

**Jacek Niedźwiecki, Guillaume Debaene, Monika Kowalik**

*Instytut Uprawy Nawożenia i Gleboznawstwa – Państwowy Instytut Badawczy  
w Puławach*

**SPEKTRALNA BIBLIOTEKA GLEB UŻYTKÓW ROLNYCH W POLSCE –  
PODSTAWOWE ZAŁOŻENIA\***

**Słowa kluczowe:** spektralna biblioteka gleb, chemometria, spektroskopia VIS-NIR,  
modele cząstkowe najmniejszych kwadratów (PLS)

---

**Wstęp**

Informacje o pokrywie glebowej świata zbierane są od dziesięcioleci przez naukowców oraz służby krajowe na różnych poziomach administracji. Szerokie możliwości ich wykorzystania w znacznym stopniu ograniczone są znacznym zróżnicowaniem w standardach, metodach analiz, definicjach oraz metodach kartowania map glebowych. Jednym z pomysłów na zgromadzenie w jednym miejscu danych spektralnych gleb na poziomie światowym jest utworzenie Globalnej Bazy Danych Spektralnych (ang. *Global Soil Spectral Library (GSSL)*). Taka baza tworzona jest od 2008. Początkowo w formie wolontariatu naukowców z 8 krajów reprezentujących Europę, Azję, Afrykę, Australię, Północną i Południową Amerykę, dziś jest to projekt naukowy wspierany różnymi źródłami finansowania (16, 7). Baza ta zawiera informacje spektralne o glebach z ponad 90 krajów świata dotyczące właściwości takich jak: zawartość żelaza, pojemność sorpcyjna, pH, tekstura, zawartość materii organicznej. Wykorzystanie spektralnej bazy danych w zakresie bliskiej podczerwieni oraz światła widzialnego przyczynia się do rozwoju spektroskopii gleb i stworzeniu nowoczesnych narzędzi badawczych (algorytmów) służących modelowaniu, monitorowaniu i mapowaniu różnych właściwości gleb (3, 5, 18, 9, 10, 12). Obszerna baza danych próbek glebowych wraz z ich georeferencjami, która jest w posiadaniu IUNG-PIB, została wykorzystana przy tworzeniu Spektralnej Biblioteki Gleb Polski (SBGP). Biblioteka jest rozbudowywana o nowe dane, które

---

\* Opracowanie wykonano w ramach zadania 1.2 w programie wieloletnim IUNG-PIB.

są stale dostarczane w wyniku prac badawczych w IUNG-PIB. Działania te wpisują się w najnowsze trendy zastosowania metod chemometrycznych w modelowaniu, monitorowaniu i mapowaniu właściwości gleb.

Spektroskopia bliskiej podczerwieni (ang. *NIRS*) jest niedestrukcyjną analityczną techniką badania interakcji pomiędzy światłem odbitym bądź pochłoniętym przez materiał a powierzchnią materiału. Technika ta jest szeroko wykorzystywana w przemyśle dzięki łatwości zastosowania, szybkości i braku konieczności skomplikowanego przygotowania próbek. Została ona po raz pierwszy użyta ponad trzy dekady temu i wykorzystana do szybkiej analizy wilgotności zboża (2). Jest ona stosowana nadal w szeregu analiz materiału roślinnego (1). Badania przy zastosowaniu techniki NIRS do analizy gleby stosowane są od niedawna. Niemniej jednak poddane już zostały analizie wybrane właściwości gleby jak np. wilgotność, węgiel organiczny, przewodność elektryczna, pojemność wymienna kationowa, pH, N, C, P, S, Ca, Mg, Na, K, Fe, Mn oraz skład granulometryczny gleb (8, 11, 13, 14, 17, 9,10). Szerokie możliwości zastosowań NIRS w analizach gleby przedstawione zostały w pracach *Viscarra Rossel i in.* (17) oraz w *Cécillon i in.* (7).

Kilkadziesiąt tysięcy próbek, szczególnie gleb piaszczystych o niskiej zawartości materii organicznej, wykorzystano do tworzenia Duńskiej Bazy Danych Spektralnych (12). Istniejące biblioteki spektralne mogą być również podstawą do rozwoju technologii – doskonalenia sensorów w nowych modelach urządzeń wykorzystujących technologię spektrometrii VIS-NIR, np. do bardziej precyzyjnego określenia korekty wilgotnościowej w analizie spektralnej gleb (19).

Najnowsze światowe badania nad rozwojem technik NIRS w badaniach gleb skupiają się na tworzeniu spektralnych bibliotek gleb. Dotychczas przeprowadzono nieliczne badania właściwości gleb za pomocą NIRS na skalę krajową (4, 18). Stworzona w Australii globalna biblioteka spektralna gleb, <http://groups.google.com/group/soil-spectroscopy>, jest stale rozbudowywana o nowe krajowe biblioteki.

Od 2010 roku również w Instytucie Uprawy Nawożenia i Gleboznawstwa – PIB tworzona jest Baza Danych Spektralnych Gleb. Początkowe zasoby informacyjne tworzyły próbki gleb zgromadzone w ramach Monitoringu Chemizmu Gleb Ornych Polski (216 próbek), prowadzonego co 5 lat oraz tzw. Dużego Monitoringu Chemizmu Gleb Ornych Polski, w którym zgromadzono kilkadziesiąt tysięcy próbek gleb z całego obszaru Polski. Ujęcie w bazie danych spektralnych istniejących już informacji o glebie w zakresie szczegółowych analiz chemicznych jest istotne również ze względów ekonomicznych – wg danych (15) koszt analizy ok. 500 próbek w zakresie podstawowych parametrów chemicznych w akredytowanym laboratorium, może przekroczyć koszty zakupu samego spektrometru.

Do budowania bazy danych spektralnych (SBGP) w różnych skalach – od skali lokalnej, przez regionalną do krajowej, porównano właściwości spektralne w celu przewidywania poszczególnych właściwości fizyko-chemicznych gleb (skład granulometryczny, całkowita zawartość węgla C, zawartość węgla organicznego

$C_{org}$ , odczyn pH, zawartość całkowitych form K, P, Mg i in.). Należy nadmienić, że tworzenie spektralnej biblioteki gleb jest procesem ciągłym i uzupełnianym o kolejne dane spektralne gleb pochodzących z różnych rejonów kraju.

## Metodyka

### a) Wybór próbek do badań spektralnych

Dotychczas zeskanowano ponad 2000 próbek z użyciem spektrofotometru pracującego w zakresie widma elektromagnetycznego w paśmie widzialnym (VIS) oraz w bliskiej podczerwieni (NIR). Powyższe próbki gleb uzyskano z zasobów zgromadzonych w IUNG-PIB (istniejąca baza danych), zawierających wyniki badań analitycznych. Część próbek została wykorzystana do walidacji modelu (pojedyncze lub serie próbek w skali lokalnej). Dodatkowo w badaniach wykorzystano próbki pochodzące z różnych projektów realizowanych w IUNG-PIB (część z nich nie zawiera informacji dot. lokalizacji punktu w terenie – GPS).

Do włączenia próbki z bazy danych IUNG-PIB i innych oraz poddania jej skanowaniu, poza jej fizyczną obecnością, niezbędne były dane otrzymane w wyniku klasycznych analiz właściwości fizycznych i chemicznych gleb. Wyniki te były niezbędne do stworzenia modeli predykcyjnych. Wybrane próbki pochodziły z kilku okresów badań monitoringowych IUNG prowadzonych w różnych skalach (skala krajowa: monitoring suszy rolniczej, monitoring chemizmu gleb ornych; skala regionalna: mapa glebowo-geologiczna; skala lokalna: Rolnicze Zakłady Doświadczalne IUNG-PIB – Baborówko, Grabów, Osiny). Większość próbek posiadała informację o odczynie gleb, zawartości węgla organicznego, składzie granulometrycznym oraz współrzędne geograficzne miejsca pobrania próbki glebowej (Tab. 1).

Wszystkie próbki zostały pobrane z warstwy ornej gleb (0-30 cm), jednak włączono również serie próbek, pochodzących z kilku głębokości profilu. Takie postępowanie może być pomocne w badaniu wpływu poziomów zróżnicowania na właściwości spektralne gleb.

Tabela 1

Opis bazy danych próbek glebowych wykorzystanych do utworzenia biblioteki spektralnej

Rodzaj bazy	Liczba próbek	Dane GPS	pH	Węgiel organiczny (SOC)	Skład granulometryczny	Inne właściwości
Lokalna RZD-IUNG	952	+	+	+	+	+
Regionalna (IUNG)	180	+	+	+	+	+
Krajowa (IUNG)	987	+	+	+	+	+
Niezależna (walidacyjna)	134	+/-	+/-	+/-	+	+/-
Suma	2253					

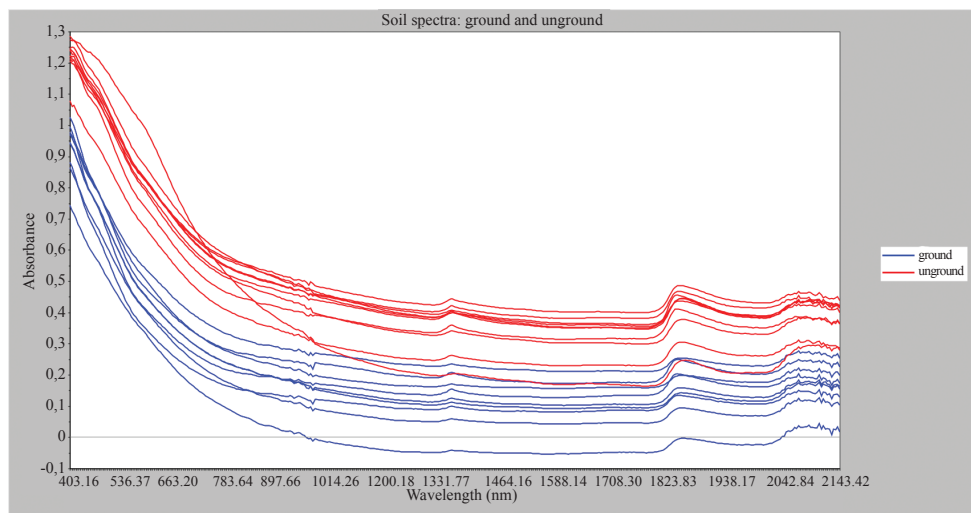
Źródło: opracowanie własne, Jacek Niedźwiecki

Z uwagi na bardzo dużą liczbę próbek (część wykresów pokrywa się ze sobą), przykłady przedstawione na wykresach reprezentują grupy próbek, a nie całą bazę danych. Niemniej jednak, modele predykcyjne opracowane zostały dla całej SBGP.

## b) Gromadzenie i skanowanie próbek z bazy danych IUNG-PIB

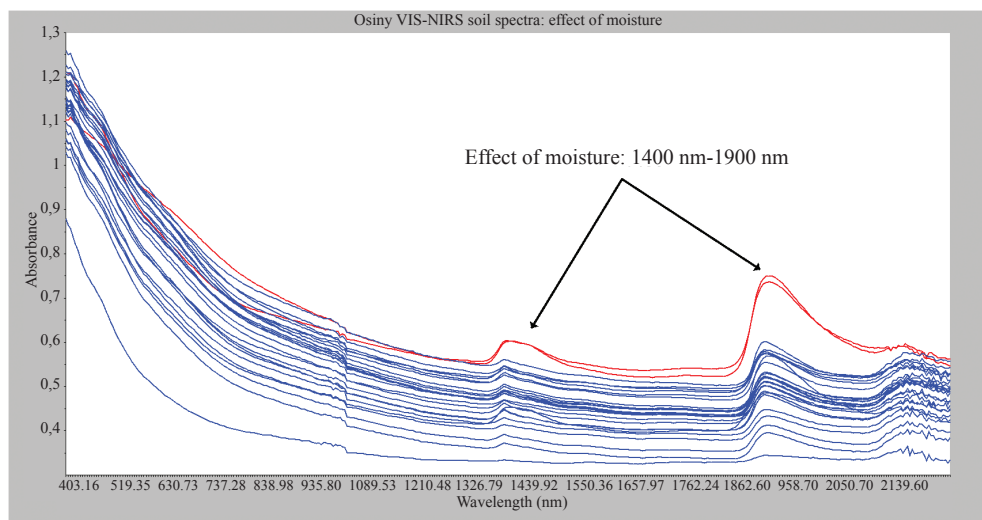
Do skanowania próbek w zakresie VIS-NIR wykorzystano spektrofotometr Veris (Veris Technologies, Salina, KS, Stany Zjednoczone). Próbki były skanowane 20-krotnie i uśredniane z użyciem oprogramowania (Veris w wersji V1.83). Rozdzielczość spektralna wynosiła 5 nm. Absorbancja skanowanej próbki została przedstawiona w formie logarytmu  $\log(1/R)$ , gdzie R oznacza współczynnik odbicia. Widma zostały pogrupowane w paśmie widzialnym i bliskiej podczerwieni w zakresie 349-2223 nm. W ten sposób otrzymano 384 dane spektralne dla każdej próbki.

W trakcie tworzenia SBGP, przebadano również wpływ sposobu przygotowania próbki glebowej na jej właściwości spektralne (Rys. 1). W tym celu niektóre próbki przed skanowaniem zmielono na młynkach agatowych, pozostałe skanowano po przesianiu przez sito  $< 2$  mm. Wszystkie skanowane próbki były w stanie powietrzno-suchym. Zmieszanie próbek miało na celu zmniejszenie efektu absorpcji VIS-NIR poprzez homogenizację próbki (światło w próbce odbija się w mniej chaotyczny sposób). Podczas skanowania, zaobserwowano jednak, że niektóre próbki (niezmielone) pochłaniały wilgoć z powietrza (Rys. 2), taka anomalia mogła wpłynąć na osłabienie zdolności predykcyjnej modeli. Z tego powodu próbki te zostały odrzucone, a do budowy SBGP wykorzystano tylko próbki przesiane przez sito.



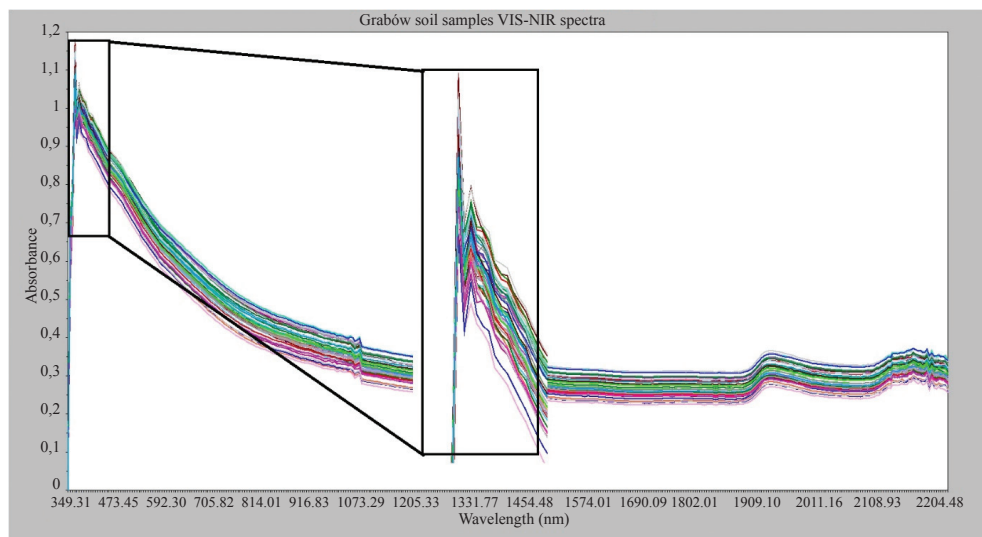
Rys. 1. Wpływ mielenia próbek na kształt widma glebowego (te same próbki pobrane przed zmieleniem – kolor czerwony i po zmieleniu- kolor niebieski)

Źródło: opracowanie własne, Guillaume Debaene



Rys. 2. Wpływ wilgotności gleby na widmo: 2 próbki wilgotne przygotowane w odmienny sposób

Źródło: opracowanie własne, Guillaume Debaene

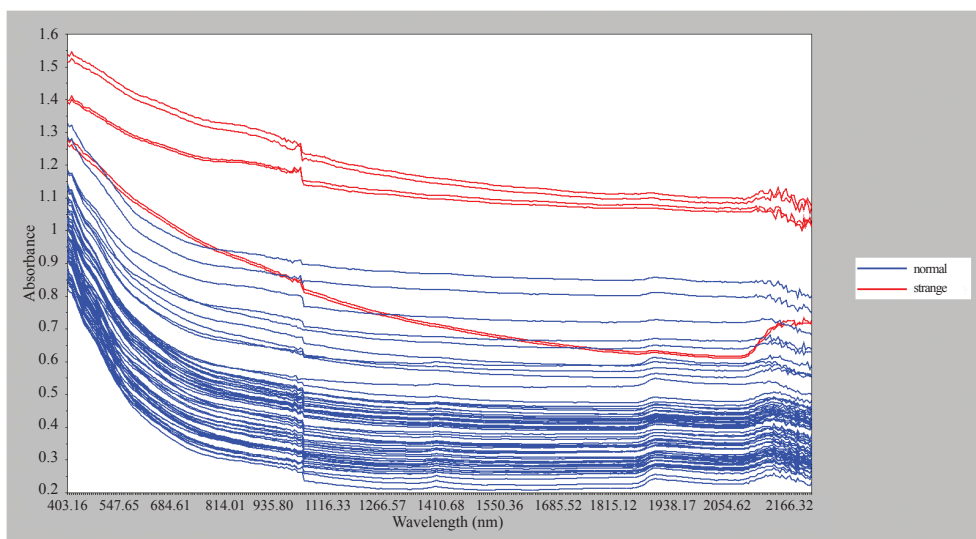


Rys. 3. Szum tła w zakresie 350-400 nm widma elektromagnetycznego

Źródło: opracowanie własne, Guillaume Debaene

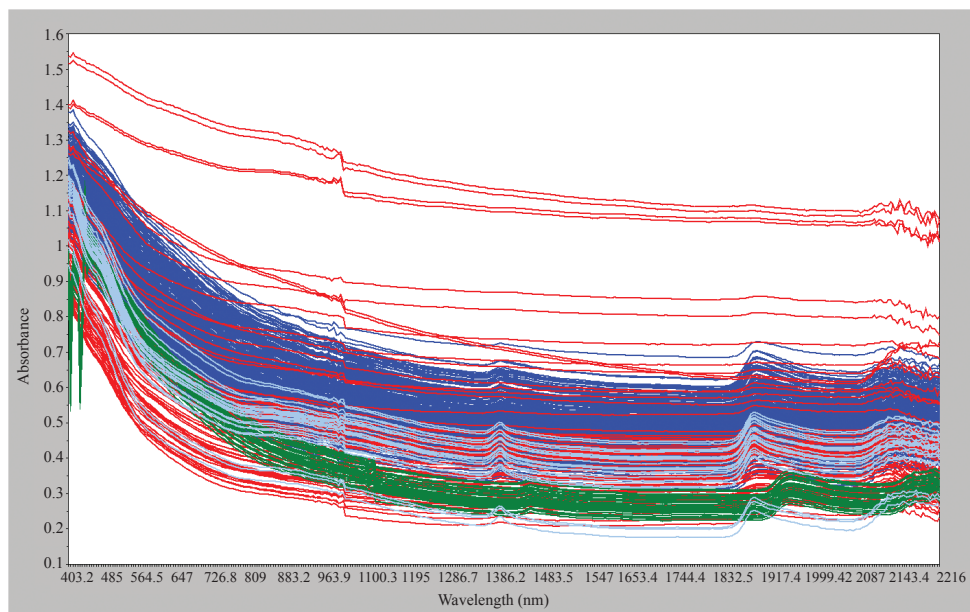
### c) Wstępna analiza widm glebowych i analiza głównych składowych (PCA)

Wygląd oraz kształt widm glebowych daje wstępną informację o wartościach odstających, szumach tła czy wilgotności próbki (Rys. 2-4). Jednak zwykle bardzo trudno jest odróżnić widmo jednej próbki od drugiej (podobieństwo kształtu widm i ich nakładanie się). Przykład podano na Rys. 5, gdzie pokazano na jednym wykresie kilka przykładowych widm glebowych (z różnych lokalizacji) – poszczególne kolory oznaczają różne lokalizacje widm glebowych. W takim przypadku konieczne jest wykorzystanie zaawansowanych metod kompresji danych (np. PCA), aby możliwe było wydobycie interesujących informacji ze zbioru danych. Dzięki analizie PCA, można w łatwy sposób obserwować np. dane odstające i grupować widma (Rys. 7-8).



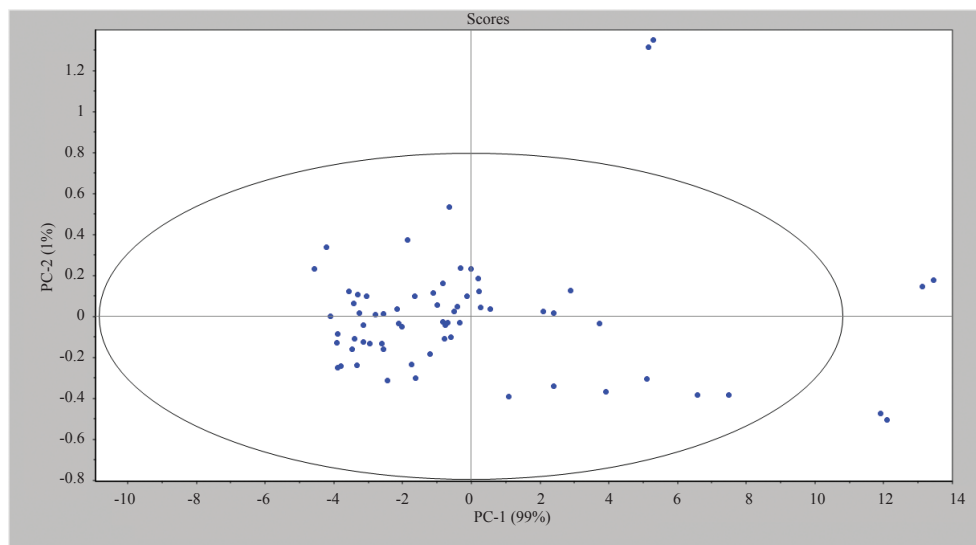
Rys. 4. Przykład widm nietypowych oraz odstających dla próbek pochodzących z obszaru G. Śląska. Dokładne badania wykazały że, były to materiały pokopalniane (kolor niebieski – widmo normalne, kolor czerwony – widmo nietypowe)

Źródło: opracowanie własne, Guillaume Debaene



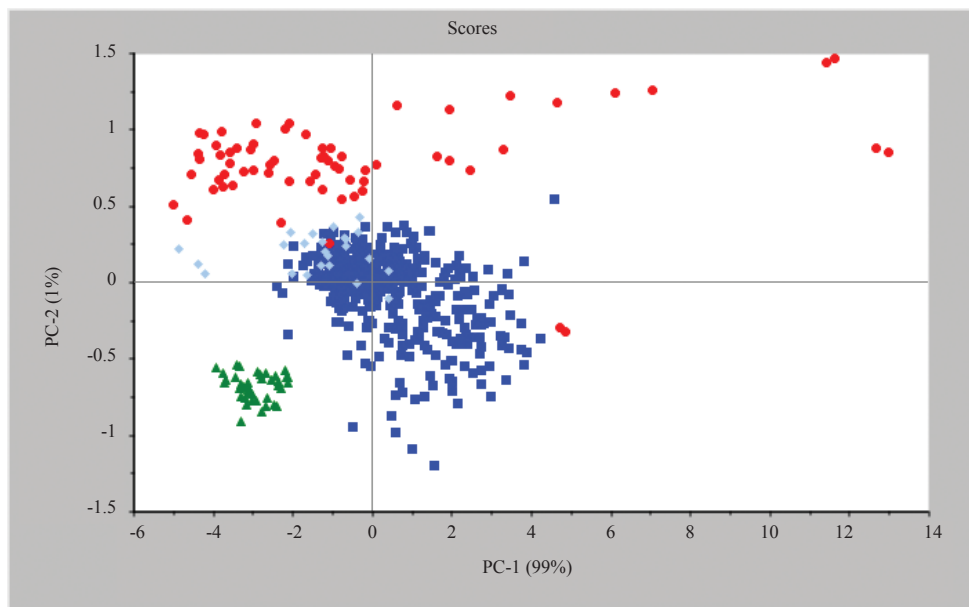
Rys. 5. Widma glebowe w zakresie VIS-NIR kilkuset próbek (528). Ze względu nakładanie się widm (2253 próbki) nie zaprezentowano wszystkich

Źródło: opracowanie własne, Guillaume Debaene



Rys. 6. Analiza PCA, test Hotelling  $T^2$  (95% przedział ufności w oparciu o próbki z Rys. 4, z widocznymi wartościami odstającymi)

Źródło: opracowanie własne, Guillaume Debaene



Rys. 7. Analiza PCA, test Hotelling  $T^2$  (95% przedział ufności w oparciu o próbki z Rys. 5). Kolor niebieskim i zielonym zaznaczono próbki pochodząc z pól, błękitnym profile glebowe, czerwonym – próbki reprezentujące różne regiony kraju.

Źródło: opracowanie własne, Guillaume Debaene

#### d) **Opracowanie metod statystycznych i modeli predykcyjnych w skali lokalnej**

Ilość promieniowania NIR, która jest absorbowana, określona jest przez właściwości i ilość wiązań obecnych w badanym materiale. Stąd widma w zakresie bliskiej podczerwieni zawierają szczegółowe informacje na temat składu chemicznego badanych substancji.

Oprócz składu chemicznego materiału, na właściwości spektralne ma również wpływ fizyczna struktura materiału. Rozmiar i kształt cząstek, puste przestrzenie pomiędzy cząstkami, gdyż wpływają na długość transmisji światła przechodzącego przez próbkę. Wszystkie te czynniki wpływają na współczynnik odbicia.

Widma NIR są bardzo złożone i niezwykle rzadko zdarza się, aby dwa różne związki chemiczne miały w całym zakresie identyczne widma, co praktycznie uniemożliwia jednoznaczną ich identyfikację. W tej sytuacji, konieczna jest wstępna obróbka matematyczna, która ułatwi identyfikację związku. Jest w tym celu specjalistyczne oprogramowanie.

Niespecyficzność widm światła odbitego i rozproszonego spowodowana jest w dużej mierze zakłóceniami i szumami wynikającymi z nakładania się widm zróżnicowanych składników glebowych, które są ze sobą powiązane, oraz tła wytwarzanego przez pracujące urządzenie.

Wszystkie te czynniki zakłócające, muszą być metodami matematycznymi odseparowane od widma tak, żeby mogły być skorelowane z konkretną właściwością gleby. Zatem, analiza spektralna gleby w zakresie VIS-NIR wymaga stosowania technik chemometrycznych i kalibracji wielowymiarowych. Chemometria dotyczy stosowania technik matematycznych lub statystycznego w analizowaniu danych chemicznych. Metody chemometrii stosowane są m.in. do: klasyfikacji obiektów badań, interpretacji widm, określania zależności pomiędzy budową chemiczną a aktywnością biologiczną związków chemicznych, modelowania przemian chemicznych. Podczas kalibracji wielowymiarowej, wykorzystuje się dane empiryczne do przewidywania nieznannej właściwości glebowej –  $y$  na podstawie wielu pomiarowych danych spektralnych  $x_1, x_2, \dots, x_K$ .

Poniższe testowanie procedur wykonano na 398 próbkach pobranych z pól produkcyjnych RZD IUNG w Baborówku. Procedury te zostały szczegółowo opisane przez Debaene i in. (9):

1. Losowy wybór próbek kalibracyjnych.
2. Wybór według stężeń chemicznych.
3. Wybór według grupowaniu metodą  $k$ -średnich.
4. Wybór według analizy głównych składowych – PCA.

Najlepszą procedurą kalibracji okazał się wybór próbek metodą  $k$ -średnich. Dzięki tej procedurze, możemy uzyskać lepsze wyniki predykcji z zestawu danych walidacji. To także pozwala nam zmniejszyć ilość próbek kalibracyjnych do budowy modelu, a tym samym zmniejszyć koszty analiz laboratoryjnych lub mapowania. Dzięki tej metodzie można było zmniejszyć liczbę kalibracji do 30% z ilości próbek i uzyskanie modeli o zadowalającej predykcji.

#### e) **Regresja metodą częściowych najmniejszych kwadratów (PLS)**

Regresja metodą cząstkowych (częściowych) najmniejszych kwadratów (PLS) dotyczy dwóch macierzy danych  $X$  i  $Y$  poprzez model liniowy wielowymiarowy. Jest metodą powszechnie stosowaną w analizie odbicia danych spektralnych. PLS rozkłada zarówno zmienną  $X$  i  $Y$ , i znajduje nowe komponenty (punkty), zwane zmiennymi ukrytymi. Następnie obliczana jest regresja pomiędzy tymi nowych składnikami zmiennych  $X$  i  $Y$ , gdzie  $X$  jest odbiciem gleby, a  $Y$  mierzoną właściwością gleby.

$$X = TP + E \quad (1)$$

$X$  tutaj to macierze absorpcji światła

$$Y = UQ + F \quad (2)$$

$Y$  to są dane chemiczne i skład granulometryczny

gleby

$$U = BT \quad (3)$$

gdzie  $T$  i  $U$  oznaczają macierze wyników,  $X$  i  $Y$  macierze zmiennych,  $P$  i  $Q$  są ładunkami czynnikowymi macierzy  $X$  i  $Y$ , odpowiednio.  $B$  jest współczynnikiem regresji pomiędzy  $U$  i  $T$ .  $E$  i  $F$  to składniki błędu (szumu).

Analizę PLS przeprowadzono dla wszystkich istniejących w bazie spektralnej danych, przy użyciu programu Unscrambler X<sup>®</sup>, wersja 10.3.

Oceny statystyk dokonano na podstawie współczynnika determinacji  $r^2$  pomiędzy wartościami przewidywanymi a mierzonymi oraz RMSE (średniego błędu kwadratowego) i RPD (ang. *ratio of performance to deviation*).

$$R^2 = \frac{\left[ \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})(y_i - \bar{y}) \right]^2}{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2 \sum_{i=1}^n (y_i - \bar{y})^2}$$

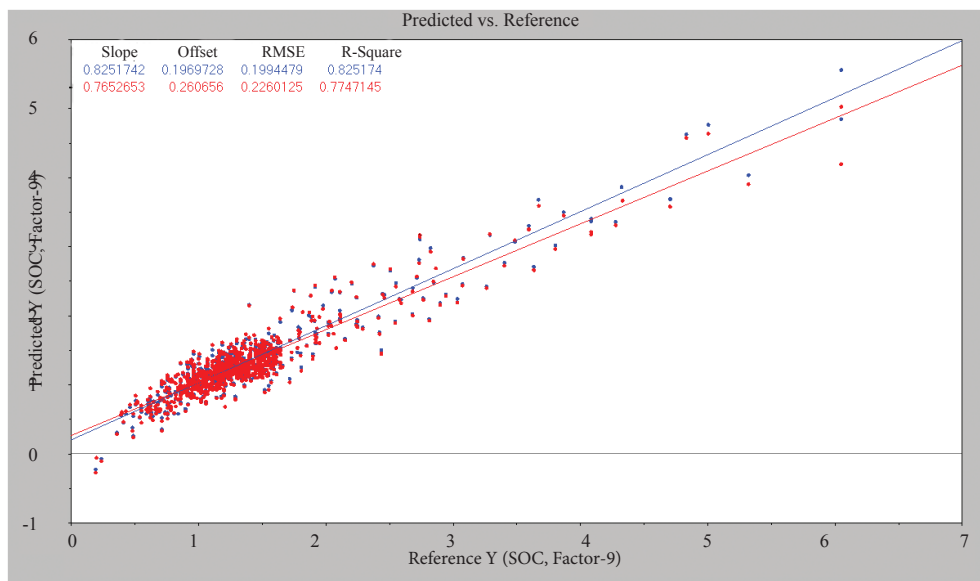
$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}$$

$$RPD = SD/RMSE$$

Rys. 8

gdzie  $\hat{y}_i$  jest wartością przewidywaną i  $y_i$  wartością zmierzoną;  $\bar{y}$  średnia z wartości zmierzonej;  $\bar{\hat{y}}$  średnia z wartości przewidywanej dla  $n$  ilość próbek.

Na rysunku 9 przedstawiono zależność regresyjną pomiędzy przewidywaną i mierzoną wartością SOC, dla zestawu danych pochodzących z pól Rolniczego Zakładu Doświadczalnego w Baborówku (10).



Rys. 9. Przewidywane vs. mierzone wartości SOC (%), niebieski – kalibracja, czerwony – walidacja (10)

Źródło: opracowanie własne, Guillaume Debaene

f) **Modele predykcyjne w skali lokalnej (RZD IUNG -PIB Baborówko)**

Wyniki modeli predykcyjnych (SOC i zawartości iłu) dla pól produkcyjnych w RZD IUNG Baborówko przedstawiono w Tab. 2 i 3. Dwieście próbek glebowych (z zestawu danych liczącego 400 próbek) było wykorzystanych do walidacji modelu. W celu zbadania wpływu ilości próbek kalibracyjnych na PLS, użyto od 10 do 100% próbek kalibracyjnych do budowy modeli węgla organicznego (SOC) oraz frakcji iłu. Wszystkie modele predykcyjne zostały zbudowane przy użyciu minimalnej liczby próbek kalibracyjnych niezbędnych do uzyskania modeli o dużej sile predykcyjnej.

Tabela 2

Metoda k-średnich w analizie PLS. Wyniki kroswalidacji i walidacji dla SOC

Ilość próbek wykorzystanych do kalibracji	Kroswalidacja		Walidacja			
	r <sup>2</sup>	RMSE	r <sup>2</sup>	RMSE	RPD	NF
10%	0,48	0,20	0,42	0,27	1,0	4
20%	0,57	0,15	0,47	0,17	1,3	10
30%	0,45	0,16	0,48	0,16	1,3	10
40%	0,53	0,14	0,63	0,13	1,5	7
50%	0,55	0,14	0,62	0,13	1,5	7
60%	0,58	0,14	0,70	0,12	1,9	7
70%	0,62	0,13	0,70	0,12	1,9	7
80%	0,58	0,14	0,68	0,12	1,8	8
90%	0,61	0,14	0,68	0,12	1,8	8
100%	0,66	0,13	0,70	0,12	1,9	8

Źródło: Debaene i in., 2014 (9)

Tabela 3

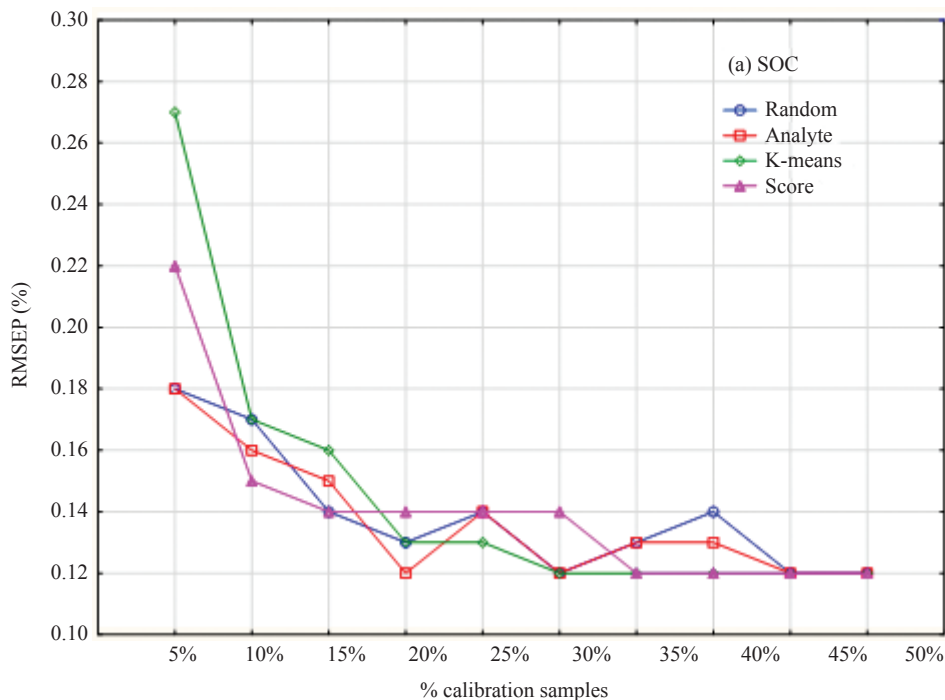
Metoda k-średnich w analizie PLS. Wyniki kroswalidacji i walidacji dla frakcji iłu

Ilość próbek wykorzystanych do kalibracji	Kroswalidacja		Walidacja			
	r <sup>2</sup>	RMSE	r <sup>2</sup>	RMSE	RPD	NF
5%	0,37	0,42	0,22	0,62	1,0	2
10%	0,47	0,48	0,50	0,43	1,4	4
15%	0,72	0,38	0,69	0,37	1,6	10
20%	0,75	0,34	0,71	0,36	1,7	9
25%	0,74	0,35	0,71	0,35	1,8	9
30%	0,73	0,34	0,72	0,33	1,9	9
35%	0,74	0,33	0,73	0,34	1,9	9
40%	0,75	0,32	0,73	0,32	1,9	9
45%	0,76	0,31	0,72	0,32	1,9	9
50%	0,77	0,31	0,73	0,32	1,9	9

NF = ilość czynników

Źródło: Debaene i in., 2014 (9)

Zgodnie z uzyskanymi wynikami, 20% próbek kalibracyjnych (z zestawu danych 400 próbek) jest niezbędnych do budowania dobrych modeli (Rys. 10). Tab. 4 przedstawia wyniki modeli predykcyjnych dla innych właściwości gleby.



Rys. 10. Wpływ liczby próbek kalibracyjnych na średni błąd kwadratowy szacowania (RMSEP)  
Źródło: Debaene i in., 2014 (9)

Tabela 4

Analiza PLS (kalibracja próbek metodą k-średnich 20% próbek kalibracyjnych i 50% próbek kalibracyjnych): pozostałe badane właściwości gleb

Właściwości	Krosvalidacja			Walidacja			
	N	r <sup>2</sup>	RMSE	r <sup>2</sup>	RMSE	RPD	NF
frakcja piasku	50%	0,79	2,60	0,79	2,52	2,2	9
	20%	0,80	2,79	0,76	2,88	1,9	9
frakcja pyłu	50%	0,78	2,33	0,79	2,25	2,2	9
	20%	0,77	2,64	0,75	2,52	2,0	7
P <sub>2</sub> O <sub>5</sub>	50%	0,36	6,50	0,39	6,27	1,3	5
	20%	0,42	5,37	0,37	6,46	1,3	7
K <sub>2</sub> O	50%	0,46	3,67	0,46	3,31	1,4	7
	20%	0,34	4,62	0,28	3,87	1,2	1
Mg	50%	0,48	1,67	0,54	1,53	1,5	4
	20%	0,44	1,68	0,53	1,54	1,4	4
pH	50%	0,46	0,36	0,51	0,34	1,4	8
	20%	0,41	0,38	0,52	0,34	1,4	8

Źródło: Debaene i in., 2014 (9), zmodyfikowano

## g) Kalibracja modeli

### *Skala regionalna*

Aby zbadać efekt głębokości próbki na modele predykcyjne węgla organicznego (SOC), zbiór danych z bazy danych mapy glebowo-geologicznej został użyty jako model regionalny. Próbki pochodzą z regionu Górnego Śląska i zostały pobrane na trzech głębokościach. Zbiór danych składa się ze 180 próbek pochodzących z 60 profili. Głębokości na jakich zostały pobrane wynosiły: 15, 30 i 50 cm. Kalibracji dokonano z wykorzystaniem dwóch procedur: procedura 1 – używając jednej głębokości do kalibracji i procedura 2 – 60 próbek wyselekcjonowanych według metody grupowania k-średnich spośród 180 próbek. Z algorytmu otrzymano 20 próbek każdej głębokości do wykorzystania do kalibracji modelu.

Wyniki kalibracji są bardzo podobne w obu procedurach ( $r^2 = 0,81$ ; RMSE = 0,41; RPD = 1,70 przy kalibracji dla jednej głębokości,  $r^2 = 0,78$ ; RMSE = 0,48; RPD = 1,62 dla próbek z różnej głębokości).

Różnice błędów (RMSE) wynikają z faktu, że próbki pochodzą z różnych głębokości. Warstwa powierzchniowa zawierająca wyższą zawartość SOC charakteryzuje się większym odchyleniem standardowym. Schemat 2 został w następnej kolejności użyty do walidacji.

### *Skala krajowa*

Przed wykonaniem kalibracji liczba próbek została zredukowana drogą wykluczenia identycznych próbek pochodzących z tej samej lokalizacji. Uśredniono dane spektralne próbek pochodzących z tej samej lokalizacji (320 próbek z RZD Baborówka i RZD Grabowa) aby otrzymać próbkę uśrednioną. Ponadto, próbki o wartościach ekstremalnych zostały usunięte, jak również próbki odstające od reszty (jedyne w swoim rodzaju) – np. próbki z koncentracją węgla organicznego > 5%, próbki o bardzo wysokiej zawartości iłu koloidalnego i inne niereprezentatywne dla SBGP (prawie 100 próbek). Z powodu zróżnicowanego charakteru gleb w skali kraju, 1200 próbek zostało użytych jako próbki kalibracyjne.

Wyniki kalibracji są następujące:

$r^2 = 0,69$ ; RMSE = 0,59%; RPD = 1,38 dla węgla organicznego,

$r^2 = 0,73$ ; RMSE = 0,89%; RPD = 1,42 dla frakcji iłu koloidalnego.

$r^2 = 0,59$ ; RMSE = 0,40 jednostki pH; RPD = 1,29 dla pH w KCl.

## h) Walidacja modeli – Zbiór danych walidacji

W Tab. 5 przedstawione zostały wyniki walidacji w skali regionalnej i krajowej z użyciem, odpowiednio, 120 i 600 próbek.

Tabela 5

Wyniki modelu walidacji dla regionalnej i krajowej bazy danych

Właściwości glebowe	Skala	Walidacja			
		$r^2$	RMSE	RPD	NF
Węgiel organiczny	R	0,74	0,48	1,60	5
	N	0,67	0,60	1,38	7
Fracja ilasta	R	0,71	0,82	1,60	5
	N	0,72	0,88	1,41	7
pH <sub>KCl</sub>	R	0,51	0,55	1,19	8
	N	0,62	0,43	1,27	8

Źródło: opracowanie własne, Guillaume Debaene

Zawartość węgla organicznego i frakcji ilastej zostały prawidłowo prognozowane, a pH przewidziano ze znaczną dokładnością.

### i) Weryfikacja modeli – Zbiór danych niezależnych

Utworzono dwa niezależne zbiory danych, [1] próbki z profilu glebowego (różne głębokości) i [2] 60 próbek z warstwy wierzchniej z niezależnego projektu. Próbki zostały zeskanowane, a prognozy części właściwości dokonano z wykorzystaniem modelu opisanego wyżej.

Rys. 11 przedstawia wyniki predykcji węgla organicznego. Model jest bardzo skuteczny dla zbioru danych z  $r^2 = 0,81$ ; RMSEP = 0,14; RPD = 1,85. Nie stworzono żadnych innych modeli dla tego zbioru.

W Tab. 6 przedstawiono wyniki predykcji z drugiego niezależnego zbioru danych (zawartość węgla organicznego, pH, frakcji ilastej). Również dla tego zbioru danych predykcja była bardzo dokładna.

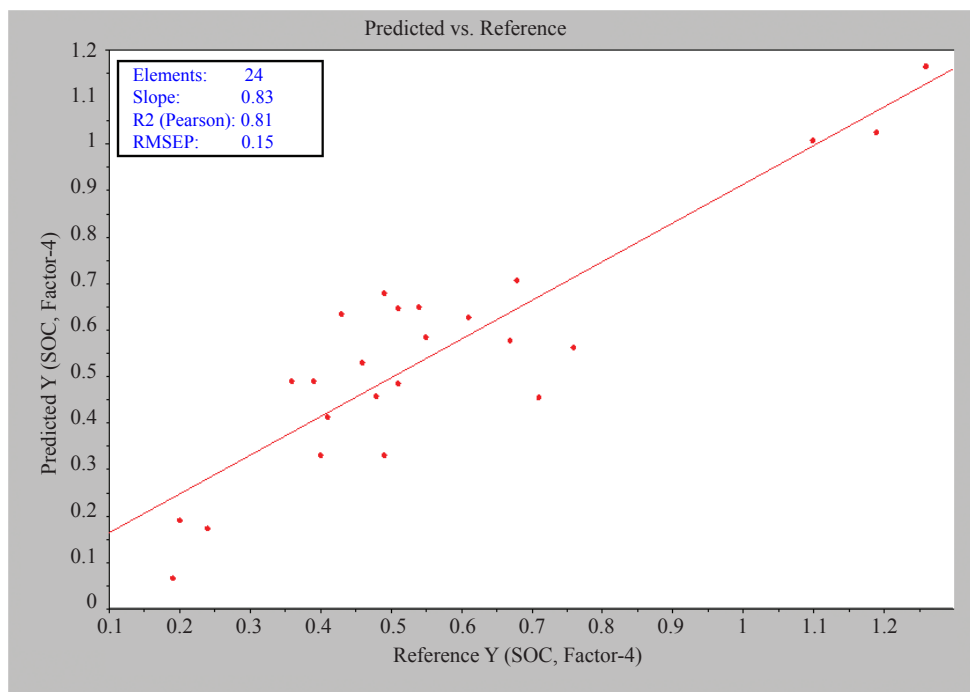
Tabela 6

Wyniki predykcji dla niezależnego zbioru danych 2

Właściwości glebowe	Walidacja			
	$r^2$	RMSE	RPD	NF
Węgiel organiczny	0,63	0,48	1,61	9
Fracja ilasta	0,67	0,78	1,69	8
pH <sub>KCl</sub>	0,55	0,75	1,39	8

Źródło: opracowanie własne, Guillaume Debaene

Co istotne, w wyniku dodania niektórych próbek z niezależnego zbioru danych do modelu kalibracji (zbiór danych 1), wyniki ulegają znacznej poprawie, np.  $r^2 = 0,71$ ; RMSE = 0,41; RPD = 1,72. Wymaga to poddania dalszej analizie.



Rys. 11. Przewidywane vs. zmierzone wartości węgla organicznego dla niezależnego zbioru danych 1

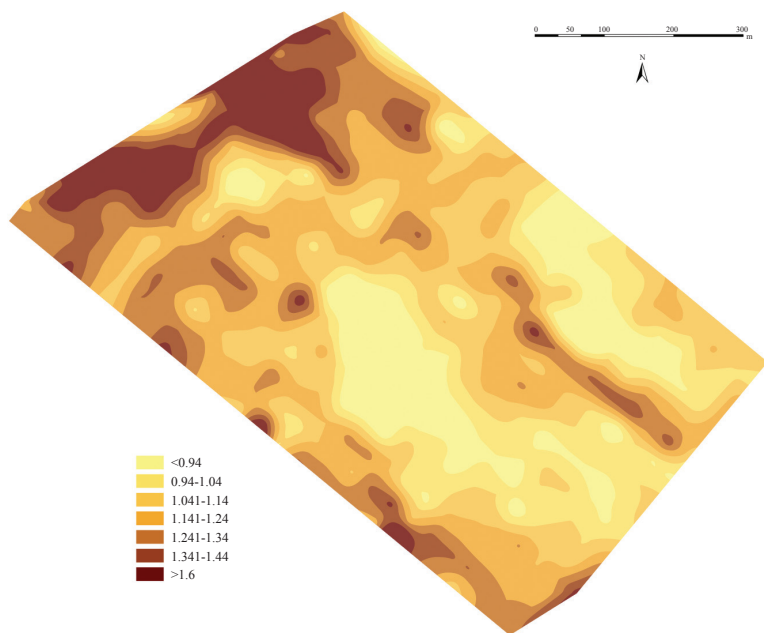
Źródło: opracowanie własne, Guillaume Debaene

Wyniki te potwierdzają użyteczność metody VIS-NIR dla prognozowania niektórych właściwości glebowych. Po przeprowadzeniu kalibracji jest to szybka i niekosztowna metoda analizy próbek glebowych.

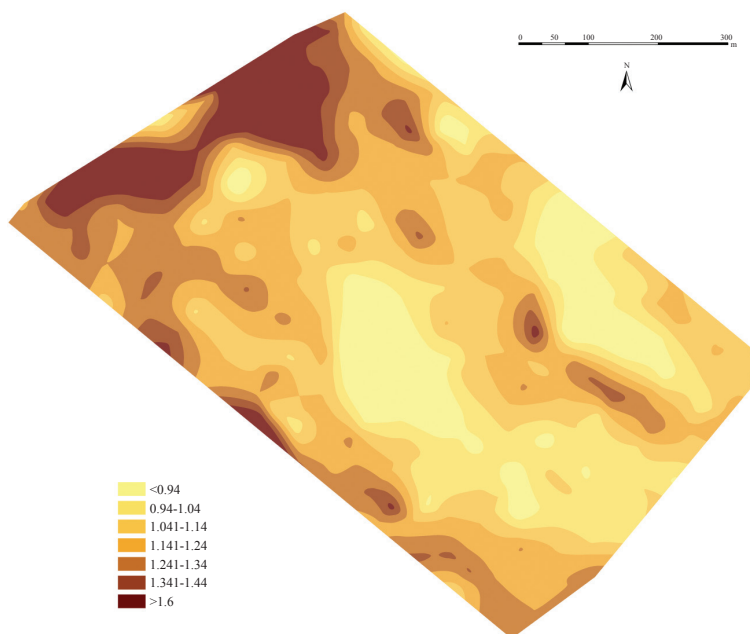
## j) Mapowanie

Niektóre mapy uzyskane na podstawie wyników naszych modeli przedstawione zostały poniżej. Mapy dla Rolniczego Zakładu Doświadczalnego (IUNG) w Baborówku opublikowane zostały przez Debaene i in. (10). W tej predykcji mapy właściwości węgla organicznego dla SB Baborówko, 93% terenu było identyczne (poniżej 0,1% różnicy w zawartości węgla organicznego) jak na mapie otrzymanej za pomocą klasycznego poboru próbek ze stałej siatki punktów (Rys. 12-14). Mapy zawartości pH i węgla organicznego zostały przedstawione w Rys. 15-16. Te dwie mapy są zgodne z naszym stanem wiedzy (badania na terenie RZD IUNG w Baborówku, (10)) i zostały sporządzone na podstawie 250 skanowanych próbek (zestaw walidacyjny), na podstawie modelu krajowego (1200 próbek glebowych).

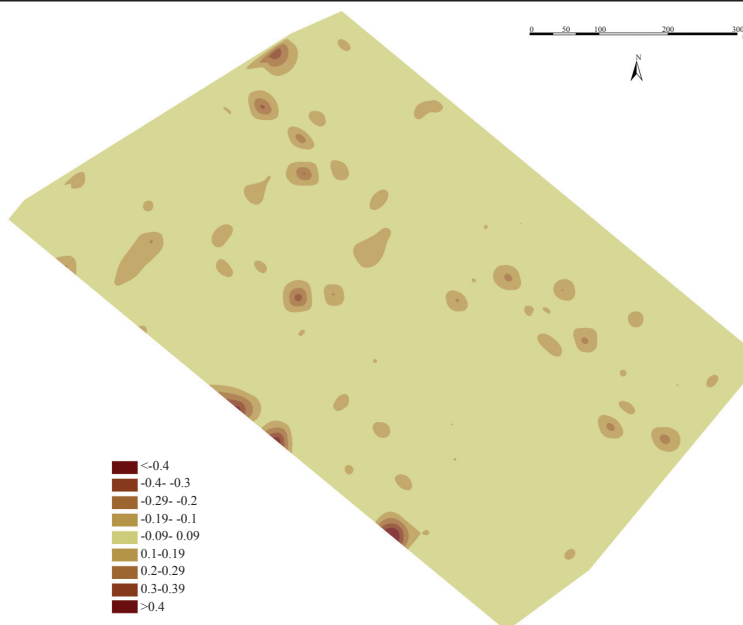
Aby zwiększyć „rozdzielczość mapy”, należy włączyć do modelu więcej próbek. Brak precyzji wynika ponadto z faktu, że model bierze pod uwagę jedynie gleby mineralne.



Rys. 12. Mapa dystrybucji przestrzennej węgla organicznego na podstawie analizy 398 próbek  
Źródło: Debaene i in., 2014 (9)

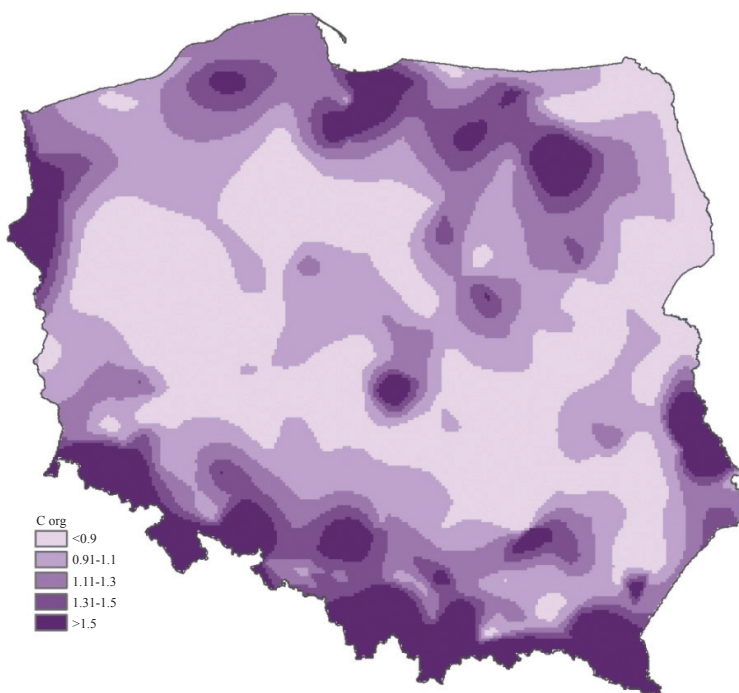


Rys. 13. Mapa dystrybucji przestrzennej węgla organicznego na podstawie próbek predykcyjnych  
VIS-NIRS (79 próbek poddanych kalibracji, 319 predykcyjnych)  
Źródło: Debaene i in., 2014 (9)



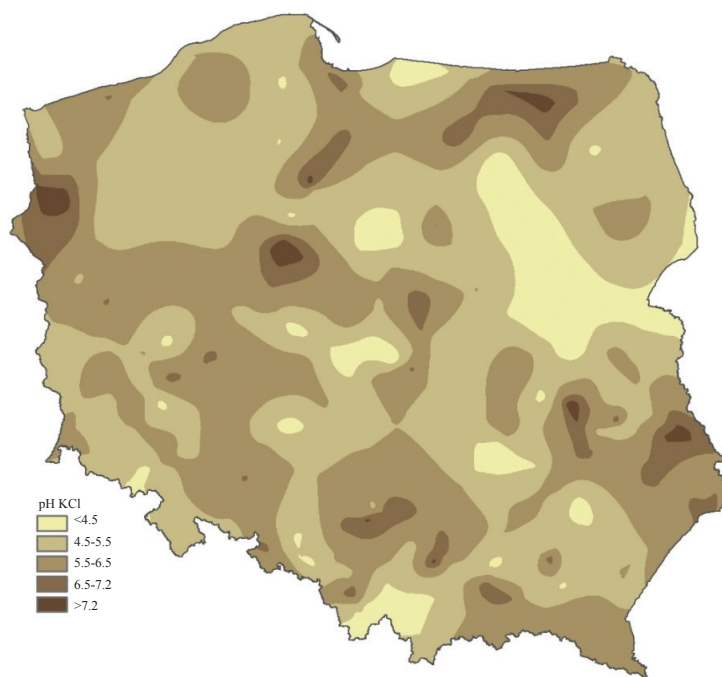
Rys. 14. Mapa dystrybucji przestrzennej różnicy pomiędzy prognozowaną a zmierzoną zawartością węgla organicznego

Źródło: Debaene i in., 2014 (9)



Rys. 15. Mapa dystrybucji przestrzennej węgla organicznego na podstawie próbek predycyjnych VIS-NIRS

Źródło: opracowanie własne, Guillaume Debaene



Rys. 16. Mapa dystrybucji przestrzennej pH w KCl na podstawie próbek predykcyjnych VIS-NIRS  
Źródło: opracowanie własne, Guillaume Debaene

## Podsumowanie

Uzyskano bardzo dokładne modele predykcyjne dla zawartości węgla organicznego i frakcji ilastej w skalach: lokalnej, regionalnej i krajowej. Wyniki predykcji (zawartość węgla organicznego i frakcji ilastej) w skali lokalnej są porównywalne z wynikami otrzymanymi klasycznymi (laboratoryjnymi) metodami analizy próbek glebowych. Niezależny nowy zbiór danych może być prognozowany z dużą dokładnością kiedy dodajemy próbki nie pochodzące z badanego zbioru danych, do istniejącego modelu kalibracyjnego.

Wykorzystanie spektralnej bazy danych w zakresie bliskiej podczerwieni oraz światła widzialnego może przyczynić się do rozwoju spektroskopii gleb i stworzenia nowoczesnych narzędzi badawczych (modeli predykcyjnych) służących modelowaniu, monitorowaniu gleb i mapowaniu właściwości glebowych. Dodatkową korzyścią z posiadania Spektralnej Biblioteki Gleb jest możliwość wykorzystania opracowanych w ramach projektu modeli kalibracyjnych do przewidywania właściwości fizykochemicznych gleb w skali kraju bądź regionalnej, oraz generowania baz przestrzennych w oparciu o sceny satelitarne, co pozwala na szersze wykorzystanie modeli

symulacyjnych i analiz przestrzennych związanych zarówno z glebą, jak i szeroko pojętym środowiskiem. Stworzona w Australii Globalna Biblioteka Spektralna Gleb <http://groups.google.com/-/group/soil-spectroscopy> jest stale rozbudowywana o nowe krajowe biblioteki. Uzyskane wyniki badań dają w perspektywie możliwość integracji krajowej bazy danych spektralnych z Globalną Biblioteką Spektralną Gleb.

### Literatura

1. Batten G.D.: Plant analysis using near infrared reflectance spectroscopy: the potential and the limitations. *Aust. J. Exp. Agric.* 1998, **38**:697–706.
2. Ben-Gera I., Norris K.H.: Direct Spectrophotometric Determination of Fat and Moisture in Meat Products. *J. Food Sci.*, 1968, **33**: 64–67.
3. Brodský L., Klement A., Penížek V., Kodešová R., Borůvka L.: Building soil spectral library of the Czech soils for quantitative digital soil mapping. *Soil Water Res*, 2011, **6(4)**: 165-172.
4. Brown D.J., Shepherd K.D., Walsh M.G., Dewayne Mays M., Reinsch T.G.: Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma*, 2006, **132**: 273–290.
5. Cambule A.H., Rossiter D.G., Stoorvogel J.J., Smaling E.M.A.: Building a near infrared spectral library for soil organic carbon estimation in the Limpopo National Park, Mozambique. *Geoderma*, 2012, **183–184**: 41-48.
6. Cécillon L., Cassagne N., Czarnes S., Gros R., Brun J.J.: Variable selection in near infrared spectra for the biological characterization of soil and earthworm casts. *Soil Biol. Biochem.*, 2008, **40**: 1975–1979.
7. Cécillon L., Barthès B. G., Gomez C., Ertlen D., Genot V., Hedde M.: Assessment and monitoring of soil quality using near-infrared reflectance spectroscopy (NIRS). w: *Eur. J. Soil Sci.*, 2009, **60(5)**: 770–784.
8. Chang C.W., Laird D.A., Mausbach M.J., Hurburgh C.R.: Near-Infrared Reflectance Spectroscopy–Principal Components Regression Analyses of Soil Properties. *Soil Sci. Soc. Am. J.* 2001, **65**: 480.
9. Debaene G., Niedźwiecki J., Pecio A., Żurek A.: Effect of the number of calibration samples on the prediction of several soil properties at the farm-scale. *Geoderma*, 2014, **214–215**: 114–125.
10. Debaene G., Pikula D., Niedźwiecki J.: 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*, 2014, **41**: 5–6.
11. Dunn B.W., Batten G.D., Beecher H.G., Ciavarella S.: 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.* 2002, **42**: 607–614.
12. Deng F., Knadel M., Thomsen A.G., Greve M.H.: Development of near infrared spectral library of Danish soils. Adamshuk V.I., Rossel R.A.V. (red.), *The Second Global Workshop on Proximal Soil Sensing: Program and Proceedings*. Monreal, Quebec, Canada, 2011, IUSS Working Group on Proximal Soil Sensing: 96-99.
13. Sørensen L.K., Dalsgaard S.: Determination of Clay and Other Soil Properties by Near Infrared Spectroscopy. *Soil Sci. Soc. Am. J.*, 2005, **69**: 159.
14. Velasquez E., Lavelle P., Barrios E., Joffre R., Reversat F.: Evaluating soil quality in tropical agroecosystems of Colombia using NIRS. *Soil Biology and Biochemistry*, 2005, **37**: 889–898.
15. Viscarra Rossel R.A., McBratney A.B.: Diffuse Reflectance Spectroscopy as a Tool for Digital Soil Mapping, w: *Digital Soil Mapping with Limited Data*, red. Alfred E. Hartemink, Alex McBratney i Mendonça-Santos, Maria de Lourdes. Dordrecht: Springer Netherlands, 2008. : 165–172
16. Viscarra Rossel R.A., Behrens T., Ben-Dor E., Brown D.J., Demattê J.A.M., Shepherd K. D.: A global spectral library to characterize the world's soil. w: *Earth-Science Reviews*, 2016, **155**: 198–230.

17. Viscarra Rossel R.A., Walvoort D.J.J., McBratney A.B., Janik L.J., Skjemstad J.O.: Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 2006, **131**: 59–75.
  18. Viscarra Rossel R.A., Webster R.: Discrimination of Australian soil horizons and classes from their visible-near infrared spectra. *Eur. J. Soil Sci.*, 2011, **62**: 637–647.
  19. Wijewardane N.K., Ge Y., Morgan C.L.S.: Prediction of soil organic and inorganic carbon at different moisture contents with dry ground VNIR: a comparative study of different approaches. *Eur. J. Soil Sci.*, 2016., **67(5)**: 605–615.
- 

Adres do korespondencji:

*dr Jacek Niedźwiecki*  
*Zakład Gleboznawstwa Erozji i Ochrony Gruntów*  
*IUNG-PIB*  
*ul. Czartoryskich 8, 24-100 Puławy*  
*tel. 81 4786 780*  
*e-mail: jacekn@iung.pulawy.pl*