Autoreferat

Dr Guillaume Debaene
Institute of Soil Science and Plant Cultivation –
State Research Institute
Department of Soil Science Erosion and Land Conservation

Title:
Application of visible and near-infrared spectroscopy for arable soils analysis

Puławy, 26th April 2019

Dr Guillaume Debaene

Institute of Soil Science and Plant Cultivation –

State Research Institute

Department of Soil Science Erosion and Land Conservation

Czartoryskich street, 8

24-100 Puławy

Poland

1. Personal data:

First name and last name: Guillaume Debaene

Date and place of birth: 18.02.1974, Dunkirk (France)

Nationality: French

Workplace: Institute of Soil Science and Plant Cultivation - State Research Institute,

Department of Soil Science Erosion and Land Protection

Contact details: +48 81 47 86 776; gdebaene@iung.pulawy.pl

2. Education:

In 1998, I received a master's degree (Maître) from the Department of Earth Sciences at the University of Lille (USTL - Université des Sciences et Technologies de Lille). Then in 1999, I received a master's degree (D.E.A) from the Department of Quaternary Geology at the Institute of Human Palaeontology in Paris (IPH – Institut de Paléontologie Humaine – Muséum National d'Histoire Naturelle – MNHN). Then, in 2000, I received a scholarship from the French Government (Bourse de cooperation scientifique et culturelle France – Pologne 2000). Hereafter, I was accepted at the doctoral studies from the Institute of Geological Sciences, Polish Academy of Sciences in Warsaw.

I defended my Doctoral Thesis „Uranium-series dating of carbonate sediments with significant detrital contamination" in 2004 at the Department of Quaternary Geology of the Polish Academy of Sciences in Warsaw, obtaining a PhD degree in Earth Sciences in the field of Geology. My thesis supervisor was Professor Helena Hercman (quaternary geology specialist). The reviewers were: Professor Jerzy Głazek (geology specialist), Professor Stanisław Hałas (physics specialist), and Professor Tim Atkinson from University College London (quaternary geology specialist).

3. Employment:

During the years 2007-2018, I have been employed as a senior researcher and technical specialist at the Department of Soil Science Erosion and Land Protection of the Institute of Soil Science and Plant Cultivation – State Research Institute in Puławy. Since 2018, I am an adjunct in the same department.

4. Indication of the scientific achievement according to article 16, paragraph 2 of the Law on scientific degrees and scientific titles as well as on degrees and title in art from March 14th 2003 (Dz. U. nr 65, poz. 595):

4.1. Title of achievement

"Application of visible and near-infrared spectroscopy for arable soils analysis"

4.2. List of publications constituting the monothematic scientific achievement [A1-6]

The scientific achievement constituting the basis for applying for the habilitation degree is documented by a series of six monothematic scientific publications with a total sum of 95 MNiSW score and a total Impact Factor index of 3.090 according to the years of publication.

1. **Debaene G.**, Niedźwiecki J., Pecio A., **2010**: Visible and near-infrared spectrophotometer for soil analysis: preliminary results. Polish Journal of Agronomy, 3: 3-9. **[A1]**

2. **Debaene G.**, Niedźwiecki J., Pecio A., Żurek A., **2014**: Effect of the number of calibration samples on the prediction of several soil properties at the farm-scale. Geoderma**,** 214: 114-125. **[A2]**

3. **Debaene G.**, Pikuła D., Niedźwiecki J., **2014**:.Use of VIS-NIRS for land management classification with a support vector machine and prediction of soil organic carbon and other soil properties. Ciencia e Investigación Agraria, 41(1), 21-32. **[A3]**

4. **Debaene G.**, Pikuła A., Niedźwiecki J., Kowalik M., **2015**: Spektroskopia bliskiej podczerwieni jako narzędzie przydatne w określaniu żyzności gleb. Studia i Raporty IUNG-PIB, 46(20): 85-103. **[A4]**

5. **Debaene G.,** Bartmiński P., Niedźwiecki J., Miturski T., **2017**: Visible and near-infrared spectroscopy as a tool for soil classification and soil profile description. Polish Journal of Soil Science, 50(1): 1-10. **[A5]**

6. Niedźwiecki J., **Debaene G.,** Kowalik M., **2017**: Spektralna biblioteka gleb użytków rolnych w Polsce – Podstawowe założenia. Studia i Raporty IUNG-PIB, 51 (5): 91-110. **[A6]**

The works presented above [**A1-A6**] were not part of any other monothematic cycle or any other habilitation degree. Detailed data regarding MNiSW scores, Impact Factor index, and citation numbers according to the Web of Science database and Scopus databases are summarised in the table below.

Table 1. List of publication with MNiSW scoring and IF with citation number and personal contribution to the paper

Article	MNiSW score (year of publication)	Impact Factor (year of publication)	5-years Impact Factor (year of publication)	Number of citations WoS / Scopus as of 18.04.2019	Personal contribution (%)
A1	6	-	-	-	85
A2	45	2.772	3.524	30 / 34	80
A3	20	0.318	0.436	0 / 1	85
A4	5	-	-	-	80
A5	14	-	-	- / 2	85
A6	5	-	-	-	80
Sum	**95**	**3.090**	**3.960**	**30 / 37**	

Moreover, I am the author or co-author of articles and monograph chapters. The complete list is provided in Annexe 8. My total score is: **IF = 12.761**, h-index = 3 according to the Web of Science database, and h-index = 4 according to the Scopus database. The number of citations according to the Web of Science database is 76 and 94 according to the Scopus database.

4.3. Discussion of the scientific goals of the research and results obtained with regards to possible applications

The main objective of the conducted research was to demonstrate that VIS-NIRS is suitable to predict several properties of arable soils (Poland). The three main hypotheses were the following:

1. Arable soils are a good material for reflectance spectroscopy and therefore the method could be used in precision agriculture.
2. Visible and near-infrared spectroscopy is a tool that can be used for mapping soils properties at low cost with a precision similar to that of classical method.

3. VIS-NIR spectra of mineral soils (arable soils) contain information's allowing soil classification.

The scientific objectives of this work were:

Objective 1. Introduction and dissemination in Poland of visible and near-infrared spectroscopy for arable soils analysis. [**A1**]

Objective 2. Building a national spectral library of arable soils for Poland and using the method to predict chemical and physical soil properties at different scales. [**A1, A2, A3, A4, A6**]

Objective 3. Reducing the cost of analysis: find the best calibration method with the lowest number of samples and use of a legacy (IUNG-PIB) soil database. [**A1, A2**]

Objective 4. Using the method for mapping at different scales (from field to national scale). [**A2, ,A4, A6**]

Objective 5. Testing if the method is suitable for soil classification [**A3, A5**]

Introduction

With the world's growing population, there is a need for a more productive and sustainable agriculture (Bongiovanni and Lowenberg-Deboer, 2004; Pimentel, 1991). Moreover, the increase of population is also related to environmental problems such as soil pollution or global warming. Precision agriculture (PA) with the use of near-infrared sensors is one of the responses to these problems. Near-infrared sensors can be laboratory based but are also promising proximal sensing tools that can be used *on-the-go* (Christy, (2008), **A4**), on-line on moving samples (Porep et al., 2015), in situ or in the laboratory for screening and monitoring purposes (**A2**; Nocita et al., 2015; Wenjun et al., 2014). Unfortunately, large numbers of samples are needed for all these tasks. The aim of visible and near-infrared (VIS-NIR) spectroscopy is to rapidly and inexpensively characterise products or samples and to detect anomalies and outliers during production or analyses. Following the seminal works by Norris and Hart (1965) and Williams (1973), the method is mostly used in the agricultural sector (cereals, grains, forages) and for food and beverages analysis (García-Sánchez et al., 2017; Roberts et al., 2004) but also within pharmaceutical (Jamrógiewicz, 2012) and chemical industries (Chung, 2007; Long et al., 2018). In the medical sector (Chou and Lan, 2013; Sakudo, 2016), the method has gained interest in recent years. The waste management sector (Smidt et al., 2011) is also actively using and developing the technique. In the last few years,

the literature involving VIS-NIR spectroscopy has greatly increased in number and in the diversity of the applications involved. Nevertheless, soil analysis is still a challenge due to the specificity and complexity of the material (Ben Dor et al., 2015; Nduwamungu et al., 2009; Roberts and Cozzolino, 2017). Regarding soils, most of the investigations have focused in ways to predict soil properties (Nocita et al., 2015) and especially soil organic carbon (SOC) (Ladoni et al., 2010). More recently, a few attempts to classify soil samples according to different parameters (soil type, SOC content, land management) have been made (**A3**; Vasques et al., 2014). The accurate estimation of soil C content (organic and inorganic) using VIS-NIR spectroscopy is of particular interest due to potential applications for carbon sequestration and soil quality research. The method has also a great potential for precision agriculture (PA) and for mapping purposes (e.g. digital soil mapping – DSM). The predictive accuracy reported in literature is often of similar level as classical laboratory analyses (Sørensen, 2002).

The basis of VIS-NIR spectroscopy is that in the NIR region, the radiation (light) is absorbed by different chemical bonds (e.g. C-H, N-H, O-H, C=O) of compounds present in the sample. The absorption is related to the concentration of these compounds. Therefore, a NIR spectrum contains information about the organic composition of that sample and there is a qualitative and quantitative relationship between the chemical composition and the spectrum. Physical properties (e.g. particle size) are also related to NIR spectrum since the shape of the spectrum is affected by reflection and scattering of the light. There are several advantages of VIS-NIR measurements over other techniques: (1) the C-H (but also C-O and N-H) associated vibrational information are repeated several times between 780 nm to 2500 nm; (2) different combination regions can be selected to improve modelling; (3) it is relatively easy to build low cost instruments with high signal-to-noise ration; (4) the method is rapid even when using low cost fibre optics, (5) the method is non-destructive, minimal (sieving) or no sample treatments are needed and (6) the method is environmental friendly since no toxic reagent are required for the analysis.

The problem with soils is that the complicated matrix generally require the use of the full spectra (780-2500 nm). Very often, the visible (VIS) region (350-780 nm) is added to the NIR spectra.

In the VIS region, most of the spectral variations between soil samples are due to sample colour caused by chromophores such as organic matter and iron oxides. In the infrared region, wavelengths are absorbed by covalent chemical bonds (e.g. C-H, O-H, and N-H). The near infrared spectra consist of overtones and combination bands of these fundamental

molecular absorptions found in the mid infrared (MIR) region. In the field of soil science, most of NIR spectrometers are working with reflectance spectroscopy. It means that a sensor captures the wavelengths reflected by the sample in the NIR region. The sensor is connected to a spectrometer that can calculate the absorbance. The absorbance is related to the chemical composition of the sample and indirectly to physical and biological soil properties. Near-infrared spectra consist of generally overlapping vibrational bands that often appear non-specific and poorly resolved in the case of soils. The most typical peaks are located at 1400, 1900, and 2200 nm and are due to light absorption by O-H bonds (water). The 2200 nm is also related to clay minerals. There are several other peaks related to organic matter functional groups but often these peaks are not clearly defined.

Due to the lack of specificity of the soil spectra and the habitual low analyte concentration in the soil matrix, qualitative and quantitative measurements with VIS-NIR spectroscopy require the use of multivariate calibration algorithms (e.g. partial least square regression – PLSR, support vector machine – SVM, neural networks – NN) and other statistical methods. Those specific methods belong to the field of chemometrics. This is due to the fact that unlike univariate calibration where one peak is directly correlated to the reference concentration, with NIR, the entire spectrum is used to correlate with the reference concentration. The goal is to relate NIR spectral response to sample chemical or physical properties used for calibration and then predict these properties on new samples. Quantitative results are possible where changes in the response of the near-infrared spectra are correlated to changes in the concentration of chemical components, or with the sample physical characteristics.

There are several steps to build a multivariate VIS-NIR model. (1) Build a matrix containing the spectral data and laboratory analyses; (2) pre-processing of the spectra (Rinnan et al, 2009); (3) choosing the spectral range (usually the full range for soils); (4) tuning of the calibration; (5) finale calibration; (6) validation with independent samples; (7) and eventually routine analysis (new samples, only the spectra are then needed).

The main subject of this study was arable soils since the development of VIS-NIR spectroscopy gives great opportunities for use in precision farming.

Objective 1. Introduction and dissemination in Poland of visible and near-infrared spectroscopy for arable soil analysis. [A1]

A quick, limited, search at the Web of Science (WoS) Core Collection can help having an insight on how the method is still new in the field of soil sciences and can underline the

originality of this study. The results also highlight a common misunderstanding about the method: the difference between proximal and remote sensing. The present work is dealing with proximal sensing, that is collecting detailed spectral information near soil surface (or in direct contact with soil) whereas remote sensing collects spectral information from satellite or airborne systems and therefore in a less detailed manner.

To simplify the analysis, the WoS query was run with two mandatory terms: (a) *near-infrared* and (b) *soil* with at least one author having a Polish address (in few cases, the author is a foreigner with two affiliations). The query returned thirty-three results. From these results, ten papers are not directly related to soil science (NIR reflectance analyses of leaf, roots, plant based food and seven articles involved with satellite imaging and remote sensing). The oldest papers (1987 to 2004) belong to the last twenty-three. Seven of these papers are focusing on the theoretical basis of NIR light reflectance in relation to soil surface roughness and are not involved with soil analyses or with the prediction of any soil properties. Of the sixteen remaining papers, there is a review and an article dealing with lignite-humus mixture and thus not soil. The last 14 are in line with the present contribution, mostly focusing on soil properties prediction. The seminal paper for Poland is from 2003 (Chodak et al., 2003), and like two-third of these papers, investigated forest soils. Only four papers (**A2**, **A3**, Bajorski et al., (2016); Siebielec et al., (2004)) studied arable soils. Many papers were authored by Polish researchers abroad e.g. (Chodak et al., 2004; Siebielec et al., 2004) due to the lack of instruments at that time in Poland (prof. Chodak and dr hab. Siebielec personal communications). Moreover, none of these papers have examined the possibility of lowering the number of samples to reduce analytical costs (**A2**), used the models for mapping (**A2**, **A4**, **A6**), nor investigated the possibility of using VIS-NIR spectra for soil classification (**A3**, **A5**).

To compare with other fields of research, the same query (within the same laps time) was undertaken with common studies in soil sciences such as soil + erosion, soil pollution then restricted to polycyclic aromatic hydrocarbon (soil + PAH) instead of near-infrared. The results are as follow: erosion (370), soil pollution (486), and PAH (183). This emphasise the novelty of the presented research that is still recent within soil studies and not popular in Poland.

The contribution (**A1**), further introduced the method to a Polish but also international readership. Most of the objectives of the scientific achievements presented above were quickly introduced as preliminary results. (1) Prediction of physical and chemical properties, (2) calibration samples selection and finally (3) the first step towards a precise mapping at a reduced cost were discussed.

The investigation was conducted at field scale at the experimental station of the Institute of Soil Science and Plant Cultivation (IUNG-PIB) in Baborówko near Poznań (Greater Poland). The whole station (132.1 ha) is divided into several fields where field trials are conducted. The field under investigation in (**A1)** is a part (9 ha) of Field 1 and belongs to a set of 3 fields with the following crop rotation: winter wheat, spring barley and winter rape. The soil type is a Luvisol according to the World reference Base for Soil Resources (WRB). This type of soil is typical for Western Poland. A soil grid was georeferenced by means of a global positioning system (GPS) where 117 top soil (0-25 cm) samples were collected. The sample size (117) seemed sufficient to build a robust model since most papers from Viscarra Rossel and McBratney review (2008) comprised between 100 and 200 samples. Two calibration schemes were investigated. The difference was in calibration sample selection. In scheme 1, out of the 117 samples, 90 were used for calibration (the first 90 samples from the reference grid). The remaining 27 samples were used as an independent validation dataset (the next samples in sequence from the reference grid). In scheme 2, 90 samples were also used for calibration but from the grid, every 4 sample was withdrawn for validation (Dunn et al., 2002). For each of the two schemes, two calibrations were built, with and without outliers. Better calibrations were obtained with scheme 2 with lower RMSE (root mean square error, subscript v is for validation and cv for cross-validation) but for SOC.

Satisfactory results were obtained for the prediction of SOC ($r^2 = 0.65$, $RMSE_V = 0.14$) and available Mg ($r^2 = 0.69$, $RMSE_V = 2.16$). These results are comparable to Dunn et al., (2002) models for Australian topsoils of very similar composition. Models with calibration samples evenly widespread across the field were more robust (lower $RMSE_V$) than models with calibration samples covering the entire field excluding the western part. Scheme 1 failed to predict several properties. This was explained by a better representation in the calibration dataset of samples similar to the ones from the validation dataset in scheme 2. In scheme 1, all validation samples were coming from an area not covered by the calibration dataset. This explains the poor validation with scheme 1 even if the calibration results (cross-validation) were acceptable. Clay is a good example to explain those results (see Fig. 1 from **A1**). In scheme 1, validation samples are coming from a highly variable area for clay content with the presence of sandy loam not well represented in the calibration dataset. In addition, the spectra have been examined also by an outlier detection algorithm that can work in multivariate space Filzmoser et al., 2005) and the calibration was re-run without outliers. There was no improvement after deleting the outliers. However, the algorithm has selected interesting samples as outliers (e.g. samples coming from flooding areas due to compaction by engines).

These samples were not differing from the other samples in the view of chemical or physical properties. The VIS-NIR method in that case can give a clue that the population is different even is the samples chemical composition is similar. This could be a hint for further research on the application of VIS-NIR spectroscopy. The fact that VIS-NIRS outliers are sometimes not outliers but interesting samples was also underlined in **A2**.

The description of the Veris® mobile platform and of the functioning of its system for spectra acquisition and data treatment was the first of an available VIS-NIR *on-the-go* system in Poland. The satisfactory prediction results obtained when scanning the samples in laboratory mode (bench top mode) suggested a potential for using the method *in situ* with the mobile platform.

With objective 1, the bases for successful predictions of several soil properties have been set down. The necessity was now to improve the prediction and robustness of the modelling and to apply the method to more chemical and physical soil properties.

Objective 2. Building a national spectral library of arable soils from Poland and using the method to predict chemical and physical soil properties at different scales. [A1, A2, A3, A4, A6]

The main advantage of the method is that many soil attributes (physical, chemical, or even biological properties) can be characterised simultaneously (Zornoza et al., 2008). In papers (**A1, A2, A3, A4, A6**), the prediction of SOC, total nitrogen (N_{tot}), sand, silt, clay, pH, hydrolytic acidity (Hh), plant available (Ca, Mg, K, Na) and exchangeable cations (P, K) were investigated. The models were built for samples coming from different scale studies: from (a) limited area in a field experiment (**A3, A5**) to (b) farm scale (**A1, A2, A6**) and from (c) regional (**A6**) to (d) national scale (**A6**).

Before building predictive models, one not only need soil spectra but also the corresponding soil properties that will be predicted latter on new samples. Therefore, it was decided to build a National Soil Spectral Database (NSSDB) with the IUNG-PIB legacy soil database. The database was enhanced with samples from different IUNG-PIB experimental stations. The advantage of using that database was that not only soils are physically available but also several chemical and physical data for each sample. The NSSDB is in constant development but to this day, more than 2500 samples are included. The state of the spectral database in 2017 is described in **A6**. The database is the result of statutory program 4.02 "Use of chemometric methods to predict mineral soil properties in Poland" The spectra were

acquired with the Veris ® spectrophotometer (Salina, KS, USA) in the 350-2220 nm range but the NSSDB is now being updated with a new instrument (improved parameters), the PSR-3500 (Spectral Evolution, MA, USA) in the 350-2500 nm range. New samples are regularly incorporated and old ones scanned with the new instrument. These new spectra are more compatible with existing spectral libraries (Knadel et al., 2012). This is important with the recent interest in building a global soil spectral library (Viscarra Rossel et al., 2016). There are very few samples from Central and Eastern Europe in the global library. Therefore, it was decided in Rome (International Workshop "Soil Spectroscopy: the present and future of soil monitoring" - 2013) to incorporate samples from the Polish Spectral Library. A spectral library is not something static but something that need constant upgrading. Therefore, new samples are added regularly to the NSSDB.

The first large (continental) spectral library was introduced by Shepherd and Walsh (2002). Authors first described a conceptual framework to use spectral libraries for the prediction of soil properties. More than 1000 archived samples from eastern and southern Africa were scanned in the VIS-NIR range and added to the library. They obtained good calibration for SOC, pH, exchangeable Ca and Mg, and particle-size distribution. That paper initiated the building of a global spectral library (ICRAF, 2015). The global library is described in details in Viscarra Rossel et al., (2016). It contains more than 23,000 spectra with several soil properties. For Europe, the LUCAS spectral library (Stevens et al., 2013) based on the LUCAS (*Land Use/Cover Area frame Survey*) survey is impressive too with more than 20,000 samples (Tóth et al., 2013). There are also few national library e.g. Denmark, France, Australia (Gogé et al., 2012; Knadel et al., 2012; Viscarra Rossel and Webster, 2011).

In the view of papers (**A2**) and (**A6**), it was determined that several steps are needed to build a good (robust) prediction model.

(1) Selection of samples to be included in the database (legacy samples, new samples…).

(2) Scanning of the samples (sieved needed or not?, how many scan replicates…).

(3) Studying the spectral characteristics of the sample population (e.g. PCA for dimension reduction). This gives a first insight about the soil population and indicates eventual grouping of the data but also can detect some outliers. That part is important for latter decision of which samples should be included in the calibration dataset.

(4) Selecting samples for calibration. The selection process is well described in (**A1**)**,** but especially in (**A2**) were four methods are described. Two of these methods for selecting calibration samples were new (schemes 3 and 4 from **A2**). The two methods are based on spectral characteristics and not on the chemical or physical soil characteristics. The best

method overall (giving the most robust models with better prediction parameters) was the one based on k-means clustering. The second method proposed was based on the PCA-scores. The two other sample selection methods were more classical and based on random selection or on sample chemical composition.

(5) The actual modelling using partial least square regression (PLSR). Other multivariate calibration methods were also investigated (e.g. Support Vector Machines).

(6) Testing on independent datasets. This was tested by using samples from the spectral library not incorporated in the calibration step or with completely new/independent samples.

When working with larger spectral database or at a higher scale of experiment (regional or national), one has to take into account that it will be difficult to incorporate in the database all types of soils with all their diversity. It is therefore very important to carefully observe the framework described above.

To test how the method (and therefore the NSSDB) is working at regional scale, samples coming from the Silesian region were used to build a predicting model for SOC. These were samples from a pedo-geological mapping project. That dataset was selected with a purpose. It was to examine the effect of sample depth on the model. 60 samples from each different depth (15, 30, 50 cm) were used for a total of 180 samples. One sample from each depth of the 60 profiles was analysed. Two calibrations schemes (with cross-validation) were tested to decide how to select calibration samples for the modelling. In the first scheme, as a calibration set, all samples (60) from one depth were used and as a second scheme, the k-mean clustering method (**A2**) was used to select 20 samples from each depths for cross-calibration (total 60 calibration samples). Using samples from each depth has given slightly improved model robustness with lower RMSEcv, therefore that method was employed for modelling SOC, clay and pH (KCl).

The model validation was completed with the rest of the dataset from each depth (120 samples – scheme 2). SOC and clay content prediction were satisfactory ($r^2 = 0.74$; RMSEv = 0.48 and $r^2 = 0.71$; RMSEv = 0.82 respectively). The prediction model for pH (KCl) was less successful ($r^2 = 0.51$; RMSEv = 0.55). This could be due to the mixture of profiles coming from arable lands but also for few of them, forest.

Finally, the testing of the method (and of the NSSDB) was undertaken with the entire dataset. 1200 samples (after outlier removal and averaging samples from field experiments) were used for the calibration of SOC, clay and pH (KCl). At the time of modelling, there were few samples with high SOC or clay content values. Therefore, sample with e.g. SOC > 5% were removed. Samples not representative of Polish soils (paper **A6**) were also removed. The

calibration results were the following: soil organic carbon ($r^2 = 0.69$; RMSEcv = 0.59), clay ($r^2 = 0.73$; RMSEcv = 0.89), and pH ($r^2 = 0.59$; RMSEcv = 0.40). Validation results were as follow: SOC ($r^2 = 0.67$; RMSEv = 0.60); clay ($r^2 = 0.72$; RMSEv = 0.88), and pH ($r^2 = 0.62$; RMSEv = 0.43).

The use of a really independent validation dataset is seldom met in the literature (Cécillon and Brun, 2007; Hodge and Woodbridge, 2010). In paper (**A6**), the models were verified on an independent dataset (samples from a profile (paper **A5)** and samples from an independent research project). These samples were scanned (since they do not belong to NSSDB) and the model for SOC, clay and pH predictions described above were applied.

The prediction results were as follow: SOC ($r^2 = 0.63$; RMSEv = 0.48), clay ($r^2 = 0.67$; RMSEv = 0.78), and pH (KCl) ($r^2 = 0.55$; RMSEv = 0.75).

Prediction results at field scale were presented in Objective 1 for one field (paper **A1**). Paper (**A2**), increased the investigated area to farm scale. The studied area was therefore much larger (53.6 ha). More than 400 samples were collected from the soil sampling grid. 200 samples were used for calibration and 200 for validation as an independent dataset. Soil properties were similar to (**A1**) as might be expected for adjacent fields.

In paper (**A2**), the same soil properties were investigated as in (**A1**). However, not only clay fraction was predicted but also sand and silt. The predictions were more robust in (**A2**) due to the increased number of calibration samples (Brown et al., 2005). Some of the prediction results were e.g. SOC ($r^2 = 0.72$; RMSEv = 0.12), clay ($r^2 = 0.73$; RMSEv = 0.32), pH (KCl) ($r^2 = 0.51$; RMSEv = 0.34), sand ($r^2 = 0.79$; RMSEv = 2.6). The other prediction results can be found in Table 5 (**A2**). These results emphasised the great potential of the method for SOC or particle-size distribution analysis (first time in Poland). The very low RMSEv for both SOC and clay content (0.12 and 0.32% respectively) also showed the high accuracy of the method.

Predictions at the smallest scale (field experiment – loamy sand) were investigated in studies (**A3**) and (**A4**). In these papers, samples from a long term experiment starting in 1980 (Pikuła and Rutkowska, 2014) were analysed for several soil properties and scanned on the VIS-NIR range. Paper (**A3**) presents results from the crop rotation 2008 (40 samples) and paper (**A5**) results from the rotations 2011-2012 (80 samples). At that scale of experiment, the results were very good for SOC and N_{tot} with RMSEv as low as 0.05 and 0.1% respectively. Such accuracy is seldom reported in the literature and is similar of even better than classical method of determination. Nevertheless, we need to take into account that such very low errors are also due to the narrow range of SOC and N_{tot} values in that field. Predictions for the

macro-elements (exchangeable P and K content) were successful in paper (**A4**). This is surely due to the fact these elements needed more sample than the 40 samples analysed in paper (**A3**). This was probably due to the fact that with 80 samples, the calibration dataset could better encompass the spread of data. The same happened with pH prediction. The dataset with the 80 samples has given more robust prediction of soil pH. Some interesting conclusions from (**A3**) are that a calibration dataset should include at least 30 samples and probably more than 40. Moreover, since the area investigated is small and with a very homogeneous soil texture, the particle-size distribution had no effect on the spectra and therefore on the modelling. The successful calibrations were the results of the changes in soil chemical properties due to the application of manure and N-fertiliser in each small cells of the field experiment.

The method has shown good predictability for SOC, N_{tot}, available and exchangeable Mg, sand, silt, and clay (**A1, A2, A3, A4, A6**). Satisfactory results were obtained for pH and available Na and promising results for Hh. To our knowledge, it was the first time Hh was predicted using VIS-NIR spectroscopy (it was not predicted ever since). For exchangeable P and K, the predictions were successful when the distribution of P and K were not too skewed towards high values. Otherwise, it was dependant on the calibration scheme (the same for exchangeable Ca and Na). That means that for all these cations, the calibration dataset need to have a normal distribution. That phenomenon was observed elsewhere (Dunn et al., 2002). This is similar to the prediction results for clay in (**A1**) (see explanation in Objective 1).

Some properties (SOC, clay, pH) were investigated at all scales. As expected, the prediction robustness is decreasing when the scale of the experiment is increasing (**A1, A2, A3, A4, A6**). More investigations are needed (e.g. neural networks, grouping of samples for calibration dataset selection).

Objective 3. Reducing the cost of analysis: find the best calibration method with the lowest number of samples and use of a legacy (IUNG-PIB) soil database. [A1, A2]

One of the goals of paper (**A2**) was to determine the optimal number of calibration samples needed to obtain the best models (high r^2 and small RMSE) and therefore reducing costs of analysis by reducing the number of calibration samples. With that goal in mind, 398 samples from the grid sampling (53.6 ha) from the experimental station of Baborówko were used. The grid cell has a size of 35 m. Samples were dived in 2 datasets (199 each, every two samples from the grid), for calibration and for validation (independent dataset). Four calibration sample selection methods were tested on the calibration dataset (random sampling,

sampling according to increasing soil properties values, sampling based on k-means clustering, and finally sampling according to PCA-scores). Cross-validation was used on the calibration dataset with an increasing number of calibration samples (20 to 199 with an increment of 20, see Tables 3 and 4 from paper **A2**). The obtained models were later validated with the 199 samples from the independent validation dataset. These several calibration datasets were used to model SOC and clay content. Results of calibration are presented in Tables 3 and 4 and in Figure 3 from paper (**A2**). The calibration results have showed that it is possible to reduce the number of samples to 80 without losing on the robustness of SOC or clay prediction models. It means that it is possible to reduce the number of calibration samples by 60%. Examples of prediction for SOC and clay using the k-means method for selecting 79 samples and by using the full dataset (199 samples) are the following: SOC_{K79} (r^2 = 0.63; RMSEv = 0.13), SOC_{F199} (r^2 = 0.72; RMSEv = 0.12), $clay_{K79}$ (r^2 = 0.71; RMSEv = 0.36), and $clay_{F199}$ (r^2 = 0.73; RMSEv = 0.32). The differences in prediction accuracy are not significant when using 79 samples instead of 199. Therefore, the other soil properties (sand, silt, pH, plant available Mg, P, K) analysed in paper (**A2**) were only predicted using 79 samples (k-means) and with the full calibration dataset for comparison. As for SOC and clay, no significant differences were observed between prediction results using these two methods.

Reducing the number of samples translates directly onto reducing the costs of analyses. A *cost analysis* was done in paper (**A2**). The reduction to 79 calibration samples instead of 199 for the prediction of 199 new samples reduces the costs by a factor 5 (80%). For SOC analyses, it means a reduction from 12,000€ to 2,300€. Similar results (Nduwamungu et al., 2009) were obtained for France and Canada with cost reduction of 94% and 63% respectively.

Another approach to reduce the cost of analyses is to use an existing soil spectral library. If a soil spectral database contains similar samples with similar properties and climate to the ones from a field to be analysed, good prediction can be obtained by building robust models using only few new samples from that field. That method is called "spiking" (Guerrero et al., 2014).

The costs of developing a soil spectral library are elevated. It is the reason why spectral library are often built from existing soil database and archived soil samples with information about soil the chemical and physical soil properties (Viscarra Rossel and Webster, 2011). Using legacy soil database is associated to errors related to the fact that several different analytical technics can be used for soil analysis, e.g. SOC can be analysed by the Tiurin method or by the Walkey-Black method (Soriano-Disla et al., 2013). This was also noticed in papers (**A1, A2, A3, A4, A6**). The larger was the scale of investigation for modelling; the

highest were the errors (RMSEv), and less precise were the models. A large database increases the chances of having non representative validation samples (this was thoroughly described in Objective 2). The ideal data distribution should be uniform but is it barely the case with real samples. Nevertheless, when compared with other large scale studies (Brown et al, 2005; Gogé et al., 2015) the obtained errors are smaller. This is due to the fact that soil samples (from the IUNG-PIB database) selected to create the models were subject to an initial selection. Bearing in mind that Polish arable soils are not very varied (in terms of SOC and clay content), samples not representative or with outlying values for SOC or clay content were rejected from the database. These outliers were not eliminated from the NSSDB. The library is in constant construction with a view to have a satisfactory representation of every type of soils even the ones occurring in small areas (paper **A6**).

Objective 4. Using the method for mapping from field scale to national scale. [A2, ,A4, A6]

Paper (**A1**) underlined the fact that VIS-NIR spectroscopy is a suitable tool for predicting SOC and clay content or even pH. The paper described for the first time the scanning method *on-the-go* but no prediction maps were presented. The first map produced by the results from a VIS-NIR SOC prediction from the 400 grid samples was presented in paper (**A2**). First, a precise SOC map was created based on real values (classical laboratory analyses). That map was used as a reference (Figure 4, **A2**) to compare with the prediction map based on predicted SOC values (Figure 5, **A2**). The prediction map was based on the calibration model from the 79 samples selected by k-means clustering (description in Objective 3). It is difficult to see any differences between the two maps. Therefore, a map of the differences between modelled SOC and reference values was created (Fig. 6, **A2**). That map shows that the differences are minimal, 93% of the map area presented differences of less than 0.1%, and the rest of the map differences of less than 0.4%. The model is therefore very successful. As a reminder, the prediction result for SOC using the k-means method to select calibration samples (79) was $r^2 = 0.63$ and RMSEv = 0.13. Surely, it is possible to obtain better prediction parameters, but it is not economically sound (see Objective 3). During the calibration selection process, the aim was to obtain small and stable RMSE (see Figure 3, paper **A2**). Figures 8 and 9 from paper **A2** demonstrate the benefits of using k-means clustering to select calibration samples in comparison to random sampling.

In the same paper (**A2**), the map of differences between predicted pH and reference values is presented (Fig. 7). As a reminder, the predicted result for pH using the k-means

clustering to select calibration samples (79) was $r^2 = 0.52$ and RMSEv = 0.34. The highest differences between predicted and reference pH values were around value 1.3. More than 70% of the map area presented differences of less than 0.3 (pH value). According to Stenberg et al., (2010), these results may not be sufficient for use in precision lime applications. However, these results give important information about the pH variability at the field scale.

In paper (**A4**), the Veris® mobile platform was used to collect a large quantity of spectra to reduce mapping costs. The platform is pulled behind a tractor. The field investigated is situated in Baborówko (IUNG-PIB) and has an area of 20 ha. This is a section of the field investigated in paper (**A2**) with 400 georeferenced position by a sampling grid. That field was chosen because a very detailed reference map for SOC was available (Figure 4, **A2**). The field was scanned with the platform at the speed of 7 km h^{-1}, 8800 spectra were recorded with their GPS localisation. Twenty representative samples were selected with a fuzzy logic algorithm (Christy, 2008) for calibration. The algorithm first groups the spectra in 20 clusters and choose the more representative spectra from each cluster. The calibration models are built according to those samples and then the 8800 spectra predicted for SOC content. It is up to the user to decide how many clusters. Since each spectrum is recorded with GPS values, it is easy then to sample topsoils from those 20 locations for SOC analysis (or any properties that need to be mapped). A map of differences of SOC content between the predicted map and the reference map was produced (Figure 6, **A4**). More than 70% of the mapped area showed less than 0.2% differences in SOC content. Therefore a model *on-the-go* calibrated with only 20 samples achieved satisfactory quality.

An attempt to use the method to build a prediction map at the national scale was undertaken in paper (**A6**). Prediction maps for SOC and pH were generated (Figures 15 and 16, paper **A6**). The two maps (SOC and clay) created from 250 scanned samples (validation dataset) are in accordance with our knowledge. To calibrate the two prediction models, 1200 samples (agricultural soils) from the NNSDB (paper **A6**) were used. The present state of the NNSDB accurately covers the diversity of Polish agricultural and mineral soils, and therefore the two prediction models only take into account those soils. Currently, the spectral library is enhanced with organic soils. This is the main objective of my actual statutory research program (program 3.16. "Use of chemometric methods to evaluate the quality of agricultural organic soils or originating from organic soils").

Objective 5. Testing if the method is suitable for soil classification [A3, A5]

Due to the diversity of soil deposits, their classification is difficult and often labour-intensive. Classification of soils, e.g. according to the WRB classification requires many analyses and significant experience. Spectral methods in the field of visible and near-infrared spectroscopy enable fast estimation of soil properties. The method has great potential to become a supporting tool in the classification process.

The inspiration for classifying soils based on spectral data was following the successful attempts to predict soil properties (SOC, soil texture, etc.) using VIS-NIR spectroscopy. Furthermore, Ben-Dor et al. (2008) and Mouazen et al. (2005), which explored the possibilities of using the method to classify soils, stimulated the attempt in paper (**A3**). Ben-Dor et al., (2008) successfully classified soil profiles taking into account spectra and laboratory analyses. Mouazen et al., (2005) were able to classify soil samples in three textural classes (sand, silt, and clay). My preliminary attempt to classify soil samples using VIS-NIR spectroscopy was carried out on samples from a field in Grabów experimental station (IUNG-PIB). A long term experiment starting in 1980 is carried out on a part of that field (Pikuła and Rutkowska, 2014). The experiment is based on two crop rotations (A and B). Crop rotation A without plants considered to enrich soil with organic carbon and crop rotation B with plants promoting the formation of organic carbon. A detailed description of the long term experiment (plants and doses of manure or N-fertiliser) is provided in papers (**A3** and **A4**). The basic hypothesis was that long term experiments (land management) have a direct effect on soil spectra. The objectives of paper (**A3**) were to demonstrate that soil spectral properties allow the classification of soil samples according to (1) crop rotation and (2) doses of manure or mineral N-fertiliser applied. Forty samples were used to build the models using support vector machines (SVM). Two modelling schemes were investigated. For each scheme, 3 prediction models were built to predict and classify samples depending on the crop rotation, the doses of manure, or the doses of N-fertiliser. The forty samples were used with cross-validation in scheme 1. For scheme 2, to be more reliable, 30 samples were used for calibration and the remaining 10 for validation (meaning that those 10 samples were classified). That procedure was repeated 3 times to remove the possible case were most of the 10 samples validation samples belonging to one micro-plot (i.e. with the same doses of manure or N-fertiliser). The most important results are presented in Table 4 (paper **A3**). The best classification accuracy was obtained when classifying samples according to the crop rotation (100% accuracy with scheme 2). Similar accuracies were obtained when classifying

samples according to the doses of manure or the doses of N-fertiliser (75% scheme 1 and 60% scheme 2). The most important conclusion from paper (**A3**) was that it is possible to classify soil samples according to land management (crop rotation, doses of manures or N-fertiliser) using spectral information alone. Using only spectra eliminate the need for expansive laboratory analyses. Moreover, using 30 calibration samples (scheme 2) seems not enough (not representative) because each wrongly classified sample reduce the accuracy of 10%.

Paper (**A5**) presents preliminary results from three independent experiments. The first experiment explores the effect of soil moisture on the spectra. Three, significantly different types of soils were selected. The samples were scanned during the drying process (from saturation to completely dry samples). Samples were scanned every three days during two weeks. The acquired spectra are presented in Figure 1 (paper **A5**). The most important conclusion is that it is possible to classify samples from spectral data but samples need to have similar moisture content. The second experiment involved one type of soil but with three level of SOC content (low, medium high). Therefore, it is the effect of SOC content on spectra that was examined. The experiment was carried out on micro-plots in Puławy (IUNG-PIB) were one initial soil is divided into three plots. The first plot contains the untouched topsoil with high SOC content, to the second plot with low SOC content was from a deeper part of the profile (the same original soil) and the third plot with medium SOC content, is a mixture of the two first plots. The results of the scanned samples are presented in Figure 4 (paper **A5**). One can see that SOC content has an important effect on the spectra. A higher SOC content is related to a higher light absorption. This allows us to think that it is surely possible to classify samples according to the organic content. The last experiment involved the scanning in the VIS-NIR range of a soil profile and investigating the possibility to classify samples according to the sample depth. From 10 cm deep (topsoil) to 150 cm deep, 3 scan were obtained (3 samples) every 20 cm. A principal component analysis (PCA) was completed on the spectra (Figure 5, **A5**) and illustrated the fact that samples from a profile are clustering according to sample depth even when no visual difference were visible. All the results from these three experiments allow us to think that VIS-NIR spectroscopy is a suitable method for soil classification and is the basis for further investigations.

Conclusions

The most important achievements of the above-mentioned works include:
1. Introduction and popularisation of visible and near-infrared spectroscopy in Poland.

2. Creation of the NNSDB soil spectral database. Defining the methodology for building a spectral library for mineral soils, and demonstration of how to use that library to predict soil properties. Demonstrating the need for further development of the database.

3. Demonstrating that a correct and accurate selection of calibration samples is of significant importance when creating a prediction model.

4. Proving that the use of spectroscopy in the visible and near-infrared range can reduce analyses costs by 80%.

5. Demonstrating that with VIS-NIR spectroscopy, it is possible to create maps of soil properties that do not differ from the maps created on the basis of expensive laboratory tests. The first prediction map was created based on the application of the *on-the-go* method.

Use of the research results

The research presented in the achievements has a high scientific value. First of all, the principle of VIS-NIR spectroscopy was methodically explained. In addition to the possibility of applying the method to predict soil properties, the investigations have shown that it can be used in other applications. It is foreseen that in the future the method will also be used to classify soil profiles.

Spectroscopic studies aimed at predicting various soil properties can be carried out at different scales, from the field to the country scale. This demonstrates that this method is universal and can successfully replace expensive laboratory analyses.

The VIS-NIR spectroscopy method can be successfully used for the creation of soil properties maps both at field and country scales. The method can therefore be a very useful tool that can play an important practical role in precision agriculture. The application of the field-scale method to predict, for example, organic carbon content, pH and other properties, will give a precise answer to the farmer as to the localisation and doses of fertilisers needed.

At country scale, my research is unique in Poland. The Spectral Soil Library of Poland can be used to quickly estimate the content of organic carbon, the content of clay fraction or soil pH. Creating maps, e.g. organic carbon content map, will allow classifying soils and designating valuable natural sites. Monitoring the abundance of organic material is also of great importance for monitoring and studying climate change. The method can become a tool used locally by local government authorities or globally by national authorities, for example to create protection plans, etc.

References

Bajorski, P., Kazmierowski, C., Cierniewski, J., Piekarczyk, J., Kusnierek, K., Krolewicz, S., Terelak, H., Stuczynski, T., Maliszewska-Kordybach, B., 2014. Use of clustering with partial least squares regression for prediction based on hyperspectral data. 6th Workshop on Hyperspectral Image and Signal Processing.

Ben-Dor, E., Heller, D., Chudnovsky, A., 2008. A Novel Method of Classifying Soil Profiles in the Field using Optical Means. Soil Science Society of America Journal 72 (4), 1113-1123.

Ben-Dor, E., Ong, C., Lau, I.C. 2015. Reflectance measurements of soils in the laboratory: Standards and Protocols. Geoderma 245-246, 112-124.

Bongiovanni, R., Lowenberg-Deboer, J., 2004. Precision Agriculture and Sustainability. Precision Agriculture 5 (4), 359–387.

Brown, D.J., Bricklemyer, R.S., Miller, P.R., 2005. Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana. Geoderma 129 (3-4), 251–267.

Cécillon, L., Brun, J.J., 2007. Near-infrared reflectance spectroscopy (NIRS): a practical tool for the assessment of soil carbon and nitrogen budget, in: Jandl, R., Olsson, M. (Eds.), Greenhouse-gas budget of soils under changing climate and land use (BurnOut). BFW, Vienna, pp. 103–110.

Chodak, M., Khanna, P., Beese, F., 2003. Hot water extractable C and N in relation to microbiological properties of soils under beech forests. Biology and Fertility of Soils 39 (2), 123–130.

Chodak, M., Khanna, P., Horvath, B., Beese, F., 2004. Near infrared spectroscopy for determination of total and exchangeable cations in geologically heterogeneous forest soils. Journal of Near Infrared Spectroscopy 12, 315–324.

Chou, P.-H., Lan, T.-H., 2013. The role of near-infrared spectroscopy in Alzheimer's disease. Journal of Clinical Gerontology and Geriatrics 4 (2), 33–36.

Christy, C.D., 2008. Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy. Computers and Electronics in Agriculture 61 (1), 10–19.

Chung, H., 2007. Applications of Near-Infrared Spectroscopy in Refineries and Important Issues to Address. Applied Spectroscopy Reviews 42 (3), 251–285.

Dunn, B.W., Batten, G.D., Beecher, H.G., Ciavarella, S., 2002. The potential of near-infrared reflectance spectroscopy for soil analysis — a case study from the Riverine Plain of south-eastern Australia. Aust. J. Exp. Agric. 42 (5), 607.

Esbensen, K., Guyot, D., Westad, F., Houmoller, L.P., 2012. Multivariate data analysis-in practice: An introduction to multivariate data analysis and experimental design, 5th ed. ed. CAMO, Oslo, xx, 598.

García-Sánchez, F., Galvez-Sola, L., Martínez-Nicolás, J.J., Muelas-Domingo, R., Nieves, M., 2017. Using Near-Infrared Spectroscopy in Agricultural Systems, in: Kyprianidis, K.G., Skvaril, J. (Eds.), Developments in Near-Infrared Spectroscopy. InTech.

Gogé, F., Joffre, R., Jolivet, C., Ross, I., Ranjard, L., 2012. Optimization criteria in sample selection step of local regression for quantitative analysis of large soil NIRS database. Chemometrics and Intelligent Laboratory Systems 110 (1), 168–176.

Guerrero, C., Stenberg, B., Wetterlind, J., Viscarra Rossel, R.A., 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.

Filzmoser, P., Garrett, R.G., Reimann, C., 2005. Multivariate outlier detection in exploration geochemistry. Computers and Geosciences 31, 579-587.

Hodge, G.R., Woodbridge, W.C., 2010. Global near Infrared Models to Predict Lignin and Cellulose Content of Pine Wood. Journal of Near Infrared Spectroscopy 18 (6), 367–380.

ICRAF, 2015. A globally distributed soil spectral library visible near infrared diffuse reflectance spectra URL

 http://worldagroforestry.org/sites/default/files/Description_ICRAFISRIC

Jamrógiewicz, M., 2012. Application of the near-infrared spectroscopy in the pharmaceutical technology. Journal of pharmaceutical and biomedical analysis 66, 1–10.

Knadel, M., Deng, F., Thomsen, A., Greve, M., 2012. Development of a Danish national Vis-NIR soil spectral library for soil organic carbon determination, in: Minasny, B., Malone, B.P., McBratney, A.B. (Eds.), Digital soil assessments and beyond. Proceedings of the 5th Global Workshop on Digital Soil Mapping, 2012, Sydney, Australia, 10-13 April 2012. CRC Press, Boca Raton, Fla., pp. 403–408.

Ladoni, M., Bahrami, H.A., Alavipanah, S.K., Norouzi, A.A., 2010. Estimating soil organic carbon from soil reflectance: A review. Precision Agriculture 11 (1), 82–99.

Long, J., Wang, K., Yang, M., Zhong, W., 2018. Rapid crude oil analysis using near-infrared reflectance spectroscopy. Petroleum Science and Technology 37 (3), 354–360.

Mouazen, A.M., Karoui, R., Baerdemaeker, J. de, Ramon, H., 2005. Classifi cation of soil texture classes by using soil visual near infrared spectroscopy and factorial discriminant analysis techniques. Journal of Near Infrared Spectroscopy 13, 231–240.

Nduwamungu, C., Ziadi, N., Parent, L.-É., Tremblay, G.F., Thuriès, L., 2009. Opportunities for, and limitations of, near infrared reflectance spectroscopy applications in soil analysis: A review. Canadian Journal of Soil Science 89 (5), 531–541.

Nocita, M., Stevens, A., van Wesemael, B., Aitkenhead, M., Bachmann, M., Barthès, B., Ben Dor, E., Brown, D.J., Clairotte, M., Csorba, A., Dardenne, P., Demattê, J.A.M., Genot, V., Guerrero, C., Knadel, M., Montanarella, L., Noon, C., Ramirez-Lopez, L., Robertson, J., Sakai, H., Soriano-Disla, J.M., Shepherd, K.D., Stenberg, B., Towett, E.K., Vargas, R., Wetterlind, J., 2015. Soil Spectroscopy: An Alternative to Wet Chemistry for Soil Monitoring, in: Sparks, D.L. (Ed.), Advances in agronomy. Volume 132, vol. 132. Advances in Agronomy 132. Academic Press, Amsterdam, pp. 139–159.

Norris, K.H., Hart, J.R., 1965. Direct spectrophotometric determination of moisture content of grain and seeds 4 (Proc. 1963 International Symposium on Humidity and Moisture), 19–25.

Pikuła, D., Rutkowska, A., 2014. Effect of leguminous crop and fertilization on soil organic carbon in 30-years field experiment. Plant, Soil and Environment 60 (11), 507-511.

Pimentel, D., 1991. Global warming, population growth, and natural resources for food production. Society & natural resources 4 (4), 347–363.

Porep, J.U., Kammerer, D.R., Carle, R., 2015. On-line application of near infrared (NIR) spectroscopy in food production. Trends in Food Science & Technology 46 (2), 211–230.

Rinnan, Å., van den Berg, F., Engelsen S.B., 2009. Review of the most common pre-processing techniques for near-infrared spectra. Trends in Analytical Chemistry 28 (10), 1201-1222,

Roberts, C.A., Workman, J., Reeves, J.B., 2004. Near-infrared spectroscopy in agriculture: Craig A. Roberts, Jerry Workman, James B. Reeves ; managing editor: Lisa Al-Amoodi. Agronomy no. 44. American Society of Agronomy, Madison, Wis.

Roberts, J.J., Cozzolino, D., 2017. Wet or dry?: The challenges of NIR to analyse soil samples. NIR news 28 (4), 3–5.

Sakudo, A., 2016. Near-infrared spectroscopy for medical applications: Current status and future perspectives. Clinica chimica acta; international journal of clinical chemistry 455, 181–188.

Shepherd, K.D., Walsh, M.G., 2002. Development of Reflectance Spectral Libraries for Characterization of Soil Properties. Soil Science Society of America Journal 66 (3), 988.

Siebielec, G., McCarty, G.W., Stuczynski, T.I., Reeves III, J.B., 2004. Near- and mid-infrared reflectance spectroscopy for measuring soil metal content. Journal of Environmental Quality 33 (6), 2056–2069.

Smidt, E., Böhm, K., Schwanninger, M., 2011. The Application of FT-IR Spectroscopy in Waste Management, Fourier Transforms - New Analytical Approaches and FTIR Strategies, in: Nikolić, G.S. (Ed.), Fourier transforms. New analytical approaches and FTIR strategies. InTech, Rijeka, Croatia.

Sørensen, L.K., 2002. True accuracy of near infrared spectroscopy and its dependence on precision of reference data. Journal of Near Infrared Spectroscopy 10, 15–25.

Soriano-Disla, J.M., Janik, L.J., Viscarra Rossel, R.A., Macdonald, L.M., McLaughlin, M.J., 2013. The Performance of Visible, Near-, and Mid-Infrared Reflectance Spectroscopy for Prediction of Soil Physical, Chemical, and Biological Properties. Applied Spectroscopy Reviews 49 (2), 139–186.

Stenberg, B., Viscarra Rossel, R.A., Mouazen, A.M., Wetterlind, J., 2010. Visible and Near Infrared Spectroscopy in Soil Science, in: Sparks, D.L. (Ed.), Advances in agronomy, vol. 107, 1st ed. ed. Advances in Agronomy. Academic Press, San Diego, CA, pp. 163–215.

Stevens, A., Nocita, M., Tóth, 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), e66409.

Tóth, G., Jones, A., Montanarella, L., 2013. The LUCAS topsoil database and derived information on the regional variability of cropland topsoil properties in the European Union. Environmental monitoring and assessment 185 (9), 7409–7425.

Vasques, G.M., Demattê, J.A.M., Viscarra Rossel, R.A., Ramírez-López, L., Terra, F.S., 2014. Soil classification using visible/near-infrared diffuse reflectance spectra from multiple depths. Geoderma 223-225, 73–78.

Viscarra Rossel, R.A., Behrens, T., Ben-Dor, E., Brown, D.J., Demattê, J.A.M., Shepherd, K.D., Shi, Z., Stenberg, B., Stevens, A., Adamchuk, V., Aïchi, H., Barthès, B.G., Bartholomeus, H.M., Bayer, A.D., Bernoux, M., Böttcher, K., Brodský, L., Du, C.W., Chappell, A., Fouad, Y., Genot, V., Gomez, C., Grunwald, S., Gubler, A., Guerrero, C., Hedley, C.B., Knadel, M., Morrás, H.J.M., Nocita, M., Ramirez-Lopez, L., Roudier, P., Campos, E.R., Sanborn, P., Sellitto, V.M., Sudduth, K.A., Rawlins, B.G., Walter, C., Winowiecki, L.A., Hong, S.Y., Ji, W., 2016. A global spectral library to characterize the world's soil. Earth-Science Reviews 155, 198–230.

Viscarra Rossel, R.A., McBratney, A.B., 2008. Diffuse reflectance spectroscopy as a tool for digital soil mapping, in: Hartemink, A.E., McBratney, A.B., Mendonça-Santos, M.d.L. (Eds.), Digital soil mapping with limited data. Springer, Dordrecht, London, pp. 165–172.

Viscarra Rossel, R.A., Webster, R., 2011. Discrimination of Australian soil horizons and classes from their visible-near infrared spectra. European Journal of Soil Science 62 (4), 637–647.

Wenjun, J., Zhou, S., Jingyi, H., Shuo, L., 2014. In situ measurement of some soil properties in paddy soil using visible and near-infrared spectroscopy. PloS one 9 (8), e105708.

Williams, P.C., 1973. The application of the Neotec grain quality analyzer to the nalyses of cereal grains and oilseeds. Cereals Science Today 18, 284–285.

Zornoza, R., Guerrero, C., Mataix-Solera, J., Scow, K.M., Arcenegui, V., Mataix-Beneyto, J., 2008. Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in Mediterranean soils. Soil Biology & Biochemistry 40 (7), 1923–1930.

5. Discussion of other scientific and research achievements

Isotopic Geochemistry

During my doctoral studies at the Institute of Geological Sciences – Polish Academy of Sciences, I was involved with studying the behaviour of U and Th isotopes in „dirty" calcites (see the list of oral presentations on the subject in appendix 3). The goal of my work was to understand U and Th sorption/desorption behaviour on clay minerals (montmorillonite, illite, and kaolinite) in lake sediments. Unlike clean carbonates (ideal material for dating with that method, Bard et al., 1990), dirty carbonates contain ^{230}Th which is not related to the *in situ* decay of ^{234}U (Schwarcz and Latham, 1989). The challenge was to use that knowledge to try dating those materials with the U-series dating method (Edwards et al., 1986). The attempt was successful and the new methodology used in papers **Gradziński et al (2002)**, **Tatur et al., (2002)** and **Debaene (2003)** to date stromatolites from Patagonia, speleothems from Slovakia and lake sediments from Poland. Three original methods for building „isochrons" were produced and used to give the correct datation of these materials.

The exhaustive investigation of clay minerals and lake sediments was latter a source of inspiration for me for me to work with organic sediments (e.g. gyttia) and write the grant proposal (Effect of soil mineralogy, organic matter, and environment on mineral soil stability, description below). In the future, I would like to use the understanding of the high mobility of ^{234}U and immobility of its daughter isotope ^{230}Th to investigate organic soil profile and eventually precisely date the different layers.

Bard E., Hamelin B., Fairbanks R.G. (1990) U-Th ages obtained by mass spectrometry in corals from Barbados: sea level during the past 130,000 years. Nature. 346, 456-458.

Debaene, G., 2003. Uranium-series dating of marly sediments: applications to Jaroszów fossil lake (SW Poland). Geochronometria, 22, 15-26.

Edwards R.L., Chen J.H., Wasserburg G.J. (1986) ^{238}U-^{234}U-^{230}Th-^{232}Th systematics and the precise measurement of time over the past 500,000 years. Earth and Planetary Science Letters. 81, 175-192.

Gradziński M., Hercman H., Bella P., **Debaene G**., Nowicki T., 2002: Tmavé laminácie v sintrových nátekoch jaskyne Domica ako indikátor akivít pravekých ľudí (Dark coloured laminae within speleothems of the Domica Cave as an indicator of the prehistoric men activity). Slovenský Kras, 40, 41-48.

Schwarcz H.P. and Latham A.G. (1989) Dirty calcites. 1. Uranium-series dating of contaminated calcite using leachate alone. Chemical Geology. 80, 35-43.

Tatur, A., del Valle, R., Bianchi, M. M., Outes, V., Villarosa, G., Niegodzisz, J., & **Debaene, G**., 2002. Late Pleistocene palaeolakes in the Andean and Extra-Andean Patagonia at mid-latitudes of South America. Quaternary International, 89(1), 135-150.

Effect of soil mineralogy, organic matter, and environment on mineral soil stability

In 2013, I obtained a funding for a research project from the National Science Centre "Influence of mineralogy of soils, organic matter and some soil environment on soil stability Agreement No. UMO-2012/07/B/ST10/04387. The project was conducted from 2013 to 2017.

The first goal of the research was to study the effect of soil pH, mineralogy of the soil colloidal fraction, and of the organic matter composition on soil stability. Soil stability was estimated by readily dispersible clay (RDC) content in function of the soil water potential. The effect of the colloidal fraction on the absorption of trace metals was also investigated. A pot experiment was carried out for these purposes. The soil collected for the experiment was a loamy sand from Osiny experimental station (IUNG-PIB) near Puławy. Half pots were treated with $CaCO_3$. As a result, two soils pH were investigated (neutral and slightly acid). To the pots were also added clay minerals (montmorillonite, kaolinite, illite) at three different concentrations (0.5%, 2%, and 5%) and trace metals (Ni, Zn, Cd, Pb). The experiment has proved that the addition of trace metals (i.e. Ni, Zn, Cd, Pb in a quantity 3 times higher than the Polish norm in the form of salt) in soil samples (without clay minerals) has no effect on the content of RDC. The addition of montmorillonite and illite (5%) increased the pH to a value of 7.5. This was not the case for kaolinite. The addition of $CaCO_3$ to the test samples as well as the addition of clay minerals caused a decrease in the hydrolytic acidity (Hh) from 2 to 0.75 cmol (+) / kg. With 5% of illite and montmorillonite, soil samples both with and without $CaCO_3$ showed a significant increase in RDC when contaminated with trace elements in comparison to clean samples (for example, an increase from 0.28 to 0.67 g/100g for a mixture of soils with 5% of montmorillonite). In the case of kaolinite, in the majority of pots, the reverse relation was observed. With 2% of kaolinite, with and without $CaCO_3$, and with a kaolinite content of 5% without calcium carbonate, the amount of RDC was higher in samples without trace elements.

The second objective of the project was to build a prediction model for RDC based on spectra acquired with VIS-NIR spectroscopy. The analysed samples were scanned in the 350-2500 nm range. After the initial pre-treatment of the spectra (e.g. moving average and SNV Standard Normal Variate – a procedure for spectral normalisation), it was possible to determine soil texture based on the spectra (after multivariate calibration, partial least square regression). This allowed building a prediction model of RDC, due to the relationship between RDC and clay minerals. The result of the PLS regression was $r^2 = 0.744$ and RMSEcv = 0.01. This means that the RDC content is provided with an accuracy of 0.01g / 100g. Moreover, it was also shown that the water potential had no effect on the model, provided that samples were scanned at the same moisture content.

The above described model has been successfully applied to the prediction of RDC in soil samples from Baborówko and Grabów, which presented a similar texture. The model described above was used but, additionally, the "spiking" method was implemented. That means that additional local samples were added to the model in order to increase the prediction accuracy.

Spatial variability of RDC was also measured. A map was obtained using the Veris® mobile platform equipped with a VIS-NIR sensor, an electrical conductivity (EC) sensor and a temperature sensor. The validation of the map was carried out on independent samples, as well as with comparison to maps of SOC, clay fraction, EC, NDVI and an orthophotography of the field. The best PLSR models were for RDC and clay fraction estimation (e.g. $r^2 = 0.74$, RMSEcv = 0.03). The models were obtained by fusing the VIS-NIR data with EC measurement (deep EC).

The results were presented in many oral and poster presentations (listed in Annexe 8) and formed the basis for writing the publications: **A3**, **A5** and **Ukalska-Jaruga et al., 2018**.

The most important outcomes of the project include:

1. Demonstrating that among the clay minerals montmorillonite and illite had the highest ability to absorb (immobilise) trace metals, while kaolinite was the lowest.

2. Demonstrating the important role of organic matter in soil stabilisation. In soils with higher SOC content, the amount of RDC was lower.

3. Demonstrating that soil stability in deciduous forests is higher than soil stability in mixed and coniferous forests. The least stable soils were found on arable land with intensive agricultural production.

4. Mapping soil stability at the field scale. Such maps may be potentially useful when planning, e.g., protective measures against unfavourable phenomena that may occur in the soil environment. Precision agriculture can be the recipient of this type of research when planning land management operations, especially in areas exposed to erosion.

5. Conducting an attempt to predict hydrolytic acidity (Hh) using VIS-NIR spectroscopy. Hydrolitic acidity is an important parameter for agriculture since it is used for calculating liming requirements.

Ukalska-Jaruga, A., **Debaene, G.**, & Smreczak, B., 2018. Particle and structure characterization of fulvic acids from agricultural soils. Journal of Soils and Sediments, 18(8), 2833-2843.

Precision agriculture

Precision agriculture (PA) or site specific crop management (SSCM) is an agriculture management system based on observing, measuring and responding to field variation. The purpose of PA is to define the decision support system (DSS) for the whole farm management in order to optimise returns on inputs while preserving resources (Pedersen and Lind, 2017). In precision agriculture, multiple sensors can be used simultaneously. One of the goals of my research was to use VIS-NIR spectroscopy for PA, since the method proved to be very sensitive to changes in soil physical and chemical properties. Soil variability can also be analysed by means of traditional remote sensing and other sensors, as described in **Pudełko et al., (2012)**, which I am a co-author.

Using aerial photographs, satellite images, electrical conductivity, and yield maps, the potential field productivity in Baborówko (IUNG-PIB experimental station) was modelled. The resulting maps represent different areas of productivity in the field. The maps have a resolution corresponding to agricultural practices and could be used by farmers. Further investigations of the uses of several sensors for PA were presented in Pecio et al., (2016). In that paper, four new fields from Baborówko experimental station were investigated (78.5 ha). 154 samples from a grid sampling were used to successfully predict SOC, pH, available (Mg, P, K), and fine particles (<0.02 mm). That fraction is important for the definition of the agronomic categories (**Jadczyszyn et al., 2017**). The methodology used was presented in paper (**A2**). The Veris® mobile platform was also used to produce *on-the-go* maps for these properties (in that case, several session of scanning were necessary due to the size of the fields). All maps produced were very similar to the original (reference) maps demonstrating that VIS-NIRS method is suitable for mapping both in laboratory and *on-the-go*.

Jadczyszyn, J., Niedźwiecki, J., & **Debaene, G.**, 2017. Analysis of Agronomic Categories in Different Soil Texture Classification Systems. Polish Journal of Soil Science, 49(1): 61-72.

Pecio A., Niedźwiecki, J., & **Debaene, G.**, 2016. Rozpoznanie pola podstawą precyzyjnego nawożenia roślin. Studia i Raporty IUNG-PIB, 48(2): 105-129.

Pedersen, S.M., Lind, K.M., 2017. Precision agriculture: Technology and economic perspectives. Progress in Precision Agriculture. Springer, Cham, 1 Online-Ressource (xii, 276 Seiten).

Pudełko, R., Niedźwiecki, J., **Debaene, G.**, Igras, J., & Kubsik, K., 2012. The remote sensing assessment of potential productivity of a field with soil spatial variability. Journal of Food, Agriculture and Environment, 10(2), 790-793.

New research interest:

1. Wettability of biological materials (and organic soils)

Since 2015, I have been conducting a scientific cooperation with the Department of Environmental Improvement at the Faculty of Civil and Environmental Engineering (Warsaw University of Life Sciences), which resulted in the following publications:

1. **Papierowska et al., 2018a** (IF: 3.740 in 2017)
2. **Papierowska et al., 2018b** (IF: 1.782 in 2017)

Ad.1

The first paper examines the comparison of methods for assessing the wettability of organic and mineral-organic soils. This article deals with the important issue of non-wettability (hydrophobicity) of soils. This is a very important concern due to the fact that the hydrophobicity has several consequences for the environment. First of all, it limits the infiltration capacity of soils, influences the process of soil water movement, significantly affects the increase of surface runoff, induces a preferential flow of water soils and leads to the occurrence of soil water erosion.

The purpose of this article was to determine the compatibility of methods for assessing soil hydrophobicity using statistical methods. The use of a weighted kappa coefficient (κw) was proposed, which allowed estimating the value of the contact angle based on simple tests such as the of water drop penetration time (WDPT) and molarity of an ethanol drop (MED). Laboratory wettability of 106 organic and mineral-organic soils collected from various locations in north-eastern Poland was measured. Soil wettability was carried out using four methods, i.e. the sessile drop method, the Wilhem plate method, the water drop penetration time (WDPT), and the alcohol molarity test (MED).

The obtained results showed that the weighted kappa coefficient, accepted as a measure of observer compliance, indicates a strong dependence ($\kappa w = 0.84$) between the mean contact angle measured with the sessile drop method and the median time value of water drop infiltration. It has also been found that hydrophilic samples with infiltration time of less than 5 seconds have a mean contact angle below 40°, while samples with extremely hydrophobic properties with an infiltration time greater than 3600s have contact angle values higher than 130°. The investigation has shown that the WDPT and MED methods (which are simple and do not require expensive equipment) can be a good estimator of contact angle value for the determination of wettability both in laboratory and field conditions.

Ad. 2

The second paper concerns the determination of wettability of leaves of deciduous trees and shrubs species commonly found in Poland.

The main objective of the research was to investigate differences in leaf wettability of 19 species of trees and shrubs (including 16 deciduous trees and 3 shrubs) commonly found in Polish forests in the temperate climate zone. Twenty undamaged leaves were selected for each species, and laboratory tests were carried out to determine the wettability based on the wet contact angle measurements using the CAM 100 optical goniometer. Measurements were done for the upper and lower side of the leaves. The values of the contact angle showed a decrease in the contact angle value during the measurement (measurement of the contact angle every 1 second during 2 minutes) and fluctuated from 60° to 140° depending on the species and the analysed leaf side. It was shown that the lower side of the leaf was more hydrophobic in 14 out of 19 analysed species. The analysed leaves were classified from highly wettable to highly non-wettable, probably depending on the strategy of plant survival. In order to investigate the influence of the species on the value of the contact angle, cluster analysis was used and these results were compared with the main component analysis (PCA) proposed by me.

Papierowska, E., Matysiak, W., Szatyłowicz, J., **Debaene, G.**, Urbanek, E., Kalisz, B., Łachacz, A. 2018a. Compatibility of methods used for soil water repellency determination for organic and organo-mineral soils. Geoderma, 314, 221-231.

Papierowska, E., Szporak-Wasilewska, S., Szewińska, J., Szatyłowicz, J., **Debaene, G.**, Utratna, M. 2018b. Contact angle measurements and water drop behavior on leaf surface for several deciduous shrub and tree species from a temperate zone. Trees, 32(5), 1253-1266.

2. Examination of soil extracts by means of visible and near-infrared spectroscopy

The effect of soil extracts investigation using VIS-NIR spectroscopy is investigated in **Ukalska-Jaruga et al., (2018)** published in the Journal of Soils and Sediments (IF 2017, 2.627).

All paper described before have clearly demonstrated that reflectance spectroscopy is a tool of choice to investigate mineral soils. The method, however, has never been used to analyse soil extract like the ones from the isolation and purification of fulvic or humic acids. With the use of a cuvette especially build to use with our spectroradiometer PSR-3500 and with fibre optics; it is now possible to measure absorbance of soil extracts. Therefore, the instrument was used to measure the spectroscopic properties of fulvic acids (FAs) from extract obtained by isolation and purification of fulvic acids (IHSS method: Hayes 1985; Swift, 1985). The spectra allowed us to illustrate that FAs behave as anionic or cationic heterogeneous ligand with many methyl, ethylene, amine, and carboxyl groups of different strength. That observation and the other analysis undertaken in this study (organic carbon

content, particle size diameter, and zeta potential) allowed us to obtain a detailed description of FAs from agricultural soils. Their complex and heterogeneous nature was confirmed by new methods (VIS-NIRS and zeta potential). Moreover, a main finding was that FAs mainly exists as small molecules that form molecular aggregates. The charge and behaviour of these aggregates will affect soil properties.

The novelty of the method and the interesting results obtained are the reason why I will continue to investigate soil extract in the future.

Hayes M.B.H. 1985. Extraction of humic substances from soil. In: Aiken, G.R., McKnight, D.M., Wershaw, R.L., MacCarthy, P. (eds) Humic substances in soil, sediment, and water. Geochemistry, isolation, and characterization. Wiley-Interscience, New York, pp 329-362.

Ukalska-Jaruga, A., **Debaene, G.**, & Smreczak, B., 2018. Particle and structure characterization of fulvic acids from agricultural soils. Journal of Soils and Sediments, 18(8), 2833-2843.

Swift, R.S. 1985. Fractionation of soil humic substances. In: Aiken, G.R., McKnight, D.M., Wershaw, R.L., MacCarthy, P. (eds) Humic substances in soil, sediment, and water. Geochemistry, isolation, and characterization. Wiley-Interscience, New York, pp 387-408.

3. Analysis of organic soils by visible and near-infrared spectroscopy

In my work so far, I have focused a lot on predicting SOC content in arable soils, because it is a key parameter for assessing soil quality. This is an interesting property because SOC content is variable both time and space. My previous work has shown that the use of spectroscopy in the wavelength range (350-2500 nm) gives good results in predicting SOC content of mineral soils. Examples of models results ($r^2 = 0.68$, RMSE = 0.11, $r^2 = 0.71$, RMSE = 0.12) show that in the case of arable soils the quality of predictions is good and similar to laboratory analyses (**Niedźwiecki and Debaene, 2013**). These results have become an inspiration to use the VIS-NIR spectroscopy to study SOC content in organic soils, i.e. peat, gyttia, etc. Currently, I am the head of a statutory research program (3.16 " Use of chemometric methods to evaluate the quality of agricultural organic soils or originating from organic soils ") on this subject. The investigation of organic soil spectra is important because there are no results for this type of soils anywhere in the national and international literature. Currently, together with my team, I am also conducting a research on the use of spectral methods for classifying organic soils for the development of the national soil spectral library (NSSDB). Preliminary results of the research were presented in **Niedźwiecki et al., 2017**. The extension of the database with organic soils is important for mapping the territory of Poland. Adding organic soils to the database will help to improve the accuracy of prediction maps presented in paper (**A6**), and which so far were based only on data for mineral soils.

Niedźwiecki, J., **Debaene G.**, 2013. Nowoczesne chemometryczne metody oznaczania substancji organicznej w glebach. Studia i Raporty IUNG-PIB, 35(9): 199-212.

Niedźwiecki J., **Debaene, G.**, Smreczak, B., Łysiak M., Ukalska-Jaruga, A., 2017. Wstępne badania nad wykorzystaniem metod spektralnych do klasyfikacji utworów organicznych na potrzeby spektralnej biblioteki gleb Polski. Studia i Raporty IUNG-PIB, 54(8): 41-55.

Summary of other scientific and research achievements

The most important achievements of the other scientific activities include:

1) Research on the wettability of materials of biological origin. The phenomenon of repellency (hydrophobicity) is common and has many consequences for the environment.

2) Studies of soil extracts with visible and near-infrared spectroscopy proved that the method can characterise fulvic acids.

3) Application visible and near-infrared spectroscopy for the study of organic soils.

4) Extending the spectral library of Polish soils (NSSDB) with organic soils to more accurately map the country (Poland).

6. Further plans for scientific development

In the future, I would like to explore the following issues:

1) Improving predictive models of soil properties by e.g. using neural network (NN) on the available spectral library.

2) Further expand the NSSDB with new spectra for mineral soils, but also adding spectra of organic soils (many spectra have already been included) and spectra from forest soils. Based on the extended database, create more accurate prediction maps of different soil properties.

3) Test whether it is possible with VIS-NIR spectroscopy to measure changes of soil organic carbon content with time and try to use the method to calculate the amount of soil organic carbon stock. This could be useful for monitoring climate change.

4) Based on the spectra, search for indicators of soil quality useful for their classification. The goal is to obtain a tool that would quickly improve the classification of profiles using e.g. the WRB method.

5) Conduct further research on the use of spectral methods for testing liquids (soil extracts). Currently, a work is underway on the application of the method for the determination of humins and humic acids in addition to fulvic acids.

6) Use VIS-NIR spectroscopy for estimating the content of polycyclic aromatic hydrocarbons (PAHs) in soil according to the allowed amount specified by the Regulation of the Ministry of the Environment (1 September 2016) with a view to assess the level of contamination (Journal of Laws of 2016, item. 1395). An article on this subject is being prepared.

7) Use VIS-NIR spectroscopy to study black carbon (a paper on this topic is currently under review).

8) Use the method to study the phenomenon of hydrophobicity of biological materials (soil, plant material).

7. List of publications not included in the scientific achievement

In addition to the publications constituting the scientific achievement, I am the author or co-author of other publications. Additionally, one article is currently under review. The list of all the papers with IF and MNiSW scores at the year of publication, number of citations according to the Web of Science and Scopus databases can be found in Annexe 8.

In addition to these articles, I am the author and co-author of 2 peer-reviewed conference materials and 2 popular science papers (Annexe 8).

8. Participation in research projects

During the years 2013-2017, I was the supervisor of a research project from the National Centre for Research (NCN) "The influence of soil mineralogy, organic matter and environment on soil stability: grant nº 2012/07/B/ST10/04387 ". I also took part in other research projects.

In the years 2001-2002, I was the recipient of research project no. 6 PO4D 012 21 KBN entitled "U-series dating of lake sediments". In the years 2000-2001 I received a French government's grant (Bourse de coopération scientifique et culturelle France - Pologne 2000).

In 2010-2012, I was a participant in the project "Methodological preparation of soil science issues resulting from the planned Directive of the European Parliament and Council establishing a framework for soil protection and amendment. Directive 2004/35/EC (COM (2006) 232, final) with the creation of IT tools for analysis of soil threats Agreement 5/2010 with PGI-NRI. I was also a participant in the NCN grant: N N310 079836 - Evaluation of spatial variability of physical and chemical soil properties with innovative *on-the-go* methods (ECOMP).

My competences were also utilised in an European Project carried out at IUNG-PIB (as a participant) FP7-ENV- Project „Environmental Optimization of irrigation management with the combined use and integration of high precision satellite data, Advanced Modelling, Process Control and Business Innovation" ENORASIS (Grant Agreement 282949).

During my work at IUNG-PIB, I actively participated in the preparation of the following proposals:

1) Stalenga J., **Debaene G**. 2014. Preparing part of the proposal for: „Building Ecological Recycling Agriculture and Societies in Burundi" (Acronym: BERAS-Burundi) submitted to "Sustainable Food Security" **Horizon 2020** (2020-SFS-2014-1).

2) Pecio A., **Debaene G**., Niedźwiecki J., Pudełko R., Wawer R. 2014. Opracowanie metod szybkiej oceny (proximal sensing) właściwości gleby i łanu roślin z wykorzystaniem metod teledetekcji i narzędzi rolnictwa precyzyjnego na potrzeby zarządzania zrównoważoną produkcją roślinną i kontroli środowiska rolniczego. Scientific partner. FARMCOM. **BIOSTRATEG** Sector Program, National Centre for Research and Development. Proposal no 270284.

3) Siebielec G., **Debaene G**. 2015. Sustainable soil-improving cropping systems - CROP4SOIL. Program **Horizon 2020**. Preparation of the proposal as Package leader for IUNG-PIB (ID: SFS-02b-2015).

4) Pecio A., **Debaene G**., Niedźwiecki J., Pudełko R., Wawer R. 2015. FARMCOM innowacyjny system zarządzania gospodarstwem rolnym i produkcją roślinną z wykorzystaniem zintegrowanych metod informatycznych, teledetekcyjnych i narzędzi rolnictwa precyzyjnego na potrzeby zrównoważonej produkcji rolniczej i kontroli środowiska. Participation in the proposal as Package Leader. **BIOSTRATEG** Sector Program, National Center for Research and Development. Proposal nᵒ 298796.

5) Kowalik M., **Debaene G**. 2015. Analysis of pore structure related to soil fauna abundance and activity using 'in-situ" MIR spectroscopy method. 2015/18/E/ST10/00281. **NCN**. SONATA Bis

6) Niedźwiecki J., **Debaene G**., Miturski T. 2015. Wykorzystanie pomiarów hiperspektralnych gleb do charakterystyki i klasyfikacji profili glebowych. **NCN**, program OPUS 8, ID 296071.

7) Kowalik M., **Debaene G**. 2016. Spectrometry VIS-NIR as a tool to study dynamic changes of soil biodiversity on different agricultural land use. **NCN**, Foundation for Polish Science, Proposal no 0052.

8) Ukalska-Jaruga, A., **Debaene G**. Application for financing a research project implemented as part of international cooperation. Title: Wpływ preparatów humusowych na fizjologię kukurydzy oraz właściwości mikrobiologiczne gleby i glebowej materii organicznej. **NCN**, program HARMONIA 9, proposal no 2018/30/M/ST10/00043.

I am currently a subcontractor in the grant (SONATA 11, 2016/21/D/ST10/01947 – **NCN**) Leader: dr Piotr Bartmiński from the Marie Curie-Skłodowska University in Lublin.

Title of the grant: "Application of hyperspectral imagery for the growing season parametrisation in the context of soil covers properties". I also took part in IUNG-PIB statutory research conducted by Piotr Koza, MA, which focused on the same subject. In both collaborations, I am/was involved with the combination of satellite images with statistical method PLS and soil spectral analysis.

In the field of remote and proximal sensing, I am currently working closely with Dr Maria Knadel from Aarhus University, Denmark and with the Polish Academy of Sciences – Space Research Centre, which resulted in a joint poster at a conference in Chile (IUSS). Poster title: Space-borne remote spectroscopy for soil properties determination. Authored by: Knadel M, Debaene G, et al. 2019.

Statutory research participation

I am and I was the supervisor of the following programs:

1. Project n° 4.02. ,,Use of chemometric methods to predict mineral soil properties from Poland", years 2012 – 2014.

2. Project n° 3.16 ,,Use of chemometric methods to evaluate the quality of agricultural organic soils or originating from organic soils", years 2017-2019.

I participated as a researcher in the following statutory programs:

1. Research project n° 4.2.5. Zakwaszenie gleb jako czynnik zagrożenia dla jakości płodów rolnych oraz wód podziemnych i powierzchniowych.

2. Research project n° 4.1.6. Ocena przydatność zaawansowanych metod do określania przestrzennej zmienność gleb użytkowanych rolniczo.

3. Research project n° 4.3.4. Ocena zawartości pierwiastków śladowych w ziarnie pszenicy w zależności od właściwości gleby i innych czynników w ujęciu przestrzennym.

4. Research project n° 2.3.9. Rozpoznanie przestrzennej zmienności właściwości fizycznych i agrochemicznych gleb na polu uprawianym metodą uproszczoną.

5. Research project n° 2.3.16. Ocena ilości i jakości glebowej materii organicznej zależnie od nawożenia oraz gatunków w zmianowaniu.

6. Research project n° 4.1.11. Doskonalenie zdalnych metod oceny glebowych i struktury krajobrazu, opartych na wykorzystaniu zdjęć satelitarnych i lotniczych.

7. Research project nᵒ 4.3.6. Wielkoskalowa parametryzacji modeli symulacyjnych jako narzędzi oceny oddziaływania rolnictwa na środowisko.

9. Reviews in scientific Journals

9.1. Reviews in Journals with Impact Factor:

2019: Precision Agriculture – 1

2018: Geoderma – 4, International Agrophysics – 1

2017: Journal of Soils and Sediments – 2; Soil Science Society America Journal – 1; Geoderma – 1; Nature – Scientific Reports – 1

2016: Computers and Electronics in Agriculture – 1; Plant and Soil – 1

2015: Journal of Environmental Management – 2

2014: Plant and Soil – 1; Geoderma – 2; European Journal of Soil Science – 1; Environmental Monitoring and Assessment – 1

2012: Philippine Agricultural Scientist – 1

9.2. Reviews in Journals without Impact Factor:

2019: Soil Science Annual – 1; Mathematics – 1

2018: Soil Science Annual – 1

2017: Polish Journal of Soil Science – 1

9.3.Review ISO standard:

Standard ISO : nᵒ ISO/DIS 17184: Soil quality – Determination of carbon and nitrogen by near-infrared spectrometry (NIRS).

10. Internships and trainings

I participated in the following trainings and field workshops:

1. „R-Data Annalyses, Machine learning – Introducion" – teaching provided by Comarch SA. 17-19.12.2018.

2. „Soils formed from gyttias and murshs", – field workshop from the Genesis Commission, Classification and Soil Cartography, Polish Society of Soil Science. Olsztyn and Brodnickie, Rytebłota, Poland, 3-6.10.2018.

3. „Technological transfer and Commercialisation of Scientific Results". – teaching provided by CoWinners Sp. Z.o.o., Puławy, Poland, 19-21.06.2018.

4. „Prediction and Analyses from Time-Series" – teaching from StatSoft Poland. 1-2.07.2015.

5. „Soil Spectroscopy: The Present and Future of Soil Monitoring", – International Workshop from F.A.O., Rome, Italy. 4-6.12.2013.

6. „Use of several regression methods with STATISTICA" – teaching from StatSoft Poland. 14.11.2013.

7. „Introduction to the processing of satellite images using the ENVI digital image analysis software" – teaching provided by SmallGis. Puławy, Poland, 24-26.09.2012.

8. „The use of statistics with results from environmental research - basic methods" – teaching provided by StatSoft Poland. 13.12.2011.

In addition, I completed an internship with Dr. Maria Knadel at the Department of Agroecology at the University of Aarhus in Denmark on 26-30 October 2011.

11. Participation in Scientific conferences

I took part in numerous international and national conferences, where I presented oral presentations and posters. A list of oral presentations and posters divided into years and into national and international conferences is available in Annexe 8.

12. Teaching and Organisational activities

From 2012 until now, I am a member of the Puławy Division of the Polish Society of Soil Science. Currently, I am the chairman of the board for the audit committee. Since 2016, I am a member of the Soil Physics Commission of the Polish Society of Soil Science, PTG. I also belong to the International Union of Soil Sciences (IUSS) and to The International Soil Carbon Network (ISCN).

On May 27, 2009, I conducted a popular science lecture entitled "Wine - tourist routes of France" at the conference "Tourism - Teaches, Plays, Educates" at the University of Social and Natural Sciences Vincent Pol in Lublin.

During the years 2012/2013 and 2013/2014, I was lecturing at the Vincent Pol University in Lublin, Poland at the Faculty of Tourism and Physical Education. I was lecturing France Geography and Tourism with emphasise on French Wine Roads.

13. Summary list of achievements

My scientific achievements, including the achievements constituting the presented cycle, include (17/04/2019):

− Number of publications according to the Web of Science database: 7

− Number of publications according to the Scopus database: 10

− Number of citations by Web of Science database: 76

− Number of citations by Scopus database: 94

− Hirsch index according to the Web of Science database is 3

− Hirsch index according to the Scopus database is 4

− Total Impact Factor of publication (by year of publication): 12,761

− Score number according to the list of Ministry of Science and Higher Education (according to the year of publication): 277.