistettu- ja metsävaratietoja voidaan käyttää myös suoraan metsäsuunnittelun laskennassa.

Laskennallisen ajantasaistuksen luotettavuuden arviointi on vaikeaa. Jo ajantasaistuksen alkuvaihe on ongelmallinen: lähtötietoina käytettävän maas- toaineiston tarkkuus vaihtelee huomattavasti, eikä vaihtelun suuruus ole yleensä tiedossa. Kuvioittai- sen arvioinnin maastotietojen keruussa tapahtui huo-

mattava muutos 1990-luvun puolivälissä. 1990-lu- vun puolen välin yli ulottuneessa TASO-tietojen keruussa koko kuviolle arvioitiin yhdistetyt, keski- määräiset puustotunnukset (Taso. Maastotyöopas... 1993). Mikäli kyseessä oli sekametsä, eri puulajien prosenttiosuudet pohjapinta-alan tai tilavuuden mu- kaan merkittiin kymmenen prosenttiyksikön tark- kuudella. Uudessa SOLMU-muotoisessa inventoin- nissa puustotiedot kerätään puulajeittain ja -jaksoit- tain (SOLMU Maastotyöopas... 1996).

Ajantasaistuksessa TASO-tiedot muutetaan SOLMU-tiedoiksi menetelmällä, jossa kaikki puu- lajiosuudet saavat samat keskitunnukset ja pohja- pinta-ala jaetaan kunkin ositteen prosenttiosuuden mukaisesti. Tästä muunnostyöstä aiheutuu virhettä sekametsikköaineistoihin, kun metsikön pohjapinta- alamediaanipuun tunnuksat (läpimitta, pituus ja ikä) yleistetään kunkin puusto-ositteen mediaanipuulle. Nämä muunnosvirheet lisäävät omalta osaltaan en- nestäänkin virheellisen inventointiaineiston hajon- ta. Ongelmasta ollaan kuitenkin vähitellen pääse- mässä eroon, kunhan viimeisiltäkin metsätiloilta, joilta on nykyisin vain TASO-tiedot, saadaan uu- det SOLMU-aineistot.

Kuviorajojen vaihtuvuus tuottaa myös ongelmia laskennallisessa ajantasaistuksessa. Mikäli kuvioin- ti muuttuu, on vaikeutena yhdistää ajantasaistetut tiedot ja uudet kuviorajat. Kuvioinnin muutosten ongelma korostuu metsiköissä, joissa entisten ku- vioiden sisällä on paljon hajontaa tai uudet kuviot on saatu yhdistelemällä vanhoja. Kuvioinnin muut- tumista ja kuvioiden rajaamisongelmaa on tutkittu melko vähän (esim. Hyppänen ym. 1996). Lasken- nallisen ajantasaistuksen käytölle on välttämätöntä vanhan ja uuden kuvioinnin läpikäynti, mikäli ha- lutaan kohdentaa tulokset kuvioille. Läpikäynnin aikana suunnittelijan on päätettävä mille kuvioil- le vanhojen tietojen ajantasaistusta voidaan käyt- tää suoraan, ja mille kuvioille joudutaan yhdistele- mään eri kuvioiden ajantasaistettuja tietoja. Lisäk- si ilmakuvulle siirretyn kuviokartan visuaalisessa tarkastelussa pitäisi löytää ongelmakuviot, kuten esimerkiksi käsitellyt kuviot ja kuviot, joiden ajan- tasaistetut tiedot näyttävät virheellisiltä, tarkempaa tarkastelua tai maastoarviointia varten.

Laskennallinen ajantasaistus ja ilmakuvien vi- suaalinen tulkinta yhdessä ovat antaneet lupaavia tuloksia (esim. Anttila 2002). Mikäli tehtyjen toi-

menpiteiden luotettava tunnistus saadaan myös liitettyä mukaan, voi menetelmän käytölle olla hyvinkin käyttöä kuvioittaisen arvioinnin inventointitietojen tuottamisessa. Laskennallista ajantasaistusta voidaan myös käyttää muidenkin kaukokartoitusmenetelmien apuna. Ajantasaistettuja tietoja voitaisiin käyttää esim. VMI:n koalatietojen päivytyksessä (esim. Kangas 1991), satelliittikuvatulkinnan referenssikoeala-aineistoina (esim. Tomppo 1990, Tokola ym. 1996) tai yksittäisen puun hahmontunnistukseen perustuvan inventointimenetelmän kalibrointitietona (esim. Uttera ym. 1998, Haara ja Nevalainen 2002).

Inventointiajankohtien välillä kuviolle tehdyt toimenpiteet, esim. harvennushakkuut, on simuloitava ajantasaistuksessa, koska muutoin kuviolla tapahtunutta kehitystä ei pystytä kuvaamaan. Käsittelyjen vaikutuksen simulointia helpottaa huomattavasti, mikäli käytettävissä on ajan tasalla olevat kuvioiden käsittelytiedot esim. luonnonvaratietokannassa. Suunnittelija joutuu kuitenkin useimmiten käyttämään aiempia toimenpide-ehdotuksia, visuaalista arviointia ja tiedustelua maanomistajalta inventointikausien välillä tehtyjen toimenpiteiden selvittämiseen. Metsänhoitoyhdistysten apua voitaisiin käyttää varsinkin välialueiden osalta tehtyjen toimenpiteiden selvityksessä. Lisäksi metsäyhtiöiden metsäpalvelusopimustiloilta löytyy tiedot myös tehdyistä toimenpiteistä, tosin tämän aineiston hyödyntäminen voi olla vaikeampaa. Satelliittikuviakin on mahdollista hyödyntää tehtyjen toimenpiteiden ja ongelmakuvioiden kartoituksessa (esim. Saukkola 1982, Varjo 1996, 1997).

Kuvioittaisessa arvioinnissa mitattavissa tunnuksissa esiintyy vaihtelevan suuruisia mittaus- ja arviointivirheitä (mm. Poso 1983, Laasasenaho ja Päivinen 1984, Ståhl 1992, Pigg 1994). Virheet voivat olla satunnaisia ja/tai systemaattisia. Ajantasaistuksen kannalta systemaattiset virheet aiheuttavat enemmän ongelmia kuin satunnaiset, mikäli niiden suuruutta ja suuntaa ei tiedetä. Koska kuvioittaisen arvioinnin virhettä ei tunneta edes keskimäärin, puhumattakaan tietystä kuviosta, on odotettavissa ongelmia, mikäli suunnittelija käyttää ajantasaistettuja tietoja kuvioittaisen arvioinnin aputietoina. Suunnittelijalla tulisi olla tietty luottamus aikaisemmin tehtyyn inventointiin ja ajantasaistukseen. Tätä luottamusta heikentää juuri tietämättömyys mittaus-

virheiden tasosta. Kokeilun arvoisena voisi pitää suunnittelualueiden ajallista jatkuvuutta; ts. sama suunnittelija kävisi uudessa inventoinnissa vanhat suunnittelualueensa läpi vanhojen ajantasaistettujen inventointitietojensa kanssa, mikäli mahdollista.

Virhettä lähtötietoihin tuovat (Gertner 1986):

- 1) Mittausvirheet: sovellustilanteissa aineiston mittausvirheet ovat yleensä suurempia kuin mallien parametrien estimointiin käytetyssä havaintoaineistossa.
- 2) Otantavirheet: kun malli estimoidaan koko populaatiolle jollakin otantamenetelmällä hankitusta osajoukosta, sisältävät saadut tulokset otantavirhettä, jonka suuruuteen vaikuttaa mm. otoksen koko, koealan koko ja otantamenetelmä.
- 3) Ennustamisvirheet: käytettäessä toisten mallien selitettäviä muuttujia selittävinä muuttujina, sisältyy näihin muuttujiin ennustevirhettä. Nämä ns. mallivirheet voivat summautua hyvinkin suuriksi.
- 4) Luokitusvirheet: mittaustulokset ilmaistaan usein luokkina, jotka voivat olla huomattavasti laajempia kuin mitä mittaustarkkuus edellyttäisi. Jos havainnot eivät ole jakautuneet symmetrisesti luokan keskipisteen ympärille, aiheuttaa tasaavan luokituksen käyttö systemaattisen ali- tai yliarvion (Päivinen ym. 1992). Vaikka havainnot olisivatkin jakaantuneet tasaisesti, luokituksesta aiheutuu systemaattista virhettä esimerkiksi läpimittojen mittauksessa, missä oletetaan, että kaikilla luokan puilla olisi yhteinen läpimitta, joka ei kuitenkaan useimmiten ole luokan aritmeettinen keskiläpimitta eikä luokan pohjapinta-alamediaanipuun keskiarvo (Loetch ym. 1973).

Laskennallisessa ajantasaistuksessa käytetään tilastollisia malleja kehitysenusteiden laadintaan. Kuitenkin mallit ovat vain todellisuuden yksinkertaistuksia, joiden tuottamat kehitysenusteet tasoittavat kasvun todellista vaihtelua. Tulevaisuutta koskevat ennusteet joudutaan laatimaan menneisyyden havaintojen perusteella. Ennusteiden epävarmuus voi sisältää systemaattista ja satunnaista vaihtelua. Ennustevirheellä on neljä päälähdettä (Kangas ja Kangas 1997):

- 1) mallin spesifioinnin virheet
- 2) mallien parametrien satunnaiset estimointivirheet
- 3) mallien jäännösvaihtelu ja
- 4) virheet mallin selittävässä muuttujissa.

Lisäksi erotetaan ns. asiantuntijavirheet, jotka syntyvät, kun mallien laadinnassa ja oletuksissa on käytetty jossain muodossa virheellistä asiantuntemusta (Alho 1990). Asiantuntemuksen käyttö ennusteiden oletuksissa korostuu varsinkin lyhyen aikavälin ennusteissa (Alho 1990).

Metsikön kehitystä kuvaavat simulointijärjestelmät koostuvat useista malleista, joissa edellisten mallien antamia estimaatteja käytetään toisten mallien selittäjinä. Pitkien malliketjujen käytössä virheet kumuloituvat. Kasvuennusteiden luotettavuuteen vaikuttaa myös ennustamisjakson pituus (Kangas 2001). Kasvua ennustetaan yleisesti peräkkäisissä viiden vuoden jaksoissa (esim. Hynynen ym. 2002), jolloin kunkin jakson ennustevirheet kertyvät myöhemmissä ennusteissa (Kangas 2001). Kasvuennusteet tulevat tällöin sitä epätarkemmiksi, mitä kaukaisempaa tulevaisuutta ne kuvaavat (Salminen 1996, Kangas 2001).

Virheiden vaikutusta on tutkittu lähinnä kokonaisvirheen kautta, koska on hankala erottaa eri virhelajien vaikutukset ja yhteisvaikutukset. Selittävien muuttujien systemaattisten mittausvirheiden vaikutukset ja mallien parametrien harhojen vaikutukset eivät summaudu keskineliövirheeseen, kuten esitetään eräissä metsällisissä julkaisuissa (esim. Kilkki 1983, 1984). Mallin ja muuttujien mittausvirheiden aiheuttamat erimerkkiset harhat voivat kuitenkin kumota toisensa keskineliövirheessä (Lappi 1993). Harhan neliö ei ole siis komponenttiensa neliöiden summa. Mittausvirheet lisäävät myös ennustevirheen varianssia (Lappi 1993).

Mowrerin (1989) mukaan simulointijärjestelmien mallien lisääntyvä monimutkaisuus johtaa ennusteiden heikentyvään tarkkuuteen. Lisäksi syöttötietojen virheet ovat merkittävämpiä kuin mallien kertoimien epävarmuudet. Kankaan (1999) mukaan eri virhelähteiden suhteellinen merkitys vaihtelee ajan suhteen: lyhyellä aikavälillä mallien jäännösvaihtelu tai aineiston laatu voivat olla vallitsevia, mutta pitkällä aikavälillä mallien väärät määritykset tullevat tärkeimmäksi virheen lähteeksi. Metsäsuunnittelun simulointi- ja laskentajärjestelmissä käytetään yleisesti lineaarista ohjelmointia (LP). Kuitenkin lineaarisen ohjelmoinnin käytön oletuksena on syöttötietojen virheettömyys, mikä ei toteudu metsävara-
tietojen osalta käytännössä koskaan.

Tämän katsauksen tavoitteena on tarkastella las-

kennallisen ajantasaistuksen luotettavuuden arviointimenetelmiä. Kankaan ja Kankaan (1997) katsauksessa käsiteltiin erilaisten puustotunnusten ennustamiseen liittyviä epävarmuuden lähteitä. Tässä katsauksessa käsitellään osaksi samoja asioita, mutta tuodaan esille myös eräitä uusia näkökulmia ja ajantasaistetaan uusilla tutkimuksilla Kankaan ja Kankaan (1997) katsausta.

2 Luotettavuuden arviointimenetelmiä

2.1 Yleistä

Yleisin käytettävissä oleva tapa arvioida luotettavuutta on laskea keskineliövirhe, joka voidaan jakaa kahteen komponenttiin: satunnaisvaihtelua mittaavaan varianssiin ja systemaattista vaihtelua mittaavaan harhaan. Metsäsuunnittelun laskentajärjestelmän luotettavuuden arvioinnissa tutkitaan joko mallien sisältämää virhettä ja sen vaikutusta tai ennusteiden epävarmuuden syitä. Mallien sisältämää virhettä etsitään ja analysoidaan yleensä herkkyysanalyysillä, joilla tutkitaan mallien yksittäisten osatekijöiden osuutta mallin tuloksissa sekä mallien herkkyyttä olosuhteissa tapahtuneille muutoksille (Salminen 1996). Ennusteiden epävarmuuden syitä tutkitaan puolestaan analysoimalla virheitä (Salminen 1996).

2.2 Ennusteiden luotettavuuden arviointi empiirisestä aineistosta

Laskentajärjestelmien ennusteiden luottamusvälien laskeminen tuottaa ongelmia, koska tällöin on otettava huomioon usean eri mallin virhemarginaalit (Kangas ja Kangas 1997). Varsinkin simuloitaessa pitkän aikavälin ennusteita, esim. metsikön kehitystä, on vaikea saada tarpeeksi kattavaa ja riippumatonta aineistoa. Luottamusväli on yksinkertaisinta laskea empiirisestä aineistosta, josta on tiedossa kehitysennusteiden virheet. Kehitysennusteen virhe on ennustemalleilla lasketun kasvun ja empiirisestä aineistosta, esimerkiksi kasvukoelamittauksilla, selvitetyn kasvun ero. Luottamusvälit voidaan las-

kea järjestelmällä vuosittaiset ennusteiden virheet suuruusjärjestykseen ja poimimalla ennusteista 2,5, 50,0 ja 97,5 prosenttipisteet; 50. prosenttipiste kuvaa ko. vuoden mediaanivirhettä. Simulointijärjestelmän virhe on 95 %:n todennäköisyydellä prosenttipisteiden 2,5 ja 97,5 välillä.

Ennusteiden luottamusväli voidaan laskea empiirisestä aineistosta myös kaavalla (Hahn ja Nelson 1973):

$$\bar{D} \pm \sqrt{1 + \frac{1}{n}} S t_{1-\alpha/2}(n-1) \quad (1)$$

missä

D = ennustusvirheiden keskiarvo

n = ennusteiden lukumäärä

S = ennustevirheiden keskihajonta

$t_{1-\alpha/2}(n-1)$ = t-jakauman $1-\alpha/2$ -fraktiili vapausasteilla $n-1$

Ennuste on tällä välillä todennäköisyydellä $1-\alpha$. Koska ennustevirheiden keskiarvo (D) ja ennustevirheiden keskihajonta (S) estimoidaan otoksesta (n), ennustevälejä on levennettävä sallimalla estimointivirhettä näille tunnuksille ($1/n$).

Joissakin tilanteissa voidaan kuitenkin haluta laskea keskimääräinen virheen arvo k :lle ennustukselle. Esimerkiksi jos koko alueen tilavuus saadaan k :n metsikön yhteenlasketusta tilavuudesta, yksittäisten metsiköiden tilavuusennusteiden ei tarvitse olla kovinkaan tarkkoja, kunhan keskimääräinen virhe jää pieneksi. Luottamusväli k :n metsikön tilavuuden ennusteen keskimääräiselle virheelle saadaan kaavalla (Hahn ja Nelson 1973):

$$\bar{D} \pm \sqrt{\frac{1}{n} + \frac{1}{k}} S t_{1-\alpha/2}(n-1) \quad (2)$$

Kaavoja 1 ja 2 voidaan käyttää mikäli virheet ovat riippumattomia. Mikäli korrelaatioita ei oteta huomioon, saadaan liian optimistinen kuva ennustevirheistä (Lappi 1993). Jos virheiden korrelaatio on täydellinen, on keskiarvon virhe myös kaavan 1 mukainen (Lappi 1993 s. 72, Kangas 1999).

Jos oletus ennustevirheiden normaalijakaumasta ei päde, ennusteväli voidaan laskea ei-parametrisillä menetelmillä. Kaikille jatkuville jakaumille pätee todennäköisyys, että kaikki havaitsemisjärjes-

tyksessä seuraavat k havaintoa ovat suurimman ja pienimmän arvon välillä ko. populaatiossa (Hahn ja Nelson 1973):

$$1 - \alpha = \frac{n(n-1)}{(n+k)(n+k-1)} \quad (3)$$

Pienin ja suurin arvo antavat todennäköisyydellä $(1-\alpha)$ ennustevälin kaikille ennusteille k . Ongelmana tässä ei-parametrisessä luottamusvälissä on sen suuri leveys empiirisen aineiston (n) ollessa pieni.

Empiiriset tarkkuusarviot lasketaan aina kuitenkin vain tietylle, menneelle ajankohdalle ja rajatulle alueelle. Niiden yleistykseen tulevaisuuteen jollekin toiselle alueelle tulee suhtautua kriittisesti. Lisäksi on otettava huomioon empiiristen aineistojen sisältämät mittausvirheet (Kangas ja Kangas 1997). Metallsisten kausiennusteiden virheet eivät useinkaan noudata normaalijakaumaa. Empiirisiä luottamusvälejä voidaan kuitenkin käyttää kausiennusteiden luotettavuuden arvioinnissa, koska luottamusvälien käyttö ei edellytä normaalijakaumaoletusta. Normaalijakauman sijasta t-jakaumaoletusta käytetään aina, kun k -arvo ja hajonta ovat otoksesta.

2.3 Ennusteiden luotettavuuden arviointi estimoimalla

Käytännössä suunnittelun tarvitsemaa useiden 10-vuotiskausien empiiristä kasvu- ja tuotosaineistoa ei ole saatavilla. Niinpä luotettavuuden arvioinnissa joudutaan käyttämään erilaisia estimointimenetelmiä. Monte Carlo -simulointimenetelmiä käytetään yleisesti arvioitaessa lähtötietoihin ja johdettuihin tunnuksiin liittyvän epävarmuuden vaikutusta (esim. Gertner ja Dzialowy 1984, Mäkelä 1988, McRoberts 1992, McRoberts ym. 1994, Mowrer 1994, McRoberts 1996, Kangas 1997, 1998, 1999). Menetelmässä simuloidaan toistuvasti harhaa ja/tai satunnaisvirhettä lähtötietoihin. Lisäksi käytettäviin malleihin lisätään satunnaistekijä, jonka hajonta saadaan mallin keskivirheestä ja jakaumaksi oletetaan yleensä normaalijakauma. Simuloinnilla saaduista tuloksista lasketaan keskivirheet, harhat ja luottamusvälit eri tunnuksien ennusteille. Monte Carlo -tarkastelussa voidaan pitäytyä mallin koko-

naisennusteeseen, jolloin mallin sisäiselle rakenteelle ei tarvitse asettaa erityisiä vaatimuksia (Salminen 1996). Monte Carlo -menetelmät sopivat erityisesti epävarmuuden ennustamiseksi tilanteissa, joissa käytetään epälineaarisia kasvumalleja ja/tai edellisten mallien selittäjiä käytetään seuraavassa vaiheessa selittävinä muuttujina.

Toinen yleisesti käytetty, epävarmuutta analysoiva menetelmä on varianssipropagointi, jolla tutkitaan mallien komponenttien varianssia mallien käytön aikana (Mowrer ja Frayer 1986, Gertner 1987, Mowrer 1990, Summers ym. 1993). Menetelmässä mallin aiheuttama varianssin likiarvo lasketaan soveltamalla Taylorin sarjakehitelmää esimerkiksi yleiseen regressiomuotoiseen kasvuyhtälöön (Gertner 1988). Sarjakehitelmän kahden ensimmäisen termin avulla voidaan estimoida tulosjakaumien keskiarvot ja varianssit (Salminen 1996, Kangas 1996). Varianssipropagoinnilla voidaan myös estimoida ennusteiden luottamusvälit (Ripley 1987). Myös eksplisiittisiä varianssipropagointimenetelmiä on käytetty (Salminen 1996), kuten Kalmanin suodatusta (Moore 1973) ja systemaattisia tilastotieteen menetelmiä kuten herkkyysanalyysia faktoriasetteluun avulla (MacNeil ym. 1985) ja frekvenssitarkasteluja (Dwyer ja Kremer 1983).

Käytettävät tilavuus- ja kasvumallit ovat yleensä epälineaarisia selittäjien suhteen. Tällöin selittävien muuttujien satunnaisvirheet (mittaus-, luokittelu-, otanta- ja mallivirheet) aiheuttavat harhaa myös ennusteisiin (Lappi 1993, Kangas ja Kangas 1997). Tätä harhaa voidaan approksimoida Taylorin sarjakehitelmään perustuvalla menetelmällä (mm. Gertner 1991, Kangas ja Kangas 1997, Kangas 2001). Myös Monte Carlo -menetelmää voidaan käyttää tämän harhan ennustamiseen (mm. Kangas ja Kangas 1997, Kangas 2001). Tiedettäessä miten herkkä malli on satunnaisvirheille, voidaan määritellä sallitut maksimivirheet. Lisäksi voidaan arvioida samaa ilmiötä kuvaavien mallien herkkyyttä lähtötiedoissa esiintyvillä virheillä.

Metsikön kasvun ja tuotoksen ennustamisen luotettavuutta voidaan tutkia myös mallittamalla kiinnostavien muuttujien, esimerkiksi tilavuuden, havaittuja virheitä (Kangas 1999). Mallissa on impliittisesti mukana kaikki menneisyydessä vaikuttaneet virhelähteet, joten virheiden taso vastaa hyvin todellisuutta (Kangas 2001). Tällä mallilla voidaan

ennakoida tulevaisuuden ennusteiden epävarmuutta, mikäli olosuhteiden oletetaan pysyvän muuttumattomina (Kangas 1999).

Metsikön kasvu- ja tuotosennusteiden luotettavuutta voidaan myös tarkastella muodostamalla ns. konservatiivinen arvio metsikön kasvulle (tai tuotokselle) (Kangas 1999). Konservatiivinen arvio ennustevirheestä saadaan ennustamalla tunnusta äärimmäisen yksinkertaisella mallilla (Alho 1990, Kangas 1999, 2001). Yksinkertaisen mallin varianssia voidaan käyttää edelleen simulointijärjestelmän epävarmuuden arvioinnissa. Tällöin oletuksena on, että yksinkertainen malli on vähemmän tarkka kuin simulointisysteemi (Kangas 1999).

Metsikön yksinkertaisen mallin varianssia voidaan käyttää ylärajana ennusteiden todelliselle varianssille (Kangas 2001). Todellisen varianssin alempi raja saadaan puolestaan Monte Carlo -menetelmän varianssista (Kangas 1999). Tämä taas perustuu oletukseen, ettei kaikkia virhelähteitä tunneta, eikä niitä sen vuoksi voida ottaa huomioon, jolloin Monte Carlo -menetelmällä estimoitu varianssi jää pienemmäksi kuin todellinen varianssi.

Alueellisesti harhaisia kuvioittaisen arvioinnin tietoja voidaan kalibroida, mikäli alueelle on tehty tarkistusmittauksia (Jonsson ym. 1993). Samoin voidaan kalibroida päivitystietoja. Kalibrointi korjaa kuitenkin vain alueellisen tason. Yksittäisten metsiköiden ja pienempien tilojen osalta kalibrointi voi huonontaa huomattavasti ”lähtötietoja”, minkä vuoksi kalibroinnin käyttöön on suhtauduttava varauksella varsinkin yksittäisen metsälön kohdalla.

Metsätalouden suunnittelussa ollaan yhä enemmän siirtymässä simulointiin ja malleihin perustuviin suunnittelulaskelmiin. Niinpä onkin entistä tärkeämpää selvittää peräkkäisten kasvu- ja tuotosennusteiden tarkkuus. Kasvumalleihin on mahdollista sisällyttää keskivirheen laskentamahdollisuus. Tähän voidaan käyttää esimerkiksi Markovin lähestymistapaa (mm. Hool 1966, Peden ym. 1973, Moser 1980). Selittävien muuttujien satunnaisvirheen huomioon ottamiseksi voidaan käyttää 2. ja 3. asteen regressiota mallin kertoimia laskettaessa (mm. Furnival ja Wilson 1971, Reich 1984).

3 Tutkimuksia metsätalouden suunnitelmalaskelmien luotettavuudesta

Mittausvirheiden vaikutusta metsikkötunnusten estimointiin ovat käsitelleet mm. Hyppönen ja Roiko-Jokela (1978), Kujala (1979), Ihalainen (1987) ja Päivinen ym. (1992). Päivisen ym. (1992) tutkimuksessa lyhyellä kasvujaksolla mittausvirheet tekivät tilavuuskasvun estimoinnin peräkkäisten mitausten avulla epäluotettavaksi. Esimerkiksi viiden vuoden jaksolla puutasolla tilavuuskasvun keskivirhe saattoi olla lähes 75 % ja koealatasollakin yli 20 %. Pitkillä puilla pituuskasvun yliarvio oli suurempi kuin lyhyillä puilla. Gertnerin ja Dzialowyn (1984) mukaan kasvupaikka ja metsikön ikä vaikuttivat virheellisistä tunnuksista eniten metsiköiden kehitystä simuloitaessa. Virheellinen iän määrittäminen vaikuttaa eniten metsikön kehityksen alkuvaiheessa (Hynynen 1995).

Kasvumallit ovat yleensä aineistoista, joissa mittausvirhettä esiintyy vähemmän kuin sovellustilanteessa. Niinpä mallien antamat kasvu- ja tuotossennusteet sovellustilanteessa ovatkin epäluotettavampia kuin mallien keskivirheet (Gustavsen 1998). Mallien sovellusalueen eroaminen muutenkin mallien laadinta-aineistosta, esim. talousmetsien kasvumallien käyttäminen luonnonmetsien kasvun ennustamisessa, lisää ennusteiden hajontaa (Gustavsen 1998). Lisäksi ennustamiskauden sääolosuhteet voivat olla toiset kuin mallitusolosuhteet. Kasvupaikkojen boniteettiluokka voi muuttua ajan mukaan esimerkiksi ennustuskauden sääolosuhteiden muuttumisen myötä (Kuusela 1997, Gustavsen 1998). Metsätalouden laskentajärjestelmissä usein käytettävät yksittäisen puun kasvumallit (mm. Hynynen ym. 2002) tuovat myös virhettä kehitysennusteisiin, koska teoreettista läpimittajakaumaa käytetään oikean sijasta (Gustavsen 1998).

Suutarlan (1985) tutkimuksessa päivitettiin 6–9 vuoden ikäiset kuviotiedot nykyhetkeen. Tilavuuden päivityksessä päästiin 17,3 %:n keskivirheeseen, kun uuden inventoinnin keskivirhe oli 14,4 %. Laasasenahon ja Päivisen (1986) mukaan päivityksessä nuorten metsiköiden tilavuusarvio heikkenee nopeimmin. Vanhoissa metsiköissä 20

vuodenkin päivityksessä syntyneet virheet jäivät melko pieniksi.

Tilakohtaisesti pienilläkin inventointivirheillä voi olla suuria vaikutuksia. Pilhjerta (1987) simuloi 122 hehtaarin tilalle kuvioittaisen arvioinnin tietoihin erisuuruisia virheitä. Jo 5 %:n aliarvio pohjapinta-alassa vähensi harvennushakkuusosuutuksia 75 %. Vastaavasti suunnitelman päätehakkuusosuutukset lisääntyivät. Pohjapinta-alan yliarvio lisäsi laaditussa suunnitelmassa harvennushakkuuta ja vähensi uudistushakkuuta.

McRobertsin ym. (1994) tutkimuksessa tarkasteltiin mittausryhmien tekemien rinnankorkeusläpimitan, latvussuhteen ja kasvupaikan satunnaisvirheiden merkitystä kasvun ennustamisessa. Simulomalla pohjapinta-alan kasvua 20 vuodelle saatiin kasvulle 7–9 % hajontaa. Tutkimuksessa käytettiin variaatiokerrointa ja luottamusvälejä.

Eid tutki systemaattisten virheiden vaikutusta herkkyysanalyysillä metsikkö- (1991a) ja tilatasolla (1991b). Kasvupaikkatyypin määrittämisen systemaattiset virheet vaikuttivat eniten hakkuuehdotuksiin ja tilan nettohykyarvoon, mutta myös systemaattiset virheet metsiköitten iässä ja pohjapinta-alassa aiheuttivat ehdotettujen hakkuiden ajoitukseen ja nettohykyarvoon selvää hajontaa. Pickens ja Dress (1988) simuloivat erilaisia virheitä suunnittelulaskelmien lähtötietoihin ja lineaarisen ohjelmoinnin ongelman rajoitteisiin. 90 prosenttia optimoinnin tuottamista ratkaisuista ei olisi ollut toteutettavissa verrattuna todellisiin tuotantomahdollisuuksiin, koska ratkaisuun valikoituivat tuotantomahdollisuuksia yliarvioivat virheet.

Eid (1993) tutki satunnaisvirheiden vaikutusta metsätilan hakkuiden optimaaliseen ajoitukseen ja nettohykyarvoon Monte Carlo -menetelmällä. Kasvupaikkaluokalle, pohjapinta-alalle ja valtapituudelle simuloitiin satunnaisvirhettä ja metsäsuunnitelmat koostettiin tilalle lineaarisella ohjelmoinnilla. Satunnaisvirheillä todettiin olevan vaikutusta, sillä 10 %:n virhetasolla päätösmuuttujista hakkuut ja nettohykyarvo vaihtelivat 2–6 %. Kasvupaikkaluokan satunnaisvirheillä oli joissain tilanteissa suurempi vaikutus kuin systemaattisilla virheillä. Satunnaisvirheet eivät aiheuttaneet systemaattista virhettä johtuen luultavimmin käytettyjen mallien lineaarisuudesta.

Kangas ja Kangas (1997) näyttivät esimerkillä

lähtötietojen epävarmuuden merkityksen suunnittelulaskelmissa ja osoittivat, että hyötyfunktion odotusarvo on sitä harhaisempi mitä suuremmat ovat lähtötietojen virheet. Kankaan ja Kankaan (1997) mukaan optimointilaskelmien tavoitefunktion arvo yliarvioidaan epävarmuuksien takia millä tahansa optimointitekniikalla. Jos maksimoidaan esimerkiksi tilavuutta, siitä saadaan yliarvio. Optimoinnissa positiiviset virheet pyrkivät tulemaan valituksi optimiratkaisuun. Sama ilmiö löytyy myös käytännön suunnittelusta (Siitonen 1996). Kangas ja Kangas (1999) tutkivat myös virheiden korrelaatorakenteiden merkitystä optimointilaskelmien tuloksiin. Tavoitemuuttujien suhteilla ja virherakenteilla oli huomattava vaikutus optimoinnin tuloksiin. Virheiden korrelaatorakenteen huomiointi pienensi tavoitefunktion harhaa.

Simulointijärjestelmän luotettavuuden arviointia varianssipropagoinnilla perinteisen Monte Carlo -menetelmän sijasta ovat käsitelleet mm. Mowrer ja Frayer (1986), Gertner 1987, Mowrer (1991) ja Kangas (1996). Mowrer ja Frayer (1986) käyttivät varianssipropagointia metsikön kasvu- ja tuotomallien estimaattien vaihtelun arvioinnissa 10-vuotiskausittain. Tutkimuksessa otettiin huomioon myös selittävien muuttujien ja regressiokertoimien yhteisvaihtelu. Vertailtaessa yhdistettyä varianssia Monte Carlo -menetelmällä saatuun huomattiin varianssipropagoinnin aliarvioivan hieman vaihtelua. Lasketut variaatiokertoimet kasvoivat jyrkästi toisen 10-vuotiskauden jälkeen.

Gertner (1987) tutki varianssipropagoinnin soveltuvuutta yksittäisen puun kasvu- ja tuotomallin simulointijärjestelmän luotettavuuden arvioimisessa. Kovarianssien regressiokertoimien ja muuttujien välillä katsottiin olevan merkityksettömän pieniä. Tutkimuksen Monte Carlo -osassa satunnaisuus sisällytettiin suoraan simulointijärjestelmän mallien estimoiuihin regressiokertoimiin. Mowrer (1991) sisällytti myös regressiokertoimien ja muuttujien kovarianssin simulointijärjestelmän varianssin estimointiin.

Kangas (1996) tutki selittävien muuttujien satunnaisvirheiden vaikutusta epälineaaristen mallien tilavuusestimaatteihin. Tutkimuksessa käytettiin kolme eri menetelmää harhan korjaamiseen: Taylorin sarjakehitelmää, rekursiivista mallittamista ja Monte Carlo -menetelmää. Tutkimuksessa käytettiin Taylo-

rin sarjakehitelmää harhan ja ennustevarianssin arviointiin. Käytetyssä Monte Carlo -menetelmässä muuttujan odotusarvo laskettiin poimimalla otos muuttujien teoreettisesta jakaumasta. Otoksen keskiarvoa pidettiin muuttujan odotusarvona. Satunnaisvirheiden aiheuttama tilavuusennusteiden harha onnistuttiin korjaamaan kullakin menetelmällä. Kankaan (1996) mukaan mallien oletusten on kuitenkin pidettävä paikkansa, sillä muuten korjaukset eivät paranna estimaatteja. Kuhunkin tilanteeseen on valittava oma menetelmänsä, koska olettamukset mallien muodosta ja informaation laadusta vaihtelevat menetelmittäin.

Stokastisen kasvun vaihtelun lisäksi on myös otettava huomioon kausittainen vaihtelu, kuten sääolojen, metsätuhojen ym. vuosittainen vaihtelu. Kangas (1998) tutki vuosittaisen kasvun vaihtelun vaikutusta jäännösvirheiden lisäksi metsikön kasvun ennusteisiin. Läpimitan kasvun vuosittainen vaihtelu lisäsi ennusteiden hajontaa selvästi harvapuustoisissa metsiköissä.

Kankaan (1999) tutkimuksessa kasvu- ja tuotosennusteiden epävarmuutta tutkittiin kahdella viisivuotiskaudella kolmella eri menetelmällä. Ensimmäisessä käytettiin Monte Carlo -simulointia yhdistäen eri mallien jäännösvirheet yhdeksi virhetermiksi. Toisessa menetelmässä mallitettiin havaittuja metsikön tilavuuden ennusteiden virheitä ja arvioitiin ennusteiden varianssia ja keskineliövirhettä saadulla mallilla. Havaitut virheet saatiin ennusteiden ja mitattujen tilavuuksien eroista. Kolmannessa menetelmässä metsikön tilavuusennusteiden epävarmuutta ennustettiin konservatiivisella arviolla, joka saatiin estimoimalla yksinkertainen malli metsikön tilavuudelle. Mallissa käytettiin selittäjänä edellisen kauden tilavuutta.

Monte Carlo -menetelmällä saadut keskivirheet jäivät selvästi empiirisiä pienemmiksi (Kangas 1999). Tällöin ennusteet sisälsivät vielä joitain virhelähteitä, joita ei otettu huomioon ja/tai oletukset eri virheiden kovarianssirakenteesta eivät pitäneet paikkaansa. Tilavuuden mallitus edellisen kauden tilavuudella tuotti hyviä arvioita simulointijärjestelmän luotettavuudesta. Tilavuuden konservatiivisella mallilla saatiin hyvin lähellä empiirisiä keskivirheitä olevia keskivirhe-estimaatteja. Empiiriset keskivirheet olivat Monte Carlo -menetelmällä ja tilavuuden naiivilla mallilla saatujen tilavuuksien keskivirhei-

den välillä. Niinpä näitä estimaatteja voitiin käyttää todellisen keskivirheen rajoina.

Laajempien alueiden kasvun ja tuotoksen simulointijärjestelmien luotettavuutta on myös tutkittu, tosin vähäisessä määrin. Soares ym. (1995) käyttivät tutkimuksessaan empiiristä menetelmää ja Gertner (1995) Taylorin sarjakehitelmää simulointijärjestelmän epävarmuuden kuvaamisessa jatkuvan inventoinnin aineistolle. Hansen ja Hahn (1983) estimoivat kasvun ennusteiden luotettavuutta ”kaksoisotannalla” (double sampling scheme). He mallittivat havaittua kasvua aliotoksen ennustetulla kasvulla. Tätä mallia käytettiin ennustetun kasvun varianssin estimoinnissa sekä yksittäisille metsiköille että koko otoksen keskiarvolle. Näillä menetelmillä luotettavuuden estimointi suurille alueille onnistui vaivattomasti. Kuitenkin empiiriset ennusteet voitiin laskea vain tietyille alueille ja tietylle menneelle ajanjaksolle. Tulevaisuuden ennusteiden epävarmuuden ennustamiseen tarvitaan mallipohjaista lähestymistapaa.

Haaran (2002) tutkimuksessa käytettiin k-lähimmän naapurin menetelmää (k-nearest neighbour classification) (Hall 1985) kasvuennusteiden luotettavuuden arvioinnissa. Metsikkökuviolle haettiin k-lähintä naapuria, joiden puustotunnusten ennustevirheiden hajonnasta saatiin puustotunnuksille luotettavuusarviot. Menetelmällä saatiin hyvin lähellä empiirisiä keskivirheitä olevia keskivirhe-estimaatteja. k-lähimmän naapurin menetelmällä voidaan estimoida valitulle metsikölle samanaikaisesti kaikki kiinnostavat tunnuksat referenssiaineiston lähimpien naapureiden vastaavista tunnuksista (Härdle 1989). Menetelmän käyttö edellyttää referenssiaineistoa, jota voidaan kuitenkin täydentää jatkuvasti. Mikäli kehitysennusteisiin käytettävän simulointijärjestelmän jokin ominaisuus muuttuu, esimerkiksi kasvumallien päivituksen vuoksi, voidaan referenssiaineisto päivittää laskemalla uudelleen referenssiaineiston toteutuneet virheet. Tämä edellyttää kuitenkin, että referenssiaineiston alkuperäinen kuvioaineisto säilytetään.

4 Johtopäätökset

Kustannusten vähentämisvaatimukset ja tehokkuuden korostaminen tuntuvat koko metsäsektorin alueella, niin myös metsäsuunnittelussa (Maa- ja metsätalousministeriön... 2001). Käytännössä inventointimenetelmät kevenevät edelleen ja jatkuvan inventoinnin käyttö lisääntyy (Maa- ja metsätalousministeriön... 2001). Toisaalta vaaditaan yhä tarkempia ja monipuolisempia inventointitietoja metsästä suunnittelun ja päätöksenteon tueksi (Varjo 1999, Kangas 2001, Karppinen ym. 2002). Esimerkiksi metsänomistajien tavoitteet ovat muuttuneet moniarvoisempaan suuntaan ja nykyisin jo puolet metsänomistajista lukeutuu monitavoitteisten ryhmään, jossa korostetaan taloudellisten tavoitteiden ohella myös metsien aineettomia hyötyjä (Karppinen ym. 2002). Tässä ristiriitaisessa tilanteessa arviot metsävaratiedon, varsinkin päivittämällä saadun aineiston, epätarkkuuden vaikutuksesta suunnittelulaskelmiin voi olla erityisen arvokasta. Tällöin uusintamittaukset voidaan kohdentaa sinne, missä on eniten epävarmuutta, ja saadaan voimavaroja uusien (esim. monimuotoisuutta kuvaavien) tunnusten mittaamiseen.

Yleensä käytännön taktisessa ja operatiivisessa metsäsuunnittelussa tulevaisuus oletetaan tunnetuksi, vaikka tämä oletus tiedetäänkin vääräksi (Pukkala 1994). Yhtäläillä deterministiset kuin stokastisetkin ennustamismenetelmät vaativat oletuksia tulevaisuudesta. Inventointitiedon virheellisyydestä johtuvan epävarmuuden kytkeminen metsälötason suunnitteluun on vaikeaa, koska virheet vaikuttavat erisuuntaisesti eri metsiköissä (Kangas ja Kangas 1997). Joitakin malleja on kuitenkin käytettävä suunnittelulaskelmien tukena. Mikään malli ei kuitenkaan kuvaa täysin oikein luonnonolosuhteita. Arvio mallien luotettavuudesta on tämänkin vuoksi tärkeää.

Luotettavuuden arviointimenetelmän valinnassa ratkaisevat mm. käytettävissä olevien aineistojen määrä ja laatu, luotettavuusennusteiden haluttu tarkkuus ja ennusteisiin käytettävissä oleva aika. Jos käytössä olisi isohko riippumaton testiaineisto ja aineiston rakenne ei vaikuttaisi tai ei muuttuisi, epäilemättä tarkin ennuste luotettavuudelle saataisiin empiirisesti laskemalla testiaineistosta keskivirhe ja harha. Jos testiaineisto olisi painottunut

vanhoihin metsiköihin, niin ennustejoukon puustojakauman muuttuessa nuoremmaksi virhetasot ovat väävät. Tätä ongelmaa ei tule mallipohjaisilla luotettavuuden arvioimismenetelmillä, sen enempää regressio- kuin k-lähimmän naapurin menetelmäpohjaisillakaan, jos virhetasot on mallitettu iän suhteen oikein.

Kattavia, empiirisiä aineistoja ei ole juurikaan mahdollista saada käyttöön, minkä vuoksi luotettavuutta arvioidaan yleensä erilaisilla estimointimenetelmillä, kuten Monte Carlo -menetelmällä ja varianssipropagoinnilla. Varianssipropagoinnilla saadaan nopeasti kasvu- ja tuotosestimaattien varianssi, mikäli varianssiyhtälöt on integroitu laskentajärjestelmään. Tällöin voidaan myös ennusteiden luottamusvälit laskea rutiinimaisesti. Kuitenkin monimutkaisissa epälineaarisisissa laskentajärjestelmissä harhan korjaus sarjoihin tai polynomisiin approksimaatiomenetelmiin perustuvilla menetelmillä voi olla hankalaa ellei mahdotontakin useissa tapauksissa (mm. Kangas 1996). Esimerkiksi Taylorin sarjaan perustuvan harhan korjauksen käyttö edellyttää, ettei mallin muoto ole liian ”käyrä”, muuten korjaukset voivat olla epäloogisia (Smolander ja Lappi 1985). Lisäksi havaitut ja oikeat arvot eivät saa poiketa liikaa toisistaan (Kangas 1996). Taylorin sarjan korjauksen käyttö edellyttää myös, että muuttujien varianssit ja kovarianssit tunnetaan ja että malli on derivoituva (Kangas 1997).

Metsikön kehitystä ja tuotosta ennustavien laskentajärjestelmien ennusteiden luotettavuuden arvioinnissa on yleensä käytetty Monte Carlo -menetelmiä yksinkertaisen toteuttamisen ja toistettavuuden vuoksi (Kangas 2001). Monte Carlo -menetelmä edellyttää virhelähteiden yhteisjakauksen tuntemista (Kangas ja Kangas 1997). Kuitenkin on vaikeaa ottaa huomioon kaikki virhelähteet. Varsinkin oikean korrelaatorakenteen selville aineistosta, joka on estimoitu useasta eri lähteestä, on vaikeahkoa, kun samalla pitäisi ottaa huomioon mallien väliset korrelaatiot samalla lähtöaineistolla sekä myös kausien väliset korrelaatiot (Kangas ja Kangas 1997). Menetelmän käyttö vie myös suhteettoman paljon aikaa ja tietokoneresursseja suurilla alueilla.

Toisaalta Monte Carlo -menetelmällä pystytään käsittelemään monimutkaisiakin ongelmia, eikä menetelmän käyttö edellytä mallien derivoituvuutta (Kangas 1996). Menetelmän käyttö onnistuu hy-

vin ilman riippumatonta testiaineistoa. Lisäksi tehokkuutta voidaan lisätä käyttämällä kehittyneempiä Monte Carlo -menetelmiä, joiden avulla laskenta-aikaa voidaan lyhentää (Rubinstein 1972). Monte Carlo -menetelmän etuna on myös, että tiettyjen oletusten tai mallien vaikutusta voidaan tutkia erikseen (Kangas 1999). Jos kokonaisvirhettä on mallitettu suoraan, eri virhelähteiden erotteleminen on hankalaa. On kuitenkin mahdollista mallittaa useita erilaisia muuttujia ja tehdä johtopäätökset näistä malleista epäsuorasti. Lisäksi Monte Carlo -menetelmällä saadaan jo simuloitulla aineistolla selville järjestelmän antamien tulosten varianssin alaraja (Kangas 1999). Tämä voi olla useissa tapauksissa jo riittävä estimaatti järjestelmän tulosten luotettavuudesta (Kangas 1999).

Havaittujen virheiden mallittamisen menetelmällä voidaan arvioida erikseen harhaa ja varianssia. Kangas (1999) mukaan menetelmän käyttö on erityisen suositeltavaa pitkän aikavälin ennusteille. Simulointijärjestelmän antamien tulosten luotettavuuden arviointi on tällä menetelmällä helpohkoa (Kangas 1999). Menetelmän heikkoutena voidaan pitää tarvetta riippumattomaan testiaineistoon.

Konservatiivisen arvion käyttö metsikkösimulaattorin halutun tunnuksen luotettavuuden arvioinnissa on antanut hyviä ennusteita (Kangas 1999). Menetelmä on myöskin melko robusti mallin rakenteiden virheille (Kangas 1999). Toisaalta se voi antaa myös liian suuria keskivirheitä, millä on vaikutusta suunnitteluun ja päätöksentekoon (Kangas 2001). Lisäksi mallilla ei voida ottaa huomioon esim. käsittelyjen tason muutoksen vaikutusta (Kangas 1999).

Samantyyppisissä metsiköissä simulointijärjestelmän antamat simulointitulokset ovat yleensä samalla luotettavuustasolla. Tähän oletukseen perustuvat mm. empiiriset menetelmät luotettavuuden arvioinnissa ja erityisesti silloin, kun tutkimusaineistoa esitellään. Tähän oletukseen perustuvalla k-lähimmän naapurin menetelmällä voidaan tuottaa luotettavuusarviot kuvion tunnuksille lähimpien naapureiden kehitysennusteiden virheiden hajonnasta. Menetelmä on hyvin joustava muuttuneille olosuhteille esimerkiksi kasvumallien päivityksessä. Lisäksi menetelmän vaatimaa referenssiaineistoa voidaan päivittää jatkuvasti (Haara 2002).

Jos halutaan hyviä ennusteita laskentajärjestelmän todellisesta laadusta on välttämätöntä käyttää

laskentajärjestelmästä riippumatonta testiaineistoa. Tällöin konservatiivisen arvion ja havaittujen virheiden mallituksen käytöllä voidaan saada laskentajärjestelmän oikeaa luotettavuutta riittävän lähellä olevia estimaatteja suunnittelun ja päätöksenteon tueksi (Kangas 1999). k-lähimmän naapurin menetelmänkin käyttö laajalla referenssiaineistolla tuottaa hyviä luotettavuusarvioita (Haara 2002). Varianssipropagoinnin käyttö monimutkaisten laskentajärjestelmien luotettavuuden arvioinnissa voi olla käytännössä liian vaikeata ja Monte Carlo -menetelmän ongelmana on kaikkien vaikuttavien tekijöiden ja niiden yhteisvaikutusten löytäminen ja huomioon ottaminen.

Kuviorajojen vaihtuvuus on myös huomioitava laskennallisen ajantasaistuksen luotettavuuden arvioinnissa. Luotettavuusarvioista ei ole kovinkaan paljon apua, mikäli ne tehdään kuviorajoiltaan ja samalla metsikkötunnuksiltaan muuttuneille metsiköille. Kuviolle ajantasaistusjakson aikana tehdyt toimenpiteet tulisi myös selvittää, sillä onhan niiden vaikutus simulointituloksiin ja sitä myötä luotettavuusarvioon merkittävä. Kuviorajojen muuttuminen ja toimenpiteiden päivitys tulisi tehdä mahdollisimman automaattiseksi esimerkiksi leimikon rajauksen yhteydessä tai hakkuukoneen liikkumisen ja hakkuupoistuman tallentamisen kautta tapahtuvaksi.

Mallien ennusteiden virheet ovat yksi monista päätöksenteossa vaikuttavista epävarmuustekijöistä. Muita epävarmuustekijöitä ovat mm. päätöksentekijän tavoitteiden ja preferenssien määrittäminen ja hintojen ja kustannusten tulevaisuuden kehitys. Ajantasaistuksessa mallien virheet ovat mukana sekä menneen kauden ennusteissa että tulevaisuuden kehitysvaihtoehtojen ennustamisessa. Mallien tuoma epävarmuus näkyy väärissä toimenpiteiden ajoituksissa ja liian optimistisissa optimiratkaisuissa.

Nykyiset metsäsuunnitelmat eivät sisällä mitään arvioita ennusteiden luotettavuudesta. Metsiköiden tarkat tilavuustiedot ja kertymäennusteet saattavat antaa liian optimistisen käsityksen suunnitelman tarkkuudesta. Metsäsuunnittelun päätöksentekotilanteessa ennustelaskelmien ja annettujen metsänkäsittelysuositusten luotettavuus tulisi havainnollistaa metsänomistajalle ja muille suunnitteluun osallistuville ymmärrettävästi. Tästä voisi olla apua päätöksentekohetkellä ja näin voitaisiin myös välttyä pettymyksiltä, kun hakkuukertymät eivät ehkä

olekaan suunnitelman mukaisia.

Miten päätöksentekijä suhtautuu tietoon päätöksen riskistä tai mahdollisesta epävarmuudesta? Todennäköisimmin erilaiset päättäjät suhtautuvat tähän tietoon eri tavoin: osa voi arvioida vaihtoehtoja varianssin alarajan mukaan (riskin suosija), osa ylärajan mukaan (riskin karttaja) ja osa on riskineutraaleja. On muistettava, että päätöksentekijän suhtautumisen riskiin on myös yksi metsäsuunnittelussa huomioonotettavista seikoista, jolla voi olla huomattava vaikutus käsittelyehdotuksiin ja suunnitelman onnistumiseen (mm. Pukkala ja Kangas 1996, Pukkala ja Miina 1997, Pukkala 1998).

Kiitokset

Tutkimus on osa Maa- ja metsätalousministeriön yhteistutkimusvaroin toteutettua Metsäsuunnittelun tietohuolto -tutkimushanketta. Käsikirjoitusta kommentoivat sen eri vaiheissa MMT Annika Kangas, MMT Mikko Kurttila, MMT Matti Maltamo ja PhD Tuula Nuutinen. Lisäksi kaksi esitarkastajaa antoivat hyödyllisiä ja huolellisia käsikirjoitusta koskevia kommentteja. Lämpimät kiitokseni kaikille yllämainituille.

Kirjallisuus

- Alho, J.M. 1990. Stochastic methods in population forecasting. *International Journal of Forecasting* 6: 521–530.
- Anttila, P. 2002. Updating stand level inventory data applying growth models and visual interpretation of aerial photographs. *Silva Fennica* 36: 549–560.
- Dwyer, R.L. & Kremer, J.N. 1983. Frequency-domain sensitivity analyses of an estuarine ecosystem simulation model. *Ecological Modelling* 18: 35–54.
- Eid, T. 1990. Long term forest planning. Economical and biological production possibilities of a forest. Agricultural University of Norway, Doctor Scientiarum Theses 1990:9. 143 s.
- 1991a. Virkninger av målefeil på taksresultater, framskrivninger og prioriteringer av tiltak i skogsbestand [Effects of erroneous measurements for the results of the inventory, the forecasts and the priorities

- of treatments in forest stands]. Research paper of Norwegian Forest Research Institute and Department of Forestry, Agricultural University of Norway 8/91: 1–39.
- 1991b. Konsekvenser av feil datagrunnlag for planlegging og gjennomføring av planer på skogeiendommer [Consequences of erroneous data basis for planning and management of forest holdings]. Research paper of Norwegian Forest Research Institute and Department of Forestry, Agricultural University of Norway. 28 s.
- 1993. Random errors and strategic planning in forestry. *Meddelelser fra Skogforskning* 46: 1–24.
- Furnival, G.M. & Wilson, R.W. Jr. 1971. Systems of equations for predicting forest growth and yield. *Julkaisussa: Statistical ecology*. Vol. 3. Pennsylvania State University Press, University Park, PA.
- Gertner, G. 1986. Postcalibration sensitivity procedure for regressor variable error. *Canadian Journal of Forest Research* 16: 1120–1123.
- 1987. Approximating precision in simulation projections: an efficient alternative to Monte Carlo methods. *Forest Science* 33: 239–244.
- 1988. Alternative method for improving the variance approximation of single tree growth and yield projections. *Julkaisussa: Ek, A.R. Shifley, S.R. & Burk, T.E. (toim.). Forest growth modelling and prediction. Proceedings of the IUFRO Conference, August 23–27, 1987, Minneapolis, Minnesota. USDA Forest Service, North Central Forest Experiment Station, General Technical Report NC-120: 739–746.*
- 1991. Prediction bias and response surface curvature. *Forest Science* 37: 755–765.
- & Dzialowy, P.J. 1984. Effects of measurement errors on an individual tree-based growth projection system. *Canadian Journal of Forest Research* 14: 311–316.
- , Cao X. & Zhu, H. 1995. A quality assessment of a Weibull based growth projection system. *Forest Ecology and Management* 71: 235–250.
- Gustavsen, H.G. 1998. Volymtillväxten och övre höjdens utveckling i talldominerade bestånd i Finland – en utvärdering av några modellens validitet i nuvarande skogar. *Metsäntutkimuslaitoksen tiedonantoja* 707. 190 s. + liitteet.
- Haara, A. 2002. Kasvuennusteiden luotettavuuden selvittäminen knn-menetelmällä ja monitavoiteoptimoinnilla. *Metsätieteen aikakauskirja* 3/2002: 391–406. (Tämä nide).
- & Nevalainen, S. 2002. Detection of dead or defoliated spruces using digital aerial data. *Forest Ecology and Management* 160: 97–107.
- Hahn, G.J. & Nelson, W. 1973. A survey of prediction intervals and their application. *Journal of Quality Technology* 5: 178–188.
- Hall, P. 1985. Near-neighbour estimation. *Julkaisussa: Kotz, S., Johnson, N.L. & Read, C.B. (toim.). Encyclopedia of statistical sciences* 6. Wiley, New York. s. 167–169.
- Hansen, M.H. & Hahn, J.T. 1983. Estimation of sampling error associated with timber change projection simulators. *Julkaisussa: Bell, J.F. & Atterbury, T. (toim.). Renewable resources inventories for monitoring changes and trends. Proceedings of an international conference, 1983, USA, Corvallis.* s. 546–549.
- Hool, J.N. 1966. A dynamic programming – Markov chain approach to forest production control. *Forest Science Monograph* 12.
- Hynynen, J. 1995. Modelling tree growth for managed stands. *Metsäntutkimuslaitoksen tiedonantoja* 576. 59 + 76 s.
- Hypänen, H., Pasanen, K. & Saramäki, J. 1996. Päätehakkuiden kuviorajojen päivitystarkkuus. *Folia Forestalia – Metsätieteen aikakauskirja* 4: 321–335.
- Hypönen, M. & Roiko-Jokela, P. 1978. Koepuiden mittauksen tarkkuus ja tehokkuus. Summary: On the accuracy and effectivity of measuring sample trees. *Folia Forestalia* 356. 25 s.
- Härdle, W. 1989. *Applied nonparametric regression*. Cambridge University, Cambridge. 333 s.
- Ihalainen, A. 1987. Puukohtaisten mittausten ja arviointien luotettavuudesta valtakunnan metsien 8. inventoinnissa pysyvillä koealoilla. *Metsäntutkimuslaitoksen tiedonantoja*. 267. 37 s.
- Jonsson, B., Jacobsson, J. & Kallur, H. 1993. *The Forest Management Planning Package. Theory and application*. *Studia Forestalia Suecica* 189. 56 s.
- Kangas, A. 1991. Updated measurement data as prior information in forest inventory. *Silva Fennica* 25: 181–191.
- 1996. On the bias and variance in tree volume predictions due to model and measurement errors. *Scandinavian Journal of Forest Research* 11: 281–290.
- 1997. On the prediction bias and variance in long-term growth projections. *Forest Ecology and Management* 96: 207–216.
- 1998. Uncertainty in growth variation and yield projections due to annual variation of diameter growth. *Forest Ecology and Management* 108: 223–230.
- 1999. Methods for assessing uncertainty of growth and yield predictions. *Canadian Journal of Forest Research* 29: 1357–1364.
- 2001. Mallien käyttö. *Julkaisussa: Maltamo, M. & Laukkanen, S. (toim.). Metsää kuvaavat mallit. Silva Carelica* 36. Joensuun yliopisto. s. 223–239.

- & Kangas, J. 1997. Mallit, ennusteet ja simulointi metsätalouden laskentajärjestelmissä. *Metsätieteen aikakauskirja – Folia Forestalia* 3: 389–404.
- & Kangas, J. 1999. Optimization bias in forest management planning solutions due to errors in forest variables. *Silva Fennica* 33: 303–315.
- Kangas, J. 2001. Metsäsuunnittelun tutkimus- ja kehittämishaasteita. Julkaisussa: Kangas, J. & Kokko, A. (toim.). *Metsän eri käyttömuotojen arvottaminen ja yhteensovittaminen. Metsän eri käyttömuotojen yhteensovittamisen tutkimusohjelman loppuraportti. Metsäntutkimuslaitoksen tiedonantoja* 800. s. 309–314.
- Karppinen, H., Hänninen, H. & Ripatti, P. 2002. Suomalainen metsänomistaja 2000. *Metsäntutkimuslaitoksen tiedonantoja* 852. 83 s.
- Kilkki, P. 1979. Outline for a data processing system in forest mensuration. Seloste: Ehdotus metsänmittaus tulosten laskentamenetelmäksi. *Silva Fennica* 13: 368–384.
- 1983. Sample trees in timber volume estimation. Seloste: Koepuut puuston tilavuuden estimoinnissa. *Acta Forestalia Fennica* 182. 35 s.
- 1984. Metsänmittausoppi. *Silva Carelica* 3. Joensuun yliopisto. 283 s.
- Kleijnen, J.P.C. 1974. *Statistical techniques in simulation*. Marcel Dekker, Inc. New York. 775 s.
- Kujala, M. 1979. Mittaustarkkuus valtakunnan metsien inventoinnissa. *Metsäntutkimuslaitos, metsänarvioimisen tutkimusosasto. Moniste*. 8 s.
- Kuusela, K. 1997. Puuston kasvun ennustamisen karoikoita. *Metsätieteen aikakauskirja – Folia Forestalia* 3: 415–417.
- Laasasenaho, J. & Päivinen, R. 1986. Kuvioittaisen arvioinnin tarkastamisesta. Summary: On the checking of inventory by compartments. *Folia Forestalia* 664. 19 s.
- Lappi, J. 1993. Metsäbiometrian menetelmiä. *Silva Carelica* 24. Joensuun yliopisto. 182 s.
- Loetch, F., Zöhrer, F. & Haller, K.E. 1973. *Forest inventory*. Vol. II. BLV Verlagsgesellschaft, München. 469 s.
- Maa- ja metsätalousministeriön metsäsuunnittelustrategia 2001–2010. Maa- ja metsätalousministeriö, Työryhmämuistio 2001:13. Helsinki 2001.
- McNeil, M.D., Skiles, J.W. & Hanson, J.D. 1985. Sensitivity analysis of a general rangeland model. *Ecological Modelling* 29: 57–76.
- McRoberts, R.E. 1992. Monte Carlo simulations of nonlinear size-age relationships. Julkaisussa: *Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*. Second International Symposium, May 21–23 1996, Fort Collins, Colorado. s. 659–666.
- 1996. Estimating variation in field crew estimates of site index. *Canadian Journal of Forest Research* 26: 560–565.
- , Hahn, J.T., Hefty, G.J. & Van Cleve, J.R. 1994. Variation in forest inventory field measurements. *Canadian Journal of Forest Research* 24: 1766–1770.
- Moore, S.F. 1973. Estimation theory applications to design of water quality monitoring systems. *Proceedings of the American Society of Civil Engineering, Journal of Hydrologic Division* 99: 815–833.
- Moser, J.W. Jr. 1980. Historical chapters in the development of modern forest growth and yield theory. Julkaisussa: *Forecasting Forest Stand Dynamics; Workshop Proceedings*. School of Forestry, Lakehead University, Thunder Bay, Ont., June 24–25, 1980.
- Mowrer, H.T. 1989. The effect of forest simulation complexity on estimate precision. Julkaisussa: Burkhardt, H.E., Rauscher, H.M. & Johann, K. (toim.). *Artificial intelligence and growth models for forest management decisions*. Proceedings of a meeting held in Vienna, Austria, September 18–22, 1989. Virginia Polytechnic Institute and State University, School of Forestry and Wildlife Resources, Blacksburg, Virginia, USA, Publication FWS-1-89: 110–118.
- 1990. Estimating components of propagated variance in growth simulation model projections. *Canadian Journal of Forest Research* 21: 379–386.
- & Frayer, W.E. 1994. Variance propagation in growth and yield projections. *Canadian Journal of Forest Research* 16: 1196–1200.
- Mäkelä, A. 1988. Performance analysis of a process-based stand growth model using Monte Carlo techniques. *Scandinavian Journal of Forest Research* 3: 315–331.
- Oksanen-Peltola, L. 1999. Metsäsuunnittelun tietotarpeet. Julkaisussa: Heikinheimo, M. (toim.). 1999. *Metsäsuunnittelun tietohuolto*. Metsäntutkimuslaitoksen julkaisuja 741. s. 3–15.
- Peden, L.M., Williams, J.S. & Frayer, W.E. 1973. Markov model for stand projection. *Forest Science* 4: 303–314.
- Pickens, J.B. & Dress, P.E. 1988. Use of stochastic production coefficients in linear programming objective function distribution, feasibility and dual activities. *Forest Science* 34: 574–591.
- Pigg, J. 1994. Keskiläpimitan ja puutavaralajijakauman sekä muiden puustotunnusten tarkkuus Metsähallituksen kuvioittaisessa arvioinnissa. *Metsänarvioimistieteen pro gradu -työ*. Helsingin yliopisto, metsänarvioimistieteen laitos. 92 s.
- Pihljerta, K. 1987. Puustotunnusten arviointivirheiden

- vaikutus kuvioitaisiin toimenpidevalintoihin MELassa. Pro gradu -työ. Joensuun yliopisto, metsätieteellinen tiedekunta. 68 s.
- Poso, S. 1983. Kuvioittaisen arvioimismenetelmän perusteita. *Silva Fennica* 17: 313–343.
- Pukkala, T. 1994. Metsäsuunnittelun perusteet. Gummerus kirjapaino Oy, Jyväskylä. 242 s.
- 1998. Multiple risks is multi-objective forest planning: integration and importance. *Forest Ecology and Management* 111: 265–284.
- & Kangas, J. 1996. A method for integrating risk and attitude toward risk into forest planning. *Forest Science* 42: 198–205.
- & Miina, J. 1997. A method for stochastic multiobjective optimization of stand management. *Forest Ecology and Management* 98: 189–203.
- Päivinen, R., Nousiainen, R. & Korhonen, K. 1992. Puutunnusten mittaamisen luotettavuus. *Folia Forestalia* 787. 18 s.
- Reich, R.M. 1984. Modeling single slash pine tree growth and spacing relationships. Ph.D. dissertation. University of Florida, Gainesville, FL.
- Reynolds, M.R. Jr. 1984. Estimating the error in model predictions. *Forest Science* 30: 454–469.
- Ripley, B.D. 1987. Stochastic simulation. John Wiley & Sons. 237 s.
- Ritchie, M.W. & Hann, D.W. 1997. Implications of disaggregation in forest growth and yield modeling. *Forest Science* 43: 223–233.
- Rubinstein, R.Y. 1981. Simulation and the Monte Carlo method. John Wiley & Sons. 278 s.
- Salminen, H. 1996. Miten arvioidaan metsää kuvaavia malleja? Julkaisussa: Hökkä, H., Salminen, H. & Varmola, M. (toim.). 1996. Pohjoisten metsien kasvu – ennen, nyt ja tulevaisuudessa. Metsäntutkimuspäivä Rovaniemellä 1996. Metsäntutkimuslaitoksen tiedonantoja 589. s. 47–58.
- Saukkola, P. 1982. Uudistushakkuiden seuranta satelliittikuvista. Valtion teknillinen tutkimuskeskus, Tutkimuksia 89. 108 s.
- Siitonen, M. 1993. Experiences in the use of forest management planning models. Tiivistelmä: Kokemuksia mallien käytöstä metsätalouden suunnittelussa. *Silva Fennica* 27: 167–178.
- 1996. MELA ja metsien kehityksen ennustaminen. Julkaisussa: Hynynen, J. & Ojansuu, R. (toim.). Puuston kehityksen ennustaminen – MELA ja vaihtoehtoja. Tutkimusseminaari Vantaalla 1996. Metsäntutkimuslaitoksen tiedonantoja 612. s. 7–19.
- Smolander, H. & Lappi, J. 1985. Integration of a nonlinear function in a changing environment: estimating photosynthesis using mean and variance of radiation. *Agricultural and Forest Meteorology* 34: 83–91.
- Soares, P., Tomé, M., Skovsgaard, J.P. & Vanclay, J.K. 1995. Evaluating a growth model for forest management using continuous forest inventory data. *Forest Ecology and Management* 71: 251–265.
- SOLMU. Maastotyöopas. 1996. Metsätalouden kehittämiskeskus Tapio, Helsinki. 1996. 80 s.
- Ståhl, G. 1992. En studie av kvalitet i skogliga avdelningsdata som insamlats med subjektiva inventeringsmetoder. Sveriges Lantbruksuniversitet, Institutionen för biometri och skogsindelning, Rapport 24.
- Summers, J.K., Wilson, H.T. & Kou, J. 1993. A method for quantifying the prediction uncertainties associated with water quality models. *Ecological Modelling* 65: 161–176.
- Suutarla, T. 1985. Kuvioittaisen inventoinnin päivitys ja sen luotettavuus. Pro gradu -työ. Helsingin yliopisto, metsänarvioimistieteen laitos. 47 s.
- TASO. Maastotyöopas. 1993. Metsätalouden kehittämiskeskus Tapio, Helsinki. 50 s.
- Tokola, T., Pitkänen, J., Partinen, S. & Muinonen, E. 1996. Point accuracy of a non-parametric method in estimation of forest characteristics with different satellite materials. *International Journal of Remote Sensing* 17: 2333–2351.
- Tomppo, E. 1990. Satellite image-based National Forest Inventory of Finland. *Photogrammetric Journal of Finland* 12: 115–120.
- Utterä, J., Haara, A., Tokola, T. & Maltamo, M., 1998. Determination of the spatial distribution of trees from digital aerial photographs. *Forest Ecology and Management* 110: 275–282.
- Varjo, J. 1996. Controlling continuously updated forest data by satellite remote sensing. *International Journal of Remote Sensing* 1: 43–67.
- 1997. Change detection and controlling forest information using multi-temporal Landsat TM imagery. *Acta Forestalia Fennica* 258. 64 s.
- 1999. Johdanto. Julkaisussa: Heikinheimo, M. (toim.). Metsäsuunnittelun tietohuolto. Metsäntutkimuslaitoksen tiedonantoja 741. s. 5–7.
- Weintraub, A. & Abramovich, A. 1995. Analysis of uncertainty of future timber yields in forest management. *Forest Science* 41: 217–234.