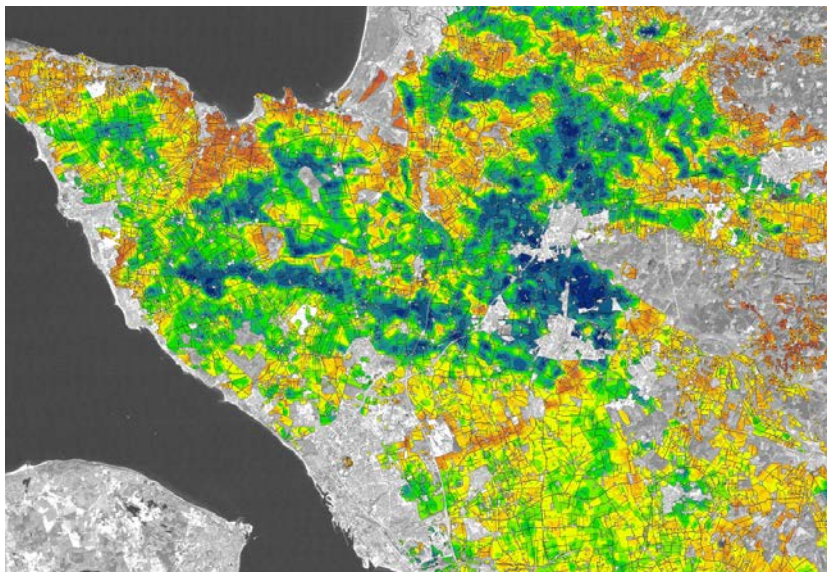




Digital markkartering av Skånes åkermark med fjärranalys



Mats Söderström

Innehållsförteckning

<i>Sammanfattning</i>	4
<i>Bakgrund</i>	5
<i>Material och metoder</i>	7
Jordprover	7
Digitala höjddata	8
Gammastrålning.....	8
Satellitdata	9
Åkermark.....	10
Modellering och mjukvaror	10
<i>Resultat</i>	12
Förklaringsvariabler	12
Responsvariabler	12
Stratifiering.....	12
Modellering.....	15
Översiktlig modell för kartläggning av hela Skåne.....	20
Alternativ metod.....	22
Andra digitala kartprodukter	22
<i>Slutsatser</i>	24
<i>Referenser</i>	25

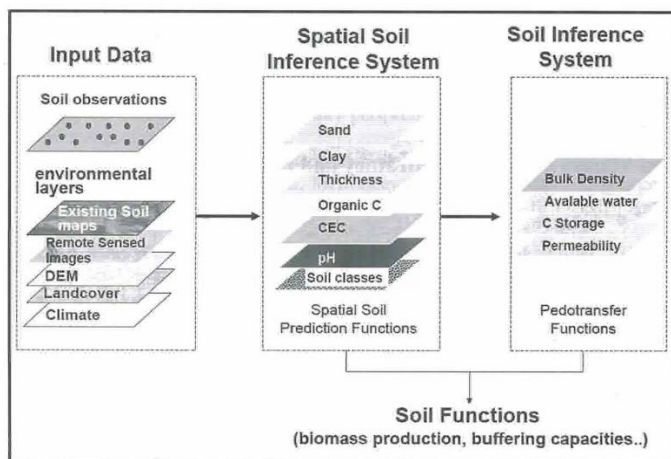
Sammanfattning

Detta är slutrapporten till projektet Applied digital soil mapping of agricultural land using remote sensing (Rymdstyrelsens dnr 199/10). I rapporten redovisas arbetet med att kombinera olika typer av fjärranalysdata (satellitdata, gammastrålningsdata och laserskannade höjddata) med jordanalyser från åkermark. Målsättningen var att testa metodik för att skapa prediktionsmodeller för ler- och mullhalt, mål-pH och buffertförmåga i åkermarkens matjord. Genom att använda metoder för *datamining* provades olika kombinationer av förklaringsvariabler, både i olika delområden och i hela Skåne. Det gick bra att ta fram relativt säkra prediktionsmodeller för lerhalt, något sämre för mål-pH och buffertförmåga, och minst lyckat var det beträffande mullhalt. Metodik för att ta fram relativt högupplösta kartor över hela Skåne presenterades och exemplifierades. Satellitdata bidrog till att skapa avsevärt bättre prediktionsmodeller jämfört med om endast gammastrålningsdata och topografiska data användes. Det redovisade arbetet var första delen av planerade två. I den andra kommande projektdelen är tanken att skapa heltäckande prediktionsmodeller och detaljerade kartor över Skånes och Västra Götalands åkermark enligt den metodik som beskrivs här. Ett sådant material skulle kunna ha stor användning i lantbruket, både inom rådgivning och direkt av lantbrukare, och leda till stora ekonomiska och miljömässiga fördelar.

Projektdeltagare var Mats Söderström (SLU, Skara), Henrik Stadig (Hushållningssällskapet i Skaraborg) och Per-Göran Andersson (Hushållningssällskapet i Malmöhus)

Bakgrund

Markkartering, d v s regelbunden jordprovtagning av åkermarken (i huvudsak av matjorden – de översta 2 dm), startade i Sverige under 1940- och 50-talet. Karteringen har tydliggjort att tillgången på fosfor, kalium och andra näringsämnen ofta varierar inom fälten. Dagens rekommendationer för markkartering utarbetas av Markkarteringsrådet och utges av Jordbruksverket (Jordbruksverket, 2012). Under senare år har nya metoder utvecklats och sensorer har börjat användas för att man ska kunna skapa mer detaljerade och bättre markkartor. Exempel på detta är mätning av markens elektriska ledningsförmåga (t ex Nehmdahl & Greve, 2001) och analys med nära infraröd (NIR) spektroskopi (Wetterlind et al., 2008). Även variation i markens naturliga radioaktiva strålning har börjat användas inom kartering av åkermark (Söderström et al., 2008). Trots allt sker den största delen av markkarteringen på vanligt sätt med 1 prov/ha, men provplatserna mäts ofta in med GPS och kartframställningen är digital. Det är lantbrukaren som står för kostnaden och man försöker av kostnadsskäl att minimera vad som analyseras och tar gärna färre prover än rekommenderat. Viktiga analyser som t ex mullhalt (motsv. organiskt material) och lerhalt (procentandel av partiklar < 0,002 mm) analyseras ofta inte alls. Under senare år har en ny disciplin utvecklats som länkar samman jordanalyser med hjälpvariabler (eng. *covariates*) som sensordata – både mätt från nära håll och med fjärranalys och andra tillgängliga kartdata. Sedan appliceras statistiska metoder där avsikten är att bygga modeller för beräkning och kostnadseffektiv kartering av olika mätvärden. På engelska går detta under benämningen *digital soil mapping DSM* (Grunwald, 2010) (Figur 1). I de flesta fall använder man billiga och heltäckande data som inte alltid är direkt korrelerade med det man egentligen vill kartera. Snarare är det indirekta samband man utnyttjar och bland en stor mängd förklaringsvariabler söker man upp statistiska samband för att skapa prediktionsmodeller (Minasny et al., 2008).



Figur 1. Principen för digital soil mapping (DSM) (Minasny et al., 2008)

Under de senaste 15 åren har lantbrukare sammanlagt tagit tusentals jordprover i Skåne. Sett över hela Skåne är dock jordproverna statistiskt klustrade i grupper på de gårdar som har karterat sin mark, och stora delar av åkermarken saknar provtagning. Spridda provtagningsprogram som syftat till att täcka in all mark pågår, framförallt Naturvårdsverkets övervakningsprogram (Eriksson et al., 2010). Antalet prover är dock få, så man får där nöja sig med att ta fram sannolika medeltal för 2,5 x 2,5 km² eller 5 x 5 km² rutor beroende hur mycket åkermark som finns i olika delar av landet. Den regionala variationen beskrivs bra men lokalt är data inte särskilt användbart (se dock även t ex Djodjic et al, 2009 och Tranter et al., 2011

för exempel på metoder för att bättre utnyttja dessa data). Jordbruksverket har under 2011-2012 tagit ett stort antal jordprover (>20 000 st.) jämnt fördelat över Sveriges åkermark för att förbättra den rumsliga täckningen. Dessa data är under sammanställning och man avser att ta fram kartor under det kommande året (Bertil Albertsson, Jordbruksverket, muntlig information 2012). Metoder för effektiv kartframställning är i det sammanhanget intressanta och det arbete som genomförts här beskriver ett tänkbart angreppssätt.

Utgångspunkten här är något annorlunda eftersom kalibreringsproverna inte har en jämn geografisk spridning, men i stället ger det en möjlighet att fånga den lokala variationen – vilket man inte kan få fram på samma sätt i de nationella programmen. Genom att kombinera jordanalyser med andra data som täcker hela området var tanken att kunna skapa en heltäckande kartbild med hög upplösning, i slutändan kanske ända ned till 20 x 20 m², men åtminstone motsvarande jordprovtagning med ett värde per 100 x 100 m². Detta kan ge både en regional bild över åkermarkens matjordsegenskaper och samtidigt kan det användas lokalt på gårdar. Om en sådan kartprodukt blir tillräckligt noggrann och korrekt skulle den kunna få flera användningsområden. Lokalt för den enskilde lantbrukaren kan man tänka sig att man kan spara in provtagnings- och analyskostnader om man t ex vill göra en kalkbehovsbestämning. I studier av läckagerisk etc. saknas ett detaljerat underlag om jordartsfördelning. Detaljerade kartor som täcker stora ytor kan även ge underlag för nya tolkningar av den kvartära utvecklingen av landskapet vilket är av intresse från ett kvartärgeologiskt och geomorfologiskt perspektiv.

Den nya nationella höjdmodellen (NNH; Lantmäteriet, Gävle) har en rumslig upplösning på 2 m x 2 m och ger helt nya möjligheter att modellera den lokala topografin. Små höjdskillnader, svackor och höjder, kan ibland sammanfalla med jordartsskillnader och det lokala dräneringsmönstret är viktigt för processer som påverkar omlagringen av jorden. NNH kan tjäna som grund för en mängd intressanta produkter eftersom dess noggrannhet är så hög. För lantbrukets del finns t ex en regel som säger att man inte får sprida gödsel om ”marklutningen är >10 % om åkermarken gränsar till sjö eller vattendrag” (24, 24 a och 24 b §§ Jordbruksverkets föreskrifter 2004:62). Även om tolkningen av denna regel i praktiken säkert kan variera så är det med befintliga papperskartor inte möjligt att med säkerhet avgränsa sådana ytor.

I det här arbetet som finansierats av Rymdstyrelsen (proj. dnr 199/10) har vi arbetat med principerna för DSM med avsikten att kartera ler- och mullhalt i åkermark huvudsakligen i Skåne, samt en mindre studie i delar av Västergötland. Målet var även att utnyttja pedotransferfunktioner för beräkning av markens buffertkapacitet och mål-pH, för att försöka kartlägga även dessa parametrar. I den här rapporten redovisas arbetet i Skåne. Projektet är den första delen i ett projekt där det är tänkt att del 1 ska innebära metodutveckling och validering, vilket innebär användning av enstaka satellitscener som därmed endast täcker delar av åkermarken, om man som nu är ute efter barmarksscener. I kommande del 2 avser vi att utnyttja metoderna i del 1 för att skapa full geografisk täckning genom att kombinera flera olika satellitscener.

Material och metoder

Jordprover

Hushållningssällskapet i Malmöhus samt i Skaraborg har sammanställt och levererat ett relativt stort antal jordanalyser av åkermark, ca 5000 i Västra Götaland och nära 30 000 i Skåne, de senare provtagna under två tidsperioder under de senaste 15 åren. Här fokuseras på data från Skåne och en sammanfattning av tillgängliga data finns i Tabell 1. Ett sådant prov representerar i normalfallet ett samlingsprov som består av upp till 10 delprover insamlade inom en 3-5 m radie. Centrpunkten för ett sådant prov är positionsbestämt med GPS. Här testades även att direkt modellera åkermarkens mål-pH och buffertförmåga, som kan beräknas med hjälp av ler- och mullhalt. Det är alltså en kombination av ler- och mullhalt som utgör grunden för mål-pH och buffertförmåga. Med sådana kartor skulle man direkt kunna beräkna kalkbehov i fält endast genom att mäta pH, som inte i sig räcker för att få fram kalkbehovet. Ett uppmätt pH-värde skulle kunna jämföras med det beräknade mål-pH-värdet i kartan och sedan kan kartvärdet för buffertkapacitet användas för att beräkna själva kalkbehovet.

Tabell 1. Sammanfattande statistik över jordanalyserdata tagna under två perioder under de senaste femton åren i Skåne (2007-2011: n = 14398; 1997-2002: n = 14663).

	År	Medel (%)	Median (%)	Std.avv. (%)	Min (%)	Max (%)
Lerhalt (%)	1997-2002	14,9	13	8,4	1	78
Lerhalt (%)	2007-2011	15,3	14	8,0	1	70
Mullhalt (%)	1997-2002	3,8	3,2	3,84	0,4	66
Mullhalt (%)	2007-2011	3,5	3,0	2,83	0,5	58
Mål-pH	1997-2002	6,1	6,1	0,15	3,8	6,7
Mål-pH	2007-2011	6,1	6,1	0,13	4,1	6,8
Buffertförmåga ^a	1997-2002	9,4	8,6	3,98	2,5	63
Buffertförmåga ^a	2007-2011	9,2	8,6	3,43	2,9	55

^a uttryckt i antal ton kalk/ha (med 50% CaO) som krävs för att höja pH en enhet

I praktiken beräknas kalkbehovet vanligen utifrån pH-värde, lerhalt och mullhalt p.g.a. högre analyskostnad för basmättnadsgrad som är ett annat sätt att mäta kalkbehov. Olika jordar har olika lämpligt pH-värde. Detta kallas mål-pH. Jordbruksverket (2012) redovisar tabeller för sambandet mellan ler- och mullhalt och mål-pH, samt även den mängd kalk (ton CaO/ha i form av kalkstensmjöl med 50 % CaO) som krävs för höjning av pH-värdet med ca en enhet. Det senare benämns här buffertförmåga. Här har vi använt ett par ekvationer i stället för beräkning av mål-pH och buffertförmåga. Mål-pH kan räknas fram enligt följande (från Kund Nissen, Lantmännen):

$$\text{pH}_{\text{mål}} = (0,01 * \text{lerhalt}) + (-0,033 * \text{mullhalt}) + 6,1 \quad (1)$$

Vid mullhalt på över 12 % kan man ev. minska det beräknade mål-pH-värdet med ytterligare 0,1 enhet. Sockerbetsodling kan kräva att mål-pH höjs med 0,5 enheter. Kalkbehovet LR (*lime requirement*) i form av kalkstensmjöl (50 % CaO) beräknas enligt Gustafsson (1999) på följande vis:

$$\text{LR (ton/ha)} = (\text{pH}_{\text{mål}} - \text{pH}_{\text{uppmätt}}) * (1,9 + (((3,5 * \text{mullhalt}) + \text{lerhalt})/3,8)) \quad (2)$$

Den kursiverade delen av ekvationen beskriver markens förmåga att stå emot ändringar av pH (buffertförmågan) och visar hur mycket kalk som måste tillföras för att höja markens pH-värde en enhet.

Digitala höjddata

Man kan förvänta sig att jordarten på en plats ofta är relaterad till dess position i landskapet. Fluvial transport och annan typ av erosion och deposition påverkas av topografin. Därför användes här höjddata som en förklaringsvariabel. Data från höjdmodellen NNH utnyttjades (Lantmäteriet, Gävle). NNH har tagits fram med hjälp av flygburen laserskanning. Man har från dessa data generat en höjdmodell som har $2 \times 2 \text{ m}^2$ spatial upplösning som anges (Lantmäteriet, Gävle) ha en noggrannhet på öppna hårdgjorda ytor på mindre än $\pm 0,2 \text{ m}$ i z-led (mindre än $\pm 0,5 \text{ m}$ generellt). För öppen åkermark sannolikt närmare det förstnämnda. Här reducerades den spatiala upplösningen till $10 \times 10 \text{ m}^2$ (medelvärdesfiltrering) för att underlätta hanteringen och för att den upplösningen överensstämmer med arean som ett jordprov ungefärligen representerar. Från denna höjdmodell togs sedan även fyra andra värden som beskriver landform fram: 1) Variabeln *Form* beräknades genom att medelhöjden inom en $0,5 \text{ ha}$ ruta jämfördes med det aktuella höjdvärdet. 2) Marklutning i varje enskild cell bestämdes (*Slope*). 3) Antalet rasterceller som dränerar till varje cell räknades ut (*Flow accumulation*). 4) Ett topografiskt vattenindex (*TWI*) beräknades enligt följande:

$$TWI = \ln(\text{Flow accumulation (i m}^2) / \tan(\text{Slope})) \quad (3)$$

Det är ett mått på potentiell markfuktighet som ofta används i studier av markegenskaper och dess relation till topografin (t ex Murphy et al., 2009).

Gammastrålning

Sveriges Geologiska Undersökningar (SGU) har sedan 1960-talet utfört flygbaserade gamma(γ)-strålningsmätningar för översiktlig kartering av markens radioaktivitet, t ex för uranprospektering. Sedan 1995 utförs mätningarna från ca 60 m flyghöjd med ett profilavstånd på 200 meter , med ca 16 m mellan datavärdena längs stråken. Detaljeringsgraden är dock inte så hög vid en sådan mätning. Ett mätvärde representerar signalen från en svårdefinierad yta på marken, huvuddelen av signalen från en yta med radien ca 200 m (IAEA, 2003). Den från luften mätbara naturliga gammastrålningen domineras av tre isotoper: kalium (^{40}K), torium (^{232}Th) och uran (^{238}U). Variation i markens naturliga radioaktiva strålning har börjat användas inom kartering av olika egenskaper hos åkermark (t ex Cook et al., 1996; Wilford & Minty, 1997; Söderström & Eriksson, 2012). Jordlagren i Sverige är i förhållande till berggrunden mycket unga. De har huvudsakligen bildats under och efter den senaste istiden. När berggrunden bryts ned eller vittrar sker en mekanisk och kemisk, relativ borttransport av U och Th. Därigenom kommer halterna av dessa ämnen att gradvis minska ju mer materialet bryts ned. De lägsta halterna har finsand som i stort sett enbart består av kvarts och fältspat (Sundevall, 2002; SSI, 2007). Det uran och torium som blivit löst i vattnet eller som finns som små korn av uranförande mineral adsorberas på lerpartiklar eller fälls på sådana. Därför har leror och leriga sediment vanligen förhöjda halter av dessa isotoper. Jordarternas modermaterial har dock stor betydelse. Förekomsten av och förhållande mellan de olika isotoperna varierar (SSI, 2007). Respektive isotopvärde omräknades till ett värde var 100:e m med hjälp av lokal kriging (Minasny et al, 2005).

För Skaraborg ordnade Hushållningssällskapet även data från sex olika gårdar där fältmätningar med fordonsburen gammastrålningsmätning utförts (metoden beskrivs t ex i Söderström, 2008). Dessa data insamlas på motsvarande sätt som den flygburna mätningen, men med en fyrhjulig motorcykel istället. Man kör med 12 eller 24 m mellan körstråken i ca $7\text{--}8 \text{ km/h}$. Ett mätvärde i sekunden samlas in och täcker en yta med $2\text{--}3 \text{ m}$ radie. Dessa data redovisas i Söderström & Stadig (2012).

Satellitdata

Ett antal olika satellitscener har införskaffats (Tabell 2). Några scener tillgängliga från tidigare projekt har även använts i projektet. I litteraturen finns redovisat en hel del olika undersökningsresultat från studier där man jämfört markegenskaper med satellitdata (några exempel: Ishida & Ando, 1999; Chen et al., 2005; Jacobsen & Söderström, 2008; Boettinger et al., 2009; Boettinger, 2010). Vanligast är det att Landsat utnyttjats, mycket beroende på tillgänglighet och täckningsgrad. Ett tidigt exempel där man testade att använda satellitdata för jordartskartering utfördes av Andersson & Gervais (1987). Den studien utfördes i nordvästra Skåne och man använde traditionell klassificeringsmetodik för att jordartskartera enligt SGU:s jordartsindelning. Även i den här studien har data från Landsat använts. Landsat 5 har varit i bruk ända sedan 1984 (Lillesand & Kiefer, 2000) men slutade leverera data i början av 2013. Det finns trots allt en mycket stor mängd data att tillgå. Landsat 7 däremot var endast fullt funktionell 1999-2003. En skillnad mellan Landsat och t ex Spot när det gäller studier av markegenskaper är att Landsat förutom banden i VNIR (synligt-nära infrarött) har två olika band i SWIR-området (short-wave IR även benämnt mid-IR, MIR) och ett band i TIR-området (termal IR). En fördel med Landsat är också att en scen täcker en större yta på marken. En nackdel är att den spatiala upplösningen är relativt låg, 30 m (15 m pankromatiskt (PAN) i Landsat 7).

Vad gäller studier av marken så har data från Aster visat sig intressant (t ex Boettinger, 2010). Aster har tre band i VNIR, sex våglängdsband i SWIR-området och fem i TIR. Den spatiala upplösningen varierar dock: banden i VNIR har 15 m spatial upplösning, 30 m i SWIR och 90 m i TIR. Emellertid slutade sensorn som registrerar SWIR (band 4-9) att fungera som den ska i april 2007, vilket betyder att satelliten endast var fullt funktionsduglig i mindre än sju år, något som begränsar tillgången på användbara scener. Vi har köpt några olika Asterscener, framförallt för detaljstudier i det kommande steg 2. Att döma av datatillgången kan man inte förvänta sig att det går att skapa en heltäckande mosaik av lämpliga Asterdata. Vi har även utnyttjat en IRS P-4-scen som redan fanns tillgänglig över Västergötland sedan ett tidigare projekt. Dessa data har hög spatial upplösning i PAN (5m) och 23 m i VNIR (dock är det endast 7 bitars data). I den föreliggande rapporten har framförallt Landsat använts för de jämförande studierna, främst beroende på att om en fungerande modellering preliminärt kan göras med enstaka Landsatscener, så är möjligheten stor att man kan skapa en heltäckande barmarksmosaik genom att kombinera flera olika Landsatscener.

Tabell 2. Satellitscener som köpts in för användning i projektet

Id	Registreringsdatum	Satellitdata	Källa, inköp	Geografisk täckning
A	1992-05-14	Landsat 5 TM	USGS ^a	Skåne
B	2000-05-06	IRS P4	^b	Skaraborg
C	2003-03-23	Landsat 7 ETM	^b	Skåne
D	2003-03-14	Landsat 7 ETM	^b	Skaraborg
E	2002-05-14	Aster VNIR, SWIR, TIR	ERSDAC ^c	Centrala Skaraborg
F	2003-03-14	Aster VNIR, SWIR, TIR	ERSDAC ^c	Centrala Skaraborg
G	2003-03-23	Aster VNIR, SWIR, TIR	ERSDAC ^c	Västra Skaraborg
H	2003-04-24	Aster VNIR, SWIR, TIR	ERSDAC ^c	Malmötrakten
I	2004-04-10	Aster VNIR, SWIR, TIR	ERSDAC ^c	Malmötrakten
J	2004-04-10	Aster VNIR, SWIR, TIR	ERSDAC ^c	Ängelholmstrakten
K	2009-04-24	Aster VNIR, TIR ^d	ERSDAC ^c	Malmötrakten
L	2009-04-24	Aster VNIR, TIR ^d	ERSDAC ^c	Ängelholmstrakten

^a United States Geological Survey, USA, nedladdad utan kostnad; ^b inköpt i tidigare projekt; ^c Earth Remote Sensing Data Analysis Center, Japan (numera Japan Space Systems); ^d SWIR-sensorn slutade fungera april 2007

Eftersom grödorna och andra åtgärder som plöjning och träda mm varierar på åkermarken mellan åren så är förhållandena ganska olika i de olika bilderna. Vi har samlat på oss vårsccener, där tanken är att kunna täcka in vissa ytor relativt väl vad gäller barmarksbilder genom att kombinera flera års bilder av olika typ. Detta kommer i huvudsak att göras i det planerade steg 2 i projektet. I steg 1 som den här rapporten avser fokuserades på att testa metoden att kombinera olika typer av data (topografi, gammastrålningsdata och satellitdata) för att prediktera åkermarkens ler- och mullhalt. Dessutom gjordes jämförelser i noggrannhet mellan olika kombinationer av indata och även undersökning av hur modellerna fungerade om området stratifierades i mindre geografiska enheter jämfört med om ett större område modellerades.

Åkermark

Jordbruksverkets blockdatabas över åkermark användes för att avgränsa åkermark. Varje polygon i blockkartan representerar åkermark som avgränsas fysiskt av hinder som t.ex diken, vägar eller skogsbyn.

Modellering och mjukvaror

En metod för s k *data-mining* (eng.) användes för att förutsättningslöst ta fram samband mellan tillgängliga förklaringsvariabler (x; ”prediktorerna”) och responsvariablerna (y). I vårt fall var y-variablerna fyra stycken: i) lerhalt, ii) mullhalt och de framräknade variablerna iii) mål-pH samt iv) buffertkapacitet. Förklaringsvariablerna som användes (som beskrivits ovan) var:

- våglängdsbanden från två Landsat-scener (band 1-7) och dess kvoter (band 1/2, 1/3, etc., 6/7),
- isotoperna ^{40}K , ^{232}Th , ^{238}U från flygburen gammastrålningsmätning och dess kvoter,
- höjd över havet från laserskanning och de härledda produkterna: form, lokalt dräneringsområde (eng. *flow accumulation*), topografiskt fuktighetsindex (TWI) samt marklutning

Här valdes att använda MARSpline som är en form av ickeparametrisk regression (Shepherd & Walsh, 2002). Med den här metoden kan man ta fram statistiska samband ur en stor datamängd, även icke-linjära samband och även om data inte är normalfördelade. Generellt kan det uttryckas:

$$y_i = m(x_{i\dots k}) + e_i \quad (4)$$

som beskriver modellsambandet $y = m(x)$ och med e_i som motsvarar felet/osäkerheten. Man kan notera att man med *data-mining* kan vaska fram rent statistiska samband som inte med nödvändighet måste ha en fysisk förklaring, vilket är en brist och risk när man sedan ska applicera modellen på andra data. Därför är det viktigt att validera modeller som tas fram mot observationer som inte ingått i själva modellkonstruktionen. Här gjordes initialt en slumpmässig uppdelning av alla tillgängliga observationer i ett kalibreringsdataset (för modellkonstruktionen) som utgjorde 80 % av tillgängliga data och ett valideringsdataset (för utvärdering) som utgjordes av resterande 20 %. För en detaljerad beskrivning av MARSpline-modellen hänvisas till Hastie et al. (2009). Metoden bygger på att data delas upp i en mängd linjära funktioner (eng. *linear splines*) $h_m(X)$ avdelade av brytpunkter (eng. *knots*) som kombineras för att producera en beräkning av y:

$$y = f(X) = \beta_0 + \sum_{m=1}^M \beta_m h_m(X) \quad (5)$$

där summeringen sker av M termer i modellen, och β_0 är intercept och β_m är viktningen av funktionerna. Allt fler funktioner läggs till för att göra en så bra prediktion som möjligt, och detta leder typiskt till att kalibreringsmodellen lätt överanpassas (eng. *overfitting*), d v s görs alltför komplicerad och specifikt anpassad till kalibreringsdata och inte sedan fungerar så bra när man applicerar den på andra data. För att hantera detta användes dock här en beskärningsfunktion (eng. *pruning*) som tar bort alla funktioner som inte på ett påtagligt sätt bidrar till prediktionen. Modellerna validerades genom att de tillämpades på den del av de tillgängliga analyserna som inte var med när modellerna togs fram. Vid validering valdes att utvärdera jämförelser mellan de observerade och predikterade mätvärdena med de statistiska måtten R^2 (determinationskoefficienten), RMSEP (medelfelet i prediktionen) och RPD (standardavvikelsen i de observerade värdena dividerat med RMSEP). För att beräkningarna ska vara bättre än medelvärdet bland de observerade värdena ska RPD-värdet vara större än ett. För interpolerade markkartor (d v s när man gör kartor på lantbrukares fält som baseras på jordprov som tagits inom fält med en täthet på ca ett prov/ha) ligger RPD ofta på 1,2 – 1,8 (se t ex Delin & Söderström, 2003 och Söderström, 2010).

Merparten av hanteringen av geografiska data och satellitbilder gjordes i det geografiska informationssystemet ArcGIS 10 (ESRI Inc., Redlands, USA) med programtilläggen Spatial Analyst, Geostatistical Analyst och Image Analysis. De statistiska analyserna och modelleringen utfördes i Statistica 10 (StatSoft Inc, Tulsa, USA).

Resultat

Förklaringsvariabler

Beträffande satellitdata fokuseras här på hur Landsatbilderna över Skåne (Tabell 2) som är de som i huvudsak kom att utnyttjas för den regionala modelleringen av markvariablerna, vilket utgjort huvuddelen av arbetet och är det som beskrivs i denna rapport. Kartor över förklaringsvariablerna hanterades i rasterform i Erdas Imagine-format eller ESRI Grid-format. Först avgränsades den del av respektive scen som bedömdes vara relativt vegetationsfri. Detta gjordes genom att åkermarken klipptes ut genom att den digitala blockkartan användes för att maska bort resterande del av scenerna. Återstoden klassificerades i 16 klasser med oövervakad klassificering. Detta gjordes med ISODATA-algoritmen (Iterative Self-Organizing Data Analysis Technique), en process som stegvis upprepar klassificerar varje pixel till ett klustercentrum så att klasserna blir så homogena som möjligt (Lillesand & Kiefer, 2000). För att bedöma vilka klasser som utgjorde barmark eller endast hade minimalt med vegetation gjordes en jämförelse mellan klasserna och dess NDVI-värde (normalized difference vegetation index; $(NIR-R)/(NIR+R)$) samt en visuell bedömning av satellitbilderna i de olika klasserna. De klasser som bedömdes vara bar åkermark användes för att maska bort den andra delen av åkermarken, som följaktligen tycktes ha ett vegetationstäck eller eventuellt annan markanvändning som kunde försvåra modelleringen av jordvariablerna. Landsat 5-scenen visas i Figur 2a. Ungefär 30 procent av åkermarken i den scenen behölls för vidare analys (sådan åkermark har rosa färgton i den bilden).

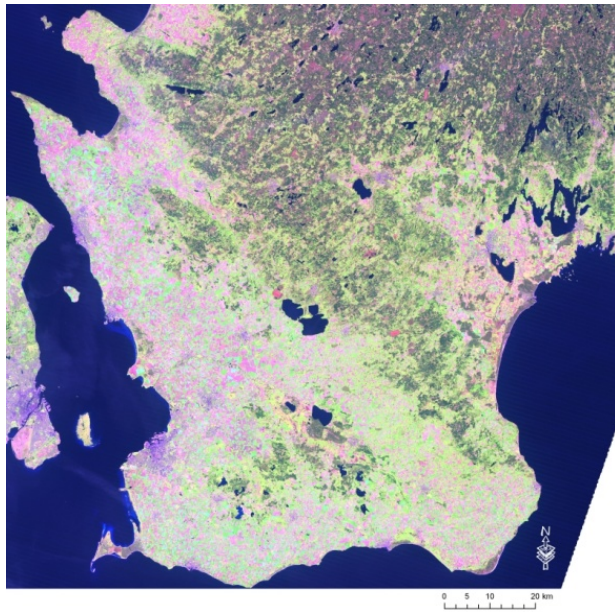
Höjddata hade initialt en spatial upplösning på $2 \times 2 \text{ m}^2$ och täckte hela Skåne. Annan mark än åkermarken maskades bort med blockkartan. En digital höjddmodell (DEM) med en reducerad upplösningen ($10 \times 10 \text{ m}^2$) togs fram genom medelvärdesfiltrering. En höjdsnitts-karta visas i Figur 2b. Form, *flow accumulation*, lutning och TWI beräknades från den höjddmodellen. På motsvarande sätt hanterades de olika gammastrålningsisotoperna ^{40}K , ^{232}Th och ^{238}U . Endast värden för åkermarken behölls, dock med en avsevärt lägre rumslig upplösning ($100 \times 100 \text{ m}^2$). En karta över torium visas i Fig. 2c.

Responsvariabler

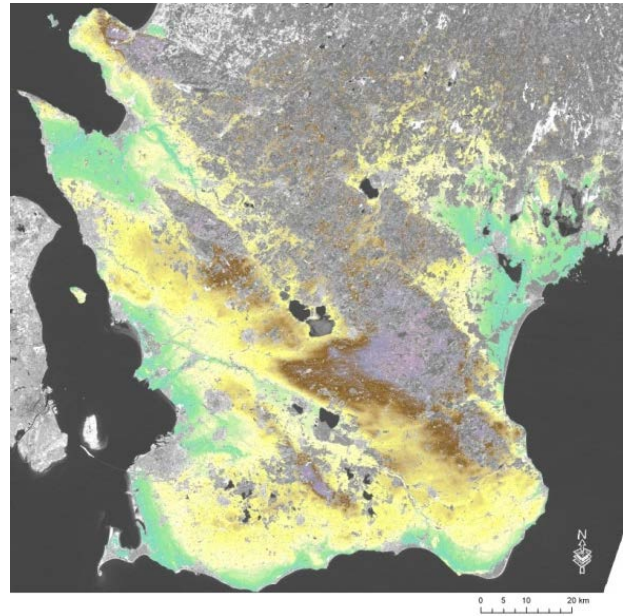
En shapefil i punktform (vektorformat) skapades med alla jordanalysdata från Skåne (Tabel 1). Alla jordanalyser, d v s från båda tidsperioderna, bedömdes statistiskt tillhöra samma population och slogs samman i ett dataset. Ler- och mullhalt förändras normalt inte särskilt mycket över den tidsperiod som det handlade om. Till denna shapefil, som innehöll responsvariablerna ler- och mullhalt, mål-pH och buffertkapacitet, extraherades värden från alla förklaringsvariablernas rasterlager. Beräkning av kvoter mellan alla vågländsband i respektive satellitbild samt mellan isotopvärden gjordes i shapefilens attributtabell.

Stratifiering

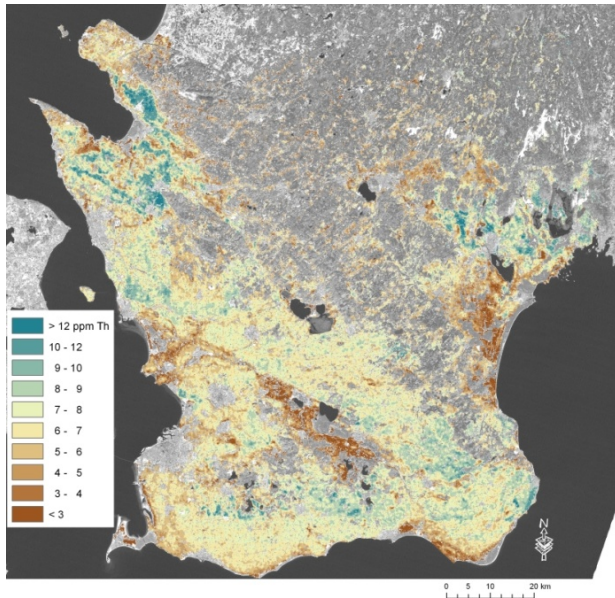
För att kunna göra en jämförelse mellan hur en modellering fungerar om man delar in hela området i mindre och mer homogena (vilket var tanken) delområden måste man skapa en rimlig geografisk indelning. Här användes Skånes olika moränområden föreslagna av Ekström redan 1936 som grund för denna stratifiering. Ekströms indelning ska skilja de olika jordarna åt på ett övergripande sätt baserat på dess genes och modermaterial. Eftersom det inte finns någon bra och detaljerad karta som direkt var användbar digitalt så digitaliserades papperskartan ur Ekström (1936) och en jämförelse gjordes sedan med isotopkartorna från gammastrålningsmätningarna.



a



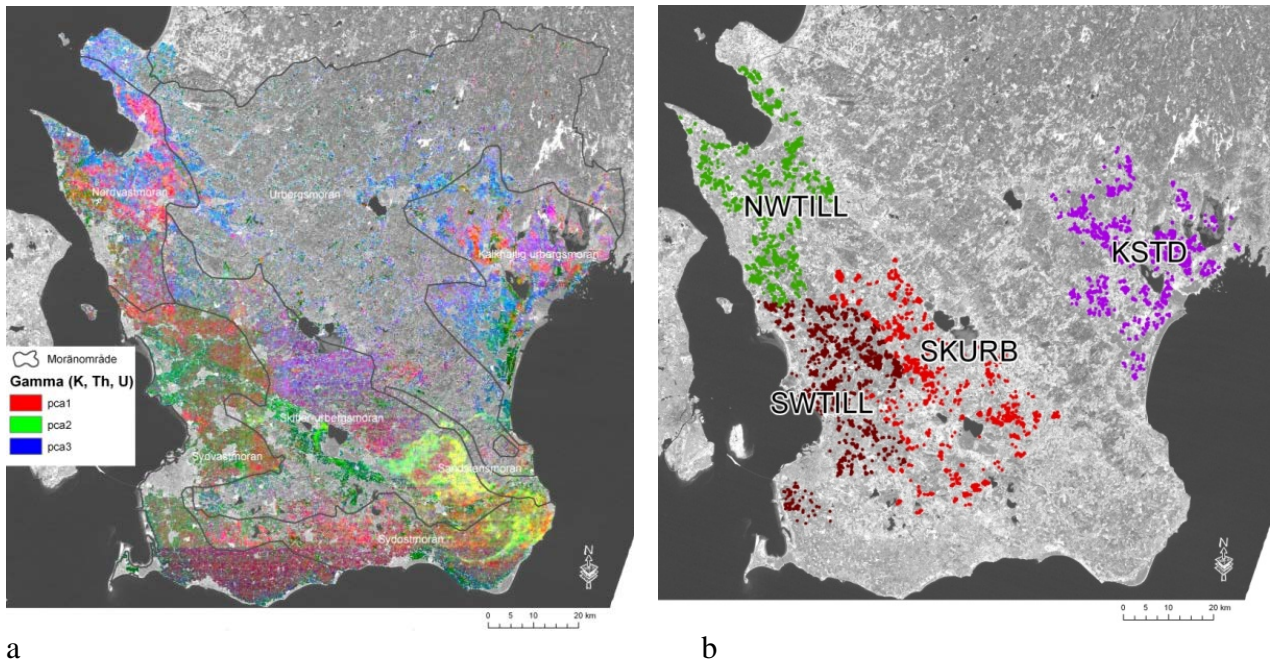
b



c

Figur 2. Kartor över några av de ingående hjälpvariablerna: a) Falskfärgad Landsat 5-bild från 1992-05-14 (RGB: SWIR band 5, NIR band 4, Red band 3), b) digitala höjddata från NNH, c) Flygburen gammastrålning (^{232}Th). I bakgrunden till Fig. b och c visas PAN-bandet från Landsat 7 scenen från 2003-03-23.

En principalkomponentkarta över isotopvärdena som överlagras av gränserna för Ekströms moränområde visas i Figur 3a. I den figuren visas PC1 som rött, PC2 som grönt och PC3 som blått. Faktum är att den gamla papperskartans gränser stämmer relativt väl med de olika färgvariationerna i PCA-kartan. Dock finns det ungefär i höjd med Falsterbonäsets nordspets och tvärs över Skåne en rak gräns i PCA-kartan som inte ser naturlig ut, och som sannolikt beror på någon icke-naturlig heterogenitet i gammastrålningsmätningarna. I stället för att försöka reda ut vad detta beror på och eventuellt räkna om gammastrålningsvärdena söder om denna gräns, så avgränsades arbetsområdet söderut här.



Figur 3. a) Gränser för Skånes moränområden (enligt Ekström, 1936) visas på en principalkomponentkarta som baseras på SGU's gammastrålningsdata, b) indelning av befintliga jordanalyser i fyra kalibreringsdataset enligt karta a (NWTILL, grönt = Nordvästmoränen, SWTILL, brunt = Sydvästmoränen, SKURB, rött = Skiffer-urbergsmoränen, KSTD, lila = Kristianstadsområdets kalkhaltiga urbergsmorän)

En indelning av befintliga jordanalyser i fyra kalibreringsdataset gjordes: Nordvästmoränen, Sydvästmoränen, Skiffer-urbergsmoränen och den kalkhaltiga urbergsmorän i Kristianstadsområdet (Figur 3b). Den sydöstra delen av området med skiffer-urbergsmorän togs bort eftersom den avvek kraftigt i PCA-kartan, sannolikt beroende på påverkan av alunskiffer från Österlen. En annan rät, nordväst-sydostlig struktur som skär genom den moränen och Sydvästmoränen i PCA-kartan, kan bero på bergartsskillnader, och här behölls data på båda sidor om denna linje.

Alla jordanalyser som hamnade inom respektive delområde fick dess tillhörighet. Som beskrivits i Material och metoder gjordes en slumpmässig uppdelning i ett kalibreringsdataset (80 %) och ett valideringsdataset (20 %) för varje delområde. En sammanfattning av responsvariablerna för respektive delområde finns i Tabel 3. Man kan utläsa vissa skillnader enligt tabellen (utan att för den skull hävda att skillnaderna är statistiskt signifikanta). Till exempel så avviker KSTD genom att det är mindre lerig jord (vilket är mycket välkänt), mer varierad mullhalt (och högst maxvärde) och mål-pH samt lägre buffertförmåga än de övriga. Lerhalten i NWTILL varierar mest och mullhalten är generellt lägst i SWTILL. Fördelningarna är i många fall skeva med en del höga värden men där huvuddelen av datamängden har lägre värden (medianvärdet lägre än medelvärde och max-värde avsevärt större än medelvärde). För mål-pH är dock data relativt normalfördelade, även om en del låga värden som nära nog kan bedömas som *outliers* förekommer i områden med hög mullhalt (då mål-pH-beräkningen kanske inte fungerar lika bra som i jordar med mer normal mullhalt). Man skulle kunna tänka sig att man borde delat upp jordarna ytterligare och hanterat jordar med de högsta mullhalterna separat. Så gjordes dock inte i detta fall.

Tabel 3. Sammanfattande statistik över jordanalysdata som använts för regional kalibreringsmodellering (antal prover: KSTD = 1579, SKURB = 1655, SWTILL = 1608, NWTILL = 1754). Förklaring till förkortningarna finns i Figur 3.

	Område	Medel	Median	Std.avv.	Min	Max
Lerhalt (%)	NWTILL	17	14	9,6	1	56
	SWTILL	17,5	17	5,9	1	47
	SKURB	15,5	14	7,5	1	52
	KSTD	10,5	7	8,5	1	47
Mullhalt (%)	NWTILL	3,8	3,6	1,5	0,9	25
	SWTILL	3,0	2,8	1,1	0,5	13
	SKURB	4,0	3,7	2,5	0,5	36
	KSTD	3,9	2,9	4,4	0,6	66
Mål-pH	NWTILL	6,1	6,1	0,11	5,5	6,5
	SWTILL	6,2	6,2	0,06	5,8	6,5
	SKURB	6,1	6,1	0,12	4,9	6,6
	KSTD	6,1	6,1	0,17	3,8	6,5
Buffertförmåga ^a	NWTILL	9,8	9,2	3,1	3,5	33
	SWTILL	9,3	9,3	2,1	3,6	22
	SKURB	9,7	9,2	3,2	3,1	38
	KSTD	8,3	7,1	4,9	2,7	63

^a uttryckt i antal ton kalk (med 50% CaO) som krävs för att höja pH en enhet

Modellering

MARSpline-modeller togs fram för alla responsvariabler och alla delområden, samt även för hela datasetet (d v s de fyra delområdena sammanslagna) vad gäller ler- och mullhalt. Modeller gjordes för olika kombinationer av kategorier av förklaringsvariabler (det är ju många variabler i respektive av de tre kategorierna satellitdata, gammastrålningsdata och höjddata). Detta för att utvärdera vilken nytta man har av de olika typerna av förklaringsvariabler i prediktionsmodelleringen.

Modellerna som togs fram applicerades sedan på det korresponderande valideringsdataset, d v s en kalibreringsmodell för t ex SKURB-området applicerades sedan på valideringsdatasetet för SKURB. Alla valideringar redovisas i Tabel 4-7 (delområdena) och i Tabel 8-9 (alla delområden sammanslagna).

Först gjordes en jämförelse mellan de olika satellitbilderna (Landsat 1992 (LS92) och Landsat 2003 (LS03)) vad gäller lerhalt och mullhalt i nordvästmoränen (NWTILL) (Tabel 4-5). I båda fallen fungerade LS92 bäst med högre R^2 och RPD och lägre RMSEP. Dessutom gjordes modeller genom att data från både LS92 och LS03 kombinerades (ej redovisat i tabellerna) men det gav inget mervärde i förbättrade valideringsvärden jämfört med om endast LS92 användes. Sålunda användes endast LS92 i det fortsatta modellarbetet.

I modellerna med de olika kategorierna av förklaringsvariablerna var utgångspunkten att använda heltäckande data, först och främst gammastrålningsvariablerna (Gamma), och sedan addera höjdvariablerna (Z) och till sist satellitvariablerna (LS92). För NWTILL kombinerades även LS92 med Z för att utvärdera effekten av den kombinationen. Man kan se i Tabel 4-7 att

det nästan alltid blir bättre att addera en förklaringsvariabelskategori, d v s LS92+Z+Gamma är bättre än Z+Gamma som är bättre än endast Gamma. Dock ger Z-variablerna sällan ett betydande tillskott. Emellertid ger LS92 ofta påtagligt bättre prediktionsmodeller. Endast LS92+Z ger mer effektiva modeller än Z+Gamma, men aldrig bättre värden än om alla variabler LS92+Z+Gamma används.

Prediktionsmodellerna för lerhalt fungerar generellt ganska bra. Modellerna för mullhalt fungerar inte särskilt bra, även om området med skiffer-urbergsmorän (SKURB) är något bättre. Prediktionerna för mål-pH är mer korrekta än de för buffertförmågan, utom för Kristianstadslätten - båda är dock avsevärt bättre än mullhalt. Mullhalten varierar ofta ganska lite samtidigt som osäkerheten i laboratorieanalysen och provtagningen medför att det redan i kalibreringsdatasetet finns en relativt hög osäkerhet. Det finns många tidigare exempel på genomförd forskning där man använt markens färgskillnader även i pankromatiska bilder (t ex Chen et al., 2005) för uppskattning av mullhalt. Det har ofta då varit fråga om enskilda fält eller mindre områden som varit relativt homogena. Här kan man tänka sig att man skulle kunna ha gjort en mer restriktiv klassning av satellitdata så att mindre och mer homogena ytor (i datarymden såväl som i geografien) modellerats. Fler våglängdsband som t ex de i Aster eller Landsatdata från flera tidpunkter över samma fält skulle kunna öka modellens träffsäkerhet.

I stort fungerar modellerna för delområdena bättre än då data från alla delområden slagits samman (jmf. Tabell 4-7 med Tabell 8-9). Skillnaden är dock inte så stor, och modellen för lerhalt för alla delområden med alla kategorier av förklaringsvariabler (Tabell 9) måste sägas vara ganska bra med $R^2 = 0,6$ och $RPD = 1,57$.

En skillnad mellan området i Västergötland (som redovisas i Söderström & Stadig, 2012) är att användning av enbart gammavariabler inte fungerar särskilt bra i Skåne enligt det som redovisas här. I Västergötland beskrev en potensfunktion sambandet mellan endast ^{232}Th från flygburen mätning och lerhalt i åkermark mycket väl (valideringsvärdena $R^2 = 0,68$ och $RPD = 1,77$), dock med ett större RMSEP (7,16). Skillnaderna beror sannolikt på att området i Skåne är betydligt mer heterogent och mer påverkat av en rad olika sedimentära bergarter vilka kanske inte på samma sätt som i Västergötland ger upphov till gammastrålningsvärden som är lika stabilt korrelerade till texturen i matjorden. Vid detaljerade gammastrålningsmätningar med fordonsburen utrustning på enskilda fält har dock sambandet mellan Th och ler varit väldigt starkt vid mätningar på två gårdar, $R^2 = 0,92$ respektive $R^2 = 0,77$ (Söderström et al., 2008 samt opublicerade data). För mullhalt är emellertid inte heller isotoper från fordonsburen gammastrålningsmätning särskilt stabila som prediktorer (Söderström et al., 2008).

Satellitdata ger ett betydande tillskott till modellernas tillförlitlighet och användningen av sådana data för generella prediktionsmodeller torde vara lämpligt att studera vidare.

Tabell 4. Validering av MARSpline-modellerna för beräkning av lerhalt (%) för respektive delområde.

Modell	N	R ²	RMSEP	RPD
a) Nordvästmoränen	439			
Gamma + LS92 + Z		0,67	5,354	1,75
Gamma + Z		0,54	6,393	1,46
Gamma		0,50	6,655	1,41
LS92 + Z		0,57	6,138	1,53
LS92*	(272)	0,54	6,382	1,47
LS03*	(272)	0,40	6,903	1,28
b) Sydvästmoränen	402			
Gamma + LS92 + Z		0,57	3,549	1,52
Gamma + Z		0,47	3,959	1,37
Gamma		0,47	3,957	1,37
c) Skiffer-urbergsmoränen	414			
Gamma + LS92 + Z		0,58	4,729	1,53
Gamma + Z		0,44	5,414	1,33
Gamma		0,31	6,025	1,20
d) Kristianstadslätten	395			
Gamma + LS92 + Z		0,58	5,623	1,53
Gamma + Z		0,41	6,613	1,30
Gamma		0,35	7,006	1,23

* Max två interaktioner i modellen, annars max en interaktion

Tabell 5. Validering av MARSpline-modellerna för beräkning av mullhalt (%) för respektive delområde.

Modell	N	R ²	RMSEP	RPD
a) Nordvästmoränen	439			
Gamma + LS92 + Z		0,15	1,504	1,08
Gamma + Z		0,05	1,586	1,03
Gamma		0,02	1,617	1,01
LS92 + Z		0,17	1,487	1,10
LS92*	(272)	0,15	1,184	1,08
LS03*	(272)	0,14	1,205	1,06
b) Sydvästmoränen	402			
Gamma + LS92 + Z		0,30	1,248	1,18
Gamma + Z		0,19	1,334	1,11
Gamma		0,12	1,387	1,06
c) Skiffer-urbergsmoränen	414			
Gamma + LS92 + Z		0,42	2,072	1,32
Gamma + Z		0,22	2,415	1,13
Gamma		0,09	2,603	1,05
d) Kristianstadslätten	395			
Gamma + LS92 + Z		0,32	3,334	1,11
Gamma + Z		0,26	3,893	0,95
Gamma		0,24	4,015	0,92

* Max två interaktioner i modellen, annars max en interaktion

Tabell 6. Validering av modellerna för beräkning av buffertförmåga (ton kalk/ha) i delområdena.

Modell	N	R ²	RMSEP	RPD
a) Nordvästmoränen	439			
Gamma + LS92 + Z		0,50	2,149	1,42
Gamma + Z		0,38	2,403	1,27
Gamma		0,36	2,450	1,24
LS92 + Z		0,41	2,333	1,30
LS92*		0,39	2,370	1,28
b) Sydvästmoränen	402			
Gamma + LS92 + Z		0,47	1,597	1,38
Gamma + Z		0,37	1,743	1,26
Gamma		0,34	1,781	1,23
c) Skiffer-urbergsmoränen	414			
Gamma + LS92 + Z		0,39	2,499	1,28
Gamma + Z		0,27	2,732	1,17
Gamma		0,13	2,994	1,06
d) Kristianstadslätten	395			
Gamma + LS92 + Z		0,37	3,715	1,20
Gamma + Z		0,31	4,123	1,08
Gamma		0,29	4,261	1,04

* Max två interaktioner i modellen, annars max en interaktion

Tabell 7. Validering av MARSpline-modellerna för beräkning av mål-pH för respektive delområde.

Modell	N	R ²	RMSEP	RPD
a) Nordvästmoränen	439			
Gamma + LS92 + Z		0,58	0,068	1,53
Gamma + Z		0,45	0,077	1,35
Gamma		0,39	0,081	1,28
LS92 + Z		0,49	0,074	1,40
LS92*		0,45	0,077	1,35
b) Sydvästmoränen	402			
Gamma + LS92 + Z		0,36	0,053	1,25
Gamma + Z		0,26	0,057	1,16
Gamma		0,24	0,057	1,14
c) Skiffer-urbergsmoränen	414			
Gamma + LS92 + Z		0,52	0,086	1,44
Gamma + Z		0,32	0,103	1,21
Gamma		0,19	0,112	1,11
d) Kristianstadslätten	395			
Gamma + LS92 + Z		0,33	0,136	1,08
Gamma + Z		0,24	0,149	0,98
Gamma		0,21	0,154	0,95

* Max två interaktioner i modellen, annars max en interaktion

Tabell 8. Validering av MARSpline-modellerna för beräkning av mullhalt (%) för alla delområden (n=1650)

Modell	R ²	RMSEP	RPD
Gamma + LS92 + Z	0,27	2,20	1,16
Gamma + Z	0,17	2,36	1,08
Gamma*	0,14	2,40	1,06
LS92 + Z*	0,22	2,25	1,13
LS92*	0,11	2,41	1,06

* Max två interaktioner i modellen, annars max en interaktion

Tabell 9. Validering av MARSpline-modellerna för beräkning av lerhalt (%) för alla delområden (n=1650)

Modell	R ²	RMSEP	RPD
Gamma + LS92 + Z*	0.60	5.25	1.57
Gamma + Z*	0.41	6.34	1.30
Gamma*	0.37	6.58	1.25
LS92 + Z*	0.51	5.78	1.43
LS92*	0.49	5.89	1.40

* Max två interaktioner i modellen, annars max en interaktion

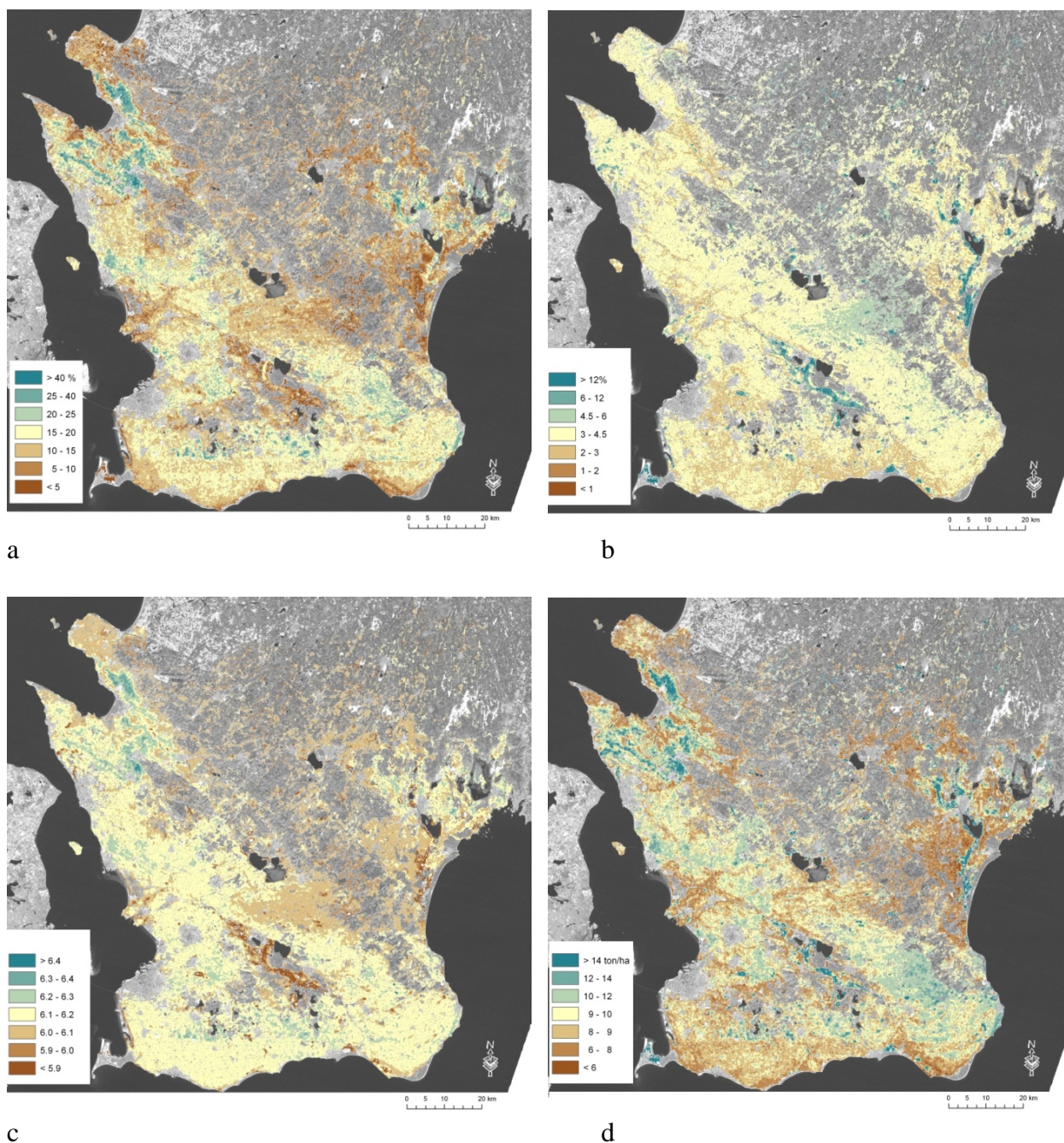
Översiktlig modell för kartläggning av hela Skåne

För att ta fram kartbilder över hela Skåne som kunde användas som exempel rent metodmässigt beträffande kartframställning från MARSpline-modellerna och även som diskussionsunderlag i det fortsatta arbetet, gjordes modeller för de fyra responsvariablerna som baserades på alla tillgängliga data, d v s alla ca 29000 jordanalysdata som redovisas i Tabell 1. För att ta fram "den bästa" heltäckande modellen gjordes här först modeller baserad på både gammastrålningsvariablerna (Gamma) samt alla höjdvariablerna (Z). Eftersom dessa data finns för hela Skåne kan en sådan modell appliceras på ingående kartdata och ge beräkningar för all åkermark i Skåne. Dessutom gjordes modeller där även satellitdata ingick, men eftersom inte den mark som klassificerades som barmark eller med mkt tunn vegetation utgjorde mer än ca 30 % av åkermarken kunde inte en heltäckande kartbild genereras av de modellerna. Den senare modelltypen hade ju visat sig mest effektiv i studierna av delområdena (Tabell 4-9) och de beräkningarna användes i första hand. I det fall inte satellitdata fanns, användes Gamma+Z-modellerna. Det betyder att för att göra en heltäckande kartbild blandades alltså de två modellerna, vilket i sin tur medförde att noggrannheten i den slutliga produkten blev varierad.

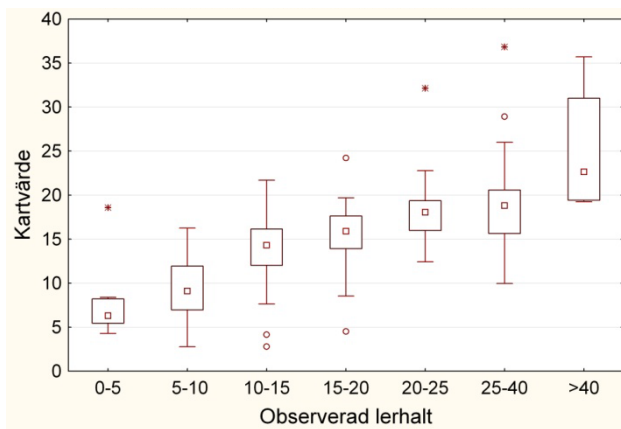
För att applicera de båda modellerna på kartdata (det är alltså en annan sak än att som tidigare applicera modellen på ett valideringsdataset), krävdes att kartdata anpassades till ett gemensamt format. Här användes den lägsta upplösningen bland ingående data ($100 \times 100 \text{ m}^2$) som avgjordes av gammastrålningsinformationen. Den geografiska upplösningen försämrades alltså för både satellitinformationen (här $30 \times 30 \text{ m}^2$) och höjdinformationen ($10 \times 10 \text{ m}^2$) vilket betyder att potentialen finns för att ta fram betydligt mer detaljerade kartprodukter. Ett punktnät skapades över all åkermark, men en buffertzona på 20 m längs fältgränserna togs bort för att undvika att satellitpixlar som influerats alltför mycket av förhållanden utanför fältet kom med. Alla data från de ingående hjälpvariablerna överfördes till punktnätets attributtabell och alla kvoter beräknades. På detta sätt kunde attributtabellen användas i Statistica som måltabell när man applicerade de framtagna MARSpline-modellerna.

Modellresultaten för den beräknade lerhalten, mullhalten, mål-pH och buffertförmågan interpolerades sedan till ett $50 \times 50 \text{ m}^2$ raster som täckte all åkermark (alltså även den tidigare borttagna buffertzonen på 20 m). Här användes interpolationsmetoden RBF (radial basis functions; Press et al, 2007) för denna procedur.

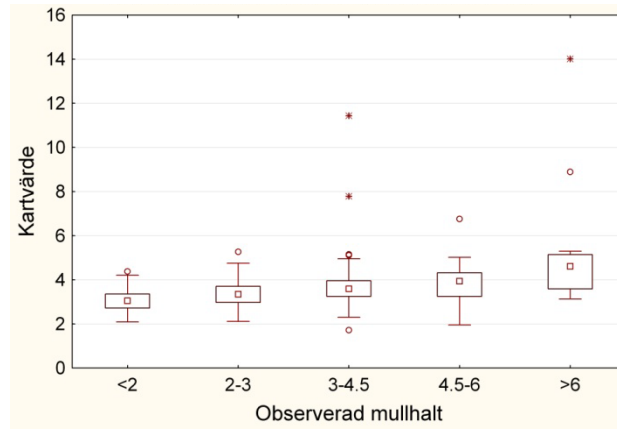
Man kan se i kartorna i Figur 4 att det går att skönja gränsen i höjd med Falsterbonäset som diskuterades i samband med PCA-kartan i Figur 3a. För att få en uppfattning om hur bra dessa kartor är trots blandningen av indata och två olika prediktionsmodeller, gjordes en jämförelse mellan kartvärden och 242 jordprover som var relativt jämnt fördelade över hela Skåne och som ingår i materialet som redovisas i Eriksson et al. (2010). I Figur 5a-d visas vilka värden i kartorna i Figur 4a-d som korresponderar med de observerade värdena. På grund av att spridningen i kartvärdena inte var så stor valdes att gruppera diagrammen efter de observerade värdena vilket inte är brukligt. Även om inte klasserna stämmer helt överens så kan man icke desto mindre se att kartvärdena på hyfsat sätt ändå generellt följer variationen i de observerade värdena, även för mullhalt som hela tiden varit svårast att prediktera. Det är tydligt att oftast, särskilt för mullhalt, så är de högsta observerade värdena för låga i kartan och omvänt, att de lägsta observerade värdena är något för höga i kartan.



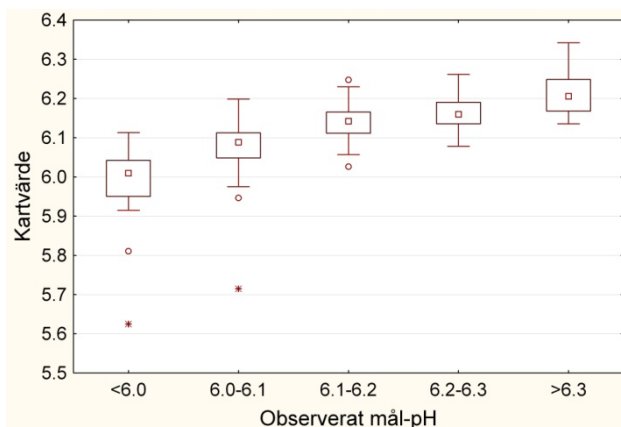
Figur 4. Kartor över Skåne, a) Lerhalt; b) Mullhalt; c) mål-pH; d) Buffertförmåga, som baseras på två globala modeller (Z+Gamma och LS92+Z+Gamma) där alla variabler från Landsat, gammastrålningsmätning och den digitala höjdmodellen ingår i den ena modellen och de två sistnämnda variabelkategorierna ingår i den andra modellen. Modellerna är kalibrerade med alla tillgängliga jordanalyser, ingen hänsyn har tagits till att det finns varierande osäkerhet i hjälpvariablerna. Utdata från modell LS92+Z+Gamma har använts där alla data funnits tillgängliga. Annars har data från modell Z+Gamma använts.



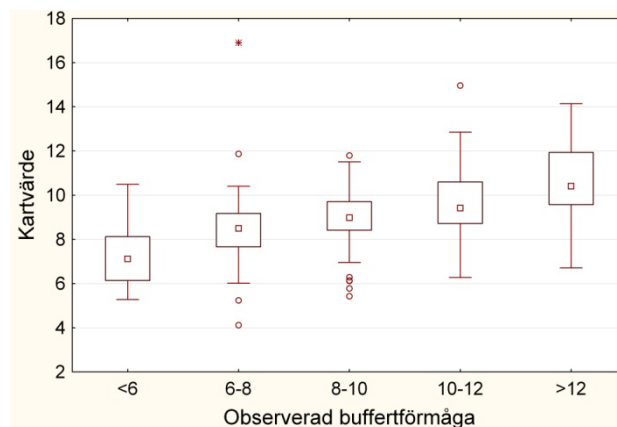
a



b



c



d

Figur 5a-d. Boxplot-diagram som visar kopplingen mellan 242 väl spridda jordprover och kartvärdena från kartorna i Fig. 4a-d. Notera att i diagrammen visas kartvärdena på y-axeln och de observerade värdena som grupperingsvariabel. Boxarnas utbredning = 25-75 percentilerna, fyrkanten i boxen = medianvärdet, staplarna = *non-outlier range*, cirkelarna = *outliers*, stjärnan = extremvärden (för definition av *outliers* och extremvärden se Hill & Lewicki, 2007)

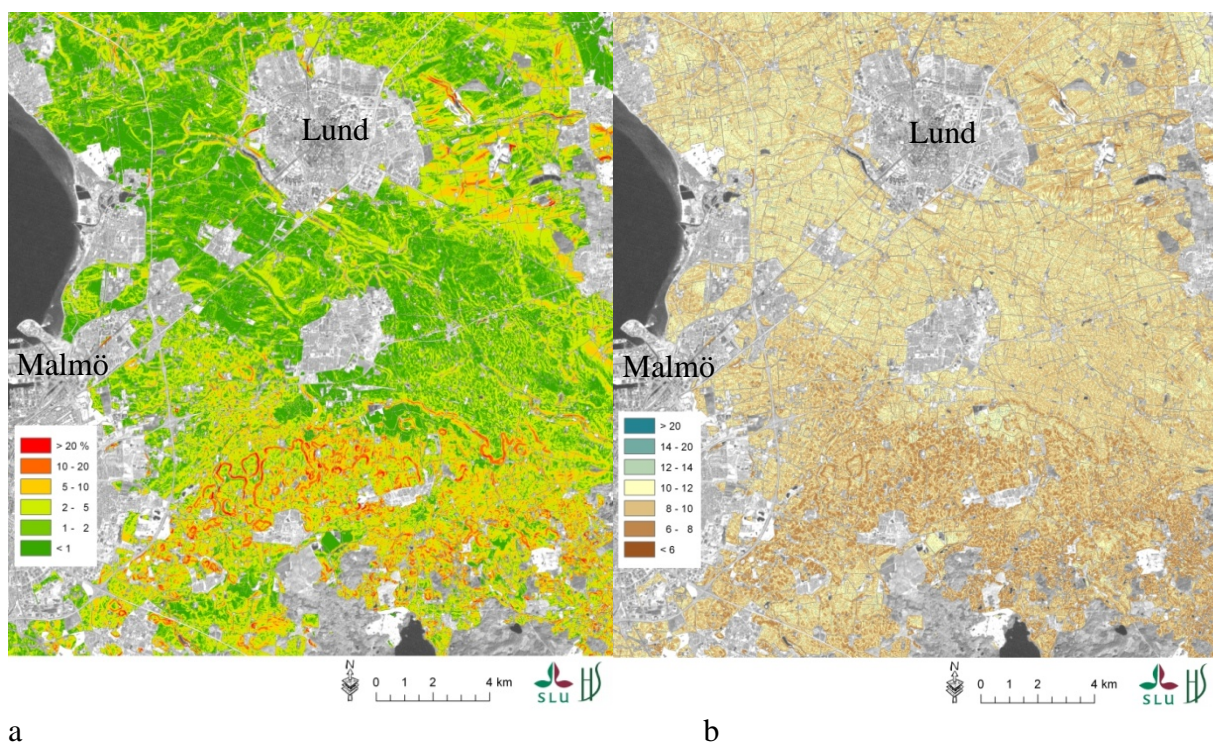
Alternativ metod

För framtida tillämpning hade det möjligen varit att föredra att applicera modellerna på det slutgiltiga rastrets upplösning (här 50 x 50 m²) i stället för att som nu gjordes först försämra upplösningen för två av de tre förklaringskategoriernas datavärden till 100 x 100 m², och sedan skapa den högre upplösningen med interpolation. I praktiken hade det alltså inneburit att man borde interpolerat ned gammastrålningsvariablerna till 50 x 50 m², reducerat upplösningen på satellit och höjddatavariablerna och sedan applicerat modellerna. Det hade i och för sig inneburit ett mycket större dataset som modellerna skulle appliceras på, men hade man kunnat hantera det kunde man möjligen erhållit högre detaljeringsgrad i de resulterande kartorna.

Andra digitala kartprodukter

Man kan tänka sig en lång rad olika kartprodukter som kan genereras av de data som använts här. Direkt användbart är data från NNH som kan ligga till grund för flera olika produkter för inomfältbruk. I Figur 6a visas en karta över Malmö-Lund-området där marklutningen är beräknad med hjälp av data från den nya nationella höjddatabasen (NNH). Jordbruksverkets regler för gödselspridning säger att man inte får sprida gödsel om

”marklutningen är >10 % om åkermarken gränsar till sjö eller vattendrag” (24, 24 a och 24 b §§ Jordbruksverkets föreskrifter 2004:62). Man kan alltså med hög noggrannhet på varje enskilt fält avgränsa de ytor som faller inom den ramen. I Jordbruksverkets skrivning anges dock att man ska beräkna lutningen på 50 x 50 m² ytor. I Figur 6 är beräkningen gjord för 10 x 10 m² ytor. Man kan eventuellt anta att Jordbruksverkets regler beror på svårigheten att beräkna marklutning för mindre ytor med hjälp av fastighetskartan i pappersform. Kanske kan dessa nya data ändra på reglerna. Figur 6b är en karta över samma område som visar TWI. Det är ett index som ger en bra fingervisning om markfuktigheten, d v s där topografien gör att vatten borde ansamlas. Man bör dock ha i åtanke att höjdvärdena på åkermark i NNH inte alltid är korrekta. Det finns exempel där det uppenbart är en grödas överyta som anges som marknivå, vilket påverkar t ex lutningsberäkning och bedömning av dräneringsbehov. Men i stor sett är NNH mycket detaljerad och felen små.



Figur 6. Exempel på produkter användbara för lantbruket: a) Kartan visar marklutning i procent framräknad från NNH. Mark med lutning över 10 % får inte gödslas om den gränsar till sjö eller vattendrag (Jordbruksverket, 2010); b) Topografiskt vattenindex (enhetslöst) som ger bild av potentiellt markvatteninnehåll – ju högre index desto större vatteninnehåll.

Slutsatser

Det gick att ta fram relativt bra lerhaltskartor med hög detaljeringsgrad (50 x 50 m²) över Skåne genom att kombinera olika hjälpvariabler med befintliga jordanalyser. Det var betydligt svårare att göra effektiva modeller för beräkning av markens mullhalt. Mullhalten varierar ofta ganska lite samtidigt som osäkerheten i laboratorieanalysen och provtagningen medför att det redan i kalibreringsdatasetet finns en relativt hög osäkerhet. Här testades även att direkt modellera åkermarkens mål-pH och buffertförmåga, som kan beräknas med hjälp av ler- och mullhalt. Det är alltså en kombination av dessa värden som utgör grunden för dessa båda variabler. Med sådana kartor skulle man direkt kunna beräkna kalkbehov i fält endast genom att mäta pH, och sedan jämföra med det beräknade mål-pH-värdet och använda kartvärdet för buffertkapacitet för att beräkna själva kalkbehovet. Både mål-pH och buffertförmåga gick att prediktera bättre än mullhalt. Metoden öppnar för en betydlig ekonomisk besparing för lantbrukare som vill få en uppfattning om kalkbehovet på sin gård.

Genom att kombinera höjddata, gammastrålningsdata satellitdata erhöles den bästa modellen. Tar man bort satellitdata försämras modellen. Det går att göra bättre modeller om man stratifierar området så att man arbetar med mindre, mer homogena områden.

Gammastrålningsdata och höjddata finns tillgängligt över större delen av Sveriges åkermark. Dock finns det vissa svårigheter att hantera dessa data eftersom dess upplösning på marken är så olika. Dessutom finns det en hel del felaktigheter i gammastrålningsdata, t ex nivåskillnader i olika delar av en region kanske beroende på att mätningarna är utförda olika år och till viss del med olika utrustning och noggrannhet. Med tanke på hur effektiv den flygburna gammastrålningsmätningen är som hjälpvariabel, borde man överväga om det är möjligt att modifiera datainsamlingen vid flygburen gammastrålningsmätning så att data blir bättre anpassade för detta ändamål. Om det är möjligt vore lägre flyghöjd och tätare mellan flygstråken över jordbruksmark intressant att prova.

Höjddata har mycket hög spatial upplösning som man snarare måste reducera för att kunna använda i detta sammanhang. I de modeller som användes här gav inte höjddata så stor förbättring i modellerna.

Däremot gav alltså satellitdata ett påtagligt positivt bidrag till bättre prediktionsmodeller. I huvudsak har vi här använt Landsatdata. Data från maj fungerade bättre än data från mars månad. Det bör vara av intresse att försöka sammanställa en mosaik av satellitscener för att skapa en bättre täckningsgrad av barmark. Med det som underlag bör man kunna skapa bra prediktionsmodeller för texturkartering med hög spatial upplösning.

För detaljkartering i vissa områden med intensivt jordbruk är det lämpligt att vidare studera möjligheten att använda andra satellitdata som ytterligare kan förbättra prediktionsmodellerna. Asterdata har delvis använts i projektet men har inte utvärderats till fullo ännu. Bildtillgången kommer dock inte att medge att man skapar heltäckande mosaiker med barmark över all åkermark med Asterdata. Flygbilder, både IR och SVV är andra datakällor som finns tillgängliga och kan vara användbara (Chen et al., 2005). Satelliter med sensorer med högre spatial upplösning skulle också kunna vara av intresse lokalt, men torde inte vara kostnadseffektiva för större områden och regioner.

Referenser

- Andersson, L. & Gervais, P. 1987. Marktypskartering i NV Skåne med satellitfjärranalys. Rapport 152, Institutionen för markvetenskap, Sveriges lantbruksuniversitet, Uppsala. 32 p.
- Boettinger, J.L. 2010. Environmental covariates for digital soil mapping in the western USA. In: J.L. Boettinger et al. (eds.), *Digital Soil Mapping, Progress in Soil Science 2*, Springer, pp. 17-28.
- Boettinger, J.L., Ramsey, R.D., Bodily, J.M., Cole, N.J., Kienast-Brown, S., Nield, S.J., Saunders, A.M. & Stum, A.K. 2009. Landsat Spectral Data for Digital Soil Mapping. In: A. E. Hartemink et al. (eds.) *Digital Soil Mapping with Limited Data*. Springer, pp. 193-202.
- Chen, F., Kissel, D. E., West L. T., Rickman, D., Luvall, J. C. & Adkins W. 2005. Mapping surface soil organic carbon for crop fields with remote sensing. *Journal of Soil and Water Conservation*, 60, 1, 51-57.
- Cook, S.E., Corner, R.J., Groves, P.R., Grealish, G.J. 1996. Use of airborne gamma radiometric data for soil mapping. *Australian Journal of Soil Research*, 34, 183-194.
- Delin, S. & Söderström, M. 2003. Potential for improving interpolation of soil data using geostatistics and soil electrical conductivity. *ACTA Agric. Scand., Sect. B, Soil and Plant Sci.* 2003, vol. 52, no. 4, pp. 127-135.
- Djodjic, F., Nisell, J., Brandt, M. & Söderström, M. 2009. Jordartskarta för jordbruksmark - jämförelsestudie mellan olika metoder för interpolation av mätpunkter samt testning av deras betydelse för PLC-beräkningar. SMED rapport 2009:25.
- Ekström, G. 1936. Skånes moränområden. Ur *Svensk geografisk årsbok*. Lund, 1936. 8:o. s. 70-78.
- Eriksson, J., Mattsson, L. & Söderström, M. 2010. Tillståndet i svensk åkermark och gröda. Naturvårdsverket, rapport 6349. 129 p.
- Grunwald, S. 2010. Current State of Digital Soil Mapping and What is Next? In: J.L. Boettinger et al. (eds.), *Digital Soil Mapping, Progress in Soil Science 2*, Springer, pp. 3-12.
- Gustafsson, K. 1999. Models for precision application of lime. Ur: Stafford J.V. (ed.): *Precision Agriculture '99. Part 1. 2nd European Conference on Precision Agriculture*, Odense Danmark. SCI, UK, pp 175-180.
- Hastie, T., Tibshirani, R. & Friedman, J. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd edition, Springer Series in Statistics, 746 p. Available at Stanford University: <http://www-stat.stanford.edu/~tibs/ElemStatLearn/>
- Hill, T. & Lewicki, P. 2007. *STATISTICS: Methods and Applications*. StatSoft, Tulsa, USA.
- IAEA, 2003. Guidelines for radioelement mapping using gamma ray spectrometry data. International Atomic Energy Agency (IAEA). IAEA-TECDOC-1363. Vienna, Austria.
- Ishida, T. & Ando, H. 1999. Use of disjunctive co-kriging to estimate soil organic matter from Landsat Thematic Mapper image. *Int. J. Remote Sensing*, 20, 1549-1565.
- Jacobsen, A. & Söderström, M. 2008. Anvendelse af geostatistik og remote sensing data til kortlægning af jordens lerindhold. *Precisionsodling Sverige, POS Teknisk rapport 16*. Institutionen för mark och miljö, Sveriges lantbruksuniversitet. 25 p.
- Jordbruksverket, 2011. Gödsel och miljö. OVR206.
- Jordbruksverket, 2012. Riktlinjer för gödsling och kalkning 2012. JO11:21.
- Lillesand, T.M. & Kiefer, R.W. 2000. *Remote sensing and image interpretation*. John Wiley and Sons, New York.
- Minasny, B., McBratney, A.B. & Whelan, B.M. 2005. VESPER version 1.62. Australian Centre for Precision Agriculture, McMillan Building A05, The University of Sydney, NSW 2006.

- Minasny, B., McBratney, A. B. & Lark, R. M. 2008. Digital Soil Mapping Technologies for Countries with Sparse Data Infrastructure. In A. E. Hartemink et al. (eds.) *Digital Soil Mapping with Limited Data*. Springer, pp. 15-30.
- Murphy, P.N.C., Ogilvie, J. & Arp, P. 2009. Topographic modelling of soil moisture conditions: a comparison and verification of two models. *European Journal of Soil Science*, 60, 94–109.
- Nehmdahl, H. & Greve, M. H. 2001. Using Soil Electrical Conductivity Measurements For De-lineating Management Zones On Highly Variable Soils in Denmark. In: Grenier, G. & Black-more, S. (eds.): *ECPA 2001. 3rd European Conference on Precision Agriculture* (vol. 1). agro Montpellier, 461-466.
- Pracilio, G., Adams, M.L., Smettem, K.R.J., Harper, R.J. 2006. Determination of spatial distribution patterns of clay and plant available potassium contents in surface soils at the farm scale using high resolution gamma ray spectrometry. *Plant & Soil*, 282, 67–82.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P. 2007. *Numerical Recipes: The Art of Scientific Computing*, 3rd edition, New York: Cambridge University Press.
- Shepherd, K.D. & Walsh, M.G. 2002. Development of reflectance spectral libraries for characterization of soil properties. *Soil Science Society of America Journal*, 66, 988–998.
- SSI, 2007. The radiation environment in Sweden. Swedish Radiation Protection Authority (SSI). Report 2007:02. (In Swedish with English summary)
- Sundevall, S.E., 2002. Metodbeskrivning och radonriskkarta över Göteborgs och Mölndals kommuner. SGU. Rapport 2002:27.
- Söderström, M., Gruvaeus, I. & Wijkmark, L. 2008. Gammastrålningsmätning för detaljerad kartering av jordarter inom fält. *Precisionsodling Sverige, POS Teknisk rapport 11*. Institutionen för mark och miljö, Sveriges lantbruksuniversitet. 30 p
- Söderström, M. 2010. Interpolerade markkartor. *POS Teknisk rapport 21*. Institutionen för mark och miljö, Sveriges Lantbruksuniversitet. 20 s.
- Söderström, M. & Eriksson, J. 2012. Gamma-ray spectrometry and geological maps as tools for cadmium risk assessment in arable soils. Submitted to *Geoderma*.
- Söderström, M. & Stadig, H. 2012. Local and regional soil clay mapping using gamma ray spectrometry. Poster paper at the 11th International Conference on Precision Agriculture (ICPA), Indianapolis, USA. 15—18 July 2012. Will be published on CD.
- Tranter, G., Jarvis, N., Moeys, J. & Söderström, M. 2011. Broad-scale digital soil mapping with geographically disparate geophysical data: a Swedish example. *Geophysical Research Abstracts*, Vol. 13, EGU2011-7538, 2011, EGU General Assembly
- Wetterlind, J., Stenberg, B. & Söderström, M. 2008. The use of near infrared (NIR) spectroscopy to improve soil mapping at the farm scale. *Precision Agriculture*, 9, 57–69.
- Wilford, J. & Minty, B., 2007. The use of airborne gamma-ray imagery for mapping soils and understanding landscape processes, in: Lagacherie, P., McBratney, A.B., Voltz M. (Eds.), *Digital soil mapping – An introductory perspective*. *Developments in Soil Science*, Vol. 31. Elsevier, pp. 207-218.
- Wong, M.T.F. & Harper, R.J. 1999. Use of on-ground gamma-ray spectrometry to measure plant-available potassium and other topsoil attributes. *Australian Journal of Soil Research* 37, 267–277.

**Förteckning över rapporter utgivna av Avdelningen för precisionsodling i serien
*Precisionsodling Sverige, Tekniska rapporter:***

- 25 Piikki, K., Söderström, M., Stenberg, M. & Roland, J. 2012. Variation i marken inom fältförsök.
- 24 Lundström, C (red). 2012. Verksamhet i AGROVÄST-projektet Precisionsodling Sverige, POS, 2011.
- 23 Lundström, C (red). 2011. Verksamhet i AGROVÄST-projektet Precisionsodling Sverige, POS, 2010.
- 22 Lundström, C (red). 2010. Verksamhet i AGROVÄST-projektet Precisionsodling Sverige, POS, 2009.
- 21 Söderström, M. 2009. Interpolerade markkartor – några riktlinjer.
- 20 Söderström, M., Börjesson, T., Pettersson, C.G., Nissen, K. & Hagner, O. 2009. Prognoser för malkornskvalitet med fjärranalys.
- 19 Börjesson, T. & Söderström, M. 2009. Bedömning av kvalitetsskillnader över tid i vallar avsedda för hösilage med Yara N-sensor.
- 18 Lundström, C (red). 2009. Verksamhet i AGROVÄST-projektet Precisionsodling Sverige, POS, 2008.
- 17 Jacobsen, A. & Söderström, M. 2008. Regional analyse af samspillet mellem satellitdata og jordbundsvariation. Delrapport 2 i SLF-projektet (dnr SLF 297/02): "Kostnadseffektiv markkartering genom stratifierad datainsamling baserad på fjärranalys"
- 16 Jacobsen, A. & Söderström, M. 2008. Anvendelse af geostatistik og remote sensing data til kortlægning af jordens lerindhold.
- 15 Söderström, M. 2008. Den traditionella markkarteringens användbarhet för precisionsodling.
- 14 Lundström, C. (red); 2008. Verksamhet i AGROVÄST-projektet Precisionsodling Sverige, POS, 2007.
- 13 Börjesson, T., Lorén, N., Larsolle, A., Söderström, M., Nilsson, J. och Nissen, K. 2008. Bildanalys som redskap för platsspecifik ogräsbekämpning.
- 12 Söderström, M., 2008. PrecisionWizard 3 – hantera precisionsodlingsdata och gör egna styrfiler till Farm Site Mate och Yara N-Sensor.
- 11 Söderström, M., Gruvaeus, I. och Wijkmark, L., 2008. Gammastrålningsmätning för detaljerad kartering av jordarter inom fält.
- 10 Söderström, M., Wijkmark, L., Martinsson, J. och Nissen, K., 2008. Avstånd mellan körspår – en jämförelse mellan traditionell spårmarkör och autostyrning med GPS.
- 9 Delin, S.(red.), 2007. Verksamhet i AGROVÄST-projektet Precisionsodling Sverige, POS, 2006
- 8 Engström, L., Börjesson, T och Lindén, B. 2007. Beståndstäthet tidigt på våren i höstvete – samband med skörd, topografi, förrådskalium och biomassa (Yara N-sensor- och NIR-mätningar)
- 7 Söderström, M., och Nissen, K., 2006. Insamling av GIS-data och navigering med GPS.
- 6 Söderström, M., 2006. PrecisionWizard - Gör styrfiler till FarmSiteMate och Yara N-sensor.

- 5 Delin, S.(red.), 2006. Dokumentation från seminariet ”Precisionsodling - avstämning av verksamhet och vision hos olika aktörer”, Skara den 19 april 2006.
- 4 Delin, S.(red.), 2006. Verksamhetsberättelse för Precisionsodling Sverige, POS, 2005.
- 3 Delin, S. 2005. Verksamhetsberättelse för Precisionsodling Sverige (POS) 2003-2004.
- 2 Börjesson, T., Åstrand, B., Engström, L. och Lindén, B., 2005. Bildanalys för att beskriva beståndsstatus i höstraps och höstvetete och ogräsförekomst i vårsäd.
- 1 Nyberg, A., Börjesson, T. och Gustavsson, A-M., 2004. Bildanalys för bedömning av klöverandel i vallar – Utvärdering av TrefoilAnalysis.

Förteckning över rapporter utgivna av Institutionen för jordbruksvetenskap Skara i serien *Precisionsodling Sverige, Tekniska rapporter* (ISSN:1651-2804):

1. Börjesson, T, Ivarsson, K., Engquist, A., Wikström, L. 2002. Kvalitetsprognoser för brödvete och malkorn med reflektansmätning i växande gröda.
2. Börjesson, T., Nyberg, A., Stenberg, M. och Wetterlind, J. 2002. Handburen Hydro sensor i vall -prediktering av torrsubstansavkastning och kvalitetsegenskaper.
3. Söderström. M. (red.). 2003. Precisionsodling Sverige 2002, Verksamhetsberättelse från arbetsgrupperna.
4. Jonsson, A. och Söderström. M. 2003. Precisionsodling - vad är det?
5. Nyberg, A., Lindén, B., Wetterlind, J. och Börjesson, T. 2003. Precisionsodling av vall: Mätningar med en handburensensor i vallförsök med nötflytgödsel på Tubbetorp i Västergötland, 2002.
6. Nyberg, A., Stenberg, M., Börjesson, T. och Stenberg, B. 2003. Precisionsodling av vall: Mätningar i växande vall med ett bärbart NIR-instrument – en pilotstudie.

Förteckning över rapporter utgivna av Institutionen för jordbruksvetenskap Skara i serien *Precisionsodling i Väst, Tekniska rapporter*:

1. Rapport från en studieresa till norra Tyskland.
2. Thylén, L & Algerbo, P-A. Teknik för växtplatsanpassad odling.
3. Seminarium och utställning i Skara den 10 mars 1998.
4. Delin, S. 2000. Hantering av geografiska data inom ett jordbruksfält.
5. Lundström, C. Delin, S. och Nissen, K. 2000. Precisionsodling - teknik och möjligheter.

AGROVÄST-projektet *Precisionsodling Sverige* syftar till att utveckla och tillämpa användbara metoder inom precisionsodlingen till nytta för det praktiska jordbruket.

I projektet arbetas med precisionsodling i form av utvärdering och tolkning av samt teknik för markkartering, kalkning, gödsling, bestämning av mark- och grödegenskaper, växtskydd samt miljöeffekter av precisionsodling.

Projektet genomförs i ett samarbete mellan bl.a. Svenska Lantmännen, Sveriges Lantbruksuniversitet (SLU), Svalöf Weibull AB, Yara AB, hushållningssällskap, Greppa Näringsen och Institutet för jordbruks- och miljöteknik (JTI).

Distribution:

Sveriges lantbruksuniversitet
Institutionen för mark och miljö
Box 234
532 23 Skara
Tel. 0511-670 00

Internet: <http://www.slu.se/>
<http://www.agrovast.se/precision>
<http://www.precisionskolan.se>