



Verktyg för den agronomiska detektiven

– Hitta den lokalt nåbara skördenivån och de
skördebegränsande faktorerna

K. Persson, B. Stenberg, O. Alshihabi, M. Söderström

Sveriges lantbruksuniversitet, SLU
Fakulteten för naturresurser och jordbruksvetenskap
Institutionen för mark och miljö
2026



Verktyg för den agronomiska detektiven – Hitta den lokalt nåbara skördenivån och de skördebegränsande faktorerna

Kristin Persson, Sveriges lantbruksuniversitet, Institutionen för mark och miljö,
<https://orcid.org/0000-0003-2120-4486>

Bo Stenberg, Sveriges lantbruksuniversitet, Institutionen för mark och miljö,
<https://orcid.org/0000-0003-3018-4990>

Omran Alshihabi, Sveriges lantbruksuniversitet, Institutionen för mark och miljö,
<https://orcid.org/0000-0002-1973-1757>

Mats Söderström Sveriges lantbruksuniversitet, Institutionen för mark och miljö,
<https://orcid.org/0000-0001-9946-0979>

Utgivare:	Sveriges lantbruksuniversitet, Fakulteten för naturresurser och jordbruksvetenskap/Institutionen för mark och miljö
Utgivningsår:	2026
Utgivningsort:	Skara
Omslagsbild:	Kristin Persson
ISBN (elektroniskt):	978-91-8124-150-1
DOI:	https://doi.org/10.54612/a.gghvf2fhv8
Nyckelord:	beslutsunderlag, fjärranalys, karta, korn, precisionsodling, prognos, skörd, vete.

© 2026 Författarna

Detta verk är licensierat under CC BY 4.0, <https://creativecommons.org/licenses/by/4.0/>. Andra licenser eller upphovsrätt kan gälla för illustrationer.

Förord

I föreliggande rapport redovisas resultat från forskningsprojektet *Verktyg för den agronomiska detektiven – Hitta den lokalt nåbara skördenivån och de skördebegränsande faktorerna*. Projektet syftade till att utveckla metoder för detaljerad skördekartering genom en kombination av optiska satellitdata och insamlade uppgifter bakom nationell skördestatistik, samt att utreda hur sådana kartor kan användas för att komma till rätta med lokala produktionsproblem. Med bättre kartunderlag kan man öka produktionen i områden där skördepotentialen i dagsläget inte utnyttjas till fullo, samtidigt som man kan dra ner på mängden insatsmedel i områden där skördepotentialen är liten. Det bör bidra till en högre och mer resurseffektiv spannmålsproduktion och tar oss mot målsättningarna i den svenska livsmedelsstrategin och EU:s jord till bord-strategi i den europeiska gröna given.

Projektet har genomförts i samverkan mellan Sveriges lantbruksuniversitet (SLU) och Lantmännen. Det finansierades av Lantmännens Forskningsstiftelse (dnr: 2023H016). Vi vill särskilt tacka Emelie Ekholm, Johan Wågstad och Andreas Persson på Lantmännen, som medverkat i projektgruppen. Vi vill även tacka Statistiska Centralbyrån (SCB), Jordbruksverket och de fyra projektgårdarna, Alnarp, Bjertorp, Hidinge och Lövsta, som bidragit med data till projektet samt goda diskussioner av möjligheter och utmaningar med satellitbaserad skördekartering.

Författarna, januari 2026



Redovisning av användning av AI: Perplexity AI har använts för översättningar, språkgranskning och språklig förbättring. All text har därefter granskats manuellt och författarna ansvarar fullt ut för innehållet.

Sammanfattning

I projektet utvecklade och utvärderade vi ett modelleringsupplägg för satellitbaserad skördekartering på regional nivå, fältnivå och inom fält. Modelleringen baserades på Sentinel-2-bilder i kombination med medelskördar per gård insamlade av SCB. Genomsnittligt medelabsolutfel (MAE) för predikterade gårdsmedelskördar var $0,7 \text{ t ha}^{-1}$ för höstvet (*Triticum aestivum* L.) och $0,8 \text{ t ha}^{-1}$ för vårkorn (*Hordeum vulgare* L.) när prediktionerna gjordes efter skörd, och något högre ($0,9$ respektive $0,8 \text{ t ha}^{-1}$) vid skördeprognoser i slutet av juni. Prognoser på regional nivå var mer träffsäkra (MAE: $0,6$ respektive $0,5 \text{ t ha}^{-1}$ för de båda grödorna), och regionala medelskördar kunde skattas med bibehållen noggrannhet även när antalet gårdsmedelskördar i kalibreringen minskades med en tredjedel.

Jämförelser mellan satellitbaserade skördekartor och skördekartor från tröskor (20 m upplösning) visade varierande överensstämmelse, vilket förmodligen till stor del förklaras av osäkerheter i tröskdata, men även i viss mån av osäkerhet i de satellitbaserade kartorna samt den utjämnande effekt som modelleringen ger. Båda datakällorna har begränsningar, men satellitbaserade kartor har fördelen att de kan tas fram utan extra utrustning och med mindre krav på databearbetning.

För den aktuella tillämpningen fungerade enkla linjära modeller med två vegetationsindex (NDWI och NDRE76) lika bra eller bättre än mer komplex maskininlärning med många index, vilken tenderade till överanpassning och sämre generalisering till nya gårdar och år. Detta betonar vikten av robust validering när kalibreringsdata är osäkra.

Satellitbaserade skördekartor från flera år lades samman och två typer av kartor togs fram: relativa skördekartor och frekvenskartor för låg skörd. Dessa användes för att identifiera stabila och instabila problemområden inom fält. Variationsmönstren stämde ofta överens med lantbrukarnas lokalkännedom om till exempel dräneringsproblem, viltskador och markanvändningshistorik, men kartorna avslöjade också tidigare mindre uppmärksammade områden med låg skörd. Kartunderlagen kan användas både för att stödja åtgärder för att höja skörden (där det är möjligt) och för att anpassa insatser till den faktiska skördepotentialen (där förbättring inte är möjlig).

Moln, molnskuggor och dis begränsar ibland möjligheten till satellitbaserad skördekartering. Ett mindre test med syntetiska molnfria Sentinel-2-tidsserier indikerade att sådana data kan vara en gångbar lösning. Testet visade också att sambandet mellan optiska vegetationsindex och spannmålsskörd ofta är starkast i senare delen av juni (sen blomning–tidig mjölkmodnad).

Mot bakgrund av resultaten rekommenderas fortsatt utveckling av användarvänliga system med: (i) förbättrad och förenklad hantering av skördedata från tröskor, (ii) integrering av skördekartering (både trösk- och satellitbaserad) i beslutsstödsystem för precisionsodling, (iii) funktioner för flerårsanalyser och manuell tolkning av dessa med stöd av kompletterande data, samt (iv) möjligheter att analysera relativa skördar inom närområden (t.ex. inom några kilometers radie) som underlag för rådgivning och en produktiv och resurseffektiv växtodling.

Nyckelord: beslutsunderlag, fjärranalys, karta, korn, precisionsodling, prognos, skörd, vete.

Abstract

The project developed and evaluated a modelling approach for satellite-based yield mapping at regional, farm, and within-field scales. The modelling used Sentinel-2 imagery in combination with average yields per farm collected by Statistics Sweden (SCB). The mean absolute error (MAE) for predicted average farm yields was 0.7 t ha⁻¹ for winter wheat (*Triticum aestivum* L.) and 0.8 t ha⁻¹ for spring barley (*Hordeum vulgare* L.) when predictions were made after harvest, and slightly higher (0.9 and 0.8 t ha⁻¹, respectively) for forecasts made at the end of June. Forecasts at the regional scale were more accurate (MAE: 0.6 and 0.5 t ha⁻¹ respectively for both crops), and regional average yields could be estimated with similar accuracy even when the number of farm yields in the calibration dataset was reduced by a third.

Comparisons between satellite-based yield maps and yield maps derived from combine harvesters (20 m spatial resolution) showed varying levels of agreement. This is likely primarily due to uncertainties in the harvester data, but also partly to uncertainties in the satellite-based maps and the smoothing effect introduced by the modelling. Both data sources have limitations, but satellite-based maps have the advantage that they can be produced without additional equipment and require less data processing.

For the intended application, simple linear models using two vegetation indices (NDWI and NDRE76) performed as well as, or better than, more complex machine learning models using many indices. The latter tended to overfit and to generalise less well to new farms and years, highlighting the importance of robust validation, especially when calibration data are uncertain.

Satellite-based yield maps from multiple years were combined, and two types of maps were produced: relative yield maps and frequency maps of low yields. These were used to identify stable and unstable low-yielding areas within fields. The spatial patterns often matched the farmers' local knowledge of, for example, drainage problems, wildlife damage, and land-use history, but the maps also revealed previously less noticed areas with low yields. The maps can be used both to support measures to increase yields (where possible) and to adapt inputs to the actual yield potential (target yield) (where improvement is not possible).

Clouds, cloud shadows and haze sometimes limit the potential for satellite-based yield mapping. A small test using synthetic cloud-free Sentinel-2 time series indicated that such data can be a viable solution and showed that the relationship between optical vegetation indices and cereal yield is often strongest in the latter part of June (from late flowering to early milk development).

In light of these results, continued development of user-friendly systems is recommended, including: (i) improved and simplified handling of yield data from combine harvesters, (ii) integration of yield mapping (both harvester- and satellite-based) into decision-support systems for precision agriculture, (iii) functions for multi-year analyses and manual interpretation supported by complementary data, and (iv) options to analyse relative yields within local areas (e.g. within a radius of a few kilometres) to support advisory services and high-yielding, resource-efficient crop production.

Keywords: barley, decision support, forecasting, mapping, precision agriculture, remote sensing, wheat, yield.

Innehållsförteckning

1. Introduktion	7
1.1 Skördekartering med tröska.....	7
1.2 Satellitbaserad skördekartering	7
1.3 Stabil och instabil skördevariation	8
1.4 Beslutsstödsystem	9
1.5 Vision och mål.....	9
2. Material och metoder.....	10
2.1 Områden, gårdar, år och grödor	10
2.2 Data och databeredning.....	11
2.2.1 Skörderegistreringar från gårdarnas tröskor.....	11
2.2.2 Satellitdata	12
2.2.3 Mikrodata från SCB	13
2.2.4 Gårdarnas markkarteringar.....	14
2.2.5 Geodatabaser	14
2.2.6 Syntetiska satellitdata.....	14
2.3 Dataanalys	15
2.3.1 Skördekartering via satellit.....	15
2.3.2 Sammanläggning av flera års skördekartor	17
2.3.3 Tolkning av skördevariation	18
3. Resultat och diskussion.....	20
3.1 Utvärdering av satellitbaserade skördekartor	20
3.1.1 Skördekartor i flera skalor.....	20
3.1.2 Jämförelse mellan enkla och komplexa modeller.....	21
3.1.3 Utvärdering av satellitkartor för gårdar och områden	22
3.1.4 Jämförelse mellan tröskkartor och satellitkartor	24
3.2 Flerårsanalys.....	26
3.2.1 Kartor över relativskörd och frekvensen av låg skörd.	26
3.2.2 Tolkning av orsaker till låg skörd	26
3.2.3 Erfarenheter från diskussioner med lantbrukargruppen	27
3.3 Modellering med ofullständiga datamängder.....	28
3.3.1 Möjliga lösningar	28
3.3.2 Test av syntetiska satellitdata.....	29
3.4 Identifierade forskningsbehov	30
3.4.1 Fältkanter är svåra att kartlägga	30
3.4.2 Satellitbaserade kartor jämnar ut skördevariationen	31
4. Konklusioner	32
Referenser.....	35

1. Introduktion

Skördekartor från flera år ger ett bra underlag för att identifiera lokala produktionsproblem och bedöma möjligheterna att åtgärda dem. De kan också användas för att fastställa en realistisk målskörd. Att åtgärda begränsande faktorer är avgörande för en förbättrad produktion, medan en realistisk målskörd är nödvändig för att optimera gödslingen och öka utnyttjandegraden av tillförd växtnäring (Raun et al., 2017). På vissa platser i världen finns t.o.m. regler som kräver att lantbrukare använder så realistiska målskördar som möjligt i växtnärringsplaneringen (Shober & Taylor, 2025).

När orsaker till låg skörd har identifierats kan olika strategier tillämpas beroende på vad som begränsar. Vissa problem, till exempel bristfällig dränering eller lågt pH-värde, kan åtgärdas genom markvårdande insatser. Om begränsningen däremot är kopplad till faktorer som jordart eller topografi bör odlingen anpassas till de naturliga förhållandena för att använda resurserna effektivt. I vissa fall kan det mest hållbara alternativet vara att ta en fältdel ur produktion.

Man bör ha i åtanke att små områden med mycket låg skörd sällan är de som har störst potential att höja den totala skörden. Det är snarare i de områden som redan ger normal eller högre skörd som man kan få störst effekt på den totala produktionen genom att åtgärda skördebegränsande faktorer. För att kunna göra detta behövs tillförlitliga skördekartor.

1.1 Skördekartering med tröska

Utrustning för skördemätning på tröskor har funnits sedan mitten av 1990-talet. Då trodde man att tekniken skulle bli grunden för platsspecifik växtodling (Schnug et al., 1993; Birrell et al., 1996; Nissen & Söderström, 1999). Så har det dock inte blivit. Många lantbrukare har utrustningen, men av olika skäl är det fortfarande svårt att ta fram tillförlitliga skördekartor, och även om man lyckas har man inget system för att lätt använda dem. Resultaten av en fransk undersökning visade att de flesta av de lantbrukare som hade skördekarteringsutrustning endast använde sina system för realtidsvisualisering under pågående skörd; mindre än 25% tog fram och använde skördekartor (Lachia et al., 2021). Tidigare undersökningar från USA, Storbritannien och Australien visade liknande trender (Lowenberg-DeBoer et al., 2019). Även i Sverige råder osäkerhet bland lantbrukare om hur man ska hantera och använda data från tröskans skördekartering (Johansson et al., 2022).

1.2 Satellitbaserad skördekartering

Med dagens goda tillgång till satellitbilder har det öppnats nya möjligheter att enkelt ta fram skördekartor med hög upplösning; satellitdata (exempelvis från

Europeiska Rymdstyrelsen ESA:s Sentinel-2-program) kan kalibreras mot uppmätta skördar med statistiska modeller. Runt om i världen pågår forsknings- och utvecklingsarbete, och resultaten är lovande (t.ex. Deines et al., 2021; Hunt et al., 2019; Perich et al., 2023; Åström et al., 2025).

I Sverige har flera studier genomförts inom området. Alshihabi et al., (2023) testade en strategi där modellerna parametriserades individuellt för varje satellitbild med hjälp av data från en välkalibrerad tröska. Studien utfördes i höstvetete (*Triticum aestivum* L.). Müller et al. (2025) använde Sentinel-2-data för att undersöka effekten av torka på skörd i regional skala för flera olika grödor, och Li et al. (2024) kombinerade olika satellitdatakällor och klimatvariabler för att göra skördeprognoser på områdesnivå. I den sistnämnda studien visades att det var möjligt att förutsäga skörden av vårkorn (*Hordeum vulgare* L.) i regional skala i Skåne redan i slutet av maj, med en felmarginal på cirka $\pm 0,5 \text{ t ha}^{-1}$.

Förutsatt att molnfria bilder och lämpliga kalibreringsdata finns tillgängliga kan satellitbaserad skördekartering även göras för tidigare år. Kartor kan också tas fram på olika aggregeringsnivåer: kartor över fältmedelvärden möjliggör analyser av produktionsbegränsningar i större skala, medan skördekartor på fältnivå ger lantbrukare möjlighet att förbättra sina metoder.

1.3 Stabil och instabil skördevariation

Så kallade skördestabilitetskarteringar (generaliserade skördenivåkartor baserade på skördekartor från många år) har tagits fram i olika forskningsprojekt sedan 1990-talet (t.ex. Blackmore, 2000). Ofta kan fält delas in i zoner med konstant hög skörd, zoner med konstant låg skörd och en instabil zon som ger olika skörd, ofta beroende på om året är blött eller torrt (Algerbo et al., 2003; Delin, 2005). Blasch et al. (2020) beskrev till exempel en metod (Multi-temporal Yield Pattern Analysis; MYPA) för att ta fram stabila och icke stabila mönster från tidsserier av skördekartor som omfattar olika grödor. Maestrini & Basso (2021) använde en statistisk ansats och testade en rangbaserad klassificeringsalgoritm (Two-Way-Outlier), vars styrka är att bättre skilja mellan fältdelar med konstant låg skörd och fältdelar med vanligtvis låg skörd. McEntee et al. (2020) visade att det är möjligt att klassificera produktionen i olika delar av fält som hög och stabil, hög och instabil, låg och stabil respektive låg och instabil, även när både spannmål och vall ingår i tidsserien av skördekartor.

1.4 Beslutsstödsystem

Digitala beslutsstödsystem (DSS) för växtodling kan göra det möjligt att skala upp satellitbaserad skördekartering och göra den tillgänglig för alla lantbrukare. I Sverige finns redan flera DSS som används för precisionsodling. Att integrera satellitbaserad skördekartering, med stöd för tolkning och som beslutsunderlag, vore därför ett naturligt nästa steg.

1.5 Vision och mål

En begränsning med de metoder för satellitbaserad skördekartering som beskrevs ovan är ofta tillgången på kalibreringsdata. Därför genomfördes denna studie, där vi undersöker möjligheten att använda mikrodata bakom nationell statistik för modelleringen. Den långsiktiga visionen är att det ska finnas lättillgängliga och användarvänliga digitala verktyg, data och metoder för att identifiera lokala skördegap i spannmålsproduktionen, samt för det detektivarbete som krävs för att hitta orsakerna till dessa. Målet är i förlängningen att öka produktionen och förbättra resurseffektiviteten inom jordbruket. De specifika målen med det aktuella projektet var att:

- a) vidareutveckla och testa metoder för skördekartering på regional nivå, fältnivå och inomfältsnivå baserat på satellitdata och skördeuppgifter på gårdsnivå från statistiska centralbyrån (SCB).
- b) utforma en praktiskt användbar metod för att avgränsa stabila och instabila problemområden med hjälp av satellitbaserade skördekartor.
- c) utveckla en strategi för att identifiera orsaker till låga skördar genom att kombinera öppet tillgängliga datamängder med lantbrukarens data och lokalkunskap.
- d) sammanfatta rekommendationer för hur ett DSS kan utformas så att lantbrukare kan utvärdera lokala skördegap inom och mellan fält samt få tillgång till relevanta data för att analysera orsakerna.

2. Material och metoder

2.1 Områden, gårdar, år och grödor

Projektområdet utgjordes av sex av SCB:s 108 skördeområden (SKO). Dessa SKO visas i Figur 1. Områdena kännetecknas av en stor andel jordbruksmark, där spannmålsodling är en viktig del: kring Uppsala (SKO 311), Örebro (SKO 1812), Vara (SKO 1622) och Lund (SKO 1213, 1214 och 1216, som slogs samman). Vi arbetade även med data från en specifik gård inom respektive område, Lövsta (Uppsala), Hidinge (Örebro), Bjertorp (Vara) och Alnarp (Lund). I projektet studerade vi åren 2017–2023. Grödorna var höstveten och vårkorn.



Figur 1. Projektområdet utgjordes av fyra delområden som innefattade sex av Statistiska Centralbyråns (SCB:s) skördeområden (orange). I texten benämns de med de blåmarkerade namnen.

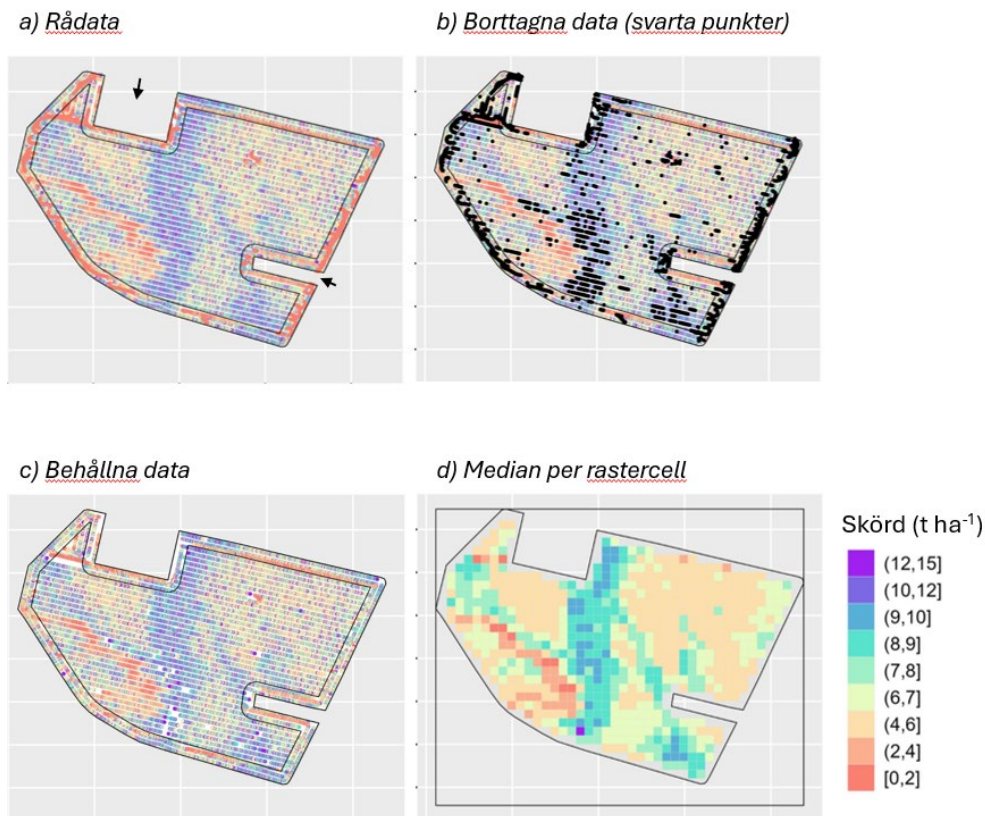
2.2 Data och databeredning

2.2.1 Skörderegistreringar från gårdarnas tröskor

Vi samlade in skördedata från tröskor från de fyra gårdarna, samt data om genomsnittlig skörd per fält. Data bereddes enligt följande protokoll (se även Figur 2):

- 1) **Urval.** Data begränsades till fält med höstvet eller vårkorn. Områden med fältförsök eller andra tydliga avvikelser (annan gröda etc.) togs bort.
- 2) **Filtrering.** Inom varje fält filtrerades skördedatan från felaktiga registreringar (punkter):
 - a) *Stopp:* De 15 sista punkterna före stopp och de 15 första efter stopp.
 - b) *Sväng:* Punkter utanför gränsvärden för ett svängindex beräknat från riktning.
 - c) *Arbetsbredd:* Punkter med mindre än 97,5 % fullt skärbord.
 - d) *Punkttäthet i kördraget:* Punkter närmare än 0,05 m från föregående punkt och punkter utanför nedre och övre gränser (0,5 m och 5 m) för medelavståndet till närmaste punkt, beräknat för de tio senaste och tio efterföljande punkterna.
 - f) *Punkttäthet per yta:* Punkter med fler än 40 observationer inom 5 m radie.
 - g) *Motorns varvtal:* Punkter där det varvtalet var under 1 RPM.
 - h) *Skörd:* Punkter med de 0,5 % lägsta och 0,5 % högsta skördarna.
- 3) **Interpolation.** Medianvärdet av skörden beräknades per 20 m × 20 m Sentinel-2-rastercell. Celler med färre än fem observationer och celler som korsade fältgränsen uteslöts.
- 4) **Nivåjustering.** Den genererade skördekartans medelvärde delades med den invägda genomsnittliga skörden för fältet, där sådan information fanns tillgänglig.

Algoritmen programmerades i R och finns tillgänglig via GitHub (https://github.com/lads-slu; target_f). De fyra gårdarna använde olika system för skördemätning (CLAAS, Harsewinkel, Tyskland; Trimble, Sunnyvale, CA, USA; Fendt, Marktoberdorf, Tyskland), och anpassningar gjordes för att passa respektive datatyp (t.ex. vilka filter och tröskelvärden som användes).



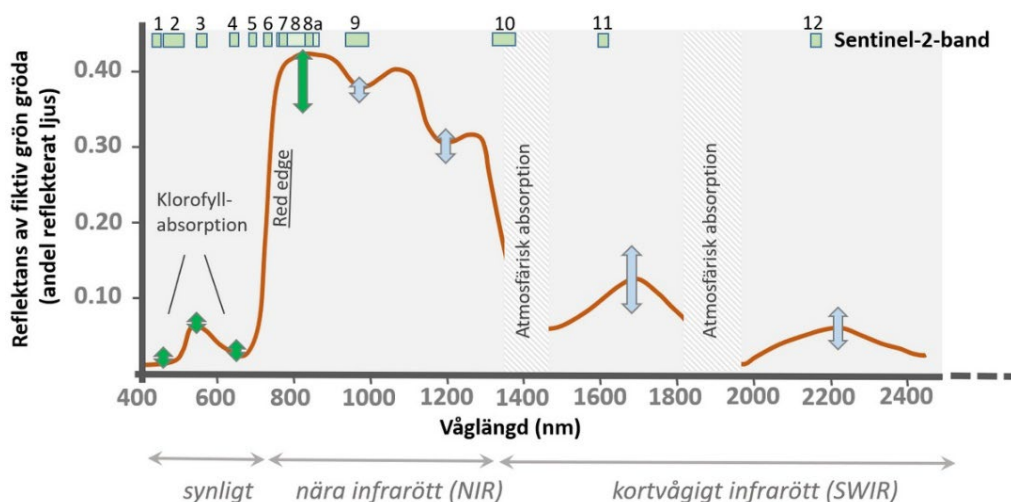
Figur 2. Exempel på filtreringsproceduren. De rektangulära områden som markeras med pilar togs bort på grund av fältförsök.

2.2.2 Satellitdata

Vi använde satellitdata från Sentinel-2. Sentinel-2 är en konstellation av två optiska satelliter som passerar Sverige ungefär två dagar av fem. En Sentinel-2-bild består av 13 olika våglängdsband, från synligt ljus till kortvågigt nära infrarött ljus (Figur 3). Relativt molnfria (enligt okulär granskning) L2A-bilder (dvs. atmosfäriskt korrigerade bilder som visar markytans reflektans) laddades ned från Copernicus Data Space Ecosystem (<https://dataspace.copernicus.eu>) för vidare hantering.

Vi använde alla våglängdsband utom band 1, 9 och 10, som i huvudsak används för moln- och disdetektion. Data från övriga band räknades om till 20 m × 20 m rasterceller (band 2–4 och 8 har ursprungligen 10 m pixlar). Exakt samma raster användes även för alla andra data i projektet för att möjliggöra analyser och jämförelser mellan olika typer av data.

Detaljerad molnmaskning (avgränsning av ytor som var täckta av moln, dis, eller molnskuggor) gjordes i huvudsak genom manuell klassning, i vissa fall i kombination med oövervakad klassificering (Iso Cluster i mjukvaran ArcGIS Pro 3.5, ESRI Inc., Redlands, CA, USA).



Figur 3 Generaliserad reflektans från en gröda i olika våglängdsband (1–12) i Sentinel-2. De blå och gröna pilarna illustrerar (förenklat) att reflektansen typiskt ändras beroende på förändringar av grödans biomassa och kväveinnehåll (gröna pilar) och vattenhalt (blå pilar). Bilden är omarbetad från Söderström et al., 2025).

För att minska eventuella atmosfäriska störningar i de molnmaskade bilderna beräknas ofta olika vegetationsindex (VI). Sentinel-2-banden 2–8, 8A och 11–12 användes för att beräkna kvotindex (QI) och normaliserade differensindex (NDI) för alla parvisa kombinationer av band. I ekvationerna 1 och 2 anger R reflektans medan i och j betecknar banden. Varje satellitbild genererade därmed ett stort antal VI-variabler, som sedan användes i skördemodelleringen.

$$QI_{ij} = \frac{R_i}{R_j} \quad (\text{Ekvation 1})$$

$$NDI_{ij} = \frac{R_i - R_j}{R_i + R_j} \quad (\text{Ekvation 2})$$

2.2.3 Mikrodata från SCB

Aktuella skördeuppgifter för höstvet och vårkorn på gårdsnivå togs fram av SCB. Det görs varje år för att producera allmän skördestatistik, och uppgifter samlas in genom webbenkät eller telefonintervjuer med ett antal lantbruksföretag inom varje SKO. Det är medelskörden för respektive lantbruk som samlas in, vilket betyder att skördesiffran kan representera flera fält. För att ta reda på vilka fält som odlats med vilka grödor på de olika gårdarna bidrog Jordbruksverket med uppgifter om skiftesidentitet. Skiftesidentiteten kunde via data från Jordbruksverkets skiftesdatabas knytas till rätt fält. Detta gjorde det möjligt att analysera satellitdata för fälten där det finns skördeuppgifter. Grunddata är sekretessbelagda och

anonymiserade. Därför redovisas inte skördeuppgifter eller skördeprediktioner i projektet på ett sådant sätt att de kan knytas till en enskild gård, utom för de fyra projektgårdar som ingick i projektet (Figur 1). Antalet lantbruksföretag för vilka det fanns skördeuppgifter uppgick inom projektområdet till ca 400 per år.

2.2.4 Gårdarnas markkarteringar

För att kunna tolka variationen i skördekartor från satellitmodellerna gjordes kartor över markanalysdata från projektgårdarnas markkarteringar. Gårdarnas markkarteringar skiljde sig något åt, i samtliga fall fanns pH, samt lättlöslig fraktion av fosfor, kalium, kalcium och magnesium (P-AL, K-AL, Ca-AL, Mg-AL) analyserade i ett jordprov per hektar. För tre av gårdarna fanns även koppar (Cu-HCl), mullhalt och lerhalt. På en gård var de senare analyserade i vart tredje prov. I ett fall fanns även analyser av sandhalt, järn och aluminium (Fe-AL, Al-AL). Vi använde den geostatistiska interpolationsmetoden block kriging (Söderström, 2010) för att beräkna markkarteringsvärden i samma raster som satellitbilderna.

2.2.5 Geodatabaser

Även små variationer i topografi inom fält kan vara viktiga för skillnader i skörd. Vi tog fram kartor över fyra olika mått som är baserade på Lantmäteriets detaljerade höjddata (Grid 2+, Lantmäteriet, Gävle) för projektgårdarnas fält. Dessa höjddata har sitt ursprung i laserskanning från flygplan. De olika måtten var: höjd över havet, topografiskt positionsindex TPI_1 och TPI_5 (höjdavvikelse jämfört med omgivande 1 ha, respektive 5 ha) (Weiss, 2001), samt markvattenindexet SMI från SLU:s Markfuktighetskarta². Det senare bygger förutom på topografi även på avstånd till vatten, avrinning och jordartsinformation (Ågren et al., 2021). Dessutom gjordes kartor över den beräknade totala mängden naturlig radioaktiv markstrålning (gammastrålning), baserade på mätdata från flyggeofysiska mätningar som görs av Sveriges Geologiska Undersökning (SGU). På Bjertorp fanns sedan tidigare gammastrålningsmätning gjord med en markburen sensor. Det har i tidigare kartläggningar av åkermark visats att gammastrålningsdata är starkt korrelerad till jordarten i matjorden (Piikki & Söderström, 2019).

2.2.6 Syntetiska satellitdata

Satellitbaserad skördekartering fungerar bäst om man har tillgång till bilder från avgörande utvecklingsstadier i grödans livscykel (Alshihabi et al., 2024). I den här studien, var det för tre fall (områdes-år) av 28 inte möjligt att få fram användbara bilder på grund av utbredd förekomst av moln, molnskuggor eller dis. Därför gjordes ett enkelt test av kommersiella syntetiska satellitdata (ClearSKY Vision,

² <https://www.slu.se/miljoanalys/statistik-och-miljodata/miljodatakatalogen/slu-markfuktighetskarter/>

Aalborg, Danmark). Data som motsvarar Sentinel-2 (spektralt och geometriskt) räknas fram från flera satellitpassager (även andra satelliter) över en viss tidsperiod, vilket resulterar i sömlösa och analysklara bilder.

Pilotstudien gjordes för Bjertorp, med en bild varannan dag under perioden april–augusti för åren 2017–2023. Sambandet mellan tröskdata och flera utvalda VI över tid utvärderades med Spearmans korrelationskoefficient (r). Spearmans r användes i denna studie i stället för Pearsons r , eftersom den baseras på rangordnade data och mäter således monotona samband (oavsett om de är linjära eller inte), medan Pearsons r endast mäter linjära samband.

2.3 Dataanalys

2.3.1 Skördekartering via satellit

Principen för satellitbaserad skördekartering som används i detta projekt illustreras i Figur 4. Förfarandet beskrivs mer detaljerat nedan.

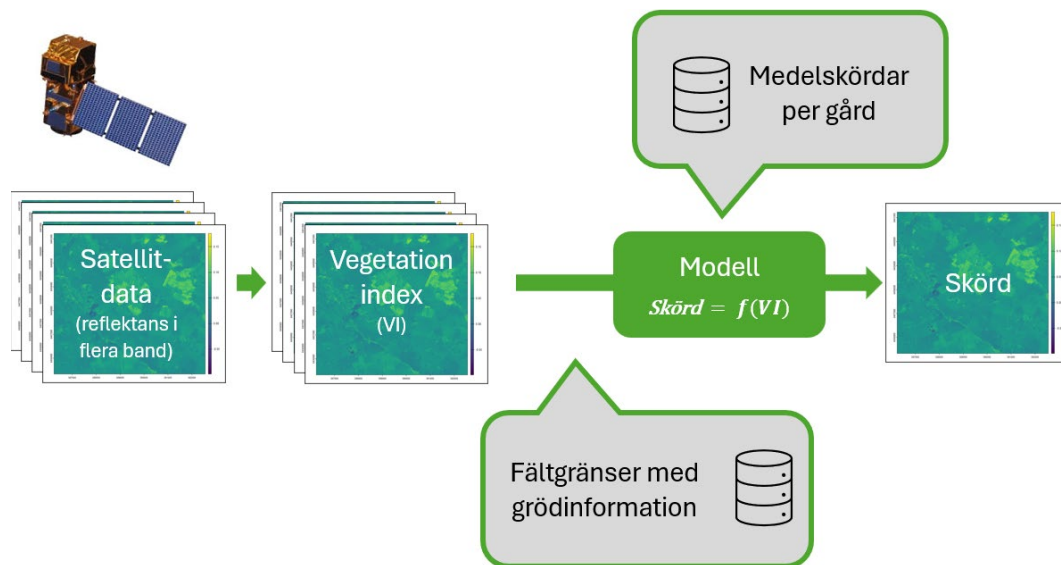
Modeller

Alla modeller använde skörd ($t\ ha^{-1}$) som responsvariabel och två eller flera VI som förklarande variabler. Modellberäkningarna gjordes med två typer av modeller:

- 1) Multipel linjär regression (MLR) med två utvalda VI.
- 2) En maskininlärningsmodell (Random Forest, RF) som inkluderade alla beräknade QI- och NDI-index som förklarande variabler.

I de linjära regressionerna användes dels NDWI (normalized difference water index), vilket är NDI baserat på band 11 och band 8 från Sentinel-2, dels NDRE76, vilket är NDI beräknat utifrån de två *red edge*-banden 7 och 6. NDWI valdes eftersom det tidigare har visat god träffsäkerhet vid prognoser av skörd för höstvet (Alshihabi et al., 2024) och NDRE76 valdes för att det påverkas av biomassans totala kväveinnehåll (Prey & Schmidhalter, 2019; Wolters et al., 2021).

Beräkningarna programmerades i statistikprogrammet R (R Core Team, 2024) och RF-kalibreringen gjordes med standardinställningarna i tilläggspaketet *randomForest* (version 4.7 1.2; Liaw & Wiener, 2002). All modellering genomfördes separat för varje gröda, område och år.



Figur 4. Principen för satellitbaserad skördekartering såsom den gjordes i detta projekt bygger på att vegetationsindex (VI) från Sentinel-2-satelliterna (ESA, Paris, Frankrike) omvandlas till skördekartor. Detta görs med hjälp av enkla eller mer avancerade modeller som har kalibrerats med mikrodatabaser (genomsnittliga skördar per gård) insamlade för nationell statistik av Statistiska centralbyrån (SCB, Örebro, Sverige). Rumslig avgränsning gjordes mha. fältgränser och grödinformation från Jordbruksverket (Jönköping, Sverige)

Modelleringsstrategier

Modellerna utvärderades med följande tester:

- 1) En k-faldig korsvalidering, där resultaten visar hur väl modellerna fungerar på gårdsmedelnivå om de tillämpas efter skörd, det vill säga när kalibreringsdata från samma år finns tillgängliga. Korsvalidering innebär att observationerna slumpvis delas upp i k antal segment där ett segment i taget utesluts vid kalibrering och istället utgör valideringsprov. Valideringsresultaten beräknas sedan på samtliga k valideringssegment sammanslagna. En slumpmässig trettiondel (k=30) eller en slumpmässig tredjedel (k=3) av gårdarna uteslöts i varje iteration. Detta beskriver kartläggning av uppnådd skörd (skördekartering).
- 2) En årsvis korsvalidering, där ett år uteslöts i varje iteration. Detta efterliknar en skördeprognos före skörd. Hela datasetet (alla gårdsmedelvärden) användes både för kalibrering och validering. Detta gjordes för att testa hur känsliga de två modelltyperna (MLR och RF) är för överanpassning.
- 3) Slutliga modeller (kalibrerade med alla data) tillämpades på satellitbildsraster (20 m upplösning) och jämfördes med skördekartor framtagna från tröskdata.

- 4) Prognoserna på gårdsnivå från (1) och (2) aggregerades också till områdesmedelvärden och jämfördes med områdesmedelvärden av rapporterade skördar.

Utvärderingsmått

Fyra valideringsmått beräknades för att kvantifiera olika aspekter av modellernas tillförlitlighet. Dessa var Nash-Sutcliffes modelleringseffektivitet (E), determinationskoefficienten för en linjär regression mellan de predikterade värdena och referensvärdena (r^2), medelabsolutfel (MAE) och medelfel (ME).

E-värdet visar hur stor del av variationen i referensskördarna, som förklaras av modellen. r^2 -värdet beskriver i vilken grad de predikterade skördarna och referensskördarna är korrelerade, även om den absoluta nivån skiljer sig. Medelabsolutfelet (MAE) anger den genomsnittliga storleken på prediktionsfelen, oavsett om de är över- eller underskattningar, medan ME visar storleken på en systematisk avvikelse (en generell överskattning eller underskattning, dvs. en skevhet).

E-värdet 1 betyder att de predikterade och rapporterade värdena är identiska – modellen ger alltså perfekta prediktioner. E-värdet 0 innebär att modellens prediktioner är lika träffsäkra som att använda medelvärdet av referensskördarna för alla fall. Negativa E-värden visar att de predikterade värdena är mindre rätt än om de alla ersatts med medelvärdet av referensvärdena, och sådana modeller är inte användbara.

Utvärderingsmått för områdesmedelvärden beräknades för hela datamängden. Mått för gårdsmedelvärden beräknades per år och område (medianvärden och interkvartilavstånd av dessa redovisas) och mått för jämförelser med tröskdata beräknades per år och fält. Datagrupperingen är viktig för E och r^2 , eftersom dessa värden utvärderar olika aspekter av hur väl variationen i referensdata fångas. För jämförelserna med tröskdata, gjordes dessa på ett urval av data (tröskkartan uppvisade inga uppenbara felaktigheter, fältet var minst 5 ha stort och standardavvikelsen i tröskans skörd var minst 0,5 t ha⁻¹ (totalt 68 fält uppfyllde dessa kriterier). För denna jämförelse beräknades endast r^2 -värden då tröskdata inte var nivåjusterade för alla gårdar. Alla mått beräknades med R-paketet *valmetrics* (Piikki et al., 2021a; b).

2.3.2 Sammanläggning av flera års skördekartor

För att kombinera skördekartorna från 2017 till 2023 beräknades först den relativa skörden genom att dividera med fältets median-skörd enligt ekvation 3, där RY_{ij} och Y_{ij} och är den relativa respektive absoluta skörden i pixel i och år j .

$$RY_{ij} = \frac{Y_{ij}}{\bar{Y}_j} \quad (\text{Ekvation 3})$$

Vid normalisering ingår ofta både centrering (t.ex. till ett medelvärde på 0 eller 1) och skalning (t.ex. till en standardavvikelse på 1). Vi valde att centrera men inte skala, detta för att behålla information om storleken på inomfältvariationen, vilken är större vissa år än andra. Ofta är variationen inom fältet mer uttalad när skördenivån är lägre (Souza et al., 2025).

Efter normalisering avgränsades det område där fleråriga analyser var möjliga. Kriterierna för detta var att data skulle finnas tillgängliga från minst tre år (oavsett om grödan var höstveten eller vårkorn) och att fältet (dvs. området i fältet med data för minst tre år) skulle vara minst fem hektar stort. Därefter togs två olika kartor fram:

- 1) **Relativa skördekartor:** beräknade som genomsnittet av årliga relativa skördar.
- 2) **Frekvenskartor för låg skörd:** beräknade som andelen år där den relativa skörden klassas som låg (tröskelvärde: 0,9).

Den första kartan visar genomsnittliga skördemönster inom fältet (högre eller lägre än median), medan den andra ger information om stabila och instabila skördemönster, eller mer specifikt, huruvida låga skördar är permanenta eller tillfälliga. Detta är ett sätt att ta fram frekvenskartor för låg skörd. Flera algoritmer för att skapa sådana kartor från tröskedata har föreslagits genom åren (Blackmore, 2000; Maestrini & Basso, 2021; Souza et al., 2025), ofta med målet att avgränsa brukningszoner med mer eller mindre homogena förhållanden. Samma metoder kan tillämpas på satellitbaserade skördekartor i rasterformat.

2.3.3 Tolkning av skördevariation

Med målet att utveckla en analysstrategi för att identifiera vilka lokala faktorer som begränsar skörden i områden med konsekvent låga skördenivåer, användes två modelleringsmetoder för att koppla samman gårdens egna markkarteringsdata och allmänt tillgängliga topografiska data (avsnitt 2.2.4 och 2.2.5) med kombinerade relativa skördekartor (avsnitt 2.3.2). De två metoderna var *Partial Least Squares Regression* (PLSR) och RF (avsnitt 2.3.1).

PLSR är en linjär multivariat regressionsmetod där flera beskrivande variabler, i detta fall från gårdens markkartering och en digital höjdmödel, representeras av ett mindre antal ortogonala latent variabler (komponenter). Dessa extraheras och roteras för att optimera modellens förmåga att förutsäga den beroende variabeln (Wold et al., 1984), i detta fall genomsnittlig relativ skörd. På detta sätt undviks problem med beroende mellan variabler, och de skattade regressionskoefficienterna

för varje beskrivande variabel i PLSR-modellen återspeglar deras relativa betydelse för sambandet med skörden.

För att eliminera artificiella enhetseffekter och ge varje variabel lika stor möjlig påverkan i de multivariata modellerna, viktades, normaliserades, de genom att divideras med sin respektive standardavvikelse. För att undvika överanpassning och hitta det optimala antalet komponenter testades modellerna med 20-faldig korsvalidering. Betydelsen av enskilda beskrivande variabler utvärderades genom deras normaliserade regressionskoefficient i den slutliga modellen.

Random Forest är en ensemblemetod som bygger flera beslutsträd under träningen (uppbyggnaden av modellen). Resultatet utgörs av den vanligaste prediktionen (vid klassificering) eller medelprediktionen (vid regression, som i denna studie) av dessa träd. Detta ökar noggrannheten och minskar risken för överanpassning genom slumpmässighet vid urval av data och variabler. Variablernas betydelse i RF-modellen bedömdes med hjälp av måttet *%incMSE* (procentuell ökning i medelkvadratfel). Detta beräknas genom permutation; hur mycket sämre blir modellens prediktioner om värdena för en viss förklaringsvariabel blandas om och därmed förlorar sitt samband med responsvariabeln? Om en variabel är viktig bör en sådan omkastning kraftigt försämra modellens tillförlitlighet (en stor ökning av felet).

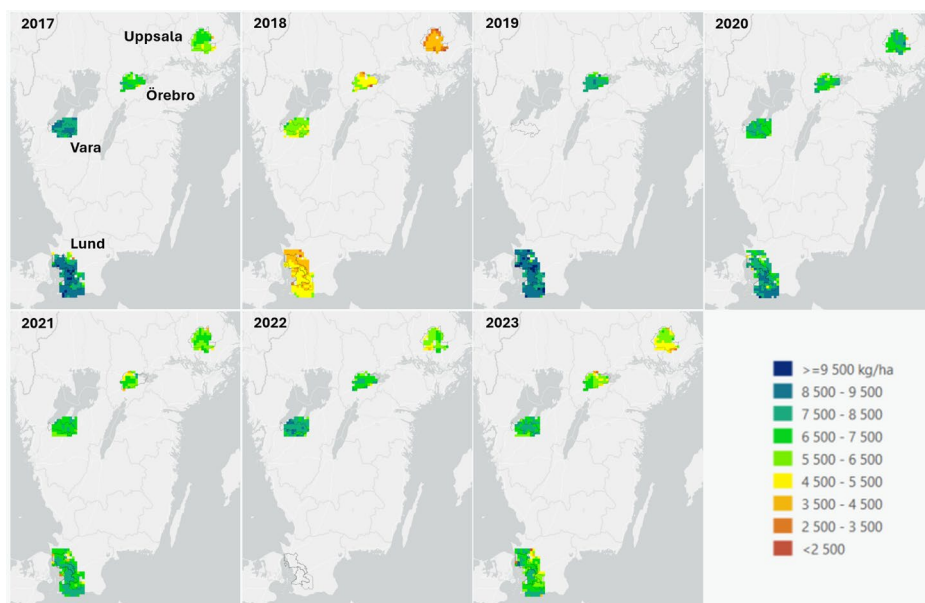
3. Resultat och diskussion

3.1 Utvärdering av satellitbaserade skördekartor

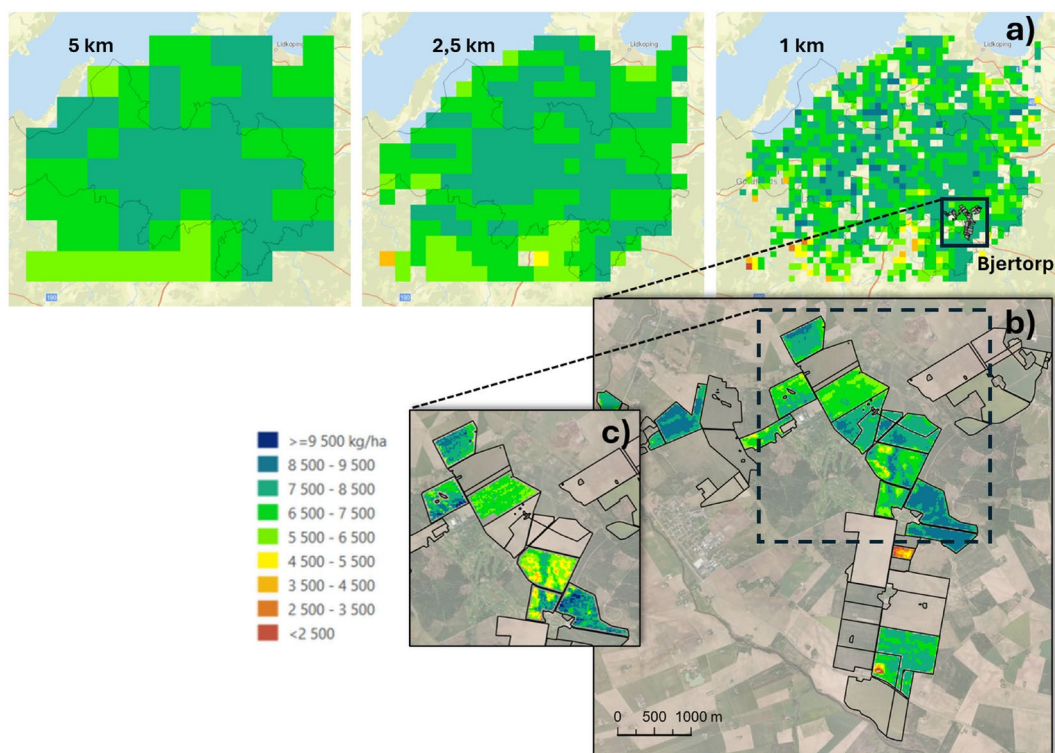
3.1.1 Skördekartor i flera skalor

Figur 5 visar framtagna skördekartor för höstveten över de fyra områdena och de sju åren, med en aggregerad upplösning om 5 km. Man kan se att skördenivåerna skiljer sig mellan år med låga skördar (2018 och i viss mån även 2023). Man kan även se att skördenivåerna generellt är olika mellan områdena, vilket är väntat, samt att skillnaden mellan områdena är större vissa år än andra. År 2021 ser exempelvis ut att vara ett år med små geografiska skillnader medan 2017 visar en tydlig trend med högst skördar i sydväst och lägre i nordost. Man kan också notera att vissa områden saknar prediktioner vissa år (Vara och Uppsala 2019 samt Lund 2022). Det beror på att inga tillräckligt molnfria satellitbilder fanns tillgängliga över dessa områden under den tidsperiod som användes i modellerna.

Figur 6 visar beräknade höstveteskördar i Vara-området med ökande detaljeringsgrad. De mest högupplösta kartorna (20 m) visar gården Bjertorp, och man kan jämföra skörden i kartor framtagna från satellit (Figur 6b) med kartor framtagna från tröskdata (Figur 6c). Det tycks som att de generella variationsmönstren stämmer relativt väl överens på många fält men variationen är betydligt mer uttalad i tröskkartorna — modelleringen tenderar att släta ut variationen (ytterligare diskussion nedan).



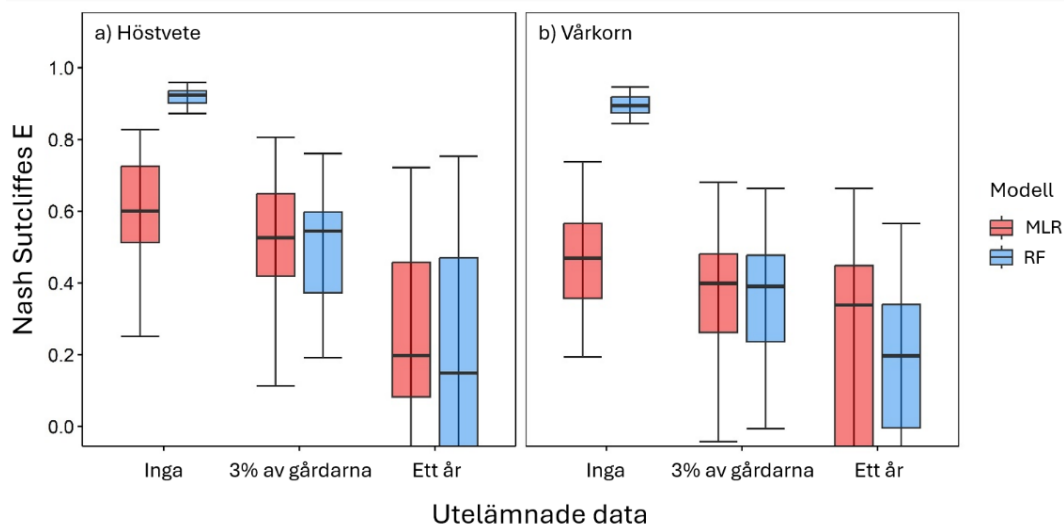
Figur 5 Höstveteskörd 2017–2023 i de fyra projektområdena enligt de regionala satellitmodellerna. Det kartlagda områdets utbredning är anpassade efter använda satellitbilder, och skiljer sig därför något från SKO-områdena. En detaljerad bild för området Vara 2023 visas i Figur 6.



Figur 6. Exempel på tillämpning av en regional satellitbaserad skördekarteringsmodell för höstvet i olika skalor 2023 i Varaområdet (SKO 1216 med omnejd): a) Pixelstorleken varierar mellan a) 1 och 5 km; b) Prediktion på fältnivå på gården Bjertorp, 20 m pixlar; c) För jämförelse – skördekartor från gårdens tröska som kalibrerats med invägd skörd. Teckenförklaringen visar skörd per ha och gäller för alla kartor a-c.

3.1.2 Jämförelse mellan enkla och komplexa modeller

Det är tydligt att den testade maskininlärningsmetoden (RF) kunde förklara variationen i kalibreringsdatamängden mycket väl (Figur 7; inga data utelämnade). När robusta modeller behövs, till exempel för att göra prognoser för ett nytt år, verkade dock den enklare regressionsmodellen (MLR) baserad på de två valda VI vara ett bättre val. För prediktioner efter skörd (Figur 7; 3% av gårdarna utelämnade) presterade MLR och RF ungefär lika bra; det fanns ingen tydlig förbättring i modellens tillförlitlighet som motiverade användningen av den mer komplexa RF-modellen. I ett modelleringsupplägg som detta, där även kalibreringsdata (skördar rapporterade av lantbrukare) sannolikt innehåller en viss osäkerhet, kan bristande robusthet på grund av överanpassning vara ett större problem. Därför valde vi att här gå vidare med de enklare MLR-modellerna.



Figur 7. Nash–Sutcliffes modelleringseffektivitet (E) för prediktioner av medelskörd per gård. Lådagrammen visar medianen (mittlinjen) och interkvartilavståndet (IQR; lådan), med spröt som sträcker sig upp till $1,5 \times \text{IQR}$ från lådan. Modellerna testades genom att i) alla data för både kalibrering och validering (inga data utelämnades), ii) i 30-faldiga korsvalideringar ($\approx 3\%$ av gårdarna utelämnades) samt iii) i korsvalidering där ett år i taget utelämnades. MLR = multipel linjär regression och RF = Random Forest-regression.

3.1.3 Utvärdering av satellitkartor för gårdar och områden

Utvärderingsmått för genomsnittlig skörd för regional nivå och gårdar visas i Tabell 1. Prediktionerna var mer träffsäkra efter skörd (när kalibreringsdata från samma år fanns tillgängliga) jämfört med prognoser under pågående säsong. Efter skörd var det genomsnittliga MAE-värdet för enskilda gårdar i genomsnitt $0,71 \text{ t ha}^{-1}$ för höstvet, medan det var något högre för vårkorn, $0,75 \text{ t ha}^{-1}$. Motsvarande storlek på felen för skördeprognoser runt slutet av juni (från sen blomning till tidig mjölkmodnad) var cirka $0,91 \text{ t ha}^{-1}$ för höstvet och $0,79 \text{ t ha}^{-1}$ för vårkorn.

Figur 8 visar att det fanns en relativt stor variation mellan enskilda gårdar (grå symboler) i hur väl prediktionerna stämmer överens med de rapporterade skördarna, även om det generella mönstret är att prediktionerna i stort följer 1:1 linjen väl. Vid tolkning av utvärderingsmått på gårdsnivå bör man komma ihåg att även de rapporterade skördarna som prediktionerna jämförs med har en viss osäkerhet, och det kan inte uteslutas att prediktionerna i vissa fall kan vara mer korrekta än de genomsnittliga skördar som rapporterats av lantbrukare i telefonintervjuer eller webbenkäter.

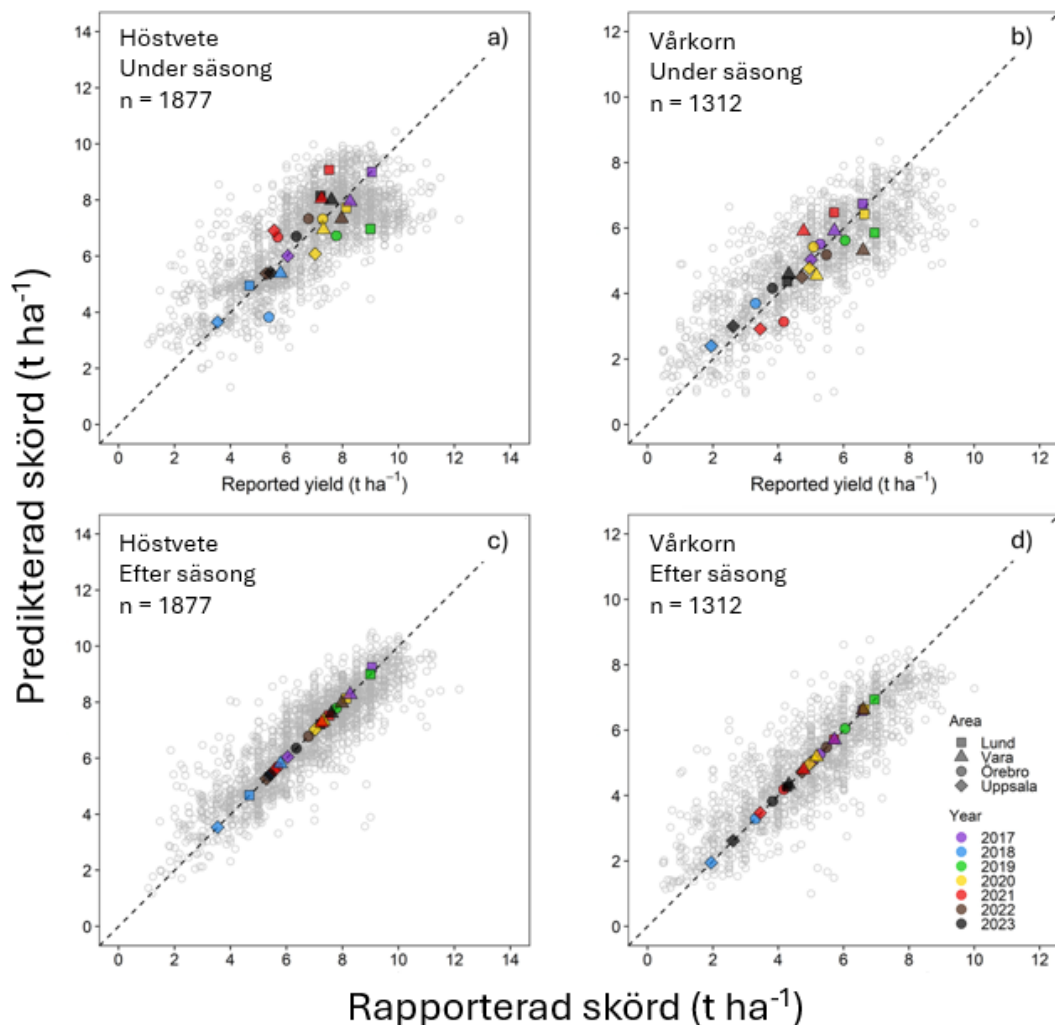
Prognoser av regionala medelskördar kunde göras under säsongen med betydligt bättre tillförlitlighet än på gårdsnivå, vilket är väntat på grund av aggregeringseffekten, där lokala variationer och mätfel delvis jämnas ut när data slås samman. I en regional prognos var MAE $0,64 \text{ t ha}^{-1}$ för höstvet och $0,46 \text{ t ha}^{-1}$

för vårkorn (Tabell 1; Figur 8a och b). Resultaten visar således att det alltså är möjligt att redan i slutet av juni göra en relativt noggrann regional skördeprognos innevarande år baserad på fjärranalys i kombination andra års rapporterade gårdsmedelskörda från telefonintervjuer.

Felnivåerna (MAE-värdena) för prediktioner av regionala skördar efter skörd anges också i Tabell 1. Dessa stämde nästan helt överens med medelvärdena för de rapporterade genomsnittliga regionala skördarna (Figur 8c–d), vilket var väntat. Om man jämför utvärderingsmått från 3-faldig och 30-faldig korsvalidering (Tabell 1) verkar det vara möjligt att minska antalet gårdsintervjuer med minst en tredjedel utan nämnvärd försämring av prognoser för regional medelskörd; fjärranalys i kombination med ett minskat antal gårdsintervjuer kan ge likvärdiga resultat som om intervjuer genomförs med betydligt fler lantbrukare.

Tabell 1. Modellutvärdering för olika prediktionstidpunkter och aggregationsnivåer. Måtten är beräknade per område och år och median (interkvartilavstånd) redovisas. Modelltypen var multipel linjär regression, CV= korsvalidering, E = Nash–Sutcliffes modellerings effektivitet, r^2 = determinationskoefficient, MAE = medelabsolut fel, ME = medelfel (skevhet).

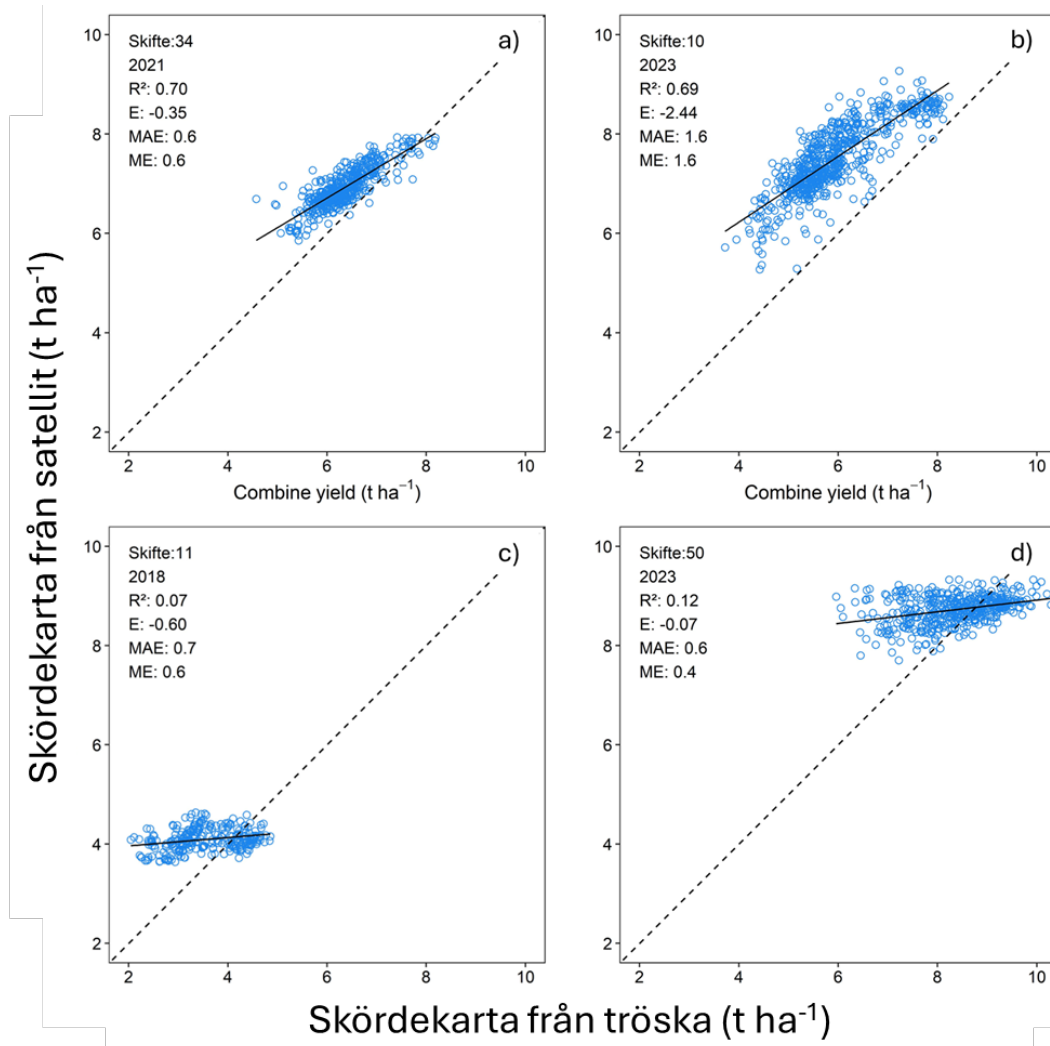
Nivå	Tidpunkt	CV	E	r^2	MAE (t ha ⁻¹)	ME (t ha ⁻¹)
<i>Höstvete</i>						
Region	Under säsong	CV (år)	0.62	0.67	0.64	-0.01
Region	Efter skörd	CV (k=3)	~1	~1	0.03	0.02
Region	Efter skörd	CV (k=30)	~1	~1	0.01	0.01
Gård	Under säsong	CV (år)	0.20 (0.38)	0.56 (0.23)	0.91 (0.39)	0.01 (0.83)
Gård	Efter skörd	CV (k=3)	0.53 (0.28)	0.54 (0.26)	0.71 (0.11)	0.01 (0.03)
Gård	Efter skörd	CV (k=30)	0.53 (0.23)	0.53 (0.23)	0.70 (0.10)	0.00 (0.01)
<i>Vårkorn</i>						
Region	Under säsong	CV (år)	0.79	0.79	0.46	-0.05
Region	Efter skörd	CV (k=3)	~1	~1	0.03	0.02
Region	Efter skörd	CV (k=30)	~1	~1	0.01	0.01
Gård	Under säsong	CV (år)	0.34 (0.56)	0.47 (0.18)	0.79 (0.27)	0.09 (0.70)
Gård	Efter skörd	CV (k=3)	0.33 (0.27)	0.35 (0.26)	0.75 (0.17)	0.00 (0.02)
Gård	Efter skörd	CV (k=30)	0.40 (0.22)	0.40 (0.21)	0.74 (0.14)	0.01 (0.01)
<i>Båda grödorna</i>						
Pixel	Efter skörd	Jfr. tröska	---	0.31 (0.43)	---	---



Figur 8. Predikterade jämfört med rapporterade skördar på gårdsnivå (grå symboler), samt medelskördar per år och område (symboler med olika form och färg). Legenden i d gäller för alla diagram. Prediktionerna gjordes med multipel linjär regression. n = antal gårdsmedelvärden.

3.1.4 Jämförelse mellan tröskkartor och satellitkartor

När man jämför skördekartor från tröskor med satellitbaserade skördekartor bör man ha i åtanke att båda metoderna har sina respektive styrkor och svagheter. I vissa fall är överensstämmelsen mellan de två metoderna relativt god (exempel i Figur 9a–b), medan den i andra fall är svag (exempel i Figur 9c–d). För tröskdata är det vanligt med problem kopplade till insamling, filtrering och kalibrering av data. När det gäller satellitbaserade kartor är en svaghet att faktorer som påverkar skörden efter att satellitbilden togs inte avspeglas i den slutliga skördekartan. En annan egenskap hos satellitbaserade skördekartor är att variationen inom fälten ofta blir mindre uttalad, eftersom modelleringen vanligtvis har en utjämnande effekt (se även Figur 6b och c).



Figur 9. Pixelvärden från satellitbaserade skördekartor (y-axeln) plottade mot pixelvärden från tröskans skördekartor (x-axeln) för fyra valda fält: (a–b) exempel med god korrelation och (c–d) exempel med svag korrelation. E = Nash–Sutcliffes-modellerings effektivitet, r^2 = determinationskoefficient, MAE = medelabsolutfel, ME = medelfel. 2018 (c) var ett torrt år med generellt låga skördar, och även 2023 (b, d) var speciellt från vädersynpunkt, med mycket torrt väder fram till juli, när det blev mycket regnigt fram till skörd.

I den här studien är kalibreringen av satellitdata gjord på regional nivå (t.ex. utan hänsyn till sortskillnader) men tillämpas på enskilda fält inom regionen. Det kan sannolikt göra att medelskördenivån avviker när man använder modeller som bygger på VI. Det bör därför inte förväntas att tröskkartor och satellitbaserade kartor alltid överensstämmer, och jämförelser bör tolkas med viss försiktighet. Tidigare forskning har visat att en minskning av den rumsliga upplösningen till 40 m (i stället för 20 m som användes här) var överensstämmelsen mellan satellitbilder och tröskdata bättre (Alshihabi et al., 2024).

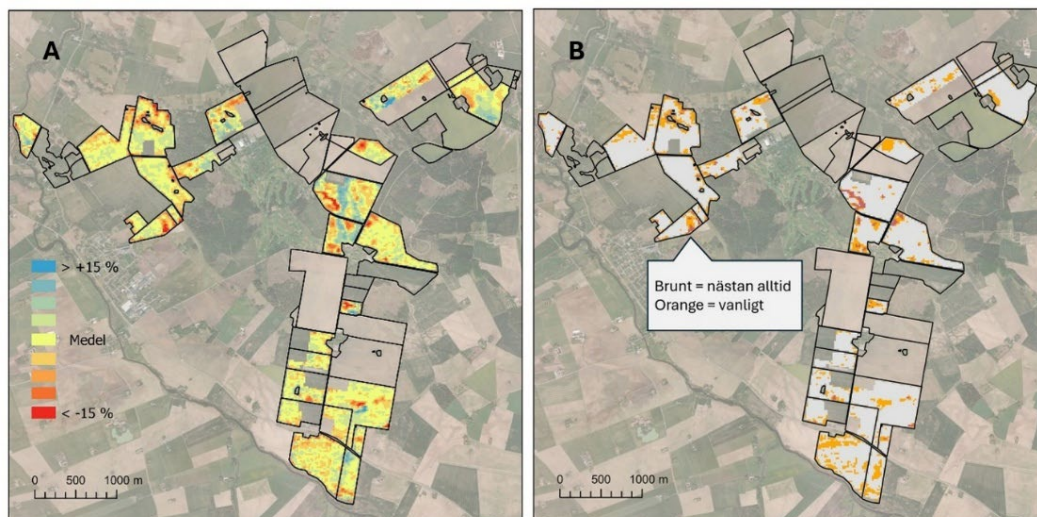
3.2 Flerårsanalys

3.2.1 Kartor över relativskörd och frekvensen av låg skörd.

Figur 9 visar exempel på en karta över relativskörd (a) och en karta över frekvensen av låg skörd (b) som tagits fram baserat på minst tre års satellitdata (höstvet eller vårkorn) för gården Bjertorp. Visualiseringen pekar tydligt ut problemområden med generellt låg skörd (de med rödast färg i karta a) och indikerar hur ofta varje pixel har en skörd betydligt under fältets medianskörd (Figur 9b). I Figur 9 blir också en svårighet med flerårsanalys av skördekartor tydlig. Den odlade arealen är inte densamma från år till år. De områden i kartan som saknar data (inom fälten med svart kant), är områden där vete eller korn odlats mindre än tre år. Olika geografiska avgränsningar mellan år blir en utmaning att hantera om man vill bygga ett automatiserat system för flerårsanalys av skördekartor. På platser som Lantmännens framtidsgård Bjertorp är förmodligen denna problematik vanligare då gården står värd för många fältförsök.

3.2.2 Tolkning av orsaker till låg skörd

Det faktum att många variabler (markegenskaper, skörd, topografi mm) samvarierar när man arbetar med pixlar inom ett fält gör det svårt att peka ut enskilda variabler som orsaken till låg skörd. Det är viktigt att komma ihåg att de statistiska samband mellan markegenskaper och topografi å ena sidan och skörd å den andra, som i varierande grad representerades av de normaliserade regressionskoefficienterna i PLSR-modellerna eller som identifieras som viktiga prediktorvariabler i RF, inte nödvändigtvis beskriver ett funktionellt samband. Det kan mycket väl vara en annan variabel som kanske inte ens är analyserad som utgör den egentliga orsaken. Till exempel ingår inga variabler som direkt beskriver markstruktur och vattendynamik. I flera fält på de fyra gårdarna kunde den relativa medelskörden beskrivas relativt väl med markkarteringsdata och topografi, med upp mot 70–80 % förklaringsgrad. Det var emellertid enbart i enstaka fall man via variabler som identifierades som viktiga i modellerna kunde separera pixlar med genomgående låg skörd. Enstaka lyckade exempel utgjordes främst av fält där stora skillnader i textur sammanföll med skördevariationer, då ofta i kombination med lokala höjdområden (även till synes små topografiska skillnader kan ha påtaglig betydelse i detta sammanhang).



Figur 10. Analys av 3–6 års satellitbaserade prediktioner av skörd (höstvet och vårkorn) på Bjertorp: a) avvikelse från respektive fälts medelskörd; b) områden med återkommande låg skörd (där låg definieras som <90 % av fältets medianskörd).

Svårigheterna att beskriva områden med lägst skörd med markkartering och topografi kan förutom vad som nämnts ovan ha flera orsaker. En kan vara att det inte handlar om enkla samband utan om komplex samverkan mellan flera faktorer. Något som RF möjligen kan vara bättre på att beskriva jämfört med PLSR, men inte enklare att tolka. En annan orsak kan helt enkelt vara att orsakerna varierar inom ett och samma fält. Något som illustrerats i tidigare studier (Engström et al., 2025). Av dessa skäl bör visuell tolkning av kartor övervägas, där lågavkastande områden enkelt kan identifieras (Figur 10), när man ska identifiera orsaker till lokalt låga skördar. Ett bidrag från de multivariata modellerna skulle kunna vara att generera en lista med troliga faktorer som man sedan bedömer utifrån lokalkännedom och agronomisk kunskap.

3.2.3 Erfarenheter från diskussioner med lantbrukargruppen

De två typerna av kartor (generella skördemönster och deras dynamik; Figur 10a och b) diskuterades i möten mellan projektgruppen och lantbrukarna. Många problemområden i kartorna kändes igen och hade kända orsaker. I andra fall var man inte säker på orsakerna till variationen. I den mån kartorna stämmer visar det att de sammanvägda skördekartorna kan ge nya insikter även om man har god kännedom om sina fält. Det blev tydligt under mötena att enbart avancerad dataanalys med normalt tillgängliga mätdata inte räcker för tolkning av skördevariationer. Lokalkunskap krävs också, särskilt när variationer beror på odlingshistorik och historiska ägarförhållanden. Exempel på kända orsaker till problem var: jordart (avvikande textur eller mullhalt), topografi (ofta lokala höjder), dålig dränering, lågt pH, nedgrävda vattenledningar, gamla vägar, viltskador (gäss

etc.) och köldhålor. Det är noterbart att flera av dessa orsaker alltså inte framkommer från kända datakällor eller etablerade analysprotokoll.

I den mån man alls använder skördekartorna från tröska, nyttjas de som ett kvitto på årets växtodling. Man ser dock stora möjligheter med skördekartering (från satellit eller tröska) och efterlyser följande för att potentialen ska kunna realiseras i praktiken:

- 1) Bättre funktionalitet för och automatik i bortfiltrering av felaktiga skörderegistreringar (från tröska).
- 2) Lättanvända funktioner för utvärdering av årets odling genom att man på ett enkelt sätt kan jämföra skördekartor med andra kartor eller annan information, t.ex. så att man kan följa hur vädret varit under säsongen samt göra jämförelser med tidigare säsonger.
- 3) Möjlighet att använda kartor över generella skördenivåer som underlag när man genererar tilldelningsfiler för anpassning av insatsmedel efter varierande skördsnivåer i fält.

Diskussionerna visade att de största hindren för att fullt ut använda skördekartor för tolkning av odlingsresultat och styrning av insatser är tekniska svårigheter och tidsbrist. Därför behöver framtida system vara kompatibla med olika fabrikat av skördemonitorer, gödningsspidare och såmaskiner, och beslutsstödet bör utformas så enkelt och användarvänligt som möjligt.

3.3 Modellering med ofullständiga datamängder

3.3.1 Möjliga lösningar

Som nämnts tidigare är det en viss utmaning att förlita sig på optiska satellitdata eftersom moln och dis kan göra bilderna oanvändbara för tolkning av grödstatus. I den aktuella studien, som omfattade fyra regioner och sju år, var det inte möjligt att få användbara data under perioden från sen blomning till tidig mjölkmodnad för tre områdes-år (Vara och Uppsala 2019 samt Lund 2022). Även under de år då data fanns tillgängliga täckte de molnfria partierna inte alltid hela området. I princip finns tre sätt att hantera detta problem:

- 1) **Borttagning** (*pruning*): Ta bort alla observationer (här: pixlar) som innehåller saknade värden och de variabler (här: VI) som i hög grad är ofullständiga (t.ex. >30% saknas).
- 2) **Ifyllnad** (*imputation*): Fyll i saknade värden med uppskattningar, t.ex. medelvärden eller prediktioner baserade på andra variabler (även från andra satelliter). Detta gör datamängden komplett men kan införa

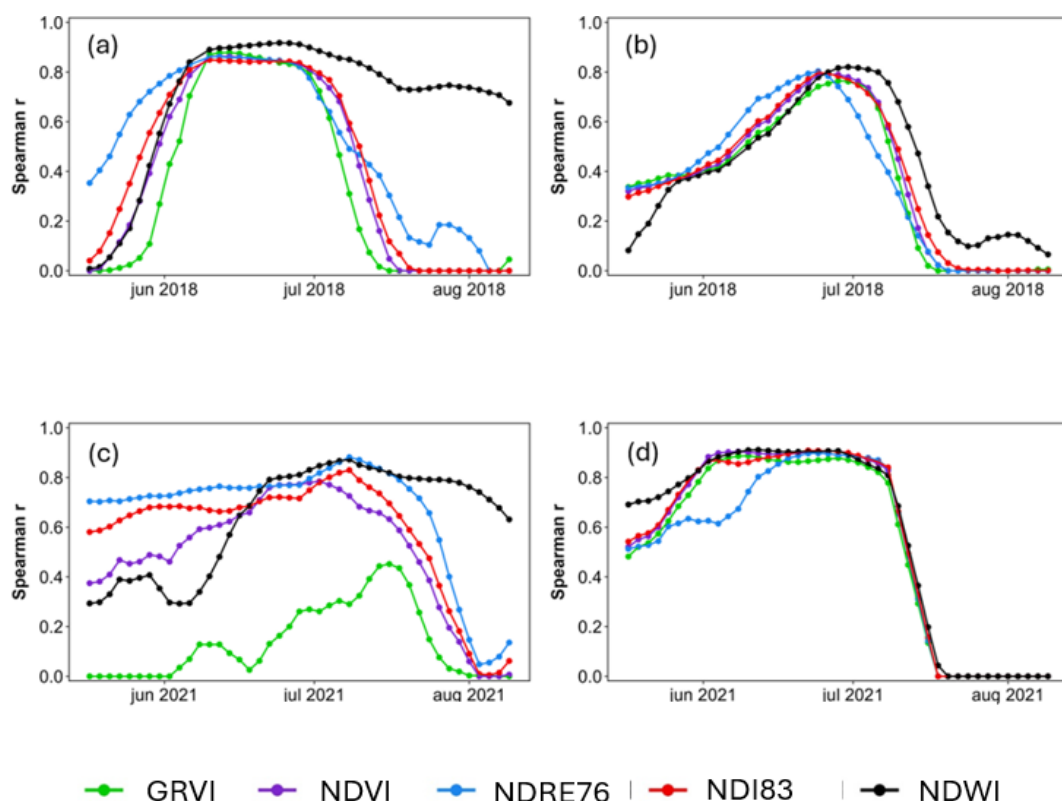
systematiska fel om databortfallet inte är slumpmässigt och metoden för ifyllnad inte kan hantera det.

- 3) **Gruppvis modellering** (*reduced modelling*): Dela in data i grupper där samma variabler finns tillgängliga och träna separata modeller för varje grupp. För nya data används den modell som motsvarar de tillgängliga variablerna (Friedman et al., 1996; Saar-Tsechansky & Provost, 2007).

Det första och enklaste alternativet användes i denna studie: områden med moln eller dis togs bort från kartläggningen och om ett fält delvis saknade data uteslöt hela fältet. En nackdel med metoden är att inte alla fält (eller ens alla områden) får en skördekartan varje år. Vi genomförde även ett pilotförsök med den andra strategin (se nedan), där en komplett molnfri tidsserie av syntetiska Sentinel-2-data genererades. Det finns mycket att vinna på att undvika automatisk molnmaskning vid implementering av satellitbaserad skördekartering i digitala system, eftersom metoderna för detta inte alltid fungerar tillfredsställande. Dessutom kan en förbehandlad tidsserie av VI ge annan information om grödors tillväxt och utveckling, till exempel *leaf area duration* (tidsintegrerade bladtyteindex), kärnfyllnadsperiodens längd eller maximal biomassa, vilka kan vara användbara prediktorvariabler i empiriska skördemodeller. Alternativt kan hela tidsserien användas för prediktion med semi-empiriska modelleringsmetoder (se exempel av Bouras et al., 2023; Federolf et al., 2025; Persson et al., 2024). Sådana strategier testades kan vara värda att utforska vidare.

3.3.2 Test av syntetiska satellitdata

Figur 11 visar hur korrelationen mellan kärnskörd och utvalda VI utvecklas över tid, baserat på syntetiska satellitbilder i kombination med skördedata från tröska för höstvet och vårkorn och två kontrasterande år: 2018, som var mycket torrt, och 2021, som var mindre extremt. Resultaten indikerar att satellitbaserade skördemodeller kan förväntas vara mer tillförlitliga från slutet av juni och framåt för de flesta VI. Korrelationerna är däremot avsevärt mer osäkra under maj och början av juni, när grödan ännu inte har avslutat sina vegetativa utvecklingsstadier, jämfört med under och efter blomning. GRVI, som endast baseras på synliga band (grönt och rött) och därför inte fångar information från nära-infraröda (NIR) och *red edge*-regionerna där grödans reflektans är särskilt känslig för beståndsstruktur, kvävestatus och vatteninnehåll, avvek från de övriga indexen för höstvet 2021. NDWI, som baseras på band 11 där reflektansen påverkas av grödans vattenstatus fungerade väl även senare på säsongen för höstvet under båda åren. Detta är en mindre pilottest som snarare visar vilka analyser som möjliggörs med kompletta tidsserier, än ger underlag för säkra slutsatser.



Figur 11. Variation i Spearmans r -värden (för korrelationen mellan vegetationsindexen och skördekartor från tröska) över perioden juni–augusti, för två grödor under två år: ett torrt år, 2018 (a–b), och ett mer normalt år, 2021 (c–d), för höstvet (a och c) respektive vårkorn (b och d) på Bjertorp. Vegetationsindexen var av NDI-typ och baserades på följande Sentinel-2-band: GRVI: band 3 och 4, NDVI: band 8 och 4, NDRE76: band 7 och 6, NDI83: band 8 och 3, samt NDWI: band 8 och 11.

3.4 Identifierade forskningsbehov

3.4.1 Fältkanter är svåra att kartlägga

En svaghet med både satellitbaserad och tröskbaserad skördekartering är att det är svårt att samla in användbar information om skördenivåer nära fältets kanter. Satellitdata påverkas av omgivande markanvändning och skörderegistrering med utrusning på tröska genererar många felaktiga värden i kanterna som måste filtreras bort (tex pga. vändningar, start, stopp och icke-fulla skärbord; se Figur 2). I ett system kan det vara enkelt att helt enkelt ta bort data längs kanten och fylla i (extrapolera) med hjälp av mer pålitliga värden längre in i fält. Möjligen bör man arbeta mer med hur man hanterar fältkanter då exempelvis Persson et al. (2025) visade att skördenivåer nära fältkanter ofta är lägre än längre in i fält. Fältkanter utgör en icke försumbar del av åkermarken (ca 25 % av arealen ligger in om 20 m från en kant *ibid.*) och om skörden är lägre där bör man precis som i resten av fältet,

åtgärda produktionsproblem där så är möjligt, alternativt anpassa sina gödningsgivor efter permanent lägre skördepotential.

3.4.2 Satellitbaserade kartor jämnar ut skördevariationen

Som nämnts ovan tenderar skördekartorna som baseras på satellitbilder att ge mindre uttalad inomfältvariation jämfört med skördekartor från tröska (se t.ex. Figur 6b och c). Det kan ha flera orsaker, bland annat att modellen i sig gör att resultatet blir mer utjämnat; höga skördar underskattas och låga skördar överskattas systematiskt. I ett diagram med predikterade skördar plottade mot observerade skördar ger detta en regressionslinje med lägre lutning jämfört med 1:1-linjen (Figur 8). En större spridning av punkter runt linjen visar i stället på större varians. Hur ska man då komma till rätta med detta? Enkla modeller (som linjär regression) är ofta stabila och pålitliga, men har vanligtvis lite större systematiskt fel och mindre varians i resultaten. Mer flexibla modeller (som RF) kan minska det systematiska felet, men riskerar å andra sidan att ge större varians i resultaten om de blir för komplexa eller om det finns för lite referensdata för kalibrering. Mer arbete kan läggas på att undersöka hur man kan balansera den utjämnande effekten mot modellens varians, men kanske ännu viktigare är att hitta sätt att hantera utjämnningen av variationen inom fältet vid användning av skördekartor som beslutsstöd för precisionsodling.

4. Konklusioner

I projektet utvecklades och utvärderades ett modelleringsupplägg för skördekartering på regional nivå, fältnivå och inom fält. Modelleringen baserades på satellitbilder och medelskördar per gård insamlade av SCB. Följande konklusioner sammanfattar projektet:

- Genomsnittligt MAE för predikterade gårdsmedelskördar var 0,7 och 0,8 t ha⁻¹ för höstveten respektive vårkorn när prediktionerna gjordes efter skörd (dvs. när rapporterade skördar från samma år kunde användas i modellutvecklingen). När prognoser gjordes under växtsäsongen (i slutet av juni) var MAE bara något högre, 0,9 och 0,8 t ha⁻¹.
- Prognoser på regional nivå var mer träffsäkra än de på gårdsnivå och hade ett MAE på 0,6 respektive 0,5 t ha⁻¹ för höstveten respektive vårkorn. Skattning av områdesskördar efter skörd, kunde göras med bibehållen noggrannhet även när antalet rapporterade genomsnittliga gårdsskördar som användes för modellkalibrering minskades med en tredjedel.
- Jämförelser mellan skördekartor från satellit och skördekartor från tröskor (båda med 20 m upplösning) visade att de i vissa fall var mycket lika, medan överensstämmelsen i andra fall var svag. Det kan ha flera orsaker, t.ex. ett litet variationsspann av skörd inom fältet, felaktiga tröskdata trots filtrering, händelser som påverkat inomfältsvariationen i skörd som inträffat mellan tidpunkten för satellitbilden och tidpunkten för skörd, samt den utjämnande effekten på skördevariationen som modelleringen naturligt ger upphov till. Sammanfattningsvis är båda typerna av data osäkra, och det är svårt att veta vilken som är mest korrekt men satellitbaserade skördekartor kan utan utrustning och besvärlig datahantering enkelt tillhandahållas åt alla, överallt.
- Vi har utvecklat en praktiskt användbar metod för att avgränsa stabila och instabila problemområden med hjälp av satellitbaserade skördekartor. Två typer av kartor togs fram: relativa skördekartor och frekvenskartor för låg skörd. Även om variationsmönstren i kartorna ofta kändes igen av lantbrukarna och hade kända orsaker, är kartläggningen av variationen viktig; för att denna typ av information ska kunna användas i beslutsstödsystem för precisionsodling behövs den i detaljerad och digital form.

- Kartorna visade också på potentiella tidigare okända (mindre uttalade) problemområden och kan därmed utgöra en startpunkt för det detektivarbete som krävs för att hitta, förstå och åtgärda produktionsproblem och därigenom höja skörden. Om detta inte är möjligt kan kartorna i stället användas som underlag för att anpassa insatser till den faktiska skördepotentialen, dvs. man låter målskörden variera inom fälten.
- Ett praktiskt problem med satellitbaserad skördekartering är att förekomst av moln och molnskuggor ger ofullständig datatillgång, och i vissa fall är dataförlusten så stor att kartering inte alls är möjlig. Syntetiska Sentinel-2-tidsserier kan vara en lösning att komma förbi detta. I en pilotstudie på Bjertorp, där vi testade sådana data såg vi också att sambandet mellan optiska VI och spannmålsskörd var starkast under senare delen av juni, vilket för både höstveten och vårkorn ofta betyder sen blomning–tidig mjölkmodnad.
- Multivariata dataanalyser (PLSR och RF) av relativa skördekartor tillsammans med tolkningsdata (resultat från jordprovtagning topografiska index mm) kunde visa vilka variabler som var statistiskt relaterade till varandra och den varierande skörden. I många fall berodde dock de verkliga orsakerna på faktorer som inte fångades av de tillgängliga tolkningsvariablerna (markanvändningshistorik, dräneringsproblem, viltskador etc.). Dataanalysmetoderna kan därför främst användas för att ta fram en förslagslista över troliga orsaker som sedan behöver utvärderas manuellt av lantbrukaren, vars lokalkännedom är nödvändig för tolkningen.
- För den satellitbaserad skördekartering fungerade enkla linjära modeller med två valda VI (NDWI och NDRE76) lika bra eller bättre än en maskininlärningsmetod som försågs med ett stort antal VI. Den sistnämnda modelleringsstrategin ledde till överanpassning; variation i kalibreringsdatasetet förklarades mycket bra men modellerna fungerade inte bättre för nya gårdar och för nya år. Det är viktigt att välja en valideringsstrategi som säkerställer robusta modeller men kanske särskilt när det som här finns en icke försumbar osäkerhet även i kalibreringsdata.
- Sammanfattningsvis rekommenderar vi, baserat på resultat från projektet, följande utvecklingsinsatser för att identifiera problemområden och därefter antingen åtgärda problemen och höja skörden, eller i stället anpassa insatser till permanent låga skördenivåer och därigenom öka resurseffektiviteten:

- De största hindren för att utnyttja skördekartor är tekniska svårigheter och tidsbrist. Därför behöver system som utvecklas vara kompatibla med olika tröskor såmaskiner och spridare från olika tillverkare och utformas så enkla och användarvänliga som möjligt.
- Gör skördekartor lätt tillgängliga för alla genom att bygga in funktionalitet för satellitbaserad kartering i beslutsstödsystem för precisionsodling. Undersök möjligheten att använda kompletta tidsserier av t.ex. syntetiska Sentinel-2-data för att säkerställa full täckning, förbättra modelleringen och förenkla datahanteringen.
- För lantbrukare med skördemätare på tröskan: utveckla funktionalitet för enklare datafiltrering (vilket är nödvändigt för att data ska vara användbara överhuvudtaget) och nivåjustering med invägda skördar.
- Tillhandahåll funktionalitet för flerårsanalyser av skördekartor. En utmaning här är hur karteringsområden ska avgränsas, eftersom växtföljd är vanlig och fält ofta delas eller ändrar geometri av andra orsaker.
- Tillhandahåll funktionalitet för enkel manuell jämförelse mellan skördekartor (för enskilda år och fleråriga) och stöddata (väder, markkartor, topografi etc.). Lägg eventuellt till en förslagslista över förklarande variabler kopplade till låg skörd, framtagen med multivariat dataanalys, men använd den inte automatiskt. Manuell tolkning med stöd av lokal kunskap är nödvändig innan beslut tas om hur låga skördar ska hanteras (åtgärdas eller accepteras).
- Funktionalitet för att analysera relativa skördar inom närområden i större skala (t.ex. inom 5 km radie) kan också vara användbar, till exempel inom rådgivningen.

Referenser

- Bouras, E.H., Olsson, P-O., Thapa, S., Diaz, J.M., Albertsson, J. & Eklundh, L. (2023). Wheat Yield Estimation at High Spatial Resolution through the Assimilation of Sentinel-2 Data into a Crop Growth Model. *Remote Sensing*, 15, 4425. <https://doi.org/10.3390/rs15184425>
- Algerbo, P. A., Mattson, L. & Thylén, L. (2003). Skörderelaterad kvävegödsling: teknik, metodik och erfarenheter. JTI – Institutet för jordbruks- och miljöteknik, Report 311. <https://www.diva-portal.org/smash/get/diva2:959553/FULLTEXT01.pdf> [Hämtad 2026-01-01].
- Alshihabi, O., Persson, K. & Söderström, M. (2023). Post-processing yield maps of winter wheat using data from satellites and combines. In J. V. Stafford (Ed.), *Precision agriculture '23* (pp. 523–530). Wageningen Academic Publishers. https://doi.org/10.3920/978-90-8686-947-3_66
- Birrell, S.J., Sudduth, K.A. & Borgelt, S.C. (1996). Comparison of sensors and techniques for crop yield mapping. *Computers and Electronics in Agriculture*, 14, 215–233. [https://doi.org/10.1016/0168-1699\(95\)00049-6](https://doi.org/10.1016/0168-1699(95)00049-6)
- Blackmore, S. (2000). The interpretation of trends from multiple yield maps. *Computers and Electronics in Agriculture*, 26, 37–51. [https://doi.org/10.1016/S0168-1699\(99\)00075-7](https://doi.org/10.1016/S0168-1699(99)00075-7)
- Blasch, G., Li, Z. & Taylor, J.A. (2020). Multi-temporal yield pattern analysis method for deriving yield zones in crop production systems. *Precision Agriculture*, 21, 1263–1290. <https://doi.org/10.1007/s11119-020-09719-1>
- Deines, J.M., Patel, R., Liang, S.Z., Dado, W., & Lobell, D.B. (2021). A million kernels of truth: Insights into scalable satellite maize yield mapping and yield gap analysis from an extensive ground dataset in the US Corn Belt. *Remote Sensing of Environment*, 253, 112174. <https://doi.org/10.1016/j.rse.2020.112174>
- Delin, S. (2005). Site-specific nitrogen fertilization demand in relation to plant available soil nitrogen and water. *Acta Universitatis Agriculturae Sueciae*, doctoral thesis 2005:6. <https://res.slu.se/id/publ/12509>
- Engström, L., Myrbeck, Å., Larsolle, A., Coucheney, E., Blombäck, K., Lewan, E., Persson, K. & Stenberg, B. (2025). Within-field variation of actual and maximum winter wheat yield in relation to crop management. *Field Crops Research*, 329, 109965. <https://doi.org/10.1016/j.fcr.2025.109965>
- Federolf, C., Reusch, S., Portz, G., Truszkowski-Graw, A. & Jasper, J. (2025). In J. V. Stafford (Ed.), *Precision agriculture '25* (pp. 673–680). Wageningen Academic, Leiden, The Netherlands. https://doi.org/10.1163/9789004725232_088
- Friedman, J. H., Kohavi, R. & Yun, Y. (1996). Lazy decision trees. In E. Horvitz & F. Jensen (Eds.), *Proceedings of the Thirteenth National Conference on Artificial*

- Intelligence (AAAI-96) (pp. 717–724). AAAI Press. <https://aaai.org/papers/107-aaai96-107-lazy-decision-trees/> [Hämtad 2026-01-01].
- Hunt, M.L., Blackburn, G.A., Carrasco, L., Redhead, J.W. & Rowland, C.S. (2019). High resolution wheat yield mapping using Sentinel-2. *Remote Sensing of Environment*, 233, 111410. <https://doi.org/10.1016/j.rse.2019.111410>
- Johansson, A., Johansson, T., & Westin, H. (2022). Hur digitaliserat är jordbruket i Västra Götaland? <https://mellanarkiv-offentlig.vgregion.se/alfresco/s/archive/stream/public/v1/source/available/sofia/rs7897-268913469-627/native/Hur%20digitaliserat%20%C3%A4r%20lantbruket%20i%20VG.pdf> [Hämtad 2026-01-01].
- Lachia, N., Pichon, L., Marcq, P., Taylor, J. & Tisseyre, B. (2021). Why are yield sensors seldom used by farmers—a French case study? In J. V. Stafford (Ed.), *Precision agriculture '21* (pp. 745–751). Wageningen Academic Publishers. https://doi.org/10.3920/978-90-8686-916-9_89
- Li, X., Jin, H., Eklundh, L., Bouras, E., Olsson, P.O., Cai, Z., Ardö, J. & Duan, Z. (2024). Estimation of district-level spring barley yield in southern Sweden using multi-source satellite data and random forest approach. *International Journal of Applied Earth Observation and Geoinformation*, 134, 104183. <https://doi.org/10.1016/j.jag.2024.104183>
- Liaw, A. & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, 2(3), 18–22. https://cran.r-project.org/doc/Rnews/Rnews_2002-3.pdf
- Lowenberg-DeBoer, J. & Erickson, B. (2019). Setting the record straight on precision agriculture adoption. *Agronomy Journal*, 111(4), 1552–1569. <https://doi.org/10.2134/agronj2018.12.0779>
- Maestrini, B. & Basso, B. (2021). Subfield crop yields and temporal stability in thousands of US Midwest fields. *Precision Agriculture*, 22, 1749–1767. <https://doi.org/10.1007/s11119-021-09810-1>
- McEntee, P.J., Bennett, S.J. & Belford, R.K. (2020). Mapping the spatial and temporal stability of production in mixed farming systems: An index that integrates crop and pasture productivity to assist in the management of variability. *Precision Agriculture*, 21(1), 77–106. <https://doi.org/10.1007/s11119-019-09658-6>
- Müller, M., Thapa, S., Bouras, E., Olsson, P.O., Jamali, S., Eklundh, L., & Ardö, J. (2025). Using sentinel-2 data to quantify the impacts of drought on crop yields at local and regional scales in Sweden. *Agricultural and Forest Meteorology*, 373, 110789. <https://doi.org/10.1016/j.agrformet.2025.110789>
- Nissen, K. & Söderström, M. (1999). Mapping in precision farming—from the farmer's perspective. In J. V. Stafford (Ed.), *Precision agriculture '99* (pp. 655–664). Sheffield Academic Press. <https://www.cabdirect.org/cabdirect/abstract/20013118464>

- Perich, G., Turkoglu, M.O., Graf, L. V., Wegner, J. D., Aasen, H., Walter, A. & Liebisch, F. (2023). Pixel-based yield mapping and prediction from Sentinel-2 using spectral indices and neural networks. *Field Crops Research*, 292, 108824. <https://doi.org/10.1016/j.fcr.2023.108824>
- Persson, K., Söderström, M., Börjesson, T., Pettersson, C-G., Lindgren, A., Adler, K., Alshihabi, O., Marzec-Schmidt, K., Forsström, S., Martinsson, J., Åsman, M. & Mickelåker, J. (2024). Baby Grain Passport–säker och hållbart producerad spannmål till barnmat. Swedish University of Agricultural Sciences (SLU), Department of Soil and Environment, report. <https://doi.org/10.54612/a.2alsqqj8c8> [Hämtad 2026-01-01].
- Persson, K., Ekholm, E., & Söderström, M. (2025). Is there reason to adapt fertilizer rates along field boundaries? In J. V. Stafford (Ed.), *Precision agriculture '25* (pp. 814–819). Wageningen Academic, Leiden, the Netherlands. https://doi.org/10.1163/9789004725232_107
- Piikki, K. & Söderström, M. (2019). Digital soil mapping of arable land in Sweden—Validation of performance at multiple scales. *Geoderma*, 352, 342–350. <https://doi.org/10.1016/j.geoderma.2017.10.049>
- Piikki, K., Wetterlind, J., Söderström, M. & Stenberg, B. (2021a). valmetrics: Metrics and Plots for Model Evaluation. R package version 1.0.0. <https://CRAN.R-project.org/package=valmetrics>
- Piikki, K., Wetterlind, J., Söderström, M. & Stenberg, B. (2021b). Perspectives on validation in digital soil mapping of continuous attributes—A review. *Soil Use and Management*, 37(1), 7–21. <https://doi.org/10.1111/sum.12694>
- Prey, L. & Schmidhalter, U. (2019). Temporal and spectral optimization of vegetation indices for estimating grain nitrogen uptake and late-seasonal nitrogen traits in wheat. *Sensors*, 19(21), 4640. <https://doi.org/10.3390/s19214640>
- Raun, W., Figueiredo, B., Dhillon, J., Fornah, A., Bushong, J., Zhang, H. & Taylor, R. (2017). Can yield goals be predicted? *Agronomy Journal*, 109(5), 2389–2395. <https://doi.org/10.2134/agronj2017.05.0279>
- R Core Team (2025). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.r-project.org>
- Saar-Tsechansky, M., & Provost, F. (2007). Handling missing values when applying classification models. <https://www.jmlr.org/papers/volume8/saar-tsechansky07a/saar-tsechansky07a.pdf> [Hämtad: 2026-01-01].
- Schnug, E., Murphy, D., Evans, E., Haneklaus, S. & Lamp, J. (1993). Yield mapping and application of yield maps to computer-aided local resource management. In Robert, P. C., Rust, R. H., & Larson, W. E. (Eds.) *Proceedings of Soil Specific Crop Management: A Workshop on Research and Development Issues* (pp. 87–93). American Society of Agronomy, Crop Science Society of America, Soil Science Society of America. <https://doi.org/10.2134/1993.soilspecificcrop.c7>

- Shober, A.L. & Taylor, W. (2025). Estimating yield goals for crops. University of Delaware, Cooperative Extension.
<https://www.udel.edu/content/dam/udelImages/canr/pdfs/extension/factsheets/Estimating-Yield-Goal-for-Crops1.pdf> [Hämtad 2025-12-29].
- Souza, E.G.D., Khosla, R., Sudduth, K.A., Johann, J.A. & Bazzi, C.L. (2025). Spatial and temporal variability in yield maps can localize field management—a case study with corn and soybean. *Agronomy*, 15(5), 1179.
<https://doi.org/10.3390/agronomy15051179>
- Söderström, M., Morandin Figueiredo, B., Engström, L., Alshihabi, O. & Persson, K. (2025). Developments in proximal crop sensing for precision agriculture. In J. V. Stafford (Ed.), *Precision agriculture for sustainability: Second edition* (pp. 1–32). Burleigh Dodds Science Publishing. <http://dx.doi.org/10.19103/AS.2025.0152.02>
- Söderström, M. (2010). Interpolerade markkartor. Precision Agriculture Sweden (POS), Technical report 21. Swedish University of Agricultural Sciences (SLU), Department of Soil and Environment, Skara, Sweden.
https://pub.epsilon.slu.se/5526/1/soderstrom_m_101206.pdf [Hämtad 2026-01-01].
- Weiss, A. (2001). Topographic position and landforms analysis. Poster presentation, ESRI User Conference, San Diego, CA. https://env761.github.io/assets/files/tpi-poster-tnc_18x22.pdf.
- Wold, S., Albano, C., Dunn, W. J., Esbensen, K., Hellberg, S. & Johansson, E. (1984). Modelling data tables by principal components and PLS: class patterns, quantitative predictive relations. *Analisis*, 12, 477–485. <http://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=8963870>
- Wolters, S., Söderström, M., Piikki, K., Reese, H. & Stenberg, M. (2021). Upscaling proximal sensor N-uptake predictions in winter wheat (*Triticum aestivum* L.) with Sentinel-2 satellite data for use in a decision support system. *Precision Agriculture*, 22(4), 1263–1283. <https://doi.org/10.1007/s11119-020-09783-7>
- Ågren, A.M., Larson, J., Paul, S.S., Laudon, H. & Lidberg, W. (2021). Use of multiple LIDAR-derived digital terrain indices and machine learning for high-resolution national-scale soil moisture mapping of the Swedish forest landscape. *Geoderma*, 404, 115280. <https://doi.org/10.1016/j.geoderma.2021.115280>
- Åström, O., Månsson, S., Lazar, I., Nilsson, M., Ekelöf, J., Oxenstierna, A., & Sopasakis, A. (2025). Predicting intra-field yield variations for winter wheat using remote sensing and Graph Attention Networks. *Computers and Electronics in Agriculture*, 237, 110499. <https://doi.org/10.1016/j.compag.2025.110499>