Validation of precision agriculture soil mapping services under practical conditions

C. Kempenaar[1, 2, 3], F. Tigchelhoff [2], J.A. Booij[1], S. Nysten[2] & C.G. Kocks[2]

[1] Wageningen University & Research, Droevendaalsesteeg 1, 6708 PB Wageningen, NL
[2] Aeres University of Applied Sciences, De Drieslag 4, 8251 JZ Dronten, NL
[3] Corresponding author: corne.kempenaar@wur.nl

Abstract

This study confirmed that soil mapping services can produce good quality soil maps showing within-field spatial variation of clay and soil organic matter contents. The results were better when there was a wider range of variation of the parameter in the field. In those cases, correlation coefficients in regression analyses between soil analysis laboratory reference data and soil mapping service data were high (R^2 between 0.7 to 0.8). A proximal sensor systems based on gamma radiation and on EC plus NIRS provided the best soil maps on spatial variation of clay and soil organic matter content. The remote sensing sensor system applying hyperspectral imaging scored a bit less good than the two proximal sensor systems mentioned, yet still useful in precision farming.

Key words

Soil mapping, digital farming, sensor, smart farming

Introduction

In the last 15 years, several soil mapping sensor systems and services have been developed for precision agriculture. Soil maps are used today for management zoning and variable rate applications (Kempenaar et al., 2018). The accuracy of the maps is important, knowing that they are a basis for decision making in precision agriculture/smart farming. Farmers of the National Field Lab for Precision Agriculture project (NPPL) asked for a comparison of soil mapping services under practical conditions. Such research was not done in The Netherlands yet, nor did we find it in literature.

Therefore, a comparative study on the accuracy of soil maps and services marketed in Dutch arable farming was carried out. We focussed on soil mapping services that deliver maps showing spatial variation in clay and soil organic matter content within fields. The companies involved were asked to provide digital soil maps of four selected arable fields. Data points from these maps were compared with reference data. In total, five types of sensor systems were studied: (1a and 1b) Electric conductivity sensing (EC and EMI; Lund, 2008), (2) Passive gamma radiation sensing (van Egmond et al. 2018), (3) Near infrared sensing (NIRS; Knadel, 2015), and (4) remote multispectral imaging (Yuzugullu, 2020).

The objective of this study was to evaluate how well data from soil maps of different commercial soil mapping services correlate with data from soil analyses in laboratories.

Materials and Methods

Study fields and soil characterization

Soil sampling and sensor measurements were done on four arable fields in The Netherlands. Field size was ca. 5 ha. Tables 1-4 show field average and statistical parameters on clay and soil organic matter (SOM) content of the soils. More than 50 top soil samples were randomly taken on each field with a soil profile sampler (diameter 3 cm, depth of sampling ca. 20 cm). Coordinates of each sampling point was determined with RTK GPS. The soil samples were individually stored in plastic bags and sent to Eurofins soil analysis laboratory in Wageningen

(https://www.eurofins.nl/en/environment/services/soil/soil-analyses/). Eurofins delivered for each soil sample an analysis report on clay and SOM content as determined with NIRS in their 'dry lab' facility.
In this way, each field had a data set of 50 reference data points for comparison with data of the soil mapping services. The 2019 study field was split in to North and South parts, and analyses were also done per sub-field. For more details on the fields, see study reports by Nysten et al. (2019) and Tigchelhoff et al. (2020 and 2021). Figure 1 gives an impression of one of the fields with a map of the clay content based on Inverse Distance Weighting (IDW) of 50 reference data points (Table 2 shows the corresponding averages and other statistics).

Table 1. Arable field in Ens, NL, Lat. – Long. 5.84 - 52.66, light clay soil, summer 2018.

Parameter	Clay%	SOM%
Average	9.5	2.5
Minimum	6	2.1
Maximum	11	3.6
RMSE (average)	1.1	0.3

Table 2. Arable field in Slootdorp, NL, L. – L. 5.03 - 52.92, marine clay soil, summer 2019.

Parameter	Clay% North	Clay% South	SOM% North	SOM% South
Average	13.9	27.2	2.1	4.3
Minimum	4	14	1.3	3.5
Maximum	20	36	3.6	5.5
RMSE (average)	9.0	7.0	1.1	1.3

Table 3. Arable field in Valthermond, NL, L. – L. 6.94 - 52.87, peat soil, spring 2020.

Parameter	Clay %	SOM %
Average	9.5	2.5
Minimum	6	2.1
Maximum	11	3.6
RMSE (average)	1.1	0.3

Table 4. Arable field in Lemelerveld, NL, L. – L. 6.38 - 52.45, sandy soil, spring 2020.

Parameter	Clay %	SOM %
Average	9.5	2.5
Minimum	6	2.1
Maximum	11	3.6
RMSE (average)	1.1	0.3

Soil mapping systems/services

Six soil mapping service companies participated in the research. They mapped the arable fields as mentioned in the previous section according to their standard practice (SOP). The mapping services differed on several aspects. In Table 5, the mapping services studied, are summarized. The service of LoonwerkGPS was transferred to Soil Masters in 2019 (1a became 1b).The services applied different sensors: electrical conductivity measurement (EC or EMI), near infrared sensor (NIRS), passive gamma radiation sensor or multispectral cameras. The sensors were mainly used as proximal sensors, either mounted just above the soil or in direct contact with the soil. The proximal sensor systems made line scans of the

field, with every 2 to 5 m a measurement and lines ca. 6 m apart. The remote sensor systems applied grid imaging (resolution ca. 10 by 10 m). Processing of the sensor data was also done according to SOP of the soil mapping service companies. Data of five soil samples per field were provided to the companies so they could do calibration at least partly on the same data. It was up to them if they used the samples.
Each company provided for the analyses a digital map of each field containing spatial data on SOM and/or clay content within 2 weeks after scanning the fields. Not all companies could deliver clay and SOM content matter maps. Some provided other soil parameters (as well), but these were not evaluated in this paper (see study reports by Nysten et al. (2019) and Tigchelhoff et al (2020, 2021). More details on the methodology of the companies is on their websites (see reference list, also accuracy of GPS applied when scanning).

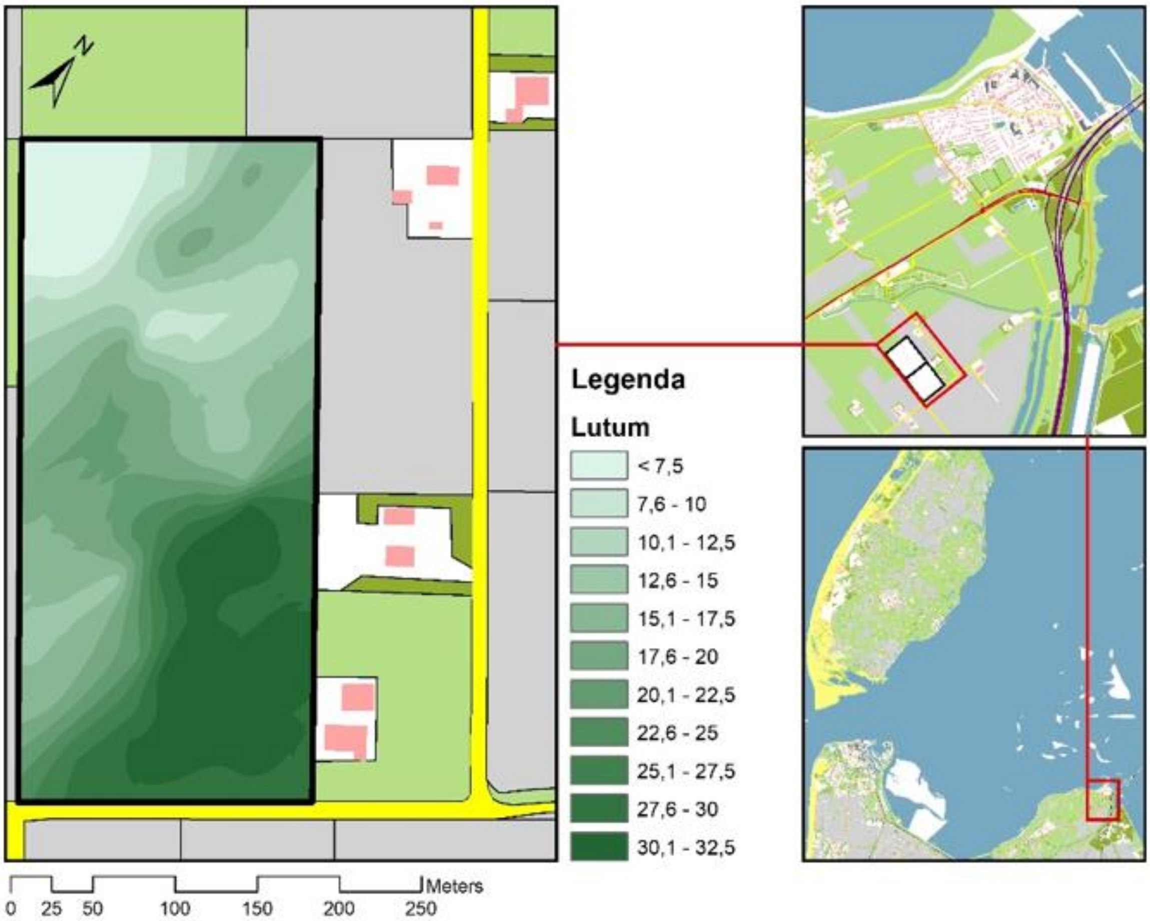

Figure 1. Clay (Lutum, %) content map (left) based on 50 soil samples analyses and IDW plus topographical information (right).

Data analysis
The data were organized in a way that reference data points could be compared with nearest data points provided by the soil mapping service companies on the basis of given latitude longitude coordinates. Microsoft Excel software was used to do so and to do statistical analyses. Linear regression analysis tool and RMSE function were applied to study correlation between reference data and mapping service data, and to determine the error between reference and mapping service data. Significant correlation was concluded in a regression analysis only if the statistical parameter $P<0.05$. Furthermore, the correlation coefficient R^2 was calculated with linear regression model of Excel. The higher this coefficient, the better the correlation. The root mean sum error (RMSE) was calculated for the

mapping service data relative to the reference data in order to determine the error between reference data and mapping service data. The smaller the RMSE, the smaller the error. And if RMSE was smaller than the standard error SE of the reference data, the map of the soil mapping service was better than the IDW map of the reference data only.

Table 5. Summary of soil mapping services tested on four arable fields in 2018-2020.

No.	Company	Sensor	Model	Orientation	Data in years 20xx	
					Clay%	OM%
1a	LoonwerkGPS	EMI	Dualem 21S	Proximal	18, 19,	18, 19
1b	Soil Masters	EMI	Dualem 21S	Proximal	19, 20	19, 20
2	Loonstra & vdW	Gamma radiation	De Mol	Proximal	18, 19, 20	18, 19, 20
3	Vantage_Agrom.	EC and NIRS	Veris MSP3	Proximal	18, 19, 20	18, 19, 20
4	CNH Agxtend	EMI	SoilXplorer	Proximal	18, 19, 20	18, 19, 20
5	Bioscope	Multispec. cam	e.g. Sentinel	Remote	19, 20	19, 20

Results

In the study reports, all individual regression analyses are shown (Nysten et al, 2019, Tigchelhoff et al., 2020, 2021). In Figure 2, as examples, regression analysis plots for clay and the 2019 North field (Table 7b) are shown. Data points of reference data are plotted against soil mapping service data of the three systems tested on that field. All relationships tested significant (P<0.05) with moderate to high correlation coefficients.

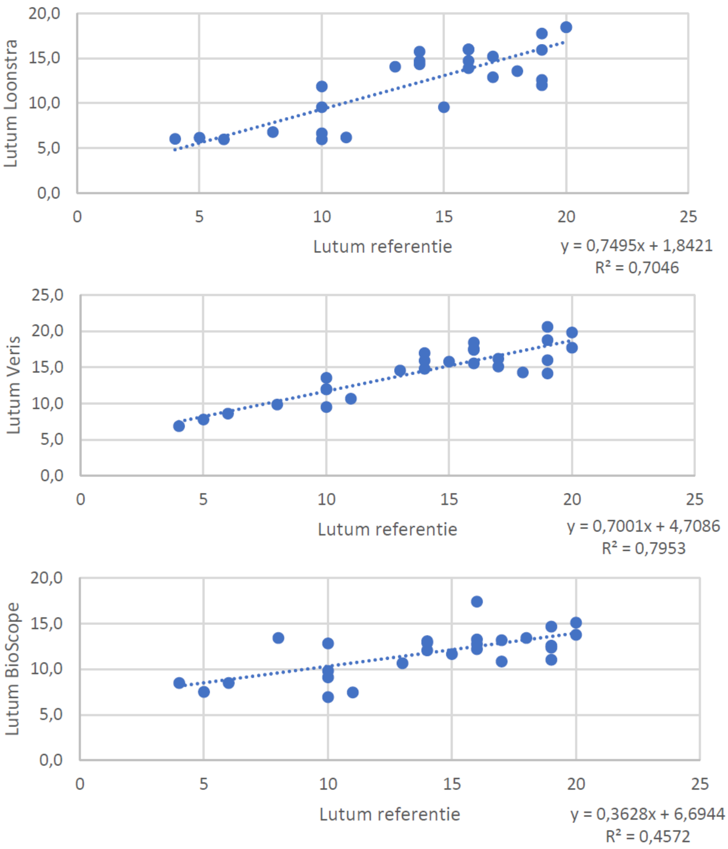

Figure 2. Regression analysis plots showing relation between clay (Lutum) reference (*referentie*) data and soil mapping service data (top is gamma radiation (2), middle is Veris (3) bottom is Bioscope (5).

2018 - Light marine clay soil

The RMSE data in Table 6 show that the soil mapping service data did not differ much from the reference data. The absolute errors of the reference data were ca. 1% for clay and 0.4% for SOM. The RMSE of the reference data were 1.1% and 0.3%, resp. The average clay and SOM content of the field were 9.5 and 2.5%, resp. The RMSE for Veris (3) was relatively high (2.34%).This might be related to a difference in additional soil sampling applied by the company in 2018 compared to the way the soil sampling was done.

In the 2018 experiment, correlations were low between reference data and soil mapping service data (Table 6). And only in three of six cases, the regression analysis tested significant. In those cases, correlation coefficients were between 0.12 and 0.36. Regression parameters were better for clay data than for SOM data. This may be explained by the fact that on this field, the range over which the content varied was larger for clay than for SOM. The Veris sensor systems had the best statistical evaluation. For clay, the gamma radiation sensor system scored in the same order.

Table 6. Statistical parameters describing the correlation and error between data provided by different soil mapping services and reference data (2018, clay soil).

No.	Sensor	Clay data			SOM data		
		R^2	P-value	RMSE	R^2	P-value	RMSE
1a	EMI	0.03	>0.05	1.29	0.00	>0.05	0.38
2	Gamma radiat.	0.34	<0.05	1.00	0.02	>0.05	0.41
3	EC and NIRS	0.36	<0.05	2.34	0.12	< 0.05	0.44

1a = LoonwerkGPS, 2 = Loonstra&vanderWeide, 3 = Veris

2019 - Marine clay soil

Data were analysed at whole field (Table 7a) and sub-field levels (Table 7b, North part of the field, and 7c South part of the field).

The RMSE data in Table 7a show that the soil mapping service data differed from the reference data ca. 3% for clay and 0.4% for SOM. For clay, the error was higher than in the 2018 experiment (3% vs. 1%). In the 2019 experiment, the variation in the clay content was also higher than in the 2018 experiment. The RMSE of the reference data for clay was 3.2% and for OM 1.2%.

Relatively high correlations between reference data and soil mapping service data were observed (Tables 7a, 7b and 7c). In all fifteen cases at whole field level and at sub-field levels, the regression analyses tested significant. At the whole field level, the correlation coefficients were ca. 0.8 for clay and a little lower for SOM. The Veris sensor systems had the best statistical evaluation. For clay, the gamma radiation sensor system scored in the same order. The Bioscope systems scored quite good on mapping the clay variation of the field. Tables 7b and 7c show that if the range in observed data is smaller, the correlations become weaker. Still, correlations were still good at the sub-field level.

Table 7a. Statistical parameters describing correlation and error between sensor data provided by different soil mapping services and the reference data 2019 South clay soil.

No.	Sensor	Clay data			SOM data		
		R^2	P-value	RMSE	R^2	P-value	RMSE
2	Gamma radiat.	0.88	<0.05	3.32	0.52	<0.05	0.87
3	EC and NIRS	0.85	<0.05	3.29	0.83	<0.05	0.51
5	Multispec. cam	0.78	<0.05	3.89			

2 = Loonstra&vanderWeide, 3 = Veris, 5 = Bioscope

Table 7b. Statistical parameters describing correlation and error between sensor data provided by different soil mapping services and the reference data 2019 North clay soil.

No.	Sensor	Clay data			SOM data		
		R^2	P-value	RMSE	R^2	P-value	RMSE
2	Gamma radiat.	0.70	< 0.05	3.02	0.24	<0.05	0.57
3	EC and NIRS	0.80	< 0.05	2.22	0.23	<0.05	0.52
5	Multispec. cam	0.46	<0.05	3.77			

2 = Loonstra&vanderWeide, 3 = Veris, 5 = Bioscope

Table 7c. Statistical parameters describing correlation and error between sensor data provided by different soil mapping services and the reference data 2019 South clay soil.

No.	Sensor	Clay data			SOM data		
		R^2	P-value	RMSE	R^2	P-value	RMSE
2	Gamma radiat.	0.69	< 0.05	3.24	0.00	>0.05	1.01
3	EC and NIRS	0.45	<0.05	3.78	0.24	<0.05	0.43
5	Multispec. cam	0.58	<0.05	4.05			

2 = Loonstra&vanderWeide, 3 = Veris, 5= Bioscope

2020 - Peat soil

The analysis was on SOM only because clay content was too low on this field. The RMSE data in Table 8 show that the soil mapping service data differed from the reference data ca. 3% for SOM. For SOM, the error was higher than in the 2018 and 2019 experiments (3% vs. 0.4%). In this experiment, the variation in the SOM content was also higher than in the 2018 and 2019 experiments. The RMSE of the reference data SOM was 4.1%.
Correlations between reference SOM data and soil mapping service data were medium to high (ca. 0.6, see Table 8). In all three cases, regression analysis tested significant. The Veris sensor system and the gamma radiation sensor system scored in the same order. The Bioscope systems also scored quite good on mapping the SOM variation of the field.

Table 8. Statistical parameters describing the correlation and error between sensor data provided by different soil mapping services and the reference data 2020 peat soil.

No.	Sensor	Clay data			SOM data		
		R^2	P-value	RMSE	R^2	P-value	RMSE
2	Gamma radiation	0.07	>0.05	0.48	0.65	<0.05	2.67
3	EC and NIRS				0.67	<0.05	2.80
5	Multispec. cam				0.57	<0.05	5.31

2 = Loonstra&vanderWeide, 3 = Veris, 5= Bioscope

2020 - Sandy soil

The analysis was also on SOM only because clay content was too low. The RMSE data in Table 9 show that the soil mapping service data differed from the reference data ca. 1.5% for SOM. For SOM, the error was higher than in the 2018 and 2019 experiments, but lower for the 2020 peat experiment. The RMSE of the reference data SOM was 0.9%.
Correlations between reference SOM data and soil mapping service data were low (ca 0.2, see Table 9). In two of three cases, regression analysis tested significant. The Veris sensor system and the gamma radiation sensor system scored in the same order. The Bioscope systems scored very poor on mapping the SOM variation in this experiment. This was due to the fact that a green manure crop was grown on the field in the winter months so that Bioscope did not have sufficient bare soil data for making the soil SOM map.

Table 9. Statistical parameters describing the correlation and error between sensor data provided by different soil mapping services and the reference data 2020 sandy soil.

No.	Sensor	Clay data			OM data		
		R^2	P-value	RMSE	R^2	P-value	RMSE
2	Gamma radiat.	0.20	<0.05	0.38	0.07	>0.05	1.63
3	EC and NIRS				0.25	<0.05	0.91
5	Multispec. cam				0.14	<0.05	1.75

2 = Loonstra&vanderWeide, 3 = Veris, 5= Bioscope

Discussion

Some specific observations and outliers were already discussed in the Result section.Some general discussion points are given here. Variable rate application of seeds, pesticides and fertilizers require accurate soil maps (Kempenaar et al, 2018). The proximal sensor systems based on EC + NIRS (Veris) and the passive gamma radiation sensors provided useful soil maps on clay content and soil organic matter for precision farming. Both systems performed better when there was a wider range of variation of the parameter in the field. Good processing and calibration is key in making accurate clay and soil organic matter maps. It is better to take an extra soil sample for calibration than minimize on costs of the service. The remote sensor system scored less good than the two aforementioned proximal sensor systems. However, the maps based on remote sensing are still of interest to end users and precision farming, knowing that their cost is lower than the costs of the proximal sensor systems. Cost and benefit of the remote and proximal soil mapping services at end user level (farmer) remains to be evaluated.

Conclusions

This research showed that soil mapping services can produce good quality soil maps of within-field spatial (2-D) variation of clay and soil organic matter of arable fields. The absolute errors of the maps compared to reference data were often smaller than the standard errors of the reference data.

This research also showed that the added value of the maps increases when there is a wider range in the variation of clay or soil organic matter content within the field. In those cases, correlation coefficients R^2 were in order of 0.7 to 0.8.

Acknowledgements

This study was carried out in the frame of the National Field Lab Precision Agriculture (NPPL: www.proeftuin precisielandbouw.nl). NPPL is a project initiated by the Dutch Ministry of Agriculture, Nature and Food. We also thank the soil mapping service companies (see ref. list) who participated in this study.

References

Egmond, F.M. van, van der Veeke, S., Knotters, M., Koomans, R.L., Walvoort, D.J.J., Limburg, J., 2018. Mapping soil texture with a gamma-ray spectrometer: comparison between UAV and proximal measurements and traditional sampling. WOT-technical report 137. Wageningen University & Research. https://edepot.wur.nl/466037. Last accessed February 2021.

Evert, F.K. van, Been, Th., Booij, A.J., Kempenaar, C., Kessel, J.G., Molendijk, P.L., 2018. Akkerweb: a platform for precision farming data, science, and practice. A paper from the Proceedings of the 14th International Conference on Precision Agriculture, June 24 – June 27, 2018. Montreal, Quebec, Canada.

Kempenaar, C., Been, Th., Booij, J.A., van Evert, F.K., Michielsen, J.M., Kocks, C.G. (2018). Advances in Variable Rate Technology Application in Potato in The Netherlands. Potato Research 60 3-4: pp 295-305 Last accessed February 2021. https://doi.org/10.1007/s11540-018-9357-4.
Knadel, M., Thomsen, A., Schelde, K., Greve, M.H., 2015. Soil organic carbon and particle sizes mapping using vis–NIR, EC and temperature mobile sensor platform. Computers and Electronics in Agriculture 114: 134-144. Last accessed February 2021. https://www.sciencedirect.com/science/article/abs/pii/S0168169915000885#!.
Lund, E.D., 2008. Soil Electrical Conductivity. Soil Science: Step-by-Step Field Analysis. Soil Science Society of America, Madison, WI, USA. Last accessed December 2020. https://www.vantage-agrometius.nl/wp-content/uploads/2020/04/ElectricalConductivity_SoilScience.pdf.
Nysten, S., Kempenaar, C., 2019. Toepassing van Bodemscans voor Smart Farming: Vergelijking van bodemscans in Ens (2018) maakt bodemeigenschappen inzichtelijk. (Application of soil scans for Precision Farming: research in Ens (2018)). Report. Aeres University of Applied Sciences, Dronten, NL. Last accessed February 2021. https://library.wur.nl/WebQuery/wurpubs/547913.
Tigchelhoff, T., Kempenaar, C., van Esssen, G., de Kleijne, G., Russcher, E., Nysten. S., 2020. Bodemscans voor Precisielandbouw: Vergelijking van sensorsystemen in Slootdorp 2019. (Soil scans for Precision Farming: research in Slootdorp (2019)). Report. Aeres University of Applied Sciences, Dronten, NL. Last accessed February 2021. https://edepot.wur.nl/529393.
Tigchelhoff, T., Nysten, S., Kempenaar, C., 2021. Toepassing van bodemscans voor Precisielandbouw: Vergelijking van sensorsystemen op dalgrond 2020. (Application of soil scans for Precision Farming: research on peat soil (2020)). Report. Aeres University of Applied Sciences, Dronten, NL. In press.
Tigchelhoff, T., Nysten, S., Kempenaar, C., 2021. Toepassing van bodemscans voor Precisielandbouw: Vergelijking van sensorsystemen op zandgrond 2020. (Application of soil scans for Precision Farming: research on sandy soil (2020)). Report. Aeres University of Applied Sciences, Dronten, NL. In press.
Yuzugullu O., Lorenz, F., Fröhlich, P. Liebisch, F., 2020. Understanding Fields by Remote Sensing: Soil Zoning and Property Mapping. Remote Sensing 12: 1116-1135. doi:10.3390/rs12071116 .

URLs to Soil mapping services/companies in this study:

- Agri-dataservices Bioscope – Remote sensing bare soil: https://bioscope.nl/
- Case New Holland - EMI: https://agxtend.com/nl/producten/soilxplorer
- Loonstra & Van der Weide – Passieve gamma radiation: http://www.loonstraenvanderweide.nl/
- Soil Masters/LoonwerkGPS - EMI: https://www.soilmasters.com/
- Vantage Agrometius – EC and NIRS: https://www.vantage-agrometius.nl/product/veris-msp3-bodemscanner/