

Hampus Holmström

Inventering genom imputering – effektiv datainsamling för småskogsbrukets planering

- Ett antal nya metoder för datainsamling går ut på att imputera, dvs. överföra, provytedata genom nyttjande av fjärranalysinformation. Imputering är på så sätt ett alternativ och komplement till den provytevisa fältinventering som planeringssystemet Indelningspaketet (IP) kräver.
- De nya metoderna är avsevärt billigare än den traditionella inventeringen. Genom att integrera olika inventeringsmetoder inom samma skogsinnehav, kan kostnaden för datainsamling sänkas med 15–50 procent. Detta gör det möjligt att även använda IP på mindre skogsinnehav.
- En utvärdering av inventeringsmetoder bör beakta både kostnader för insamling av data, och eventuella framtida förluster på grund av felaktiga beslut orsakade av fel i tillståndsbeskrivningarna. Sådana totalkostnadsanalyser visar att imputeringsmetoderna kan användas i flertalet fall. Mer noggranna fältinventeringar bör genomföras främst i värdefulla bestånd där en åtgärd, t.ex. gallring eller slutavverkning, är nära förestående.



FIGUR 1. De nya metoderna går ut på att imputera, eller överföra, data från redan inventerade provytor till "okända" områden. Vilka provytor som imputeras avgörs med stöd från någon gemensam informationskälla, t.ex. flygbilder. Genom att imputera kompletta ytor, med uppgifter på trädnivå, kan avancerade planeringsmodeller användas utan att fältinventering genomförts inom aktuell avdelning.

Att dagens handlingar är väl genomtänkta är en förutsättning för en bra, eller till och med bättre, framtid. Inom skogsbruket understryks planeringens betydelse av att omloppstiden är lång, och att åtgärder får relativt drastiska följder – ett avverkat träd görs svårigen oavverkat. Skoglig planering ska främst försöka besvara frågorna när (nu, om några år eller aldrig?), var (i vilka skogar/bestånd?) och hur åtgärder, t.ex. gallring eller slutavverkning, bör utföras för att uppnå eller närma sig ett visst mål i skogsbruket.

Traditionella skogsbruksplaner har hamnat i något av en ond cirkel. När resurserna för datainsamling är begränsade, blir kvaliteten på data så dåliga att uppgifterna knappast håller för mer långsiktiga prognoser och analyser. För en skogsägare upplevs därför beslutsstödet från skogsbruksplanen som rätt begränsat, och många vill inte lägga mer resurser på planframställan. I detta Fakta skog beskrivs nya metoder för kostnadseffektiv dataförsörjning i den skogliga planeringen.

Indelningspaketet

Under 1970-talet utvecklades planeringssystemet Indelningspaketet (IP), som sedan dess flitigt nyttjats av större skogsbolag i Sverige. Systemet är ett verktyg för ekonomisk optimering i ett skogsbruk fokuserat på främst virkesproduktion. Idag används systemet bl.a. för att ta fram strategiska beslutsunderlag. Ett sådant underlag kan ge svar på hur stora uttagsnivåer som kan förväntas under vissa krav på jämnhet i intäkter över tiden, samt vilken skogs-skötselstrategi som leder till målet.

När IP används, erhålls uttagsnivåer och skötselprogram för de närmaste hundra åren – men tidshorisonten är i princip oändlig. Långsiktighet är central, och därför ställs stora krav på att skogstillståndet från början är korrekt beskrivet. Traditionellt görs en stickprovsinventering i fält, där enskilda träd mäts inom ett antal provytor. Då skogsbolagens innehav är stora finns troligtvis många avdelningar som liknar varandra, och resultat från en avdelning kan representera flera liknande. En representativ bild av hela

skogsinnehavet är dock nödvändigt, och för mindre skogsfastigheter blir kostnaderna för en sådan datainsamling orimligt höga. Användandet av systemet har följaktligen varit begränsat inom småskogsbruket.

Hushållningsprojektet

Hushållningsprojektet har bedrivits som temaforskning vid SLU och SkogForsk med syfte att vidareutveckla IP. Att under vissa förutsättningar ersätta traditionell fältinventering med billigare och effektivare metoder, kan öka användbarheten av IP inom småskaligt familjeskogsbruk. Genom att kombinera olika typer av översiktlig information, exempelvis från olika fjärranalysensorer, kan detaljerade uppgifter "imputeras" till områden vars tillstånd behöver beskrivas. Imputering kan jämföras med dess motsats – amputering. Istället för att avlägsna något, tillförs här information från en del av skogslandskapet till en annan för att möjliggöra en viss funktionalitet. T.ex. kan enskilda trädmodeller användas för tillväxtprognoser, men även vid tester av olika skötselgrepp. Imputering är inte något nytt inom skoglig planering, ex-

empelvis imputeras trädåldrar från provträd (som borrats för åldersbestämning) till klavträd (som initialt bara beskrivits med trädslag och brösthöjdsdiameter).

För att IP ska fungera, krävs att en mängd variabler beskrivs i detalj. Imputering kan möta dessa högt ställda krav. När provytedata från fältinventeringar återanvänds och imputeras till andra avdelningar, kan planeringssystemet "luras att tro" att dessa data samlats in i den aktuella avdelningen.

Vid imputering används den så kallade k NN-metoden (faktaruta). Provytedata från de k närmaste referensytorna, "grannarna", imputeras till den aktuella avdelningen. För att kunna utse de närmaste referensytorna, krävs dock först en definition på vad som menas med närhet.

Närhet i flera dimensioner

Givetvis vill vi imputera ytor som så mycket som möjligt liknar den avdelning där tillståndet ska skattas. Vi söker då efter närliggande värden i bärardata, t.ex. i flygbildstolkningar

k NN-metoden och fjärranalysinformation

k NN (k Nearest Neighbour)-metoden går ut på att imputera – överföra – provytedata från de k närmaste referensytorna till ett från början okänt område (yta/avdelning). Därefter kan Indelningspaketet användas för den aktuella avdelningen. Eftersom metoden aldrig kan skatta värden utanför data i referensmaterialet, är den starkt beroende av att materialet är representativt i förhållande till de skogar vars tillstånd ska skattas.

En okänd yta eller avdelning skattas som ett viktat medelvärde av värden hos de k närmaste referensprovyterna. Viktningen görs normalt i proportion till närheten – den närmaste provytan erhåller en större vikt i förhållande till den provyta som är mer olik avdelningen. Värden på k mellan 1 och 20 har utvärderats i dessa studier. Om $k = 1$ imputeras bara den närmaste referensytans värden, som då får vikten "1".

Eftersom fjärranalysinformationen i detta sammanhang bär uppgifter (provytedata) från en viss plats (eller från en viss tid) till en annan, kallas informationen för *bärardata*. I skogliga sammanhang har k NN-metoden främst använts tillsammans med satellitbildsinformation. I detta projekt, med krav på direkt praktisk tillämpbarhet, bedömdes det vara intressant att använda flygbilder. Inom det svenska skogsbruket finns en stor erfarenhet av flygbilder, och tillgången är förhållandevis god. Den information som användes var främst tolkade trädhöjder och slutenhet – variabler som kan relateras direkt till fältmätningar, och förväntas motsvara den information som framtida laserbaserade fjärranalysensorer kan leverera. Dessa sensorer används redan idag på försöksstadiet inom skogsbruket. Även satellitdata som bärare, i kombination med radardata från en flygburen sensor, utvärderades inom projektet.

och/eller satellitbilsdata. För att definiera närhet i våra studier har vi först sammanställt referensytor inom en viss radie från det område som ska skattas. Därefter har vi jämfört bärardata hos aktuellt område med samtliga provtyper i referensmaterialet, för att på så sätt mäta deras inbördes närhet.

Att sammanfatta begreppet "skogstillstånd" i ett enda värde är inte helt enkelt. I den skogliga planeringen vet vi dock att t.ex. virkesförråd är av central betydelse. När virkesförrådet skattas med regressionsfunktioner från bärardata, för både referensyta och aktuell yta/avdelning, är avståndet dem emellan skillnaden mellan dessa skattningar. Det går sedan att gå vidare med skillnader i skattningar av t.ex. ålder och ståndortsindex, och lägga samman dessa avstånd till ett totalavstånd. Allt för att ytorna som imputeras ska vara så representativa som möjligt, med avseende på det skogstillstånd som ska beskrivas.

Noggrannhet i skattningar

För att beskriva skogliga tillstånd används olika metoder i olika situationer. I IP används i normalfallet objektiva stickprovsinventeringar, där mätningar görs på cirka tio provtyper per avdelning, med ett medelfel motsvarande ungefär tio procent vid volymsskattningar på beståndsnivå. Noggrannheten vid kNN-skattningar påverkas av hur avstånden mäts, hur närhet i olika dimensioner viktas och hur referensmaterialet är sammansatt. Störst betydelse har dock vilken mängd information som finns i det bärardata som används.

I denna studie utvärderades ett antal olika informationskällor som bärardata, enskilda informationskällor eller i kombination med varandra. I tabell 1 visas exempel på metodernas uppmätta fel. Skattningar gjordes för två centrala skogliga variabler – virkesförråd och ålder. Med flygbildstolkningar som bärardata, hamnar medelfelet mellan 13,7 och 21,4 procent för virkesvolym på beståndsnivå. Skillnaderna i felens storlek beror främst på skogliga förutsättningar – skog är generellt tätare i södra Sverige (Remningstorp), vilket innebär en svårare tolkning av

TABELL 1. Det totala medelfelet (RMSE, Root Mean Square Error) vid kNN-skattningar av virkesförråd respektive ålder på avdelningsnivå. Felen minskar när flygbildstolkning kombineras med information från beståndsregister. Samma effekt erhålls vid kombination av optisk data (från satellitbild) och radardata. Av utvärderade informationskällor uppvisade flygbilder tillsammans med beståndsregister största noggrannheten i skattningarna.

Bärardata	RMSE	
	Virkesförråd, m ³ sk/ha	Ålder, år
Flygbildstolkning ^a	37,7 (21,4 %)	9,7 (21,2 %)
Flygbildstolkning och beståndsregister ^a	35,5 (20,1 %)	7,5 (16,3 %)
Flygbildstolkning och beståndsregister ^b	19,1 (13,7 %)	13,7 (16,6 %)
Optisk data ^{a,c}	50,5 (29,9 %)	13,3 (29,6 %)
Radardata ^{a,d}	64,6 (38,2 %)	16,9 (37,6 %)
Optisk data och radardata ^a	37,0 (21,9 %)	11,2 (24,9 %)

^a Försöksområde Remningstorp (Västra Götaland).

^b Försöksområde Brattåker (Västerbotten).

^c Satellitbild SPOT-4 XS.

^d Flygburen sensor CARABAS-II VHF SAR.

trädhöjd och slutenhet i bilderna. Även satellitbilsdata (optisk data) i kombination med radardata utvärderades, och felet (21,9 procent) blev mindre när dessa källor användes tillsammans i kNN-skattningarna. Med imputering nås således inte noggrannheter i klass med fältinventeringar, men förmodligen är noggrannheten tillräckligt bra för att duga i vissa situationer.

Konsekvenser av fel

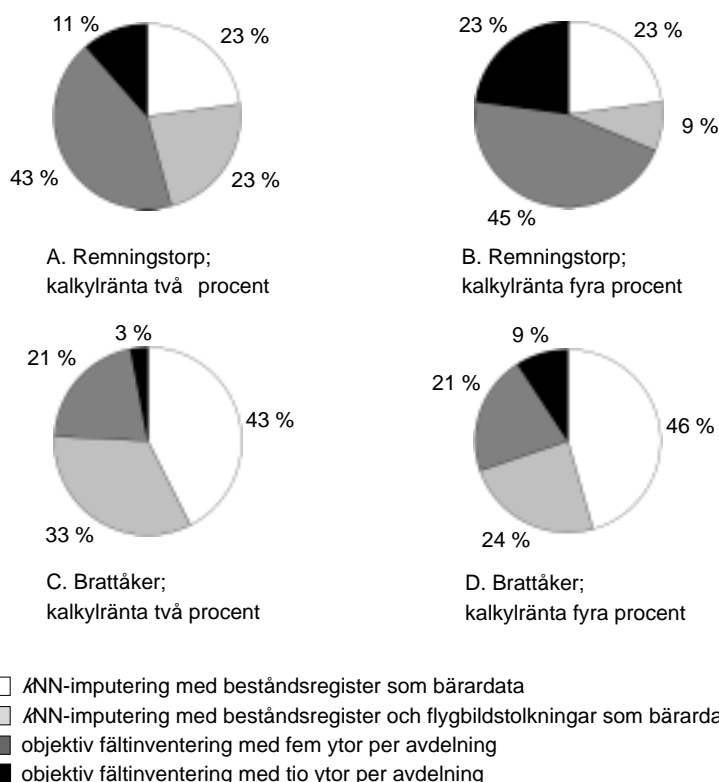
När går det då att acceptera liknande skattningsfel? Så kallade totalkostnadsanalyser utvärderar följderna av att använda inventeringsmetoder med vissa skattningsfel. Totalkostnaden för en viss metod är summan av värdet av förväntade framtida förluster (orsakade av felaktiga beslut tagna på felaktiga data) och kostnaden för att genomföra datainsamlingen. Hur mycket pengar som ska satsas på datainsamling är alltid en avvägningsfråga. Förlusterna beror, som tidigare nämnts, på att felaktiga tillståndsbeskrivningar i början leder till val av inoptimala åtgärder. Dessa förluster har beräknats som skillnaden mellan en avdelnings maximala nuvärde, dvs. dagens värde av framtida intäkter och utgifter i skogsbruket, och det nuvärde som uppstår när avdelningen sköts inoptimalt.

För två skogsfastigheter, Remningstorp i södra Sverige och Brattåker i norr, simulerades imputering av

provtyper. Två typer av bärardata användes – ett beståndsregister, motsvarande informationen i en vanlig skogsbruksplan, samt beståndsregister i kombination med flygbildstolkningar. Dessa jämfördes med IP:s vanliga stickprovsinventering med två olika intensiteter: fem respektive tio ytor per avdelning. Utifrån data från varje inventeringsmetod, användes IP därefter för att få fram optimala skötselprogram för varje avdelning inom de två fastigheterna.

Analys genomfördes där avdelningarnas arealer och räntan varierades (en till tio hektar resp. två och fyra procent), medan övriga förutsättningar hölls konstanta. Det visade sig att totalkostnaden för dataförsörjningen för de två fastigheterna kunde sänkas med 15–50 procent genom att nyttja olika inventeringsmetoder inom olika delar av innehavet. Detta jämfört med när bara en metod, den med lägst totalkostnad, användes för samtliga avdelningar i innehavet.

Optimala dataförsörjningsstrategier exemplifieras för de två skogsinnehaven i figur 2. Södra Sveriges högre tillväxt medför en högre intensitet i skogsbruket, exempelvis kortare omloppstider och mer frekventa gallringar. Därför ställs också krav på bättre tillståndsbeskrivningar för att "pricka rätt" i åtgärdsförslagen. En liknande effekt uppstår när räntan höjs. Resultatet är att en högre ränta leder till att en större del av skogen sköts inoptimalt.



FIGUR 2. Andel i procent av fastigheternas areal som, för minimering av totalkostnad för dataförsörjning, inventeras med respektive metod. Genom att integrera olika inventeringsmetoder inom samma innehav kan kostnaden för datainsamling sänkas med 15–50 procent. Figuren visar att en stor andel av arealen kan beskrivas med tillräcklig noggrannhet genom imputering. Behovet av bättre precision i tillståndsbeskrivningarna genom provyteinventering är större i södra Sverige (A, B) samt vid en högre kalkylränta (B, D).

tatet innebär att imputeringsmetodik kan tillämpas för en relativt stor andel av arealen i samband med IP-analyser. Mer noggranna fältinventeringar bör dock nyttjas i värdefulla bestånd med en åtgärd, t.ex. gallring och slutavverkning, nära förestående.

Databaser av referensprovtytor finns sammanställda för hela landet, så imputering kan användas och används redan idag på flera håll. Intresset är stort för metoden då det finns mängder av vanliga skogsbruksplaner inom småskogsbruket som saknar be-

skrivningar av konsekvenser på längre sikt (mer än tio år) av ett visst brukande. Även storskogsbruket har visat intresse då resultat, t.ex. optimala skötselprogram, kan erhållas för varje avdelning i innehavet, och inte bara för stickprovsavdelningar som tidigare.

Heureka i framtiden

Till skillnad från Hushållningsprojektet, vars huvudsyfte var att vidareutveckla IP, är forskningsprogrammet Heureka inriktat på att utveckla nästa generations skogliga planeringssystem.

tem. Tyngdpunkten ligger i att hantera och planera för ett flermålsinriktat skogsbruk, vilket är relevant även för det småskaliga skogsbruket. Heureka kommer sannolikt att kräva både heltäckande och högupplösta indata för att kunna tillämpa mer avancerade modeller för virkesproduktion och natur-, jakt- och upplevelsevärden etc. Här kan imputering vara en värdefull dataförsörjningsmetod.

Ämnesord

Skoglig planering, inventeringsmetodik, fjärranalysinformation, ÅNN, icke-parametrisk regression

Läs mer

Holmström, H. 2001. Data Acquisition for Forestry Planning by Remote Sensing Based Sample Plot Imputation. *Acta Universitatis Agriculturae Sueciae, Silvestria 201*. SLU, Umeå.

Författare



Hampus Holmström är SkogD och forskare vid institutionen för skoglig resurshushållning och geomatik, SLU, 901 83 UMEÅ. Pågående forsknings- och utvecklingsarbete behandlar främst småskogsbrukets planering inom ramen för Heureka-programmet. Tel: 090-786 59 14, 070-476 26 88. E-post: Hampus.Holmstrom@resgeom.slu.se

Tack

Torleif Carlsson, SkogForsk, och Hans Kallur, ÖKA Skogsplan, båda drivande krafter bakom IP:s vidareutveckling. Allt hade dock varit intet utan planeringssystemets upphovsman, professor emeritus Bengt Jonsson.



Ansvarig utgivare:
Redaktör:

Internet:
Prenumeration och lösnnummer:

Prenumerationspris:
Tryck:

Göran Hallsby, SLU, institutionen för skogsskötsel, 901 83 Umeå
Camilla Nilsson, SLU Informationsavdelningen, Box 7077, 750 07 Uppsala
Telefon: 018-67 21 34 • Telefax: 018-67 35 20
E-post: Camilla.Nilsson@info.slu.se
www.slu.se/forskning/fakta
SLU Publikationstjänst, Box 7075, 750 07 Uppsala
Telefon: 018-67 11 00 • Telefax: 018-67 35 00
E-post: Publikationstjanst@slu.se
320 kr + moms
SLU Repro, Uppsala 2002
ISSN 1400-7789 © SLU