

Billigare jordanalyser oavsett gårdsstorlek med en nationell NIR-databas

Syfte

Det övergripande målet med projektet var att ta fram en metod för bra bestämningar av inomfältvariationer i marken till ett rimligt pris oavsett gårdsstorlek. För att kunna göra detta skapades en stor nationell databas med över 12 000 jordbruksjordar analyserade med nära infraröd reflektans (NIR) - spektroskopi. Med en så stor och differentierad databas täcks de flesta typer av svenska jordbruksjordar in vilket skulle kunna göra det möjligt att ta fram mer generella kalibreringsmodeller för bestämning av t ex jordart och mullhalt, något som avsevärt ökar effektiviteten och minskar kostnaderna för analyserna.

Vidare syftade projektet till att utvärdera en metod där ett fåtal prov från den lokala gården läggs till den stora databasen för att bättre beskriva variationen inom en gård.

Billigare analyser med NIR-spektroskopi

För att högupplösta markkartor ska användas i praktiken får inte nyttan överskuggas av alltför höga provtagnings- och analyskostnader. Under flera år har därför mycket forskning ägnats åt sensorer och alternativa billigare strategier och analysmetoder, både i Sverige och internationellt, i många fall finansierad av SLF (t ex Piikki et al., 2015; Wetterlind et al., 2015). Hit hör nära infraröd reflektans (NIR) -spektroskopi, ofta inkluderande det synliga våglängdsområdet (vis-NIR), som i flera studier visat på stora möjligheter till billigare markkarteringsanalyser och tack vare nya algoritmer och kalibreringsstrategier kommer vi allt närmare en kommersiell användning (se omfattande sammanfattning av forskningen på området fram till 2010 av Stenberg et al. (2010). En fördel med NIR-spektroskopi är att den är bra på att bestämma både lerhalt och mullhalt.

Vis-NIR-spektroskopi kräver att man gör kalibreringar, d.v.s. relaterar NIR-mätningarna till den markparameter man är intresserad av. För detta krävs ett antal traditionellt analyserade prov till kalibreringen. En stor del av kostnadseffektiviseringen ligger i att minimera antalet kalibreringsprov som behövs vid en gårdskartering. Allra mest effektivt är det med en generell, nationell, kalibreringsmodell direkt kan användas på prov från den enskilda gården utan ytterligare analyser med den traditionella metoden. Flera länder har byggt upp nationella vis-NIR-databaser och det finns även större databaser som innefattar flera länder, t.ex. över Europa (Stevens et al., 2013). Det är viktigt att jordproven som ingår i databasen representerar variationen i de jordar som kan komma att analyseras med metoden. Den oerhörda variation som olika jordtyper uppvisar innebär att databasernas omfattning kan behöva vara mycket stor. Jordbruksmark i Sverige har emellertid en relativt homogen mineralogi vilket borde göra förutsättningarna för nationella kalibreringar relativt goda.

Problemet med de stora kalibreringarna är att de visat sig fungera dåligt för att beskriva variationer på inomfält- och gårdsnivå, där lokala gårds- eller fältkalibreringar givit betydligt bättre resultat (Wetterlind and Stenberg, 2010). Men med lokala kalibreringar följer också lokala kalibreringsprov analyserade med traditionella metoder vilket ökar kostnaderna. En metod som ofta går under beteckningen "spiking" i den internationella litteraturen innebär att man genom att lägga till prov från den lokala gården/området till den stora databasen får betydligt bättre resultat än med enbart den stora databasen och samma eller till och med bättre precision än med en lokal kalibrering (Brown, 2007; Gogé et al., 2013; Guerrero et al., 2010; Sankey et al., 2008; Wetterlind and Stenberg, 2010). Guerrero, Stenberg, Wetterlind et al., (2014) visade att man genom att ge de lokala proven extra viktning i förhållande till proven i den nationella databasen, kunde få mycket goda

bestämningar av organiskt kol med endast åtta lokala prov. Med så få lokala kalibreringsprov blir metoden intressanta även för mindre gårdar med färre hektar att slå ut kostnaderna på. Naturligtvis kan även större gårdar klara sig med ett mindre antal prov.

Material och metoder

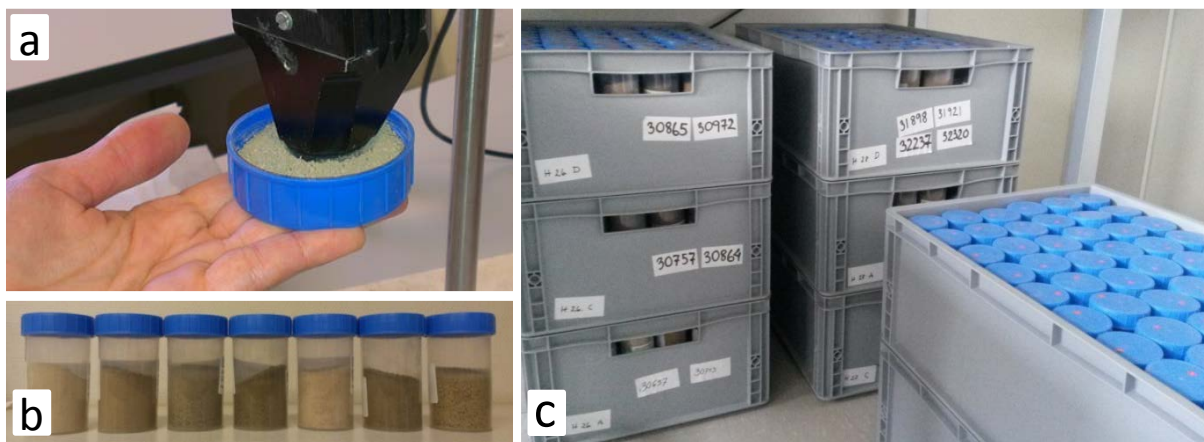
Under 2011-2012 finansierade jordbruksverket en omfattande jordprovtagning av svensk jordbruksmark där ca 12 500 matjordsprov (motsvarande 1 prov/km²) samlades in och analyserades bland annat på textur och mullhalt.

Vis-NIR-mätningar

I projektet analyserades 5 000 av de nationella jordproven med vis-NIR spektroskopi under 2014 (figur 1). 2015 fick vi ytterligare finansiering från jordbruksverket för att kunna analysera de resterande jordproven och det tota datasettet på ca 12 500 jordprov har använts i analyserna i den här rapporten.

Vis-NIR-mätningarna gjordes på torkade och sållade (<2 mm) prov. Mätningarna gjordes med ett FieldSpec Pro FR instrument (Analytical Spectral Devices Inc., Boulder, CO, USA, www.asdi.com) utrustat med en kontaktprob kopplat till instrumentet via en fiberoptisk kabel. Reflektansmätningarna kalibrerades mot en extern vit referens (Spektralon). Upplägget är en anpassning av det mätsystem som används i Australien. Ett tillägg är att det vid varje mätningstillfälle även gjordes mätningar på två standarder (två olika sandprov distribuerade bland internationella forskargrupper som jobbar med vis-NIR-mätningar på jord). De två standarderna är till för att underlätta användandet av data från olika instrument och/eller olika mätupplägg (Ben Dor et al. 2015). För att få så representativa mätningar som möjligt utan att tidsåtgången blev allt för stor gjordes två reflektansmätningar per prov där jorden blandades om mellan mätningarna (figur 1 a). Ett medeltal av de två mätningarna användes sedan i analyserna. Instrumentet registrerar reflektansen i våglängdsområdet 350-2500 nm med 1 nm intervall och 3-10 nm upplösning.

Vis-NIR mätningarna uttrycktes som absorbans genom $\log(1/\text{reflektans})$ och varje spektrum transformerades med förstaderivatan vilket förstärker svagare signaler och medför en indirekt baslinjekorrektion. För att reducera mängden brus som också förstärks filterades spektrumen med 11 punkters interpoleringsteknik (Savitzky & Golay, 1964) före analyserna. Våglängderna upp till 460 nm togs bort från vidare analyser på grund av allt för mycket brus. Vidare analyser i studien begränsades till de 11 425 jordar med mindre än 12 % mull.



Figur 1. Vis-NIR-mätning med kontaktprob a), jordprov från den nationella markkarteringen b och c).

Nationell kalibrering

Två tredjedelar, 7 619 prov, av de 11 425 jordproven med en mullhalt under 12 % valdes ut till kalibreringsprov med hjälp av Kennard-Stone algoritmen (Kennard and Stone, 1969). Kennard-Stone algoritmen väljer ut prov så att kalibreringsproven blir jämnt utspridda i den variation som finns i vis-NIR spektrumet. Resterande 3 806 prov användes till validering och ingick inte på något sätt i kalibreringen av modellerna (tabell 1). Genom att använda den multivariata kalibreringstekniken PLS (Partial Least Squares) gjordes globala nationella kalibreringsmodeller med vis-NIR-spektrum för bestämning av mull- och lerhalt (N_{PLS}). En så kallad "memory based learning" teknik, ibland kallad "LOCAL", användes också (N_{MBL}). Det är ingen global modell utan de kalibreringsprov som mest liknar de enskilda valideringsproven i vis-NIR-spektrum identifieras och "lokala" PLS-kalibreringsmodeller tas fram för varje valideringsprov. Det vill säga, alla valideringsprov bestäms med olika kalibreringsmodeller, men utifrån samma principer och baserat på samma databas. Kalibreringarna för ler respektive mull validerades med valideringsproven.

Tabell 1. Mull- och lerhaltsvariationen i de nationella kalibrerings- och valideringsproven.

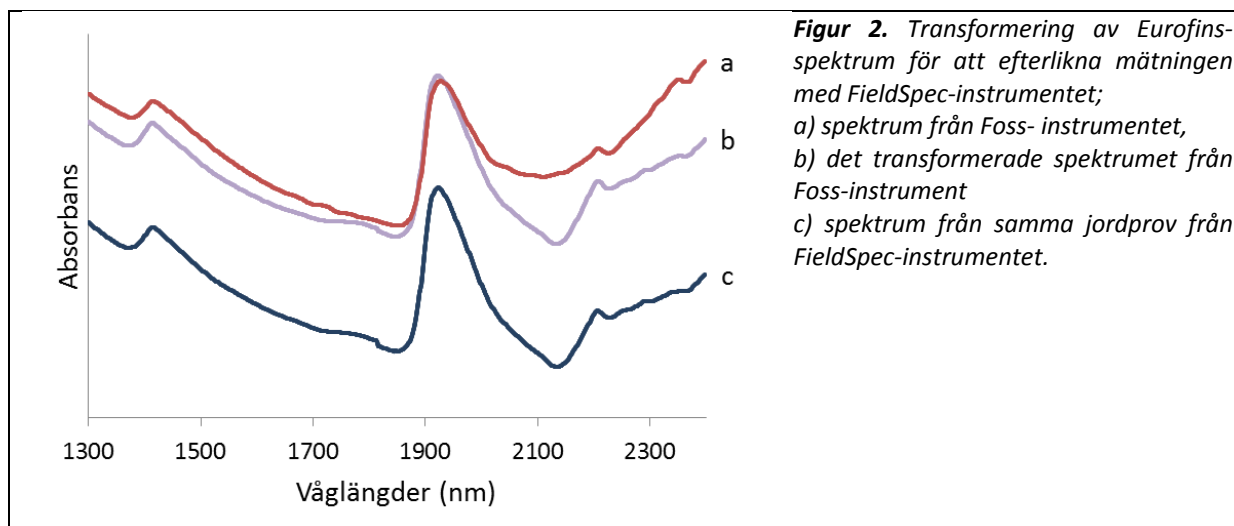
	NIR-instrument	antal	Mullhalt (min - max, %)	Lerhalt (min - max, %)
Kalibreringsprov	ASD	7 619	<0,1 - 11,9	<2 - 74
Valideringsprov	ASD	3 806	0,7 - 11,7	<2 - 67

Utvärdering av modellerna på gårdsnivå

Eftersom jordartsbestämningarna är tänkta att användas på gårdsnivå validerades de nationella modellerna N_{PLS} och N_{MBL} också med jordprov från 11 gårdar (tabell 2). Fyra av gårdarna (Hacksta, Kärrtorp, Bränneberg och Sjöstorp) ingick även i det jordbruksverks-finansierade projektet. NIR-mätningarna på de gårdarna gjordes med samma instrument som de nationella proven och i analyserna med de gårdarna användes hela våglängdsområdet (460 - 2500 nm). Jordproven från de övriga sju gårdarna ingick i ett tidigare SLF-finansierat projekt (Validering och implementering av NIR-teknik i jordlaboratoriers analysrutiner för billigare markkartering; H1033307) och analyserades med NIR-spektroskopi på Eurofins labb i Kristianstad med ett FOSS NIRSystems 5000 mellan 1300 och 2400 nm med en analysintervall av 2 nm och en upplösning av 5 nm. Våglängdsområdet som användes i analyserna med de gårdarna begränsades därför till 1300 - 2400 även för de nationella proven. Men det är inte bara våglängdsområdet som varierar mellan instrumenten, olika instrument kan ge upphov till andra systematiska skillnader i spektrumet också. Därför transformerades spektrumet på de sju gårdarna för att efterlikna de i det nationella datasettet genom att subtrahera med den genomsnittliga skillnaden mellan 17 jordprov som mätts med båda instrumenten (figur 2).

Tabell 2. Gårdarna som användes för validering av NIR-modellerna.

Gård	NIR-instrument	Län	Areal (ha)	Antal valideringsprov	Mullhalt (min - max, %)	Lerhalt (min - max, %)
Bränneberg	FieldSpec	Västra Götaland	69	58	3,2 - 4,8	37 - 58
Hacksta	FieldSpec	Uppsala län	97	110	2,3 - 7,7	25 - 66
Kärrtorp	FieldSpec	Västra Götaland	62	71	5,4 - 10,8	11 - 34
Sjöstorp	FieldSpec	Skåne län	78	98	2,0 - 7,4	12 - 52
Askersund	FOSS	Örebro län	165	167	1,7 - 8,3	1 - 26
Eskilstuna	FOSS	Södermanlands län	110	109	2,1 - 5,8	41 - 70
Linköping	FOSS	Östergötlands län	135	135	2,4 - 8,3	16 - 69
Mariestad	FOSS	Västra Götalands län	42	41	1,4 - 8,7	11 - 64
Storfors	FOSS	Värmlands län	148	148	2,7 - 9,2	4 - 60
Sjöbo	FOSS	Skåne län	130	128	2,8 - 9,2	12 - 47
Västerås	FOSS	Västmanlands län	125	125	2,1 - 8,0	27 - 56



Förutom att använda de nationella modellerna som de var "spetsades" också den globala nationella PLS-modellen med 10, 20 eller 40 lokala prov från den aktuella gården (S10, S20 och S40). För att kompensera för att de lokala proven var så få i förhållande till det nationella datasettet gavs de ökad vikt; 100 ggr i S10, 50 ggr i S20 och 25 ggr i S40.

De nationella modellerna (N_{PLS} och N_{MBL}) och de spetsade nationella modellerna (S10, S20 och S40) jämfördes även med kalibreringsmodeller gjorda på enbart prov från gårdarna med 10, 20 och 40 prov (L10, L20 och L40). S40 och L40 kunde bara göras på de sju gårdarna som ingått i det tidigare SLF-projektet och S20 och L20 var mellan 20 och 25 prov på de fyra gårdarna mätta med Field Spec-instrumentet.

Kalibreringsmodellerna utvärderades med r^2 -värdet mellan beräknad och uppmät mull- och lerhalt som anger kalibreringens precision och feltermen RMSEP (root mean squared error of prediction) som återger kalibreringens träffsäkerhet och har samma enhet som den beroende variabeln (d v s mull- och lerhalt). Analyser och statistik gjordes i the Unscrambler X 10.3 (CAMO PROCESS AS, Oslo, Norway) och i R (R Development Core Team, 2012).

Resultat och diskussion

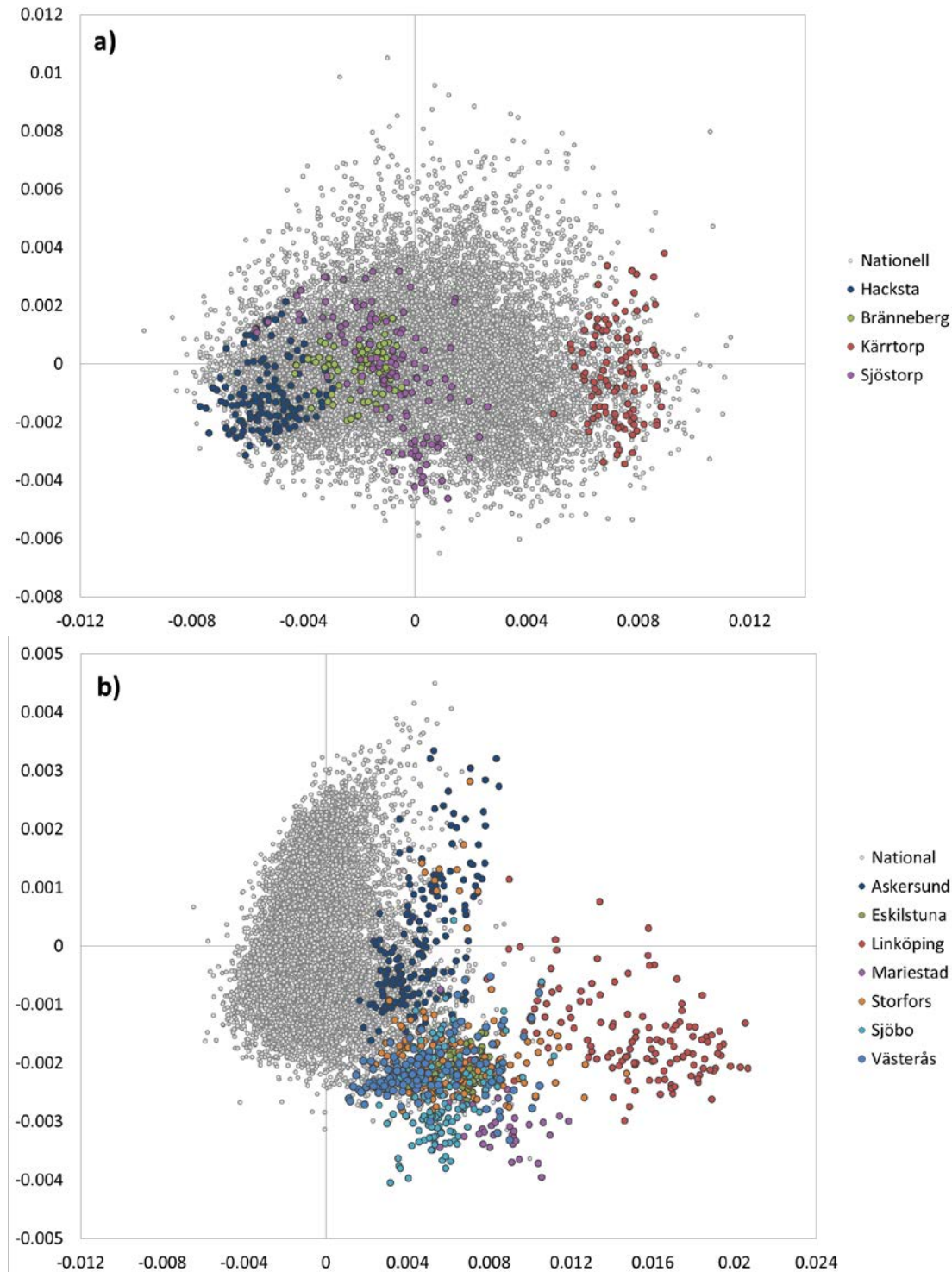
Variationen bland spektrumen av de nationella jordproven kan visas i en principalkomponentanalys (PCA). Figur 3 visar spridningen i de första två principalkomponenterna för de 11 425 jordarna med en mullhalt under 12 %. Även spektrumen från gårdarna är passivt projicerade i datarymden. De fyra gårdarna mätta med FieldSpec-instrument rymdes inom variationen i de nationella jordproven (figur 3a). Men det är intressant att se att även några fält kan uppvisa stor variation i förhållande till den totala variationen bland de nationella jordproven. De sju gårdarna mätta med ett annat NIR-instrument visar trots transformeringen en viss systematisk skillnad jämfört med de nationella proven (Figur 3b), speciellt i den första komponenten (x-axeln).

Nationell kalibrering

Tabell 3 visar resultaten av de nationella mull- och lerhaltsmodellerna (N_{PLS} och N_{MBL}) validerade med de nationella valideringsproven. Lerhalten gick bäst att bestämma med vis-NIR än mullhalten och både ler- och mullhaltsbestämningarna förbättrades med MBL-modellen jämfört med PLS-modellen. I MBL-modellen görs olika kalibreringsmodeller för varje enskilt valideringsprov där prov som är väldigt olika det aktuella valideringsprovet (och ofta tillför en onödigt stor variation till modellen) inte ingår i modellen för det provet. Fram för allt för mullhaltsbestämningen tycks detta ha stor betydelse och stora mullhaltmodeller har visat sig svåra även internationellt (Stenberg et al., 2010).

Tabell 3. De nationella mull- och lerhaltmodellerna validerade med nationella valideringsprov.

	N_{PLS}		N_{MBL}	
	r^2	RMSEP	r^2	RMSEP
Lerhalt (%)	0,90	4,5	0,95	3,5
Mullhalt (%)	0,51	1,13	0,75	0,80

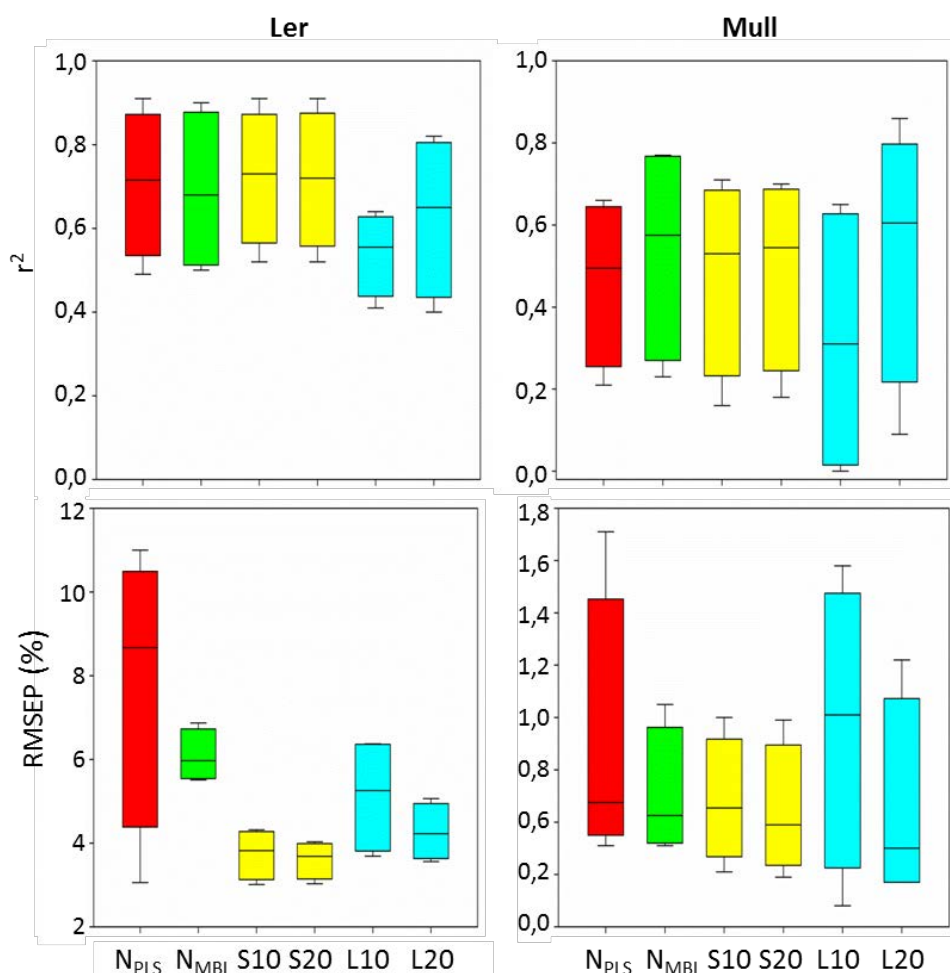


Figur 3. Plottar av de två första principalkomponenterna från PCA-analyser där de enskilda gårdarna är projicerade på de nationella jordproven. a) de nationella proven och gårdarna mätta med samma instrument (460-2500 nm våglängdsområde) och b) gårdarna mätta på Eurofins med spektrumen korrigerade för att likna de nationella (1310 – 2388 nm våglängdsområde).

Utvärdering av modellerna på gårdsnivå

Figur 4 och 5 visar sammanställningar av resultaten för valideringarna på gårdsnivå för de fyra gårdarna där NIR-mätningarna gjorts med samma instrument som de nationella proven (figur 4) och de sju gårdarna där NIR-mätningarna gjorts med ett annat instrument (figur 5). Att använda de nationella modellerna för att bestämma inomfältvariationen i ler- och mullhalt på de olika gårdarna resulterade oftast i stora systematiska fel. Generellt fungerade de nationella modellerna sämre på de gårdar där NIR-mätningarna gjorts med ett annat instrument. Detta kan förklaras med att spektrumerna från det andra instrumentet, trots transformeringen, visade på systematiska skillnader mot de nationella spektrumerna (figur 3b).

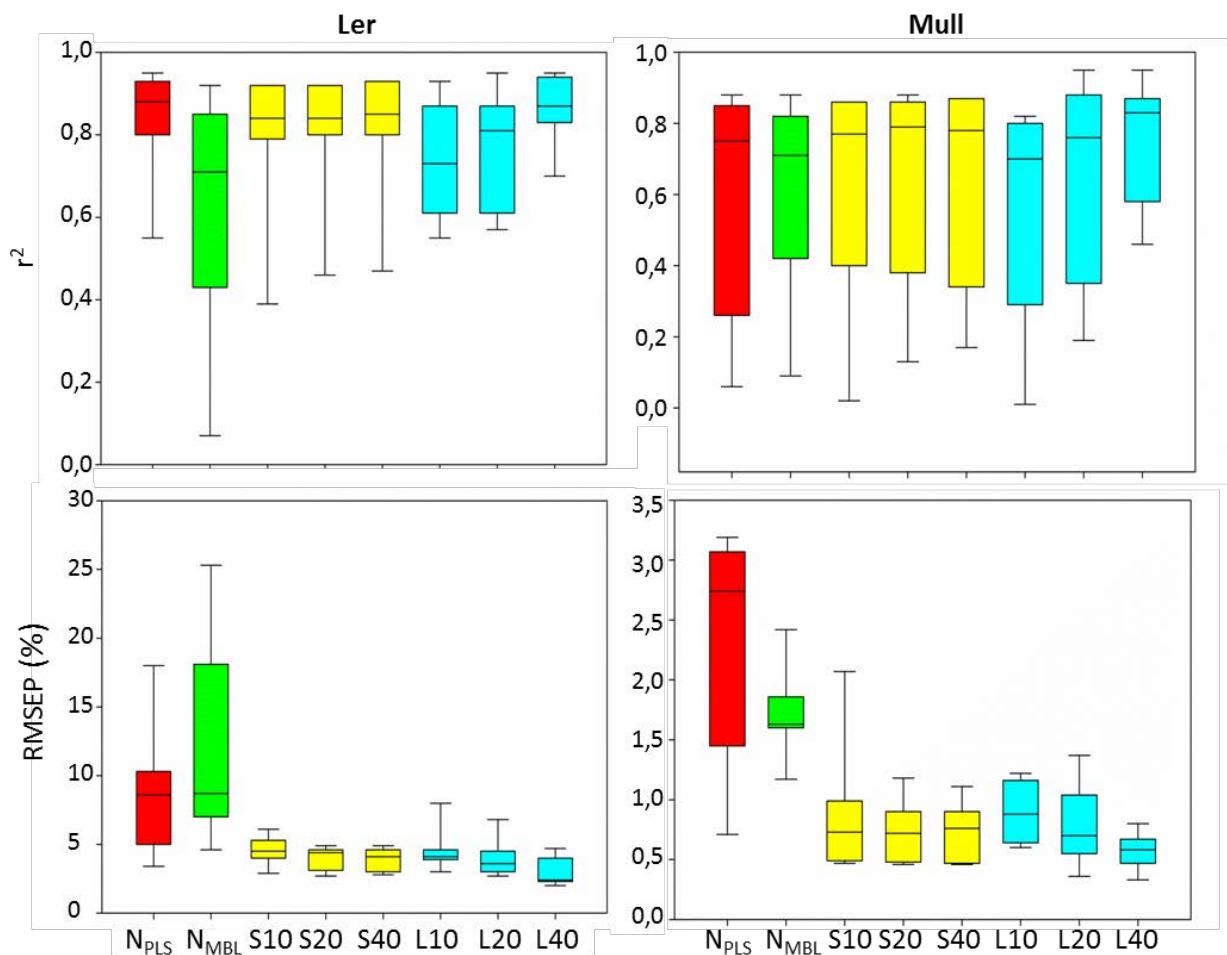
N_{MBL} -modellerna resulterade i bättre bestämningar av både ler- och mullhalterna på gårdarna jämfört med N_{PLS} , med undantag för lerhaltsbestämningar på de sju gårdarna med NIR-mätningar från FOSS. Att lägga till lokala prov från gården till de nationella proven (S10, S20 och S40) resulterade i en kraftig minskning av felen. Den kraftiga minskningen beror framförallt på en



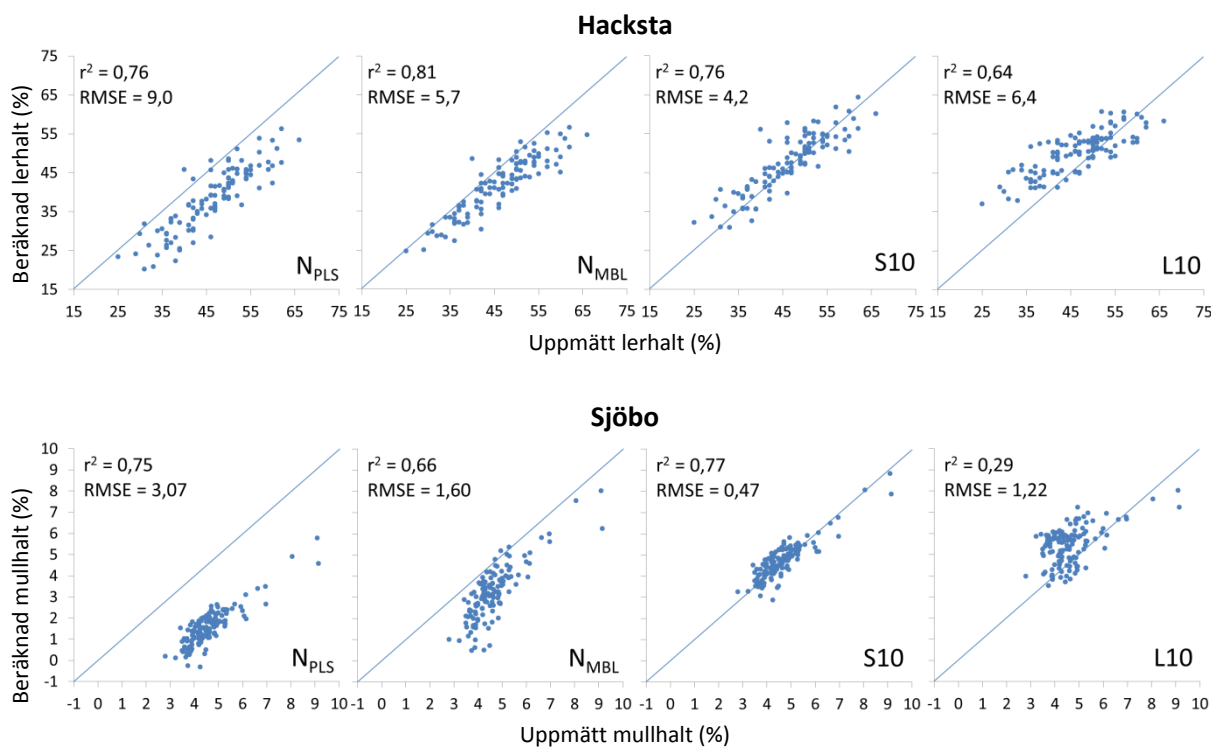
Figur 4. Sammanställning av valideringsresultaten för ler- och mullhalt bestämda med nationell modell (N_{PLS}), memory based learning-modell (N_{MBL}), nationell modell spikad med 10 och 20 lokala prov (S10 och S20) och modeller med enbart 10 och 20 lokala prov (L10 och L20) på de fyra gårdarna där NIR-mätningarna gjorts med samma instrument som de nationella proven. Boxarna representerar 25-75 percentilerna och strecket anger medianen.

minskning av det systematiska felet vilket exemplifieras med två gårdar i figur 6. Att det rör sig om systematiska fel, bias, syns också på r^2 -värdena i figur 4 och 5 där r^2 -värdena i stort sett är lika bra i de nationella modellerna som för de "spetsade" modellerna. Det vill säga, de nationella modellerna klara av att bestämma de relativa skillnaderna i lerhalt och mullhalt inom gårdarna, men hamnar fel i absoluta tal, dvs precisionen är hög, men inte träffsäkerheten.

Lokala modeller med enbart 10, 20 eller 40 prov från den enskilda gården resulterade också i små fel jämfört med de nationella modellerna (figur 4 och 5). På gårdar med NIR-mätningar från FiledSpec-instrumentet var både ler- och mullhaltsbestämningarna bäste med de spetsade modellerna. Modeller baserade på väldigt få prov, vilket är fallet med 10 och även 20 prov, är ofta mer osäkra, vilket avspeglas i de lägre r^2 -värdena och den stora spridningen i RMSEP mellan gårdarna (figur 4). På de gårdar där NIR-mätningarna gjorts med FOSS-instrument var de lokala modellerna istället lika bra eller bättre än de spetsade modellerna sett till RMSEP. Men ser man till r^2 -värdena var bilden den samma som för de 4 gårdarna med NIR-mätningar från FieldSpec-instrumentet, d.v.s. lägre för L10 och L20 än för de nationella och spetsade modellerna (Figur 5).



Figur 5. Sammanställning av valideringsresultaten för ler- och mullhalt bestämda med nationell modell (N_{PLS}), memory based learning-modell (N_{MBL}), nationell modell spikad med 10, 20 och 40 lokala prov (S_{10} , S_{20} och S_{40}) och modeller med enbart 10, 20 och 40 lokala prov (L_{10} , L_{20} och L_{40}) på de sju gårdarna där NIR-mätningarna gjorts med FOSS-instrumentet. Boxarna representerar Boxarna representerar 25-75 percentilerna och strecket anger medianen.



Figur 6. NIR-bestämd ler- och mullhalt mot uppmätt ler- och mullhalt för nationell PLS-modell (N_{PLS}), nationell memory based lerner-modell (N_{MBL}), nationell modell spikad med 10 lokala prov (S10) och enbart 10 loka prov (L10) på två gårdar. En gård där NIR-är mätt med samma instrument som de nationella proven (lerhalt på Hacksta) och en gård där NIR-mätningarna gjordes på FOSS-instrumentet (mullhalt på Sjöbo).

Bäst ler- och mullhaltsbestämningar med FOSS-instrumentet gav L40 både sett till RMSEP och r^2 värdena. Men det finns en tendens, fram för allt för mull, att det är möjligt att få likvärdiga resultat med färre lokala prov med spetsade modeller jämfört med lokala även när NIR-mätningarna är gjorda med olika instrument. Strategin att spetsa den nationella databasen med lokala prov klarade alltså av att hantera spektrum från olika instrument, vilket N_{MBL} -modellen däremot inte gjorde. En förklaring till det är att N_{MBL} -modellen enbart har spektrumen i den nationella databasen att välja mellan. Finns det inga spektrum som är tillräckligt lika gårdens spektrum, t.ex. på grund av att mätningarna på gården är gjorda med ett annat instrument, blir inte modellen tillräckligt bra. Det samma gäller om gårdens jordtyp är dåligt representerad i den nationella databasen.

Slutsatser

Nationella modeller, oavsett kalibreringsmetod (N_{PLS} eller N_{MBL}) resulterade i ler och mullhaltsbestämningar med stora systematiska fel när modellerna validerades på gårdsnivå.

När NIR-mätningarna på gården gjorts med ett annat instrument än det som användes för de nationella jordproven blev felen ännu större. Något som kan förklaras av de systematiska skillnader mellan spektrumen mätta med olika instrument som fanns kvar även efter en matematisk korrektion av gårdsspektrumen.

Genom att spetsa den nationella databasen med så få som 10 lokala prov från gården minskades det systematiska felet i både ler- och mullhaltsbestämningarna kraftigt vilket resulterade i nästan lika bra bestämningar som de gjorda med en lokal gårdsmodell baserade på 40 prov.

Att göra traditionella mull- och lerhaltsanalyser på bara 10 lokala jordprov på den enskilda gården och sedan endast göra NIR mätningar på resterande jordprov skulle innebära betydande kostnadsbesparingar och göra tekniken attraktiv även för gårdar med mindre areal.

Vidare erbjuder den nationella vis-NIR-databasen tillsammans med de magasinerade jordproven fantastiska möjligheter till vidare studier, t ex kopplat till kvalitén på det organiska material och mineralogin i svensk jordbruksmark.

Referenser

- Ben Dor, E. , Ong, C., Lau, I. C. 2015. Reflectance measurements of soils in the laboratory: Brook Standards and protocols. *Geoderma*, 245, 112-124.
- Brown, D.J. 2007. Using a global VNIR soil-spectral library for local soil characterization and landscape modeling in a 2nd-order Uganda watershed. *Geoderma* 140(4), 444-453.
- Gogé, F., Gomez, C., Jolivet, C., Joffre, R. 2013. Which strategy is best to predict soil properties of a local site from a national Vis-NIR database? *Geoderma* 213(2014), 1-9.
- Guerrero, C., Zornoza, R., Gomez, I., Mataix-Beneyto, J. 2010. Spiking of NIR regional models using samples from target sites: Effect of model size on prediction accuracy. *Geoderma* 158(1-2), 66-77.
- Guerrero, C., Stenberg, B., Wetterlind, J., Rossel, R.A.V., Maestre, F.T., Mouazen, A.M., Zornoza, R., Ruiz-Sinoga, J.D., Kuang, B. 2014. Assessment of soil organic carbon at local scale with spiked NIR calibrations: effects of selection and extra-weighting on the spiking subset. *European Journal of Soil Science* 65(2), 248-263.
- Kennard, R.W., Stone, L.A. 1969. Computer aided design of experiments. *Technometrics* 11, 137-148.
- Piikki, K., Wetterlind, J., Söderström, M., Stenberg, B. 2015. Three-dimensional digital soil mapping of agricultural fields by integration of multiple proximal sensor data obtained from different sensing methods. *Precision Agriculture*, 16(1), 29-45
- R Development Core Team 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. ISBN 3-900051-07-0 [WWW document]. URL <http://www.R-project.org/>.
- Sankey, J.B., Brown, D.J., Bernard, M.L., Lawrence, R.L. 2008. Comparing local vs. global visible and near-infrared (VisNIR) diffuse reflectance spectroscopy (DRS) calibrations for the prediction of soil clay, organic C and inorganic C. *Geoderma* 148(2), 149-158.
- Savitzky, A., Golay, M., 1964. Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry* 36, 1627-1639.
- Stenberg, B., Rossel, R.A.V., Mouazen, A.M., Wetterlind, J. 2010. Visible and near infrared spectroscopy in soil science. In: D.L. Sparks (Ed.), *Advances in Agronomy*, Vol 107. *Advances in Agronomy*, pp. 163-215.
- Stevens, A., Nocita, M., Toth, G., Montanarella, L., van Wesemael, B. 2013. Prediction of Soil Organic Carbon at the European Scale by Visible and Near InfraRed Reflectance Spectroscopy. *Plos One* 8(6).
- Wetterlind, J. Piikki, K., Söderström, M., Stenberg, B. 2015. Exploring the predictability of soil texture and organic matter content with a commercial integrated soil profiling tool. *European Journal of Soil Science*, 66(4), 631-638
- Wetterlind, J., Stenberg, B., 2010. Near-infrared spectroscopy for within-field soil characterization: small local calibrations compared with national libraries spiked with local samples. *European Journal of Soil Science* 61(6), 823-843.