

Juha Alho ja Kari Korhonen

Metsien tulevan kehityksen ennakointi

Paljonko Suomessa on puuta 10 vuoden kuluttua? Entä 20 tai 50 vuoden kuluttua? Tuntuu selvältä, että näihin kysymyksiin ei voida antaa tarkkoja vastauksia. Ennusteiden toteutuva virhe voi johtua huonosti valitusta ennustemenetelmästä, mutta tilastotieteelliseltä kannalta on ilmeistä, että käytettiinpä ennustamiseen mitä menetelmää tahansa, tuloksiin liittyy jonkin tasoinen virhe.

Tilanne on samantapainen muillakin aloilla. Esimerkiksi pörssikauppaan liittyvässä taloustieteessä lähdetään usein siitä, että osakekurssien muutoksia ei voi ennustaa. Sen sijaan esimerkiksi optiokaupassa pyritään arvioimaan osakekurssien muutosten todennäköistä suuruutta (eli ns. volatiliiteettia). Tämän pohjalta voidaan päätellä optioiden oikeat arvot.

Itse asiassa metsien inventoinneissa tulee vastaan tilanne, jossa joudutaan myös ratkaisemaan, kuinka paljon resursseja kulutetaan tiedon hankintaan. Inventoinnithan ovat (suuralueilla) otantapohjaisia. Valitsemalla suurempi otos, voitaisiin inventoinnin virhettä pienentää. Tilastotieteellisestä näkökulmasta inventoinnit ovatkin eräänlaisia ennusteita, joihin liittyy samaan tapaan virhe kuin tulevaisuutta koskeviin ennusteisiin. Tärkeätä olisi tällöin tietää, miten suurta virhettä on syytä odottaa (vaikkei sitä, onko virhe positiivinen vai negatiivinen voisikaan ennakoida!).

Analogia: väestön kasvun ennusteet

Metsissä olevan puun volyymin muutoksia voi-

daan kuvata yksinkertaisella kirjanpitoyhtälöllä: vuoden $t + 1$ volyyymi = vuoden t volyyymi + kasvu – poistuma. Yhtälön vasempaan puoleen liittyvä ennustevirhe seuraa oikean puolen termien epävarmuudesta. Aivan vastaavaa tilannetta on tutkittu melko paljon tilastollisen väestötieteen puolella. Jos $V(t)$ on vektori, jonka komponenteissa on väestön koko eri ikäryhmissä (esim. iät 0, 1, 2, ...), niin väestön muuttuminen voidaan ilmaista seuraavanlaisena yhtälönä:

$$V(t + 1) = M(t)V(t) + N(t) \quad (1)$$

missä matriisi $M(t)$ sisältää ikäryhmittäiset eloonjäämistodennäköisyydet vuonna t ensimmäisellä alidiagonaalilla, ikäryhmittäiset hedelmällisyysluvut vuonna t ensimmäisellä vaakarivillä ja muuten matriisin elementit ovat nolliä. (Tämä versio vastaa tarkkaan ottaen vain naisväestön kehittymistä. Yhtälöä (1) vastaava esitys on kuitenkin hyvin yleisessä tapauksessa mahdollinen, kun vain $M(t)$:n elementit määritellään tarkoituksenmukaisesti.) Vektori $N(t)$ kuvaa vuorostaan nettomuuttoliikettä.

Väestön ennustetta tehtäessä otetaan lähtökohdaksi esim. vuoden $t = 0$ väestö. Vuosille $t > 0$ muodostetaan matriisien $M(t)$ ennusteet $\hat{M}(t)$ ja yhtälöä (1) sovelletaan toistuvasti, niin että esim. vuoden $t = 2$ väestön ennuste on

$$\hat{V}(2) = \hat{M}(1)\hat{M}(0)V(0) + \hat{M}(1)\hat{N}(0) + \hat{N}(1) \quad (2)$$

(Yhtälössä (2) voi hyvin merkitä $V(0)$:n paikalle $\hat{V}(0)$, jos halutaan korostaa sitä, että lähtöväestö-

kin on usein vain tietyllä tarkkuudella tunnettu.) Yhtälöstä nähdään, että vuotta $t = 2$ koskeva ennustevirhe koostuu vuosien $t = 0$ ja $t = 1$ ikäryhmittäistä hedelmällisyyttä ja kuolevuutta koskevien ennusten virheistä. Yhtälö on epälineaarinen ja sen analyysissa on käytetty sekä analyyttisiä approksimaatioita että simulointia.

Näillä tarkasteluilla on merkitystä metsätieteiden kannalta, sillä yhtälöä (1) vastaavia malleja on käytetty puiden siirtymiseen läpimittaluokasta seuraavaan.

Kokonaisvirheanalyysi

Väestön kehitystä tarkasteltaessa on otettu käyttöön lähestymistapa, jota kutsutaan kokonaisvirheanalyysiksi. Tämä tarkoittaa yksinkertaisesti sitä, että kaikki väestön kehitykseen vaikuttavat virhelähteet otetaan huomioon. Itsestään selvältä tuntuvan periaatteen noudattaminen on kuitenkin käytännössä vaativaa. Väestötieteilijät ovat hyvin perillä hedelmällisyyden ja kuolevuuden muutoksista ja niiden ennustamiseksi nähdään paljon vaivaa. Muuttoliike käyttäytyy sen sijaan hyvin arvaamatomalla tavalla, ja sitä koskevat tiedot ovat useimmissa maissa puutteellisia. Tästä syystä muuttoliikkeen ennusteen virheen arviointi on hankalampaa kuin hedelmällisyyden ja kuolevuuden.

Tällä ilmiöllä on analogia kaavan (1) metsätieteellisissä tulkinnoissa. Niissä $N(t)$ vastaisi joko uusien metsäalueiden synnyttämistä esimerkiksi soita ojittamalla tai metsien siirtymistä talouskäytön ulkopuolelle (suo-ojien tukkiminen!). Näitä ilmiöitä on vaikea ennustaa, mutta ne voivat silti olla merkittäviä epävarmuuden lähteitä metsiä koskevissa ennusteissa.

Toinen kokonaisvirheanalyysiin liittyvä vaatimus on ns. mallivirheen huomioonottaminen. Jos metsien kehitystä pyritään ennakoimaan esim. Melajärjestelmän avulla, niin Melan pohjana olevin mutkikkaisiin yhtälöryhmiin välttämättä liittyvät yksinkertaistukset täytyisi osata ottaa virheanalyysissa myös huomioon. Laskentaan vaikuttavat mallivirheet kumuloituvat samaan tapaan kuin yhtälössä (2) ja niillä voi olla vaikutusta myös siihen, millainen on todennäköisin tuleva kehityskulku.

Tarvitaanko metsien ennusteiden luotettavuuden arviointia?

Edellä kuvatut näkökohdat olivat motiivina, kun Joensuussa järjestettiin 9.5.1997 seminaari teemana ”Metsien kehityksen ennakointiin liittyvä epävarmuus”. Seminaariesitelmien pohjalta kirjoitetut artikkelit ilmestyvät ohessa. Seminaarin päätteeksi keskusteltiin mahdollisuudesta käynnistää tutkimushanke, jossa pyrittäisiin laatimaan perusteltu arvio siitä, miten tarkasti metsien kehitystä voidaan ennakoida Suomen oloissa. Olennaista tällaisen hankkeen onnistumiselle on se, että kaikki ennusteiden eri komponenttien parissa työskentelevät tutkimusalat ovat siinä tavalla tai toisella mukana. Yksityiskohtaisemmin aihetta on tarkoitus pohtia loppuvuodesta järjestettävässä kokoontumisessa, johon pyritään saamaan kaikki kiinnostuneet mukaan.

■ Alho toimii Joensuun yliopiston tilastotieteen laitoksella ja Korhonen Joensuun yliopiston metsätieteellisessä tiedekunnassa.