

dr Krzysztof Lamorski
Zakład Metrologii i Modelowania Procesów Agrofizycznych
Instytut Agrofizyki im. B. Dobrzańskiego, PAN

AUTOREFERAT

1. Imię i nazwisko

Krzysztof Lamorski

2. Posiadane dyplomy, stopnie naukowe – z podaniem nazwy, miejsca i roku ich uzyskania oraz tytuł rozprawy doktorskiej

2005 Instytut Agrofizyki PAN w Lublinie, Zakład Metrologii i Modelowania Procesów Agrofizycznych. Stopień doktora nauk rolniczych w zakresie agronomii, specjalność agrofizyka.

Tytuł rozprawy doktorskiej: „Modelowe badania pola temperatury w profilu glebowym z lokalnym zaburzeniem właściwości cieplnych”.

1997 Tytuł magistra fizyki, magisterskie studia stacjonarne, Uniwersytet Marii Curie Skłodowskiej w Lublinie, Wydział Matematyki i Fizyki, Kierunek Fizyka Teoretyczna.

Tytuł pracy magisterskiej: „Modele kosmologiczne z zależnym od czasu członem Λ ”.

3. Informacje o dotychczasowym zatrudnieniu w jednostkach naukowych

od 11/2014	Instytut Agrofizyki im. B. Dobrzańskiego PAN, starszy fizyk.
11/2005 – 10/2014	Instytut Agrofizyki im. B. Dobrzańskiego PAN, adiunkt.
11/2000 – 10/2004	Instytut Agrofizyki im. B. Dobrzańskiego PAN, asystent.
04/1998 – 08/2000	Instytut Agrofizyki im. B. Dobrzańskiego PAN, asystent.
11/1997 – 03/1998	Instytut Agrofizyki im. B. Dobrzańskiego PAN, młodszy fizyk.

od 10.2007	Instytut Nauk Rolniczych, Państwowa Wyższa Szkoła Zawodowa w Chełmie, starszy wykładowca
10.2001 - 09.2007	Instytut Matematyki i Informatyki, Państwowa Wyższa Szkoła Zawodowa w Chełmie, wykładowca

4. Wskazanie osiągnięcia wynikającego z art. 16 ust. 2 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. nr 65, poz. 595 ze zm.)

a) tytuł osiągnięcia naukowego

Modelowanie retencji i transportu wody w glebie z wykorzystaniem technik sztucznej inteligencji.

b) autorzy, tytuły, rok wydania, nazwa wydawnictwa

1. **Lamorski K.**, Pachepsky Y., Sławiński C., Walczak R.: *Using support vector machines to develop pedotransfer functions for water retention of soils in Poland*, Soil Science Society of America Journal, 2008, vol. 72, no. 5, pp. 1243-1247, doi:10.2136/sssaj2007.0280N. **(30 pkt. MNiSW, IF: 2,104, cytowania: 36)**
2. **Lamorski K.**, Sławiński C., Moreno F., Barna G., Skierucha W., Arrue J.: *Modelling soil water retention using support vector machines with genetic algorithm optimisation*, 2014, Scientific World Journal, vol. 2014, Article ID 740521, <http://dx.doi.org/10.1155/2014/740521>. **(30 pkt. MNiSW, IF: 1,219, cytowania: 2)**
3. **Lamorski K.**, Pastuszka T., Krzyszczyk J., Sławiński C., Witkowska-Walczak B.: *Soil water dynamic modeling using the physical and support vector machine methods*, Vadose Zone Journal, 2013, vol.12, no.4, doi:10.2136/vzj2013.05.0085. **(35 pkt. MNiSW, IF: 2,2)**
4. **Lamorski K.**, Bieganowski A., Ryzak M., Sochan A., Sławiński C., Stelmach W.: *Assessment of the usefulness of particle size distribution measured by laser diffraction for soil water retention modelling*. Journal of Plant Nutrition and Soil Science, 2014, vol. 177, pp. 803-813, doi: 10.1002/jpln.201300594. **(30 pkt. MNiSW, IF: 1,663)**

Liczba punktów MNiSW publikacji wchodzących w skład osiągnięcia naukowego wg. wykazu MNiSW z dnia 31 grudnia 2014: **125**.

Całkowity IF publikacji wchodzących w skład osiągnięcia naukowego wg. daty opublikowania: **7,186**.

Mój wkład jako autora wiodącego w przypadku wszystkich artykułów dotyczył ustalenia koncepcji badań oraz opracowania wyników i wniosków. Samo przeprowadzenie badań, niezbędnych obliczeń czy przygotowania narzędzi programistycznych niezbędnych do wykonania badań (tam gdzie to było niezbędne) również było wykonane przeze mnie.

Całość modelowania wykorzystującego techniki sztucznej inteligencji była wykonana przeze mnie. Zadanie modelowania zmian wilgotności w profilu glebowym z wykorzystaniem klasycznej techniki bazującej na równaniu Richardsa (publikacja 2) było wykonane przez jednego ze współautorów.

Byłem głównym autorem przygotowującym manuskrypt w przypadku dwóch prac, w przypadku dwóch pozostałych współuczestniczyłem w przygotowaniu manuskryptu.

- c) omówienie celu naukowego ww. pracy i osiągniętych wyników wraz z omówieniem ich ewentualnego wykorzystania:

Wstęp

Woda zawarta w glebie jest jednym z ważniejszych czynników wpływających na szereg procesów zachodzących w naturalnym środowisku. Wzrost roślin, plonowanie jest nierozdzielnie związane z dostępnością wody glebowej, gdzie zarówno jej nadmiar jak i niedobór niekorzystnie wpływa na roślinę. Woda glebowa i jej transport odgrywa ważną rolę nie tylko w praktyce rolniczej ale i w innych obszarach takich jak np.: inżynieria środowiska, hydrologia zlewni czy inżynieria budownictwa. Retencja i transport wody

glebowej jest również istotnym składnikiem ogólnego obiegu wody w przyrodzie który może być analizowany w mikro, mezo i makro skali w kontekście zmian klimatu.

Dlatego też badania dotyczące zjawisk transportu wody glebowej są wciąż prowadzone. Badania te mają charakter zarówno badań podstawowych jak i badań stosowanych.

W zakresie badań podstawowych analizuje się fizyczne podstawy zjawisk transportu wody w glebie, a sam transport wody glebowej jest traktowany jako element szerszego problemu - zjawisk transportu w ciałach porowatych. Istnieją wciąż otwarte zagadnienia badawcze o fundamentalnym znaczeniu dla opisu stanu wody w glebie, nie rozwiązane lub dyskusyjne. Można zaliczyć do nich między innymi: poprawność definicji potencjału wody glebowej, zjawisko retencji wodnej czy zjawisko dynamicznego nierównowagowego transportu wody (Diamantopoulos i Durner, 2012). Nowe techniki eksperymentalne oraz nowe metody modelowania które można obecnie stosować do transportu w ciałach porowatych dają szansę na lepszy wgląd w fizyczne podstawy zjawisk transportu wody w glebie.

Natomiast badania stosowane dotyczące wody glebowej koncentrują się głównie na rozwoju praktycznie wykorzystywanych technik modelowania i monitorowania zmian wilgotności ośrodka glebowego. Zwiększenie efektywności i dokładności technik modelowania pozwala na bardziej racjonalne gospodarowanie dostępnymi zasobami wody, poprzez możliwość predykcji wilgotności gleby, powiązanie zmian wilgotności gleby z rozwojem i plonami upraw, itp.

Badania przeprowadzone przez habilitanta, prezentowane jako osiągnięcie, dotyczyły rozwoju metod modelowania: transportu wody w ośrodku glebowym i krzywej retencji wodnej. Badania te mają zarówno aspekt praktyczny jak i poznawczy.

Cel

Celem przeprowadzonych badań, przedstawionych w monotematycznym cyklu publikacji, było zbadanie możliwości zastosowania nowych technik sztucznej inteligencji do modelowania krzywej retencji wodnej gleb oraz do modelowania zjawiska transportu wody w ośrodku glebowym.

Poszukiwano metody modelowania pozwalającej na zwiększenie dokładności predykcji przebiegu krzywej retencji oraz pozwalającej wykorzystać informacje o glebie pozyskiwane nowymi metodami pomiarowymi. Celem pracy było również opracowanie alternatywnej do klasycznej - bazującej na rozwiązaniu równania Richardsa - metody modelowania zmian wilgotności ośrodka glebowego zachodzących w warunkach naturalnych.

Metody

Krzywa retencji wodnej jest podstawową charakterystyką hydrofizyczną gleby. Opisuje ona zależność pomiędzy potencjałem wody glebowej a wilgotnością ośrodka glebowego. Zjawisko retencji wodnej samo w sobie jest wciąż obiektem badań

podstawowych i stosowanych. Znajomość krzywej retencji wodnej jest niezbędna między innymi do modelowania transportu wody w ośrodku glebowym.

Istnieją dwie podstawowe metody wyznaczenia krzywej retencji wodnej: pomiar laboratoryjny lub estymacja jej przebiegu przy pomocy odpowiednich modeli. Laboratoryjne wyznaczenie krzywej retencji wodnej daje najdokładniejsze wyniki lecz jest procedurą długotrwałą (typowo trwa kilka miesięcy) i w związku z tym nie zawsze może być stosowana. Stąd popularność różnych metod pośredniego wyznaczania krzywej retencji wodnej, tzw. "pedotransfer functions" (Bouma, 1989). Krzywą retencji wodnej można estymować na dwa sposoby: bezpośrednio - estymując wilgotności gleby dla wybranych, z góry określonych wartości potencjału wody glebowej, lub pośrednio (metoda parametryczna) - estymując parametry funkcyjnego przedstawienia krzywej retencji wodnej, obecnie najczęściej jest w tym celu wykorzystywana parametryzacja bazująca na funkcji Mualema- van Genuchtena (M-vG).

Modele krzywej retencji wodnej budowane są w oparciu o eksperymentalne informacje dot. retencji różnych gleb i ich charakterystyk fizykochemicznych. Opracowany model retencji pozwala na estymację krzywej retencji wodnej na podstawie łatwiej mierzalnych charakterystyk gleby. Do parametrów gleby najczęściej wykorzystywanych w tym celu należą: rozkład granulometryczny, zagęszczenie oraz zawartość materii organicznej.

Z uwagi na ich dużą rolę praktyczną wciąż są rozwijane techniki modelowania krzywej retencji.

Techniki sztucznej inteligencji są klasą algorytmów pozwalających na zaawansowane przetwarzanie danych, w szczególności pozwalających na wyszukiwanie i wykorzystywanie do budowy modeli predykcyjnych ukrytych informacji o zależnościach pomiędzy zmiennymi w analizowanych zbiorach danych. Spośród różnych algorytmów należących do tej kategorii, w prezentowanych pracach wykorzystywane są: Sztuczne Sieci Neuronowe (SSN) oraz Support Vector Machines (SVM). Do kategorii technik sztucznej inteligencji należą również algorytmy optymalizacyjne, w szczególności stosowane w pracach Algorytmy Genetyczne (AG).

Obie metody: SVM i SSN są technikami eksploracji danych pozwalającymi na modelowanie złożonych nieliniowych zależności pomiędzy różnymi zmiennymi w analizowanych zbiorach danych.

Podstawowa metodyka stosowania algorytmów eksploracji danych bazuje na kilku etapach: normalizacji dostępnych danych, podziale danych na zbiór uczący i testowy, opracowaniu modelu na podstawie zbioru uczącego, testowaniu poprawności działania modelu na podstawie zbioru testowego.

Normalizacja danych jest niezbędnym krokiem wstępnym polegającym na takiej transformacji danych aby wartości wszystkich przetwarzanych danych (zmiennych) miały wartości mniej więcej z tego samego zakresu. Procedura ta ma na celu uniknięcie sytuacji, kiedy jedna ze zmiennych w zbiorze danych, ze względu na nieproporcjonalnie większe wartości w porównaniu z pozostałymi determinuje działanie opracowywanych

modeli. Różne metody wstępnej normalizacji danych są stosowane standardowo w technikach sztucznej inteligencji.

Technika SVM podobnie jak sztuczne sieci neuronowe jest jedną z metod sztucznej inteligencji. SVM została opracowana w 1995 roku jako metoda nieliniowej klasyfikacji (Vapnik, 1995). Dwa lata później rozszerzono SVM do zastosowań typu regresyjnego (Ducker et al., 1997). Pomimo zupełnie innych matematycznych podstaw działania SSN i regresyjnych SVM, z punktu widzenia aplikacyjnego, SVM przypominają szczególny rodzaj SSN - sieci jednokierunkowe.

Modele bazujące na technice SVM zależą w swoim działaniu od wybranej funkcji jądra. Znanych jest kilka różnych form funkcji jądra. W prezentowanych pracach stosowano: funkcję jądra bazującą na radialnych funkcjach rozkładu wyrażającą się wzorem: $K(\bar{x}^i, \bar{x}^j) = e^{-\gamma \|\bar{x}^i - \bar{x}^j\|^2}$ gdzie: \bar{x}^i, \bar{x}^j są wektorami ze zbioru danych natomiast γ jest dodatkowym dobieralnym parametrem modelu oraz liniową funkcję jądra wyrażającą się następującym wzorem: $K(\bar{x}^i, \bar{x}^j) = \bar{x}^i \cdot \bar{x}^j$.

Uczenie regresyjnego modelu bazującego na algorytmie SVM można w wielkim uproszczeniu opisać jako procedurę wyboru spośród dostępnego (stosowanego do uczenia) zbioru danych, najbardziej reprezentatywnych danych. Te wyselekcjonowane w procedurze uczenia dane nazywane są wektorami wspierającymi (SV - support vectors) i stają się bezpośrednio podstawowymi elementami opracowywanego modelu.

Wytrenowany model, na podstawie wektorów wspierających i innych składowych modelu, np. funkcji jądra i parametrów modelu może być wykorzystywany do estymacji wartości na podstawie nowych danych wejściowych.

Opis badań i otrzymanych wyników

Przedstawione poniżej badania są wynikiem mojej pracy nad wykorzystaniem różnych nowych technik eksploracji danych i technik sztucznej inteligencji w hydrologii gleby do modelowania retencji wodnej i transportu wody w ośrodku glebowym. Częściowo publikacje te powstały dzięki współpracy międzynarodowej, z naukowcami z Węgier, Hiszpanii i USA.

Publikacja 1

Problem badawczy

W pierwszym etapie badań zbadano możliwość wykorzystania nowej metody sztucznej inteligencji - Support Vector Machines (SVM) do modelowania krzywej retencji wodnej, oraz porównano wyniki uzyskiwane metodą SVM z wynikami uzyskanymi przy wykorzystaniu Sztucznych Sieci Neuronowych (SSN) (I.B.1).

W czasie gdy rozpoczęto prace nad tym artykułem do modelowania krzywej retencji wodnej standardowo stosowano technikę regresji nieliniowej lub SSN. Głównymi przesłankami które skłoniły mnie do próby zastosowania nowej metody - SVM była chęć poszukiwania bardziej dokładnej i efektywnej techniki modelowania krzywej retencji wodnej oraz praktyczne problemy związane z modelowaniem krzywej retencji przy pomocy SSN.

Opisywana publikacja (I.B.1) była pierwszą opublikowaną pracą stosującą nową technikę SVM do modelowania krzywej retencji wodnej.

Badania własne, metodyka

Pierwsze prace dotyczące estymacji krzywej retencji wodnej wykorzystywały proste techniki regresyjne, głównie wieloczynnikową regresję nieliniową (Ahuja i in. 1985; Walczak i in. 2001). Jednak z uwagi na złożone zależności pomiędzy fizycznymi charakterystykami gleb a ich krzywymi retencji dokładność technik regresyjnych pozostawiała wiele do życzenia.

Wraz z popularyzacją komputerowych technik przetwarzania danych i rozwojem technik sztucznej inteligencji sięgnięto po sztuczne sieci neuronowe w celu modelowania krzywej retencji wodnej. W tym celu były wykorzystywane wielowarstwowe sieci jednokierunkowe z uwagi na ich możliwości odtwarzania nieliniowych zależności pomiędzy wielkościami estymowanymi a zmiennymi niezależnymi. SSN były z powodzeniem wykorzystywane do modelowania krzywej retencji (Baker i in. 2008, Pachepsky i in. 1996), a ich dokładność predykcji przewyższała dokładność modeli bazujących na technikach regresyjnych.

Jednak ze stosowaniem SSN do modelowania retencji wiązały się pewne problemy mające wpływ na dokładność ich działania.

Jednym z nich był problem poprawnego wyboru architektury samej sieci neuronowej, który w przypadku sieci jednokierunkowych sprowadza się do określenia odpowiedniej ilości neuronów w warstwie ukrytej perceptronu. Istnieją co prawda reguły heurystyczne pozwalające oszacować ilość neuronów w warstwie ukrytej, jednak nie gwarantują one wyznaczenia optymalnej dla danego zbioru danych i modelowanych charakterystyk ilości neuronów w warstwie ukrytej. W praktyce, autorzy wykorzystujący SSN do modelowania krzywej retencji, po prostu próbowali różnych architektur sieci w poszukiwaniu tej najlepszej, co komplikowało etap opracowywania modelu krzywej retencji. Alternatywą był arbitralny wybór architektury sieci - która mogła nie gwarantować najlepszych wyników estymacji krzywej retencji.

W odróżnieniu do SSN, ze względu na odmienny sposób działania, SVM nie wymagają określania obliczeniowej struktury modelu z góry. W pewnym sensie, odpowiednikiem neuronów z warstwy ukrytej SSN w przypadku modelu SVM jest ilość tak zwanych wektorów wspierających, które wykorzystane zostały do budowy modelu. Interesującą cechą algorytmu SVM jest to, że ilość wektorów wspierających jest automatycznie optymalizowana w trakcie procedury uczenia modelu SVM, w związku z czym, problem arbitralności wyboru architektury modelu przestaje istnieć.

Inny problem dotyczący SSN dotyczy efektywnego uczenia sieci neuronowych. Algorytmy wykorzystywane do uczenia sieci neuronowej, optymalizujące jej parametry, mają tendencję do "utykania" w minimach lokalnych hiperpłaszczyzny funkcji celu - co prowadzi do wytrenowania SSN która estymuje krzywą retencji w sposób nieoptymalny. Efekt ten ma związek z niejednoznacznością zależności pomiędzy parametrami ośrodka glebowego a krzywą retencji. Dodatkowo funkcja celu zależy od bardzo dużej liczby

parametrów charakteryzujących sieć neuronową których optymalizacji wartości służy. Na przykład dla najprostszego nawet perceptronu, wykorzystywanego do estymacji parametrów krzywej Mualema-van Genuchtena tych parametrów będzie około 50-ciu. Tak skomplikowane zadanie optymalizacyjne wykonywane w trakcie uczenia sieci neuronowej stosowanej jako model retencji często prowadziło do osiągnięcia wyników nieoptymalnych. Autorzy radzili sobie z tym problemem poprzez opracowywanie wielu sieci neuronowych modelujących krzywą retencji, które trenowane były rozpoczynając poszukiwania parametrów sieci neuronowej z różnych punktów startowych w przestrzeni możliwych wartości parametrów (Minasny i Perfect 2004; Baker i in. 2008). Po czym w celu estymacji krzywej retencji wykorzystywane były wszystkie wytrenowane sieci neuronowe których wyniki uśredniano. Dodatkowo stosowane były też techniki walidacji krzyżowej - poszczególne podmodele SSN opracowane były na innym podzbiórze dostępnych danych.

W przypadku algorytmu SVM nie ma konieczności optymalizacji tak dużej liczby parametrów. W najgorszym przypadku istnieje konieczność poprawnego doboru tylko trzech parametrów modelu, co powoduje że praktycznie problem "utykania" w minimach lokalnych nie istnieje. Przy czym jak się okazuje, liczba tych parametrów może jeszcze być zmniejszona bez większej szkody dla dokładności opracowanego modelu retencji, co było jednym z tematów odrębnej publikacji (I.B.2).

Omawiane tutaj przesłanki wskazujące na potencjalne korzyści ze stosowania metody SVM, zamiast modelowania bazującego na sztucznych sieciach neuronowych, nie przesądzają czy rzeczywiście wyniki estymacji krzywej retencji przez SVM będą porównywalne lub lepsze od tych generowanych przez SSN, co jest hipotezą roboczą tej pracy.

W celu weryfikacji tej hipotezy przeprowadzono badania polegające na opracowaniu dwóch różnych modeli krzywej retencji, jednego bazującego na technice SVM i drugiego wykorzystującego SSN. Aby można było wnioskować o efektywności działania SVM w porównaniu z SSN, oba modele opracowywane były w oparciu o ten sam zbiór danych glebowych.

Źródłem informacji o glebach wykorzystywanych w tej pracy była baza danych BiGLEB (Bieganowski i in. 2013), opracowana w IA PAN, grupująca informacje odnośnie reprezentatywnych polskich gleb mineralnych. Spośród dostępnych danych wybrano 806 poziomów pochodzących z 290 profili glebowych, dla których dostępne były informacje o: krzywej retencji, rozkładzie granulometrycznym i zagęszczeniu.

Celem było opracowanie modeli krzywej retencji wodnej estymujących bezpośrednio wilgotność objętościową gleby dla wybranych 11 wartości potencjału wody glebowej: -0.1 , -1.0 , -3.2 , -10.0 , -15.8 , -31.6 , -50.1 , -100.0 , -158.5 , -501.2 , oraz -1584.9 kPa na podstawie informacji o rozkładzie granulometrycznym i zagęszczeniu gleby.

Dostępne dane glebowe zostały znormalizowane następnie były losowo dzielone w stosunku 2:1, gdzie większa część danych wykorzystywana była jako zbiór uczący - zastosowany do opracowania modelu krzywej retencji SVM i SSN, natomiast mniejsza część była wykorzystywana do testowania opracowanego modelu. Taki losowy podział

danych oraz uczenie na jego podstawie był przeprowadzony pięć razy - co doprowadziło do opracowania pięciu niezależnych pod modeli każdego rodzaju (SVM i SSN). Ostatecznie jako wyniki predykcji modeli retencji przedstawiane były średnie z wyników poszczególnych pod modeli. Dzięki takiej procedurze można było skonfrontować SVM z optymalną techniką SSN pozwalającą na uniknięcie problemu minimum lokalnego funkcji celu.

W przypadku modelu SSN architektura sieci neuronowej składała się z trzech warstw: wejściowej, ukrytej i wyjściowej. W warstwie wejściowej były 3 neurony odpowiadające: zagęszczeniu, frakcji piaski i frakcji iltu. W warstwie wyjściowej był jeden neuron - odpowiadający estymowanej wilgotności gleby dla danego potencjału wody glebowej. Warstwa ukryta zawierała różne liczby neuronów - od 2 do 15 - gdyż autorzy chcieli mieć pewność zbudowania najlepszego modelu bazującego na SSN. Minimalna zastosowana liczba neuronów równa 2, była wartością zgodną z powszechnie akceptowaną rekomendacją (Maren i in. 1990) wyrażającą się wzorem $H = \sqrt{I \cdot O}$ gdzie: H - jest ilością neuronów w warstwie ukrytej, I - jest ilością neuronów w warstwie wejściowej a O - jest ilością neuronów w warstwie wyjściowej. Do uczenia sieci neuronowych wykorzystano dedykowane ogólnodostępne oprogramowanie SNNS (Stuttgart Neural Network Simulator). Ostatecznie do porównań z modelem SVM wybrano sieć o najlepszej architekturze. W każdym przypadku okazało się że siecią tą jest sieć o tylko dwóch neuronach w warstwie ukrytej.

W ramach tej pracy, w modelowaniu SVM, spośród kilku istniejących funkcji jądra, zastosowano funkcję jądra bazującą na radialnych funkcjach rozkładu. Wybór konkretnej funkcji jądra był wyborem arbitralnym jednak podyktowanym praktyką wykorzystywania SVM w obszarach innych niż modelowanie krzywej retencji - jest to najczęściej wybierana forma funkcji jądra w zagadnieniach typu regresyjnego, pozwala ona na osiągnięcie najlepszych rezultatów dla dużych zbiorów danych.

Ostatecznie model krzywej retencji wykorzystujący SVM zależał od trzech parametrów: C, ϵ i γ , które należało optymalnie dobrać tak, by dokładność predykcji modelu dla zbiorów uczących była najlepsza. Do tego celu zastosowano technikę walidacji krzyżowej Leave One Out (LOO).

W celu opracowania modeli retencji wykorzystujących SVM opracowano specjalnie do tego celu przeznaczone oprogramowanie bazujące na bibliotece programistycznej *SVMLIGHT* pozwalające na zastosowanie walidacji krzyżowej metodą LOO.

Po opracowaniu, oba modele krzywej retencji wodnej, bazujące na SVM i SSN, były porównane. Podstawowymi wskaźnikami statystycznymi stosowanymi jako kryterium porównań były: średni błąd kwadratowy pomiędzy wartościami wilgotności estymowanymi przez modele a wartościami pochodzącymi z pomiarów, oraz współczynnik determinacji R^2 .

Poniższa tabela pokazuje wyniki analizy statystycznej wskaźników poprawności estymacji wykonywanych przez modele SVM i SSN dla każdego z 11 potencjałów wody glebowej dla których estymowana i mierzona była wilgotność gleby.

Potencjał [kPa]	RMSE [cm ³ /cm ³]			R ² [-]		
	ANN	SVM	p	ANN	SVM	p
-0.1	0.004	0.006	0.003	0.994	0.992	0.065
-1	0.019	0.018	0.068	0.913	0.916	0.002
-3.2	0.041	0.039	0.000	0.743	0.770	0.000
-10	0.040	0.039	0.077	0.853	0.858	0.212
-15.8	0.041	0.041	0.979	0.853	0.852	0.385
-31.6	0.045	0.044	0.529	0.823	0.824	0.877
-50.1	0.051	0.050	0.535	0.777	0.783	0.451
-100	0.050	0.047	0.000	0.776	0.800	0.000
-158.5	0.051	0.047	0.000	0.757	0.787	0.000
-501.2	0.048	0.045	0.004	0.715	0.752	0.001
-1584.9	0.046	0.044	0.030	0.652	0.678	0.010

Wyniki, podsumowanie i wnioski

Wykazano że zaproponowana nowa technika modelowania może być z powodzeniem stosowana do modelowania krzywej retencji wodnej.

Analiza współczynnika determinacji R² pokazuje że model bazujący na SVM daje wyniki lepiej się korelujące w wartościami zmierzonymi niż model zbudowany w oparciu o sztuczne sieci neuronowe.

Na podstawie uzyskanych wyników można stwierdzić że model retencji wodnej bazujący na nowej metodzie modelowania SVM osiąga wyniki lepsze lub porównywalne z wynikami uzyskanymi przy pomocy starszej metody SSN. Analizując średni błąd kwadratowy predykcji widać że model SVM pracuje lepiej niż model SSN dla niższych potencjałów wody glebowej. Natomiast SSN daje lepszą predykcję dla wyższych potencjałów.

Biorąc pod uwagę te wyniki, można stwierdzić że opracowana metoda modelowania krzywej retencji wodnej bazująca na wykorzystaniu techniki SVM jest godna polecenia do tego typu studiów. Będąc pozbawiona wad metod bazujących na SSN daje wyniki porównywalne lub lepsze, co zresztą zostało potwierdzone w innych, późniejszych pracach (Twarakavi i in., 2009).

Publikacja 2

Problem badawczy

Doświadczenia zebrane w trakcie prac nad pierwszą publikacją cyklu (I.B.1), pokazały że istnieje konieczność rozwoju metod modelowania krzywej retencji wodnej bazujących na technice SVM. Aczkolwiek biorąc pod uwagę dokładność estymacji opracowanego modelu samo zastosowanie techniki SVM było sukcesem, to problemem okazała się zastosowana metoda poszukiwania optymalnych parametrów budowanego modelu. Choć prowadziła ona do poprawnego rozwiązania, okupiona była nieakceptowalnym w praktyce, bardzo długim czasem niezbędnym do opracowania

modelu - metoda optymalizacji LOO wymagała ponad miesięcznych obliczeń w środowisku rozproszonym.

Celem tej publikacji było opracowanie lepszej, bardziej efektywnej metody optymalizacji parametrów modelu retencji bazującego na SVM dzięki której dokładność estymacji opracowanego modelu będzie największa z możliwych dla danego zbioru danych glebowych.

Publikacja ta, jest drugą publikacją w cyklu i tak przebiegała nad nią praca, aczkolwiek chronologicznie ukazała się jako trzecia.

Badania własne, metodyka

W trakcie opracowywania modeli retencji wodnej z wykorzystaniem techniki SVM istnieje konieczność poprawnego doboru parametrów modelu, odpowiedniego dla wykorzystywanego zbioru danych uczących. Zasadniczo można stosować dwa podejścia. Wybrać najlepszy zestaw parametrów spośród pewnego skończonego zbioru przetestowanych parametrów (Twarakavi i in., 2009), co w praktyce może oznaczać wybór nieoptymalny. Ponadto takie podejście zawsze naraża na zarzut arbitralności wyboru parametrów modelu.

Alternatywą jest stosowanie technik optymalizacyjnych do tego celu, co jest rozwiązaniem metodologicznie poprawnym i przejrzystym. W opisywanej pracy podjęto próbę zastosowania kolejnej techniki sztucznej inteligencji - algorytmów genetycznych - w celu poszukiwania optymalnego zestawu parametrów modeli krzywej retencji wodnej. Algorytmy genetyczne są techniką optymalizacyjną która dobrze sprawdza się w przypadkach gdy funkcja celu posiada minima lokalne a taka sytuacja w ogólności ma miejsce w przypadku modelowania krzywej retencji. Z drugiej strony, technika ta jest relatywnie tania w sensie złożoności obliczeniowej.

Podjęto więc próbę jej zastosowania do poszukiwania parametrów modelu krzywej retencji bazującego na technice SVM. Według zasad eksploracji danych parametry modelu dobiera się dla zbioru uczącego, podczas gdy weryfikuje się poprawność predykcji modelu w oparciu o zbiór testowy. W tym przypadku wybrane przez AG parametry modelu pozwoliły na osiągnięcie prawie zerowego RMSE dla zbioru uczącego, jednak zdolność predykcyjna modelu - która decyduje o jego użyteczności - była bardzo niska. Proste zastosowanie algorytmów genetycznych, w tym przypadku doprowadziło do efektu znanego jako przetrenowanie modelu. W przypadku SVM oznacza to że prawie wszystkie dane ze zbioru uczącego były zastosowane w modelu jako wektory wspierające. Podczas gdy teoria SVM mówi że wartość ta powinna wynosić około połowy (Schölkopf i in., 2000). Prosta próba zastosowanie AG okazała się być błędem.

Te wstępne wyniki wymogły poszukiwania właściwej metody stosowania algorytmów genetycznych do optymalizacji parametrów modelu krzywej retencji wodnej. Skoro problemem było przetrenowanie modelu manifestujące się zbyt dużą liczbą wektorów wspierających szukano technik kontrolujących ich liczbę. Zaproponowano dwa rozwiązania tego problemu:

- Zastosowanie zmodyfikowanego algorytmu SVM w którym na ilość wektorów wspierających wprost wpływa jeden z parametrów algorytmu. Algorytm SVM oryginalnie zaproponowany przez Vapnika (1995) znany jest obecnie jako C-SVM, zależy od dwóch parametrów: C i ϵ , plus od ewentualnych parametrów zastosowanej funkcji jądra. Żaden z tych parametrów nie ma bezpośrednio wpływu na ilość wektorów wspierających w otrzymanym modelu. W roku 2000 zaproponowano (Schölkopf i in., 2000) zmodyfikowaną wersję algorytmu: ν -SVM który zależy od dwóch parametrów: C i ν , oraz od parametrów stosowanego jądra. Ciekawą cechą tej wersji algorytmu jest bezpośrednia zależność pomiędzy wartością parametru ν a ilością wektorów wspierających. W praktyce jego wartość określa wprost, jaki ułamek zbioru uczącego mają stanowić wektory wspierające. W niniejszej pracy określono wobec tego wartość parametru $\nu=0,5$. Założenie to pozwoliło dodatkowo na zmniejszenie ilości parametrów modelu.
- Modyfikację funkcji celu tak, by w sposób jawny dodatkowo uwzględniała oczekiwaną ilość wektorów wspierających. To rozwiązanie jest rozwiązaniem autorskim. Zamiast standardowo stosowanego jako funkcja celu średniego błędu kwadratowego zastosowano funkcję celu wg. zaproponowanego równania:

$$F_{aim}(RMSE, nSV) = 1 - e^{-\frac{RMSE^2}{2\sigma_{rmse}^2} - \frac{(nSV - nSV_{exp})^2}{2\sigma_{nsv}^2}}$$

gdzie: RMSE - bieżący średni błąd kwadratowy, nSV - bieżąca ilość wektorów wspierających, nSV_{exp} - oczekiwana ilość wektorów wspierających, σ_{rmse} , σ_{nsv} - są parametrami określanymi arbitralnie, aczkolwiek w określony sposób dla danego zagadnienia optymalizacyjnego. Na podstawie przeprowadzonych analiz dla dostępnych danych opracowano rekomendacje dotyczące wartości tych parametrów. Poprawną wartość σ_{rmse} oszacowano jako równą wartości maksymalnej średniego błędu kwadratowego który może wystąpić w procesie optymalizacji parametrów. Odpowiednia wartość σ_{nsv} oszacowana została na $0,1 \cdot nSV_{max}$.

Ostatecznie do przetestowania wytypowano cztery różne rodzaje modeli: bazujących na dwóch różnych algorytmach C-SVM i ν -SVM, dla których zastosowano dwa różne rodzaje funkcji jądra - liniowe i bazujące na radialnych funkcjach rozkładu. Tabela poniższa zawiera podstawowe charakterystyki testowanych modeli.

oznaczenie modelu	algorytm SVM	funkcja jądra	parametry modelu
C-radial	C-SVM	radialna	C, ϵ , γ
C-linear	C-SVM	linowa	C, ϵ
nu-radial	ν -SVM	radialna	C, γ
nu-linear	ν -SVM	linowa	C

Parametry każdego z czterech modeli optymalizowane były przy pomocy algorytmów genetycznych z wykorzystaniem dwóch różnych funkcji celu: standardowej - będącej średnim błędem kwadratowym i nowej - zaproponowanej w pracy. Algorytmy genetyczne były wykorzystane do optymalizacji wartości parametrów w granicach zdefiniowanych następująco: $1 < C < 1200$, $0.0001 < \varepsilon < 1$, i $0.00001 < \gamma < 10$. Działanie algorytmów genetycznych jest określone przez trzy podstawowe parametry, których wartości były następujące: wielkość populacji - 100, prawdopodobieństwo mutacji - 0,02 oraz współczynnik elitaryzmu - 0,2.

Źródłem informacji o glebach wykorzystywanych w tej pracy była baza danych BiGLEB, opracowana w IA PAN, grupująca informacje odnośnie reprezentatywnych profili mineralnych gleb ornich polskich. Spośród dostępnych danych wybrano 639 poziomów pochodzących z 290 profili glebowych, dla których dostępne były informacje o: krzywej retencji, rozkładzie granulometrycznym, porowatości ogólnej i zagęszczeniu.

Celem było opracowanie modeli krzywej retencji wodnej estymujących bezpośrednio wilgotność objętościową gleby dla wybranych 6 wartości potencjału wody glebowej: -0.98 , -3.10 , -9.81 , -31.02 , -491.66 and -1554.78 kPa, na podstawie dostępnych informacji o glebach.

Do opracowania modeli zastosowano walidację krzyżowej metodą k-fold. W tym celu po uprzedniej normalizacji, dostępne dane glebowe podzielono na 10 równolicznych (w przybliżeniu) podzbiorów, z których cyklicznie łączono 9 tworząc zbiór uczący i pozostawiając pozostały podzbiór w charakterze zbioru testowego. Modele regresji każdą metodą opracowywano dla każdej z 10 par zbiorów uczący - zbiór testowy. Ostateczne analizowane wyniki były średnimi z wartości generowanych przez poszczególne podmodele.

Przykładowe wyniki analizy statystycznej wyników testowanych modeli są przedstawione w poniższej tabeli. Przedstawiono tutaj dane dla wybranego potencjału wody glebowej dla którego opracowane były modele retencji. Zależności występujące dla wybranego potencjału są podobne dla pozostałych potencjałów wody glebowej dla których opracowywano modele retencji.

potencjał [kPa]	Rodzaj modelu	ilość SV		Zbiór uczący				Zbiór testowy			
				R ² [-]		RMSE [cm ³ /cm ³]		R ² [-]		RMSE [cm ³ /cm ³]	
		aim 1	aim 2	aim 1	aim 2	aim 1	aim 2	aim 1	aim 2	aim 1	aim 2
-9.81	C-linear	134.0 (14.34)	133.7 (111.77)	0.76 (0.004)	0.75 (0.011)	0.0516 (0.0007)	0.0529 (0.0016)	0.72	0.72	0.0497	0.0499
	nu-linear	189.7 (1.49)	189.7 (1.77)	0.76 (0.004)	0.76 (0.004)	0.0517 (0.0007)	0.0517 (0.0007)	0.72	0.72	0.0498	0.0498
	C-radial	360.2 (8.46)	187.7 (1.77)	0.98 (0.003)	0.93 (0.003)	0.0151 (0.0014)	0.0282 (0.0006)	0.35	0.63	0.0959	0.0577
	nu-radial	372.4 (2.95)	194.8 (2.49)	0.98 (0.004)	0.85 (0.007)	0.0152 (0.0015)	0.0416 (0.0010)	0.33	0.82	0.0985	0.0395

aim 1 - RMSE jako funkcja celu, aim 2 - funkcja celu zaproponowana w artykule

Prezentowane wyniki pozwalają zauważyć że zastosowanie zmodyfikowanej funkcji celu pozwoliło na drastyczną poprawę jakości oszacowań modeli retencji w przypadku gdy wykorzystywana była radialna funkcja rozkładu. W tym przypadku zarówno RMSE jak i współczynnik determinacji mają wartości akceptowalne dla modeli krzywej retencji wodnej. Z drugiej strony, widać że doszło do przetrenowania w przypadku stosowania RMSE jako funkcji celu dla modeli SVM wykorzystujących radialne funkcje rozkładu jako funkcje jądra. Stąd wniosek że proponowana w pracy funkcja celu spełniła swoje zadanie i pozwoliła na znalezienie optymalnych wartości parametrów modelu.

Warto zwrócić uwagę że z przykładami tak drastycznego przetrenowania mamy do czynienia wyłącznie w przypadku radialnej postaci funkcji jądra. Dalsza analiza prezentowana w artykule pokazała że jest to spowodowane dużą czułością modelu na wartość parametru γ stosowanej funkcji jądra.

Wyniki, podsumowanie i wnioski

W wyniku prac, prezentowanych w tej publikacji opracowano skuteczną metodę budowania modeli krzywej retencji wodnej gleb: bazującą na technice SVM i wspieraną przez algorytmy genetyczne.

W szczególności zaproponowano nową formę funkcji celu pozwalającą na uzyskanie optymalnych wyników modelowania krzywej retencji wodnej. Wskazano również kryteria doboru parametrów zaproponowanej funkcji celu.

Wykazano użyteczność nowej techniki modelowania - algorytmu ν -SVM do modelowania krzywej retencji. Algorytm ten pozwolił na osiągnięcie lepszych wyników niż stosowany dotychczas klasyczny algorytm C-SVM.

Zastosowanie algorytmu ν -SVM pozwoliło również zmniejszyć ilość parametrów modelu co upraszcza opracowywanie modeli retencji wodnej.

Wykazano użyteczność alternatywnej - liniowej - formy funkcji jądra, zamiast funkcji jądra bazującej na radialnych funkcjach rozkładu do opracowywania modeli retencji. Jądro liniowe pozwoliło na opracowanie modeli retencji o porównywalnej zdolnościach predykcyjnych przy jednoczesnym zmniejszeniu ilości parametrów tych modeli o jeden w porównaniu z modelami stosującymi funkcję jądra bazującą na radialnych funkcjach rozkładu.

Publikacja 3

Problem badawczy

Trzecia praca z cyklu jest poświęcona modelowaniu dynamiki zmian wilgotności w profilu glebowym w warunkach naturalnych. W pracy tej podjęto próbę wykorzystania techniki modelowania SVM do predykcji zmian wilgotności w glebie oraz porównania jej wyników z wynikami uzyskanymi przy pomocy klasycznej metody modelowania bazującej na równaniu Richardsa.

Równanie Richardsa bazuje na teoretycznie poprawnym fizycznym opisie zjawiska transportu wody w ciałach porowatych. Jednak w warunkach naturalnych, głównie ze względu na przybliżenia wprowadzane w opisie modelowanego ośrodka i jego interakcji z

atmosferą stosowanie równania Richardsa wymaga w praktyce etapu walidacji modelu w oparciu o znane z pomiarów dane o wilgotności ośrodka glebowego i warunkach meteorologicznych.

Oznacza to, że w przypadku modelowania fizycznego też musimy dysponować danymi pomiarowymi niezbędnymi do przygotowania modelu do pracy (walidacja) podobnie jak niezbędny jest taki zbiór do opracowania modelu bazującego na SVM (trening). Dopiero zwalidowany model można stosować do predykcji zmian wilgotności w danej lokalizacji.

To podobieństwo pomiędzy procesem walidacji i uczenia zasugerowało mi zastosowanie techniki SVM do modelowania zmian wilgotności w profilu glebowym w warunkach naturalnych wymuszeń.

Modelowanie bazujące na SVM, mimo że nie pozwala na zrozumienie fizycznych podstaw przebiegających procesów, pozwolić powinno w praktyce na predykcję ich przebiegu. Stwierdzenie to stało się też hipotezą roboczą prezentowanej publikacji.

Badania własne, metodyka

W ramach pracy dokonano porównania dwóch modeli: bazującego na metodyce SVM, z fizycznym modelem wykorzystującym równanie Richardsa. Do uczenia modelu (w przypadku SVM) lub walidacji modelu (w przypadku modelu fizycznego), oraz do testowania modeli wykorzystano dane z dwuletniego eksperymentu polowego przeprowadzonego na stacji agrometeorologicznej przy IA PAN w Lublinie. W wyniku tego eksperymentu zarejestrowano: wilgotności i temperaturę gleby na 6 głębokościach: 5, 10, 25, 50, 75 i 150 cm, temperaturę powietrza (przy powierzchni gruntu i na 2 metrach), wilgotność powietrza (przy powierzchni gruntu i na 2 metrach), ciśnienie atmosferyczne, promieniowanie krótko i długo falowe, nasłonecznienie, albedo, kierunek i prędkość wiatru, temperaturę gleby, opad oraz ewapotranspirację potencjalną. Dostępne były dane z dwóch okresów: 1 maj do 15 listopad 2006 (ten zbiór danych był stosowany jako zbiór uczący/walidacyjny) oraz 19 kwiecień do 11 październik 2007 (ten zbiór danych stosowany był jako zbiór testowy).

Jeżeli chodzi o metodykę modelowania z wykorzystaniem SVM w pracy tej rozwinięto techniki opracowane i przedstawione w publikacji poprzedniej (I.B.2), dzięki czemu można było z góry wytypować poprawny sposób budowy modelu pod kątem stosowanych algorytmów SVM i metod optymalizacji parametrów modelu. W oparciu o poprzednie badania zastosowano algorytm ν -SVM z liniową funkcją jądra. Parametry modelu były: $\nu=0.5$ natomiast $C=100$. Wybór wartości parametru ν był podyktowany przesłankami dyskutowanymi w pracy poprzedniej. Arbitralny wybór parametru C został natomiast dokonany w oparciu o obserwacje poczynione w pracy (I.B.2) dotyczące małej czułości modeli ν -SVM z liniową funkcją jądra na wartość parametru C . Wybór stosowanej funkcji jądra był podyktowany nie tylko poprzednimi doświadczeniami habilitanta lecz również tym, że przez niektórych autorów jest ona zalecana w przypadku niezbyt licznych zbiorów stosowanych do uczenia modelu, a z takim przypadkiem tutaj miałem do czynienia.

Otwartą kwestią był wybór odpowiednich charakterystyk meteorologicznych do modelowania SVM. Ostatecznie po analizie korelacyjnej pomiędzy wytypowanymi charakterystykami meteorologicznymi a wilgotnością gleby rejestrowaną na poszczególnych głębokościach, wybrano następujące wielkości jako zmienne niezależne modelu SVM: maksymalna dzienna temperatura, dzienny opad, średnia dzienna prędkość wiatru, średnia dzienna wilgotność powietrza, oraz średnia dzienna wartość promieniowania krótkofalowego.

Na podstawie przygotowanych danych wytrenowano model SVM który dla bieżącego dnia estymował wilgotność gleby na wybranych głębokościach (5, 10, 25, 50, 75 i 150 cm) na podstawie wilgotności z dnia poprzedniego i bieżących wartości zmiennych meteorologicznych. Model ten był następnie zastosowany do predykcji zmian wilgotności na wybranych głębokościach w profilu glebowym wyłącznie w oparciu o informacje o warunkach meteorologicznych w kolejnych dniach i informację o wilgotności początkowej w profilu glebowym na początku okresu testowego.

Jako model fizyczny wykorzystujący równanie Richardsa, zastosowano zaimplementowany w komercyjnym oprogramowaniu HYDRUS-1D model podwójnej porowatości (ang. dual-porosity) (Šimůnek i van Genuchten 2008).

2008a. Model ten estymował wilgotności gleby w profilu glebowym na podstawie charakterystyk meteorologicznych oraz ustalonych parametrów charakteryzujących ośrodki glebowy. Model był walidowany w oparciu o dane pomiarowe z roku 2006, te same które zastosowane były do opracowania modelu SVM.

Dokładność predykcji obu modeli SVM i fizycznego była testowana na drugiej części zbioru danych pomiarowych - pomiar z 2007 roku.

Porównanie wskaźników statystycznych pokazujących jakość estymacji obu modeli pokazuje poniższa tabela.

Głębokość [cm]	Model fizyczny		Model SVM		
	R ² [-]	RMSE [cm ³ /cm ³]	Ilość SV/ całk. ilość danych	R ² [-]	RMSE [cm ³ /cm ³]
5	0.31	0.059	77/145	0.86	0.035
10	0.44	0.048	70/145	0.83	0.030
25	0.46	0.045	74/145	0.87	0.021
50	0.46	0.030	74/145	0.44	0.019
75	0.35	0.034	77/145	0.64	0.012
150	0.40	0.006	72/145	0.61	0.004

Wyniki, podsumowanie i wnioski

Wykazano że technikę SVM można z powodzeniem stosować do predykcji wilgotności w profilu glebowym.

Widać że w przypadku wszystkich głębokości średni błąd kwadratowy estymacji wilgotności przez model SVM jest niższy niż błąd modelu fizycznego. Podobnie rzecz się ma jeżeli weźmiemy pod uwagę wartości współczynnika determinacji. Wartości R² wyższe

dla modelu SVM niż modelu fizycznego ponownie wskazują na lepsze działanie tego pierwszego.

Pozwala to sformułować twierdzenie że dokładność estymacji opracowanego modelu bazującego na SVM przewyższa dokładność estymacji modelu bazującego na klasycznym równaniu Richardsa.

W praktyce technika SVM stosowana do modelowania zmian wilgotności w profilu glebowym jest atrakcyjną alternatywą dla fizycznych modeli bazujących na równaniu Richardsa.

Publikacja 4

Problem badawczy

Ostania praca cyklu dotyczy wykorzystania techniki SVM do modelowania krzywej retencji wodnej gleb. W pracy tej wykorzystano dotychczasowe doświadczenia do próby oceny użyteczności danych o rozkładzie granulometrycznym zmierzonych metodą dyfrakcji laserowej do modelowania retencji.

Standardowo wykorzystywaną metodą pomiaru rozkładu granulometrycznego gleb jest metoda sitowo-sedymentacyjna (SHM). W ostatnich latach popularyzuje się alternatywna metoda pomiaru - metoda dyfrakcji laserowej (LDM). LDM jest metodą szybką, nie wymaga dużej ilości materiału glebowego do analizy i daje powtarzalne wyniki (Eshel i in. 2004; Arriaga i in 2006).

Rozkład granulometryczny jest jedną z podstawowych charakterystyk ośrodka glebowego wykorzystywaną w modelowaniu retencji wodnej. Do tej pory, autorzy wykorzystujący rozkłady zmierzone LDM do modelowania retencji najpierw przeliczali wyniki rozkładów z LDM na SHM i te dopiero stosowali do budowy swoich modeli (Puhlmann i von Wilpert 2012).

Problem w tym, że rozkłady granulometryczne mierzone obiema metodami różnią się, natomiast metoda SHM jest metodą referencyjną. W związku z tym wielu autorów próbowało opracować metodę przeliczania wyników pochodzących z LDM na ich odpowiedniki SHM. Niestety taka poprawna metoda przeliczania nie została jak na razie opracowana.

W związku z tym że ten dodatkowy krok konwersji danych rozkładów pomiędzy metodami pomiarowymi jest dyskusyjny i może wprowadzać dodatkowe błędy podjęto próbę bezpośredniego wykorzystania danych o rozkładach pochodzących z metody LDM do modelowania retencji wodnej - co było podstawowym celem tej pracy.

Z uwagi na fakt że metoda LDM standardowo pozwala na pozyskanie dodatkowych charakterystyk rozkładu granulometrycznego w porównaniu z SHM podjęto również próbę sprawdzenia czy wykorzystanie tych dodatkowych informacji do modelowania retencji wpłynie korzystnie na jakość estymacji modelu retencji.

Badania własne, metodyka

Jako metodę modelowania wykorzystywaną w tym zadaniu wytypowano wersję ν -SVM algorytmu, który bazował na liniowej funkcji jądra. Parametr ν algorytmu SVM

zgodnie z wcześniejszymi doświadczeniami (I.B.2 i I.B.3) wynosił 0,5. Pozostały parametr modelu C wyznaczany był przy pomocy algorytmów genetycznych. Tak sformułowany model gwarantował osiągnięcie optymalnych wyników estymacji retencji.

Zbiór danych wykorzystywany do uczenia i testowania modeli składał się z danych dotyczących 84 próbek pobranych z poziomu A różnych profili glebowych. Te profile były podzbiorem 100 profili glebowych wytypowanych przez twórców bazy Gleb Mineralnych Polski BIGLEB jako profile reprezentatywne.

Dla próbek wyznaczono krzywe retencji wodnej dla 10 wartości potencjału wody glebowej (-0.10, -0.98, -3.10, -9.81, -31.02, -49.17, -98.1, -155.48, -491.66 i -1554.78 kPa). Oznaczono zagęszczenie, oraz rozkład granulometryczny dwiema metodami LDM i SHM. Przy czym metodą SHM wyznaczono wyłącznie 3 podstawowe frakcje: piasek, pył i il. Metoda LDM pozwoliła na wyznaczenie dodatkowych charakterystyk rozkładu granulometrycznego takich jak: szczegółowe informacje o frakcjach ($fr(<2 \mu\text{m})$, $fr(2-5 \mu\text{m})$, $fr(5-20 \mu\text{m})$, $fr(20-50 \mu\text{m})$, $fr(50-100 \mu\text{m})$, $fr(100-250 \mu\text{m})$, $fr(250-500 \mu\text{m})$, $fr(500-1000 \mu\text{m})$, $fr(1000-2000 \mu\text{m})$), decyle rozkładu ($d(0.1)$: $d(0.5)$ i $d(0.9)$), powierzchniowy średni moment rozkładu (Surface Area Moment Mean - $D(3,2)$) i objęściowy średni moment rozkładu (Volume Moment Mean - $D(4,3)$).

Dane glebowe przed dalszym wykorzystaniem zostały znormalizowane co jest standardową techniką stosowaną w modelowaniu SVM.

Ze względu na relatywnie małą ilość danych glebowych zdecydowano się na zastosowanie metody walidacji krzyżowej typu k-fold. Dostępne dane podzielono losowo na 6 podzbiorów. Następnie rotacyjnie dane pięciu z nich stosowano jako zbiór uczący (70 elementów) natomiast pozostały podzbiór był zbiorem testowym (14 elementów) danego modelu. Jako ostateczny wynik modelu stosowano średnią wartość estymacji wilgotności dokonywanych przez poszczególne podmodele.

W pracy opracowano cztery modele retencji. Zestawienie parametrów od których zależały poszczególne modele, stosowana metoda pomiarowa dot. rozkładu granulometrycznego oraz ich oznaczenia zawiera poniższa tabela.

Oznaczenie modelu	Charakterystyki fizyczne gleb stosowane jako zmienne wejściowe modelu	Ilość zmiennych wejściowych/ilość zmiennych PCA	Stosowane źródło danych o rozkładzie granulometrycznym
M_SHM	bd, fr_{silt} , fr_{clay}	3/-	SHM
M_LDM1	bd, fr_{silt} , fr_{clay}	3/-	LDM
M_LDM2	bd, fr_{silt} , fr_{clay} $d(0.1)$: $d(0.9)$	5/5	LDM
M_LDM3	bd, $fr(<2 \mu\text{m})$: $fr(2-5 \mu\text{m})$: $fr(5-20 \mu\text{m})$: $fr(20-50 \mu\text{m})$: $fr(50-100 \mu\text{m})$: $fr(100-250 \mu\text{m})$: $fr(250-500 \mu\text{m})$: $fr(500-1000 \mu\text{m})$: $fr(1000-2000 \mu\text{m})$	10/6	LDM

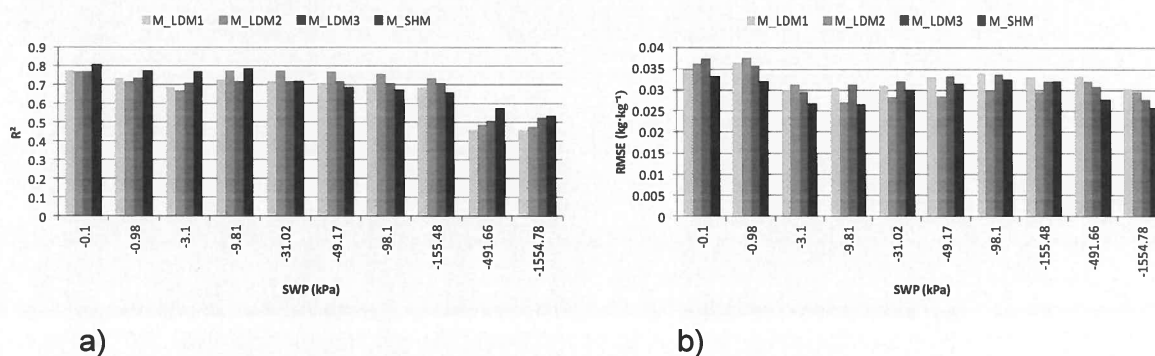
Model M_LDM1 był prostym odpowiednikiem modelu M_SHM, oba bazowały na podstawowych informacjach o glebach: zagęszczeniu, zawartości frakcji pyłu oraz

zawartości frakcji iłu, z tym że do pozyskania informacji o rozkładach stosowano różne techniki pomiarowe, LDM i SHM.

W przypadku modeli M_LDM2 i M_LDM3 podjęto próbę wykorzystania dodatkowych informacji o rozkładzie granulometrycznym. W przypadku modelu M_LDM2, po dodatkowej analizie korelacyjnej, spośród dostępnych dodatkowych charakterystyk rozkładu: $d(0.1)$: $d(0.5)$, $d(0.9)$, $D(3,2)$ i $D(4,3)$ wybrano dwie ($d(0.1)$, $d(0.9)$), które były najmniej skorelowane z już wybranymi parametrami modelu (bd , fr_{silt} , fr_{clay}). W przypadku modelu M_LDM3, oprócz zagęszczenia zastosowano informacje o wszystkich dostępnych pod-frakcjach wyznaczonych metodą LDM.

Z uwagi na stosunkowo dużą ilość parametrów modeli M_LDM2 i M_LDM3, które dodatkowo są ze sobą częściowo skorelowane zastosowano metodę analizy głównych składowych (PCA). W wyniku analizy PCA uzyskano zbiór wektorów własnych PCA równoważnych wejściowym parametrom fizycznym. Aby ograniczyć ostateczną ilość zmiennych wejściowych modeli M_LDM2 i M_LDM3 tylko do tych, które mogą mieć istotny wpływ na ich działanie, odrzucono te wektory własne PCA, których standardowe odchylenia wynosiły mniej niż 0,1 standardowego odchylenia głównego wektora własnego. W przypadku modelu M_LDM2 procedura ta nie spowodowała zmniejszenia ilości zmiennych wejściowych modelu, w przypadku modelu M_LDM3 ilość zmiennych wejściowych modelu uległa zmniejszeniu z 10 do 6.

W oparciu o przyjęte założenia opracowano modele retencji. Rysunki poniższe przedstawiają zbiorczo statystyczne wskaźniki jakości estymacji opracowanych modeli: RMSE i R^2 .



Wyniki, podsumowanie i wnioski

W oparciu o otrzymane wyniki można stwierdzić, że informacje o rozkładzie granulometrycznym pozyskane metodą LDM można z powodzeniem stosować bezpośrednio do modelowania krzywej retencji wodnej.

Wykazano że przeliczania rozkładów z LDM do SHM, stosowane przez innych autorów, nie jest wymogiem osiągnięcia dobrych wyników modelowania krzywej retencji.

Wykazano również że w celu osiągnięcia najlepszej możliwej dokładności predykcji krzywej retencji wodnej należy rozważyć stosowanie jako zmiennych wejściowych modeli retencji innych niż podstawowe frakcje (piasek, pył, ił) informacji o rozkładzie granulometrycznym dostępnych w technice LDM.

Podsumowanie i wnioski

Przeprowadzone badania wykazały użyteczność technik sztucznej inteligencji: SVM i AG do modelowania krzywej retencji wodnej i modelowania zmian wilgotności w profilu glebowym.

Opracowane metody pozwalały na osiągnięcie wyników lepszych lub co najmniej tak samo dobrych jak metody alternatywne, przy czym opracowana metodyka budowania modeli pozwala na ich efektywniejsze opracowywanie.

W przypadku modelowania zmian wilgotności ośrodka glebowego zaproponowanymi metodami, osiągnięto wyniki wyraźnie lepsze niż wyniki uzyskane metodami klasycznymi.

Bibliografia

1. Ahuja L. R., Naney J. W., Williams R. D.: 1985, Estimating soil water characteristics from simpler properties or limited data, *Soil Science Society of America Journal*, vol. 49, no. 5, pp. 1100–1105.
2. Arriaga, F. J., Lowery, B., Mays, D. W.: 2006, A fast method for determining soil particle size distribution using a laser instruments. *Soil Sci.* 171, 663–674.
3. Baker L., Ellison D.: 2008, Optimisation of pedotransfer functions using an artificial neural network ensemble method, *Geoderma*, vol. 144, no. 1-2, pp. 212–224.
4. Bieganowski A., Witkowska-Walczak B., Gliński J. i in.: 2013, Database of Polish arable mineral soils: a review, *International Agrophysics*, vol. 27, no. 3, pp. 335–350.
5. Bouma J.: 1989, Using soil survey data for quantitative land evaluation, *Advances in Soil Science* 9: 177–213.
6. Diamantopoulos E., Durner W.: 2012, Dynamic Nonequilibrium of Water Flow in Porous Media: A Review, *Vadose Zone Journal*, Vol. 11 No. 3, doi:10.2136/vzj2011.0197.
7. Drucker H., Burges C., Kaufman L., Smola A., Vapnik V.N.: 1997, Support Vector Regression Machines, in *Advances in Neural Information Processing Systems 9*, NIPS 1996, 155–161, MIT Press.
8. Eshel, G., Levy, G. J., Mingelgrin, U., Singer, M. J.: 2004, Critical evaluation of the use of laser diffraction for particle-size distribution analysis. *Soil Sci. Soc. Am. J.* 68, 736–743.
9. Maren, A.J., Harston C.T., Pap R.M.: 1990. *Handbook of neural computing applications*. Academic Press, San Diego.
10. Minasny, B., Perfect E.: 2004. Solute adsorption and transport parameters. p. 195–224. w Ya. Pachepsky and W.J. Rawls (ed.) *Development of pedotransfer functions in hydrology. Developments in Soil Science* 30. Elsevier, Amsterdam.
11. Pachepsky Y. A., Timlin D., Varallyay G.: 1996, Artificial neural networks to estimate soil water retention from easily measurable data, *Soil Science Society of America Journal*, vol. 60, no. 3, pp. 727–733, 1996.
12. Puhmann, H., von, Wilpert, K.: 2012, Pedotransfer functions for water retention and unsaturated hydraulic conductivity of forest soils. *J. Plant Nutr. Soil Sci.* 175, 221–235.
13. Schölkopf, B., Smola, A. J., Williamson, R. C., Bartlett, P. L.: 2000, New support vector algorithms. *Neural Comput.* 12, 1207–1245.
14. Šimůnek, J., van Genuchten M.Th., Šejna M.: 2008. The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media. Version 4.0. *Dep. Environ. Sci., Univ. of California, Riverside*.

15. Twarakavi, N.K.C., Šimůnek, J., Schaap, M.G: 2009, Development of pedotransfer functions for estimation of soil hydraulic parameters using support vector machines, Soil Science Society of America Journal, Vol. 73, no. 5, pp. 1443-1452.
16. Vapnik V.N.: 1995, The Nature of Statistical Learning Theory, Springer, New York, NY, USA.
17. Walczak R., Witkowska-Walczak B., and Sławiński C.: 2001, Comparison of correlation models for the estimation of the water retention characteristics of soil, International Agrophysics, vol.16, no. 1, pp. 79–82.

5. Omówienie pozostałych osiągnięć naukowo-badawczych

Po ukończeniu studiów fizycznych na Wydziale Matematyki i Fizyki UMCS w roku 1997 rozpocząłem pracę w Zakładzie Hydrotermofizyki Środowiska Glebowego i Materiałów Roślinnych Instytutu Agrofizyki PAN w Lublinie w zespole badawczym prof. dr hab. Ryszarda Walczaka.

Od początku moje zainteresowania badawcze ogniskowały się na modelowaniu procesów transportu przebiegających w ośrodku glebowym. Jednym z pierwszych wykonanych zadań badawczych było opracowanie na podstawie studiów literaturowych modelu transportu wody i energii oraz jego implementacja w postaci oprogramowania. Stworzyłem i przetestowałem oprogramowanie, które pozwalało na opis zmian wilgotności i temperatury w warstwowym ośrodku glebowym. W wyniku realizacji tego tematu powstała publikacja (II.D.21) wyniki były też prezentowane na dwóch konferencjach (III.B.29, III.B.30).

W 1998 roku nawiązana została współpraca pomiędzy Zakładzie Hydrotermofizyki Środowiska Glebowego i Materiałów Roślinnych IA PAN a Wojskowym Instytutem Uzbrojenia w Zielonce k. Warszawy (WITU) reprezentowanym przez ppłk dr. Waldemara Świdarskiego. W wyniku tej współpracy zostałem podwykonawcą w projekcie KBN realizowanym przez WITU poświęconym wykrywaniu min przeciwczołgowych z wykorzystaniem technik termografii aktywnej. Zadanie moje polegało na wsparciu badań eksperymentalnych modelowaniem zmian temperatury w niejednorodnym ośrodku glebowym. Ostatecznym celem modelowania było opracowanie zaleceń dla eksperymentatorów pozwalających na zwiększenie prawdopodobieństwa detekcji miny. W celu wykonania tego zadania należało zmodyfikować opracowane wcześniej oprogramowanie tak, by można było modelować heterogeniczny ośrodek glebowy z wtrąceniami o innych właściwościach cieplnych i hydrologicznych niż gleba. Opracowano również wersję modelu transportu opisującą przebieg zjawisk w osiowo symetrycznym układzie odniesienia, co było niezbędne do realizacji tego zadania. Wyniki badań były publikowane (II.A.6, II.D.18, II.D.19, II.D.20) oraz prezentowane na licznych międzynarodowych konferencjach (III.B.23, III.B.24, III.B.25, III.B.26, III.B.27, III.B.28).

Odrębnym zainteresowaniem badawczym była metodyka opracowania, analizy i weryfikacji modeli agrofizycznych. W szczególności podjąłem próby zastosowania nowej popularyzującej się w innych dziedzinach nauki metody obliczeniowej: algorytmów genetycznych do tego celu. Pierwszą z prac z tego cyklu, była praca (II.D.17) poświęcona zastosowaniu algorytmów genetycznych (AG) do walidacji modeli numerycznych

transportu ciepła w glebie i rozwiązania zagadnienia odwrotnego. W pracy tej pokazałem że AG mogą być z powodzeniem stosowane do tego celu, pozwalając na osiągnięcie optymalnych wyników przez walidowany model fizyczny. W pracach (II.D.15, II.D.13) dyskutowano metody weryfikacji modeli oraz stosowano je w kontekście konkretnych modeli.

Podsumowaniem tych zainteresowań badawczych była rozprawa doktorska, obroniona w 2005 roku, wykorzystująca w praktyce opracowane oprogramowanie, poznane techniki modelowania i doświadczenie do zagadnienia analizy dokładności pomiarów wykonywanych strumieniomierzem płytkowym (II.D.9).

Oprócz zainteresowania technikami modelowania procesów transportu zachodzących w profilu glebowym w roku 2002 zainteresowałem się metodami modelowania i parametryzacji krzywej retencji wodnej gleb. Przy czym, w dalszym ciągu próbowałem stosować w badaniach nowe techniki obliczeniowe. Pierwsza praca która dotyczyła tego tematu (II.D.16) była poświęcona zastosowaniu sieci neuronowych do wyznaczania wartości parametrów modelu M-vG krzywej retencji wodnej gleb. W pracy tej wykorzystano dane 39 różnych gleb o różnych zagęszczeniach (gleby były sztucznie zagęszczane). W sumie zbiór danych podlegający analizie zawierał dane o 370 próbkach glebowych, dla których wyznaczono krzywe retencji oraz fizyczne charakterystyki gleb: zagęszczenie, rozkład granulometryczny. Opracowano bazujący na sztucznych sieciach neuronowych model który estymował parametry M-vG krzywej retencji na podstawie informacji o zagęszczeniu i informacji o rozkładzie granulometrycznym gleby. Dokładność estymacji opracowanego modelu była wysoka, średni błąd kwadratowy estymacji wilgotności objętościowej wynosił od 0.0013 do 0,04 m³m⁻³. Innym zagadnieniem z tej tematyki badawczej w którym uczestniczyłem, było modelowanie krzywej retencji i współczynnika przewodnictwa nienasyconego w oparciu o statystyczne metody regresji (II.D.14). W innej pracy (II.D.10), analizowano zależności pomiędzy współczynnikami modelu retencji Mualema-van Genuchtena (M-vG) a charakterystykami gleb: rozkładem granulometrycznym, zagęszczeniem oraz poziomem z którego pobrano glebę. Analizowano dane 78 gleb z 27 profili. Zaobserwowano zależności pomiędzy niektórymi parametrami modelu M-vG: n i θ_s a rozkładem granulometrycznym gleb. Wartość parametru n zwiększała się razem z zawartością frakcji piasku. Wilgotności nasycona θ_s była dodatnio skorelowana z zawartością frakcji pyłu.

Innym rodzajem zadań badawczych którymi się również interesowałem było opracowywanie metod pomiarowych i eksperymentalnych.

W 2002 roku w wyniku współpracy z pracownikami IA PAN z innych jednostek organizacyjnych (dr Tadeusz Rudko i mgr. Jarosław Piekarczyk) opracowano koncepcję i prototyp urządzenia do oceny podatności na pękanie łuszczyń rzepaku (II.D.12). Byłem współautorem koncepcji działania tego urządzenia oraz twórcą samego urządzenia w zakresie oprogramowania niezbędnego do wykonywania pomiarów i sterowania urządzeniem przy pomocy komputera. W 2003 roku dokonano zgłoszenia patentowego w wyniku którego w 2009 roku uzyskano patent i ochronę wynalazku (II.B.1).

Współpracując z zespołem prof. dr hab. Wojciecha Skieruchy opracowałem koncepcję organizacji akwizycji danych w sieci pomiarowej urządzeń TDR opracowanych przez ten zespół, monitorujących wilgotność i temperaturę gleby. Sieć tych urządzeń komunikuje się przy pomocy Internetu z serwerem centralnym, gdzie dane pomiarowe są przesyłane i przetwarzane. Byłem również twórcą oprogramowania wykorzystywanego do akwizycji i przetwarzania rejestrowanych danych. Wyniki tych badań prezentowane są w dwóch publikacjach (II.A.4, II.D.7).

W roku 2011 rozpoczęła się realizacja projektu "Opracowanie modeli PTF krzywej retencji wodnej z uwzględnieniem efektu histerezy" którego byłem współpomysłodawcą i jednym z głównych wykonawców (II.I.2). W ramach realizacji tego projektu opracowałem koncepcję urządzenia pomiarowego pozwalającego na zautomatyzowany pomiar krzywej retencji gleb w procesach nawilżania i osuszania. Pomiar wilgotności w tym urządzeniu bazuje na technice TDR, jednak ze względu na specyfikę konstrukcji urządzenia pomiaru wilgotności można dokonać jedynie z zastosowaniem jednej sondy pomiarowej. Praca (II.A.2) analizuje dokładność i reprezentatywność pomiaru wilgotności standardowej próbki gleby (100 cm³) przy pomocy jedynie jednej sondy w porównaniu z ogólnie zaakceptowaną procedurą (wykorzystującą 3 sondy pomiarowe). W pracy wykazano równoważność obu metod pomiarowych.

Innym tematem badawczym w realizacji którego uczestniczyłem były badania dotyczące wpływu uprawy na właściwości hydrologiczne gleby oraz kształtowanie się stosunków wodnych w profilu glebowym pod uprawą. W pracy (II.A.5) analizowano wyniki 3 letniego eksperymentu poletkowego który dotyczył wpływu rodzaju uprawy: uproszczonej lub tradycyjnej na właściwości hydrologiczne gleby oraz zmiany wilgotności w profilu glebowym. Wykazano korzystny wpływ uprawy uproszczonej na właściwości hydrologiczne gleby co odzwierciedlało się w lepszych warunkach wodnych w przypadku uprawy uproszczonej. Tematem pracy (II.D.8) był wpływ uprawianych roślin na stosunki wodne w profilu glebowym. Uprawianymi roślinami były popularne rośliny energetyczne: wierzba krzewiasta i ślazo wiec pensylwański. Bezpośrednio do poletek doświadczalnych przylegały ugorowane poletka kontrolne. Eksperyment był prowadzony przez 2 kolejne sezony wegetacyjne w latach 2007 - 2008. Monitorowano zmiany wilgotności w profilu glebowym, oraz charakterystyki wodne gleb: krzywą retencji. Zaobserwowano wpływ uprawianych roślin na kształtowanie wilgotności w profilu glebowym. Na podstawie analizy krzywych retencji wodnej można stwierdzić, że zasoby wody łatwo dostępnej dla roślin były wyższe na poletkach kontrolnych w porównaniu z poletkami doświadczalnymi, zarówno pod wierzbą jak i ślazo wiecem.

Współuczestniczyłem również w badaniach poświęconych zależności współczynnika nasyconego przewodnictwa dla gleb modyfikowanymi związkami powierzchniowo czynnymi. Taka procedura wykorzystująca kationowe związki powierzchniowo czynne jest wskazywana jako potencjalnie wspierająca rekultywację gleb zanieczyszczonych związkami ropopochodnymi. Współczynnik przewodnictwa nasyconego gleb ma wpływ na ten proces. Wnioski pracy (II.D.4) gdzie badano eksperymentalnie współczynnik przewodnictwa nasyconego dla substancji ropopochodnej (DUNASOL) 4 różnych gleb,

wskazują że współczynnik przewodnictwa nasyconego gleb traktowanych związkami powierzchniowo czynnymi jest mniejszy niż gleby nie modyfikowanej.

Uczestniczyłem również w badaniach dotyczących metodyki wyznaczania wymiaru fraktalnego gleb na podstawie informacji o rozkładzie granulometrycznym wyznaczanym metodą dyfrakcji laserowej. Wymiar fraktalny gleb jest jedną z charakterystyk gleb która jest wykorzystywana w jej badaniach. Jest on wielkością korelowaną z innymi właściwościami gleb, np. z krzywą retencji wodnej. Obecnie popularyzuje się nowa metoda pomiaru rozkładu granulometrycznego wykorzystująca zjawisko dyfrakcji. W pracy (II.A.3) wykazano silną zależność pomiędzy sposobem wyznaczania wymiaru fraktalnego a uzyskiwanymi jego wartościami w przypadku gdy korzysta się z informacji o rozkładzie granulometrycznym mierzonym metodą dyfrakcji laserowej. Wybór ilości i zakresów wyznaczanych frakcji wpływa istotnie na uzyskiwane wyniki. W wyniku przeprowadzonych badań zaproponowano standardową procedurę wyznaczania wymiaru fraktalnego gleb minimalizującą błąd mogący być wprowadzony przez nieodpowiedni podział rozkładu granulometrycznego na frakcje.

W 2014 roku rozpocząłem współpracę z zespołem dr. hab. Andrzeja Bieganowskiego dotyczącą modelowania zjawiska rozbryzgu materiału glebowego wywoływanego przez krople deszczu, w kontekście erozji wodnej gleb. W wyniku tej współpracy powstała publikacja (II.A.1) dotycząca opracowania i walidacji systemu pomiarowego wykorzystywanego do pomiaru energii kinetycznej spadającej kropli i sił jej oddziaływania z podłożem.

Mój dorobek publikacyjny obejmuje **70** publikacji naukowych i komunikatów konferencyjnych. Jestem autorem lub współautorem **33** oryginalnych prac twórczych i przeglądowych, w tym **10** w czasopismach z IF. Aktualnie opracowuję kolejne publikacje naukowe.

Jestem również współautorem jednego wynalazku posiadającego ochronę patentową (II.B.1).

Mój całościowy dorobek naukowy wg punktacji MNiSW, wg. wykazu MNiSW z dnia 31 grudnia 2014 wynosi **446** punkty, z czego **97** uzyskałem przed obroną doktoratu. Sumaryczny Impact Factor dla opublikowanych przeze mnie publikacji wynosi **15,36** (jedna opublikowane przed uzyskaniem stopnia doktora), liczba cytowań według bazy Web of Science wynosi **60** (54 bez autocytowań), a indeks Hirscha według bazy Web of Science wynosi **4**.

W ciągu całego okresu pracy zawodowej brałem czynny udział w realizacji **4** projektów badawczych (II.I.1-II.I.4), w czym kierowałem **1** projektem (II.I.4) naukowym.

Podczas pracy naukowej brałem czynny udział w wielu konferencjach międzynarodowych i krajowych (III.B.1.-III.B.30.), prezentując wyniki w formie ustnej lub w formie posterów.

Uczestniczyłem również w organizacji konferencji, będąc przewodniczącym komitetu organizacyjnego konferencji międzynarodowej (III.C.1), oraz członkiem rady programowej konferencji krajowej (III.C.2.).

W ramach pracy naukowej współpracuję aktywnie z ośrodkami naukowymi za granicą: Environmental Microbial and Food Safety Lab, USDA, ARS w Beltsville, MD, z USA; Department of Crop Production and Soil Science, Georgikon Faculty, University of Pannonia z Węgier oraz Department of Civil Engineering, National Chiao Tung University, Hsinchu, Taiwan. W wyniku tej współpracy powstały prace naukowe lub złożone były wspólne aplikacje na konkursy organizowane przez NCBIr oraz NCN.

Byłem również ekspertem współuczestniczącym w organizacji i opracowaniu europejskiej bazy danych charakterystyk hydrologicznych gleb "European Hydropedological Data Inventory (EU-HYDI)" (II.E.1). Praca ta była koordynowana przez Institute for Environment and Sustainability, Joint Research Centre w Isprze, Włochy.

Obecnie jestem członkiem European Geosciences Union, w latach 1999 - 2012 byłem członkiem Polskiego Towarzystwa Agrofizycznego.

Od 2009 roku wykonałem **24** recenzje artykułów naukowych dla **9** czasopism o zasięgu międzynarodowym (III.P.1) i jedną recenzję wydawniczą monografii (III.P.2).

Obecnie jestem pomocniczym opiekunem naukowym dwóch doktorantów (III.K.1) realizujących swoje prace doktorskie w Zakładzie Metrologii i Modelowania Procesów Agrofizycznych IA PAN. Od 2001 roku jestem również wykładowcą akademickim zatrudnionym w Państwowej Wyższej Szkole Zawodowej w Chełmie (III.I.1), gdzie byłem promotorem **33** prac licencjackich (III.J.1).

Oprócz pracy naukowej w IA PAN udzielam się również w zadaniach administracyjnych. Byłem organizatorem pracowni tomografii komputerowej w Laboratorium Oceny, Ulepszania i Wykorzystywania Osadów Pofermentacyjnych. Pełnię również obowiązki Pełnomocnika Dyrektora IA PAN do Spraw Informatyzacji Instytutu Agrofizyki PAN.

Lamorski Krzysztof