

Prof. dr hab. inż. Jerzy Cierniewski
Zakład Gleboznawstwa i Teledetekcji Gleb
Instytut Geografii Fizycznej i Kształtowania
Środowiska Przyrodniczego
Uniwersytet im. Adama Mickiewicza w Poznaniu

**Recenzja rozprawy doktorskiej mgr inż. Saida Nawara pt.
„Digital soil mapping using spectroscopy and remote sensing:
a case study from Egypt”**

Niniejsza rozprawa przygotowana została pod opieką prof. UJ dr hab. Jacka Kozaka. Do oceny przedstawiono ją w formie 54-stronicowego maszynopisu omawiającego:

- obecny stan wiedzy na temat numerycznego kartowania gleb wykorzystującego techniki bliższej teledetekcji (*proximal sensing*) i teledetekcji satelitarnej do pozyskiwania danych o przestrzennej zmienności właściwości pokrywy glebowej,
- a przede wszystkim efekty badań przeprowadzonych przez Autora niniejszej rozprawy, publikowane w trzech oddzielnych współautorskich pracach, poszerzające wiedzę na temat efektywności wykorzystania kartografii numerycznej do określenia zmienności przestrzennej i czasowej zasolenia gleb, zawartości materii organicznej (OM) i frakcji ilastej.

Do tego maszynopisu, zawierającego bardzo liczne odwołania do literatury, obejmującej 142 pozycje, i 7 rycin z odpowiednimi wykresami i fotografiami odnoszącymi się do badanych przez Autora gleb, dołączono również kopie wspomnianych publikacji oraz deklaracje współautorów o ich procentowym wkładzie w przygotowanie poszczególnych prac do druku. Mgr. inż. Said Nawar we wszystkich tych publikacjach jest pierwszym autorem. Wszystkie trzy zostały napisane z prof. Joachimem Hillem i dr Henningiem Buddenbaumem jako współautorami z Environmental Remote Sensing and Geoinformatics Uniwersytetu Trir w Niemczech. Pierwsza z tych prac, zatytułowana: „Estimation of soil salinity using three quantitative methods based on visible and near infrared reflectance spectroscopy: a case study from Egypt”, opublikowana została w 2014 r. online przez Wydawnictwo Springer na 14 stronach w *Arabian Journal of Geoscience* z Impact Factor - 1,152, wycenianym przez MNiSW na 20 punktów. Druga, pt.: „Modeling and mapping of soil salinity with reflectance spectroscopy and Landsat data using two quantitative methods (PLSR and MARS)”, została napisana także z udziałem Promotora prof. dr hab. Jacka Kozaka. Ta druga, licząca 22 strony,

i trzecia 25-stronicowa, zatytułowana: "Digital mapping of soil properties using multivariate statistical analysis and ASRER data in an arid region" opublikowane zostały odpowiednio w 2014 i 2015 r. w czasopiśmie *Remote Sensing* o swobodnym dostępie (*Open Access*) z Impact Factor - 2,729; i 35 punktami wg. MNiSW.

Jak wynika z oświadczeń współautorów tych trzech publikacji, wkład mgr. inż. Saida Nawara wynosi odpowiednio: do stworzenia ich koncepcji – 80-90%, opracowania przeglądu literatury – 100%, zaprojektowania badań – 85-90%, zebrania danych – 100%, analizy danych – 80-85%, prezentacji wyników – 95%, interpretacji i dyskusji wyników – 80-85% i redakcji tekstu – 80-85%.

Niniejsza rozprawa spełnia wszystkie wymogi określone w siedmiu punktach Uchwały 01 Rady Instytutu Geografii i Gospodarki Przestrzennej Uniwersytetu Jagiellońskiego z dnia 11 grudnia 2014 r. w sprawie szczegółowych wytycznych dla rozpraw doktorskich tworzonych na podstawie spójnego tematycznie zbioru opublikowanych artykułów.

Uwagi o stanie wiedzy o numerycznym kartowaniu gleb

Informacje dotyczące klasycznej kartografii gleboznawczej, wspomaganej nawet fotointerpretacją, zawarte są tu w kilku zdaniach, a w trzech wymienionych wyżej publikacjach nie ma ich wcale. Słusznie, bo dociekania Autora ocenianej pracy doktorskiej skupiają się na znalezieniu na przykładzie gleb słonych w Egipcie najefektywniejszych metod numerycznego kartowania zmienności przestrzennej i czasowej zasolenia gleb, a także zawartości w nich materii organicznej i frakcji ilastej. Mgr. inż. Said Nawar posłużył się teledetekcyjnymi technikami: spektroskopii hiperspektralnej gleb z dystansu kilkunastu cm w laboratorium i kilkudziesięciu cm w polu oraz technikami wykorzystującymi multispektralne dane satelitarne uzyskiwane z pułapu ponad 700 km w zakresie widzialnym, bliskiej i środkowej podczerwieni. Niezależnie od wielkości tego dystansu, do wydobycia ilościowych informacji o badanych właściwościach gleb za pośrednictwem promieniowania odbitego w tym szerokim zakresie spektrum od gleby *ex situ* w warunkach laboratoryjnych czy *in situ* w polu, użył odpowiednich funkcji matematycznych i metod statystycznych. Funkcje te na wstępnym etapie przetwarzania danych spektralnych miały na celu redukcję szumów zarejestrowanego promieniowania, wygładzenie przebiegu tego promieniowania i wzmocnienia słabszego sygnału we funkcji długości fali. Na etapie finalnym, już inne funkcje zawarte w liniowych i nieliniowych modelach, pozwalały obliczyć z tych przetworzonych

danych spektralnych szukane właściwości gleby ilościowo wyrażone, oszacowując jednocześnie ich statystyczną poprawność. Takie postępowanie było już wcześniej wykorzystywane do pozyskiwania ilościowych informacji o glebach. Zostało ono obszernie i wyczerpująco przedstawione w rozdziale 2 maszynopisu niniejszej rozprawy, omawiającym obecny stan wiedzy na temat kartowania numerycznego gleb. Nie przetestowano jednak do tej pory efektywności modeli nieliniowych do kartowania numerycznego zasolonych gleb przy udziale teledetekcji bliższej i teledetekcji satelitarnej. **Zatem, badanie efektywności modelu nieliniowego w kontekście gleb zasolonych stanowi niewątpliwie novum tej dysertacji.**

W rozdziale tym zawarto też informacje o tym co rozumie się obecnie pod pojęciem bliższej teledetekcji gleb (*proximal soil sensing*). **Ja dodałbym do nich, za McBratney'em i in. (2011), że jest to nowa bardzo dynamicznie rozwijająca się dyscyplina, gdzie nico arbitralnie traktuje się słowo „bliższej” (*proximal*) w pojęciu bliższej teledetekcji. Sensor może dotyczyć badany obiekt a odległość pomiędzy badanym obiektem a obserwującym (wyczuwającym) go sensorem może być rzędu milimetrów, centymetrów lub metrów, ale nie dziesiątków metrów, setek metrów czy jeszcze więcej.** Wystarczająco dokładnie omówiono też w tym rozdziale zagadnienia dotyczące: dyfuzyjnej spektroskopii gleb w zakresie widzialnym i bliskiej podczerwieni. Omawiając krzywe charakteryzujące refleksyjność hiperspektralną gleb, które gromadzi się od 2006 roku razem z wynikami analiz laboratoryjnych gleb w bibliotekach spektralnych gleb na świecie, zwrócono szczególną uwagę na odcinki spektrum, gdzie obserwowane są wyraźne efekty absorpcji powodowane obecnością materii organicznej, tlenków żelaza, węglanów i wody higroskopowej oraz specyficznych minerałów zawartych w materiale glebowym.

Powyższe zagadnienia intensywnie zajmują naukowców zaledwie od około dwóch dekad. Stąd, odniesienia do literatury wykorzystywanej w tym rozdziale maszynopisu, ale także i w tych trzech wspomnianych publikacjach uwzględnianych w niniejszej rozprawie, dotyczą w większości prac publikowanych w tym okresie. Prace opublikowane przed 2000 rokiem stanowią tu tylko około 10-20% ich całkowitej liczby, a prace opublikowane między 2000 i 2010 r. oraz po 2010 r., a więc w czasie o dwukrotnie krótszym, - odpowiednio 45-50% i 30-40%.

Mgr. inż. Said Nawar w podsumowaniu słusznie wyraża opinię, że wykorzystanie hiperspektralnych danych uzyskanych za pomocą teledetekcji bliższej w laboratorium czy w polu w połączeniu z teledetekcyjnymi multispektralnymi danymi pozyskanymi za pośrednictwem satelitów w zakresie optycznym pozwala znacznie wydajniej i bardziej

opłacalnie ekonomicznie kartować numerycznie właściwości gleb, niż w sposób tradycyjny oparty na uzyskiwaniu ilościowych informacji o właściwościach fizycznych i chemicznych gleb tylko poprzez analizy laboratoryjne tych właściwości z bardzo wielu prób pobranych w polu. Podkreśla, że stosując metody teledetekcji do kartowania gleb nie możemy jeszcze korzystać z opracowanego standardowego postępowania. Brakuje bowiem odpowiedniego doświadczenia i wiedzy wynikającej ze zbyt małej liczby przetestowanych do tej pory warunków glebowych i zastosowanych sposobów przetwarzania danych spektralnych.

Uwagi dotyczące badań przeprowadzonych przez Autora rozprawy

Zatem, zasadniczym celem niniejszej pracy doktorskiej, przedstawionej w trzech wymienionych wyżej publikacjach, stało się przetestowanie efektywności metod bliższej teledetekcji i teledetekcji satelitarnej w zakresie optycznym dla przedstawienia zmienności przestrzennej i czasowej wybranych właściwości gleb słonych. Tę efektywność oceniano, skupiając się przede wszystkim na wyborze najodpowiedniejszych funkcji matematycznych i metod statystycznych umożliwiających wydobycie jak najdokładniejszych informacji o wybranych właściwościach gleb z ich danych spektralnych. Wszystkie te publikacje łączy to samo miejsce badań, położone w północno-zachodniej części Półwyspu Synaj. Dwie pierwsze z nich opierają się na danych glebowych zebranych z tej samej powierzchni o areale 175 km², natomiast trzecia – z odpowiednio rozszerzonej tej powierzchni do 380 km².

Pierwsza praca, zatytułowana: „Estimation of soil salinity using three quantitative methods based on visible and near infrared reflectance spectroscopy: a case study from Egypt”, opublikowana w *Arabian Journal of Geoscience* w 2014, koncentrowała się na szukaniu najodpowiedniejszego sposobu wstępnego przetwarzania hiperspektralnych widm reflektancji tych gleb zmierzonych w warunkach laboratoryjnych, który zapewniłby możliwie najściślejsze ich skorelowanie ze stanem zasolenia analizowanych gleb wyrażonym za pomocą przewodności elektrycznej (EC_e) zmierzonej też w laboratorium. Wykorzystano widma 94 prób glebowych, wcześniej powietrznie wysuszonych, roztartych, przesianych przez sito o oczkach 2 mm. Próbkę o grubości 2 cm, o wyrównanej powierzchni, umieszczono na tacach (**niestety o nie określonej wielkości**). Oświetlono je wolframową lampą halogenową, oddaloną od próbek o 45 cm, a ich radiancję mierzono z odległości 15 cm pod kątem zenitalnym 30° za pomocą hiperspektralnego radiometru w zakresie od 350 do 2500 nm. **Brakuje mi w tym miejscu dokładniejszej informacji o położeniu zenitalnym lampy**

nad próbką i też o usytuowaniu azymutalnym sensora radiometru względem źródła światła. Ponieważ, wielkość odbitego promieniowania od tak przygotowanych próbek docierająca do sensora zależy zarówno od kierunku obserwacji, jak i kierunku oświetlenia próbki opisanych kątami zenitalnymi i azymutalnymi. Nie dość poprawne wydaje się też sformułowanie odnoszące się do sposobu obliczania absolutnej reflektancji tych prób glebowych przy wykorzystaniu wzorca bieli. Bardziej poprawnie byłoby, gdyby wyrazić to w następujący sposób: *“To calculate the absolute reflectance of the samples, the radiance from each samples was divided by the radiance from the white reference panel, both under the same illumination and viewing conditions, and multiplied by the correction factor of the reference panel”*, zastępując zdanie kończące podrozdział *“Pomiary spektralne”* (na stronie 3). To samo zdanie zawarto także w publikacji nr 2 (na str. 10817), a więc i tam wymagało ono takiej korekty.

Z widm nieprzetworzonych oraz przetworzonych za pomocą pierwszej pochodnej wygładzonej filtrem Savitzky'ego-Golaya (FD-SG) oraz metodą *Continuum-removed* (CR) obliczano trzy wskaźniki spektralne dla wszystkich możliwych kombinacji par kanałów spektralnych, korelując je metodą regresji liniowej z wartościami EC_e . Najściślejszą korelację ustalono dla wskaźnika wyrażającego stosunek reflektancji badanych gleb dla długości fali 1483 nm do ich reflektancji dla fali o długości 1918 nm, odnoszącego się do widma przetworzonego za pomocą funkcji FD-SG. Dla przewidywania wielkości zasolenia gleb wyrażonego przez EC_e za pośrednictwem odpowiednio przetworzonych danych spektralnych badanych gleb posłużono się modelem cząstkowych najmniejszych kwadratów (*partial least-square regression* – PLSR) i modelem wielozmiennej regresji adaptacyjnej z użyciem funkcji sklejonych (*multivariate adaptive regression splines* - MARS), wcześniej nie wykorzystywanym do kartografii gleb zasolonych.

Liniowy model PLSR jest powszechnie wykorzystywany w chemometrii i ilościowej analizie spektralnej. Opisuje zależność między dwoma macierzami danych, danych niezależnych X dotyczących widma i danych zależnych Y odnoszących się w tym wypadku do zmierzonych właściwości gleby. Algorytm PLSR za pomocą wieloczynnikowego modelu liniowego wybiera kolejno następujące po sobie czynniki ukryte, które maksymalizują zależność liniową między zmiennymi X i Y. Nieliniowy model MARS, stosowany ostatnio w wielu dyscyplinach, ustala zależności pomiędzy zmiennymi X i Y, wykorzystując regresję liniową w krótszych zdefiniowanych odcinkach zależności między nimi. Jego nieliniowość jest symulowana przez zmieniające się w tych odcinkach nachylenie tej zależności.

W omawianej pracy ustalono, że model MARS dawał większą poprawność wyrażoną współczynnikiem determinacji (R^2), osiągającym wartość 0,81, niż model PLSR z $R^2=0,77$.

W drugiej z kolei pracy, zatytułowanej: „Modeling and mapping of soil salinity with reflectance spectroscopy and Landsat data using two quantitative methods (PLSR and MARS)”, opublikowanej w *Remote Sensing* w 2014 r., modele PLSR i MARS wykorzystane do sporządzenia map numerycznych zasolenia gleb, zbudowano w taki sposób jak w pracy poprzedniej, korzystając z laboratoryjnych pomiarów spektralnych prób glebowych i ich przewodności elektrycznej ustalonej też w laboratorium. Zmienność przestrzenną zasolenia gleb, badanych w tym samym miejscu Półwyspu Synaj jak w pracy poprzedniej, określono za pośrednictwem odpowiednio przetworzonych danych satelitarnych Landsat 5 *Thematic Mapper* (TM) z 2006 r. i Landsat 7 *Enhanced Thematic Mapper Plus* (ETM+) z 2012. Wykorzystanie tych danych, zarejestrowanych w odstępie 6 lat, pozwoliło również ocenić zmienność czasową zasolenia badanych gleb. **Współrzędne opisujące położenie obszaru badań w niniejszej pracy i poprzedniej są takie same, choć rysunki (nr 1 w obu publikacjach) ujawniają wyraźne rozbieżności. W niniejszej pracy pisze się o 70 pobranych próbkach w 2012 r. a na rysunku 1 widnieje 81 punktów odnoszących się według legendy do tego okresu. Dostrzegam na nim też tylko 32 punkty, zamiast 37, przeznaczone do walidacji rozkładu zasolenia gleb w 2006 r. W poprzedniej pracy nie wspomniano kiedy zgromadzono analizowane w niej próby glebowe. Porównując wyniki ich zasolenia, wyrażone przez EC_e , i innych ich właściwości ($CaCO_3$, OM, odczynu i tekstury) z odpowiadającymi im wynikami prób pobranych w 2012, prezentowanych w niniejszej pracy, można sądzić, że są to te same próby. Potwierdza to też ich identyczna lokalizacja przedstawiona na rysunkach 1 w obu pracach. W niniejszej pracy, tak jak i w poprzedniej i kolejnej nie poinformowano w jakim laboratorium analizowano właściwości fizyczne i chemiczne gleb, a przecież instytucja, w której laboratorium funkcjonuje może świadczyć o poziomie poprawności wyników. Wspomina się tu tylko o laboratorium Uniwersytetu Trier jako miejscu wykonywania pomiarów spektralnych gleb.**

Nie we wszystkich miejscach, skąd pobierano próby glebowe, mierzono też refleksyjność gleb w warunkach polowych przy użyciu radiometru polowego. Ze względu na zgromadzenie niedostatecznej liczby polowych pomiarów spektralnych, użyto je tylko do kalibracji i walidacji danych analizowanych obrazów satelitarnych. Modele PLSR i MARS, wykorzystywane w tej pracy dostosowano do rozdzielczości multispektralnej radiometrów

satelitarnych TM i ETM+, dokonując także korekcji tych danych ze względu na wpływ atmosfery.

Tak jak we wcześniej opublikowanej pracy, tak i w niniejszej ustalono, że wykorzystanie nieliniowego modelu MARS do predykcji zasolenia gleb wyrażonej za pomocą EC_e dawał poprawniejsze wyniki niż liniowy model PLSR. Poprawność prognozowania EC_e za pomocą modelu MARS została opisana współczynnikiem determinacji $R^2=0,73$ i błędem średniokwadratowym $RMS=6,53 \text{ dS m}^{-1}$, natomiast modelu PLSR odpowiednio wartościami $R^2=0,70$ i $RMS=6,95 \text{ dS m}^{-1}$. Dokonując walidacji tych modeli za pomocą 37 i 26 punktów sprawdzających, odnoszących się odpowiednio do stanu zasolenia w 2006 i 2012 r., model MARS zapewnił rozkład EC_e przy wykorzystaniu danych Landsat 5 TM i Landsat 7 ETM+, z poprawnością $R^2=0,72$ i $R^2=0,74$, natomiast model PLSR z nieznacznie niższymi wartościami $R^2=0,71$ i $R^2=0,73$.

W trzeciej z kolei pracy, zatytułowanej: "Digital mapping of soil properties using multivariate statistical analysis and ASRER data in an arid region", opublikowanej w 2015 r. w *Remote Sensing*, efektywność powyższych modeli analizowano w kontekście kartowania nie tylko zmienność przestrzennej zasolenia gleb, ale również zawartości materii organicznej (OM) i frakcji ilastej (F_c) gleb strefy suchej. Wykorzystano dane z radiometru satelitarnego *Advanced Spaceborne Thermal Emission and Reflection Radiometer* (ASTER), wcześniej nie testowane przez Autora rozprawy doktorskiej, i wyniki analiz tych właściwości uzyskanych z prób glebowych pobranych w czasie zbliżonym do czasu rejestracji danych satelitarnych. Użyto obrazy ASTER z 31 maja 2006 r. oraz próby glebowe pobrane w szczycie pory suchej pomiędzy 26 maja i 8 czerwca 2006 r., kiedy pokrycie roślinnością jest najmniejsze. Badaniem objęto obszar rozszerzony istotnie na południe w stosunku do omawianego w dwóch pierwszych publikacjach, o areale 380 km^2 przeszło dwukrotnie większym niż w pracach wyżej omawianych.

Obrazy ASTER poddano korekcji atmosferycznej, a piksele o rozdzielczości przestrzennej 30 m w kanałach z zakresu *Short Wave Infra-Red* (SWIR) przetworzono do rozdzielczości 15 m, jaką posiadają piksele w kanałach *Visible Near Infra-Red* (VNIR). W modelach predykcyjnych MARS i PLSR rolę zmiennych niezależnych (leżących na osi X tych modeli) pełniły dane spektralne ASTER, a rolę zmiennych zależnych (leżących na osi Y) – dane dotyczące prognozowanych właściwości gleb, ujęte ilościowo (EC_e , OM i F_c). Do zbudowania tych modeli skorelowano dane glebowe charakteryzujące te trzy cechy glebowe oznaczone w 86 pobranych próbach z właściwościami spektralnymi miejsc skąd pobrano te próby

zarejestrowanymi w 3 kanałach VNIR i 6 kanałach SWIR obrazu ASTER. Dla wszystkich trzech analizowanych właściwości gleb, posłużenie się modelem MARS dało lepsze wyniki niż modelem PLSR. Świadczą o tym wyższe wskaźniki determinacji (R^2) dla MARS niż PLSR, wynoszące odpowiednio 0,85 i 0,80 dla przewodności elektrycznej EC_e , 0,94 i 0,90 dla zawartości frakcji ilastej (F_c) oraz 0,79 i 0,73 dla zawartości materii organicznej (OM). Niezależna ocena poprawności map przedstawiających zmienność EC_e , F_c i OM przy użyciu tych modeli, przeprowadzona z użyciem 32 próbek gleby też dowiodła wyższą efektywność modelu MARS, z wartościami R^2 wynoszącym dla nich odpowiednio 0,81, 0,89 i 0,73, w porównaniu do wartości $R^2 = 0,78, 0,87$ i $0,71$ dla modeli PLSR. Powyższe wyniki wskazują dobitnie, że użycie modelu MARS jest właściwsze niż modelu PLSR w kartowaniu gleb w strefie półpustynnej.

W niniejszej publikacji nie dostrzegłem innych braków poza tym, że nie poinformowano w jakim laboratorium analizowano właściwości fizyczne i chemiczne badanych gleb, o czym wspominałem w omówieniu drugiej publikacji.

Wszystkie krytyczne uwagi dotyczące ocenianej rozprawy doktorskiej uwypukliłem posługując się wytłuszczonym tekstem.

Oceniona przeze mnie rozprawa doktorska spełnia warunki określone w artykule 13 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki. Niniejsza praca stanowi oryginalne rozwiązanie problemu naukowego. Testowanie efektywności nieliniowego modelu MARS w kontekście gleb zasolonych strefy suchej stanowi niewątpliwie novum tej dysertacji. Stan wiedzy prezentowany w maszynopisie tej pracy jak i trzech opublikowanych pracach świadczy o bardzo głębokiej wiedzy teoretycznej mgr. inż. Saida Nawara i jego odpowiednich umiejętnościach do samodzielnego prowadzenia pracy naukowej, poświadczonych dodatkowo Jego procentowym wkładem w przygotowanie trzech opublikowanych pracach dołączonych do niniejszej rozprawy doktorskiej.

Poznań, 21 lutego 2015

