

Slutrapport

Projektrubrik: Ajourhållning av skogliga laserskattningar med skördardata

Huvudsökande: Erik Willén, Skogforsk

Projektets löptid: 2018-04-01 – 2019-12-31

Populärvetenskaplig sammanfattning

Under 2015 avslutade Skogsstyrelsen arbetet med första versionen av Skogliga Grunddata - en nationell kartering av skogliga data baserat på Lantmäteriets laserdata och Riksskogstaxeringens inventeringsdata. Skogliga Grunddata används bla för att uppdatera beståndsregister, hitta intressanta köpobjekt och för att identifiera hänsyn. Laserskanningen som ligger till grund för skattningarna är från 2009-2014, vilket gör att data för stora delar av landet börjar bli inaktuellt. En ny laserskanning har startat, men det kan ta uppemot 8 år innan den är komplett.

Fjärranalysskattningar av skogliga data kräver referensdata från fält. Skördardata som fältreferens ger betydligt lägre kostnader än fältinventeringar och standardiserade skördardata samlas in kontinuerligt. De beskriver den avverkade skogen och genom att använda den informationen med Skogliga Grunddata kan skattningen över ännu icke avverkade skogar ajourhållas.

Projektet har utvärderat imputering som en metod för att ajourhålla skogliga skattningar med skördardata. Imputering används för utbytesprognoser från skördardata och kan implementeras i skördardatabaser. Idag är utvecklingen inom Artificiell Intelligens och Machine Learning mycket snabb, och projektet har parallellt med imputering även utforskat möjligheter att använda Deep Learning (DL, djupa neurala nätverk) som en alternativ metod. Jämfört med imputering är detta en mycket mer komplex metod som kräver mycket stora datamängder och processkapacitet, men har potentialen att automatiskt hitta och utnyttja mycket komplexa samband för att skatta variabler som är oåtkomliga med andra metoder. Tillämpningar med skördardata i kombination med nationellt tillgängliga fjärranalysdata är sannolikt ovanligt väl lämpade för Deep Learning eftersom datamängderna är så stora och därför att värdet av att skatta de komplexa målvariabler som behövs för utbytesprognoser och virkesegenskaper är stora.

Resultaten visar att ajourhållning med skördardata är möjligt och mer lämpligt i skogar där gallring utförts. Dels för att det saknas skördardata i skogar innan första gallring, men även att det i samband med ingreppen blir svårt att effektivt ajourhålla på annat sätt, Medelfelet ligger på över 30 % vilket är betydligt högre än vad som är förväntat direkt efter skanningstidpunkten, men behöver jämföras med alternativa metoder (exempelvis framskrivning med funktioner) för att bedöma om det är tillräckligt bra. Denna jämförelse blir möjlig redan under våren 2020 då Skogsstyrelsen lanserar framräknade skogliga skattningar.

Projektet har preliminärt visat att DL är en framkomlig metod för denna tillämpning, men kräver mycket resurser för att sammanställa stora datamängder, konstruera nätverk som är effektiva och inte överanpassade, samt processkraft för att träna nätverk. Det finns ännu inga givna lösningar, men fortsatt utveckling kommer sannolikt att kunna ge bra skattningar för många fler målvariabler än projektet har utvärderat.

Resultat

Analysen bygger på tillgängliga skogliga grunddata från Skogsstyrelsen som ska ajourhållas med hjälp av skördardata. Skogforsk har byggt upp en skördardatabas för forskningsändamål som täcker stora delar av landet, bilaga 1, figur 1. Skördardata har använts från tidpunkter runt fältinventeringstidpunkt för utvärderingsdata, år 2018. I databasen fanns då cirka 20 000 avverkningar.

För att utvärdera ajourhållningen har Sveaskog ställt sin kontrolltaxering till förfogande. Det rör sig om drygt 2000 provbestånd över landet med koordinatsatta provytor i varje bestånd. Av dessa sorterades ett antal bort, på grund av inkompleta data eller en taxerad medelhöjd (hgv) under 10 meter, då skördare sällan används innan träden nått denna höjd. Detta ledde till att 1282 provbestånd användes för att skatta skogstillståndet vid taxeringstidpunkten. Genom att korsreferera provbestånden med utförda åtgärder identifierades 153 bestånd som gallrats mellan laserskanning och utförd taxering. Från skördardata skapades homogena referensytor, i medeltal 0,8 ha stora, baserat på de avverkade trädens positioner och höjd. Skogens tillstånd beräknades för varje referensyta utifrån skördarens mätning av de avverkade träden. Uppskattad volym, grundyta, hgv och dgv samt skanningsdatum sammanställdes från Skogliga grunddata för alla provbestånd respektive referensytor. För referensytorna beräknades ålder och ståndortsindex (SI) med hjälp av temperatursumma och hgv. För varje provbestånd valdes de 1000 geografiskt närmaste referensytorna, som hade laserskannats samt avverkats inom plus/minus ett år från provbeståndets skanningsdatum respektive taxeringsdatum ut för imputering.

Skattningsmetoden, Most Similar Neighbor (k-MSN), som användes är en s.k. imputeringsteknik som väljer ut ett antal (k) referensytor som är mest lika provbeståndet med avseende på ett antal oberoende variabler, i detta fall: hgv och grundyta från skogliga grunddata samt ålder och SI. Skattningen av det skogliga tillståndet vid taxeringstidpunkten bestäms genom medelvärdesbildning av volym, grundyta, hgv och dgv från de (k) utvalda referensytorna.

Resultaten av skattningarna utvärderades genom att beräkna Root Mean Square Error (RMSE) och systematiska avvikelser, utvärderingen presenteras för hela landet (Sverige) samt uppdelat på Götaland, Svealand och södra Norrland (M Sverige) och de fem nordligaste länen (N Norrland) (bilaga 1, tab 1, fig 2). För volym och grundyta blev RMSE högt, 34 % resp. 32 %, medan hgv och dgv skattades bättre 10 % resp. 17 %. I figur 2 ses att volym och grundyta systematiskt överskattades, framförallt i norra Norrland, samtidigt som hgv och dgv visade marginella systematiska avvikelser på 2-3 %.

För de 153 bestånd som gallrats efter laserskanningen, imputerades en skattad beskrivning av kvarvarande bestånd baserat på en modell för automatisk gallringsuppföljning utvecklad av Skogforsk. Dessa resultat var lovande, då RMSE var lägre och utan de stora systematiska fel för volym och grundyta som redovisats i bilaga 1, figur 2. Även skattningar för äldre skog, bestånd över 70 år, hade betydligt lägre systematiska fel än vi ser i figur 2.

Skattningar för skog under 70 år däremot gav störst spridning och störst systematiska fel, speciellt norra Norrland, där volym och grundyta systematiskt överskattades med 35 % resp. 33 % jämfört med 12 % resp. 16 % för övriga landet. Medelavståndet till de imputerade referensytorna var 50 % längre i norra Norrland jämfört med övriga landet vilket antyder att referensmaterialet var för litet i detta område, vilket kan förklara en del av de stora skillnaderna. Åldersberäkningen för skördardata är en svag länk, om beståndsålder för referensdata hade varit tillgängligt skulle SI kunna beräknas bättre och skattningarna sannolikt förbättras.

Inom projektet utvecklades djupa neurala nätverk (DL) för att skatta variablerna med laserdata baserat på samma träningsdata som i imputeringen. För varje referensyta extraherades laserdata och från dessa beräknades ett hundratal beskrivande variabler (metriker) som användes som oberoende

variabler i nätverken. Arbetet inriktades mot att hitta lämpligt stora nätverk som levererar tillförlitliga skattningar utan överanpassning men ändå tillräckligt djupa för ändamålet. Metodiken visade sig vara klart användbar för denna problemställning, levererade likartad precision som imputering (se några exempel i Tab. 3-5, Fig. 2-4, bilaga 2), men också skattningar av fler målvariabler inklusive sortimentsutbyte (Tab. 5, Fig. 4, bilaga 2). DL kräver däremot mycket mer utveckling och processkraft än imputering. Beroendet av mycket stora datamängder är tydligt, där risken för överanpassning visade sig stor. Enkla nät, med ett djup på ca 4-5 gömda lager fungerade approximativt lika bra som djupare. Fortsatt utveckling kräver framförallt stor processkraft för att kunna träna komplexa nät tillräckligt länge, vilket för närvarande sker i samarbete med AI Innovation Sweden i Göteborg.

Målbeskrivning

Projektets mål var att ta fram kostnadseffektiva metoder att ajourhålla skogliga laserskattningar med skördardata. Två metoder utvecklades, en som bygger på tillgängliga registerdata och en annan för användning på landskapsnivå utan krav på registerdata.

Metoderna utvärderades tillsammans med skogsföretag mot objektiva stickprovsinventeringar.

Projektets arbetsrapport (publiceras under våren 2020) ska innehålla tydligt beskrivna metoder för ajourhållning som direkt ska kunna implementeras i en databasmiljö som flera skogsbolag bygger upp med skördardata.

De planerade aktiviteterna har genomförts, men metodutvecklingen tog längre tid än förväntat. Detta beroende på att studieområdet skalades upp till hela landet för att få tillgång till de stora mängder träningsdata som AI-metoder kräver. Större mängder skördardata fick bearbetas för att kunna nyttjas i Deep Learning. Projektet upplevde även att analysen av framtagna resultat krävde mer tid än planerat då metoderna kan kräva detta. Vid ajourhållning av skogstyper som inte är vanliga i träningsdata (skördardata) kunde det ibland bli stora geografiska avstånd för att hitta liknande skogar och då kan ajourhållningen försämrats.

Detta leder till att all planerad kommunikation inom ramen för projektiden inte genomförts, men planeras för 2020. Samtidigt har projektet lett till kompletterande finansiering för att kunna fortsätta med forskningen. Flera skogsföretag har under projektiden uppmärksammat möjligheten att använda skördardata inom AI och våra erfarenheter kan stötta vidare insatser kring användningen av skördardata för ajourhållning.

Under våren 2020 kommer Skogsstyrelsen att publicera en ajourhållning av skogliga grunddata med tillväxtfunktioner och vi kommer att jämföra våra resultat med deras för att visa på träffsäkerheten i olika typer av ajourhållning av skogliga fjärranalysskattningar

Kommunikation och nyttiggörande av resultat

Skogforsk har informerat om att projektet pågår inom samverkansgruppen för skoglig planering (PSG) och kommer redovisa projektresultaten i samband med inplanerade möten. Det senaste PSG-mötet hade ett tema kring artificiell intelligens i skoglig planering och kommande möte våren 2020 har ett tema kring skördardata för ajourhållning där våra resultat ska presenteras.

Delresultaten har kommit under hösten 2019 varför vi sätter samman en mer komplett populärredovisning (arbetsrapport) under våren 2020. I samband med arbetsrapporten kommer även webbnyheter om projektet. Projektet har haft en referensgrupp med representanter från Sveaskog, Holmen, Södra och Skogssällskapet som hållit möten där projektresultat redovisats och har ett avslutande möte sker i vår.

Under 2019 har Jon Söderberg på Skogforsk erhållit medel för en forskarutbildning där den planerade vetenskapliga artikeln passar väl in i utbildningen och manus kan skrivas under 2020. Den detaljerade planeringen för forskarutbildningen är emellertid inte fastställd vid detta rapporteringstillfälle. Oavsett detta planeras för en vetenskaplig publikation under 2020.

Slutligen drar planeringen för RIU konferensen 2020 igång under våren där en presentation av hur skördardata kan komma in i ajourhållningen av skogliga data kan inkluderas.