

dr hab. inż. Maria Strzelczyk, prof. ITP-PIB
Instytut Technologiczno-Przyrodniczy
Państwowy Instytut Badawczy
Oddział we Wrocławiu
Ul. Heleny Motykówny 7, 51-209 Wrocław

Wrocław, 27.03.2023 r.

RECENZJA ROZPRAWY DOKTORSKIEJ

mgr inż. Grzegorza Chrobaka

**pt. „System wspierania oceny stanu ekologicznego ekosystemów jeziornych
wspomagany algorytmami uczenia maszynowego” wykonanej w ramach studiów
doktoranckich na kierunku Ochrona i Kształtowanie Środowiska w, Wydziale
Inżynierii Kształtowania Środowiska i Geodezji Uniwersytecie Przyrodniczym we
Wrocławiu**

Promotor: dr hab. inż. Tomasz Kowalczyk, prof. UPWr

Drugi promotor: Professor Thomas B. Fischer, PhD

Oceniana rozprawa doktorska została przygotowana jako spójny tematycznie zbiór 3 artykułów, których suma punktów wg wykazu czasopism MNISW z 1 grudnia 2022 r. wynosi 420 pkt., IF 17,997, opublikowanych w renomowanych czasopismach naukowych. Wszystkie prace są opracowaniami zbiorowymi, w których pierwszym autorem jest Doktorant Pan mgr inż. Grzegorz Chrobak:

- Chrobak G. (75%), Kowalczyk T., Fischer T.B., Szewrański Sz., Chrobak K., Wąsowicz B., Kazak J.K. (2022) First, do no harm - Missing data treatment to support lake ecological condition assessment. *Environmental Modelling & Software*, 2022, vol. 158, Numer artykułu: 105558. DOI: 10.1016/j.envsoft.2022.1555.
- Chrobak G. (80%), Kowalczyk T., Fischer T.B., Szewrański Sz., Chrobak K., Kazak J.K. (2021) Ecological state evaluation of lake ecosystems revisited: Latent variables with kSVM algorithm approach for assessment automatization and data comprehension. *Ecological Indicators*, 2021, vol. 125, s.1-15, Numer artykułu: 107567. DOI: 10.1016/j.ecolind.2021.107567.
- Chrobak G. (70%), Kowalczyk T., Fischer T.B., Szewrański Sz., Chrobak K., Kazak J.K. (2021) Combining indicators for better decisions – Algorithms vs experts on lakes ecological status assessment. *Ecological Indicators*, 2021, vol. 132, s.1-11, Numer artykułu: 108318. DOI: 10.1016/j.ecolind.2021.108318.

Zgodnie z informacją zawartą w podrozdziale „Struktura rozprawy” udział pierwszego autora Pana mgr inż. Grzegorza Chrobaka w wymienionych publikacjach wynosi 315 pkt., IF: 13,494 pkt. Dołączone oświadczenia współautorów publikacji dotyczące ich roli w publikacji świadczą o wiodącej roli Doktoranta w ich przygotowaniu. Tematyka rozprawy dotyczy usprawnienia procesu oceny stanu ekologicznego jezior w oparciu o trzy etapy: imputację brakujących danych, modelowanie procesu oceny oraz priorytetyzację ekosystemów wskazanych do działań naprawczych i/lub ochronnych przy wykorzystaniu algorytmów uczenia maszynowego. Każdy z elementów systemu oceny omawianego w rozprawie jest opisany w oddzielnej publikacji stanowiącej podstawę rozprawy. Tytuł pracy doktorskiej odpowiada tematyce analizowanych publikacji.

Układ rozprawy doktorskiej

Rozprawa, wraz z załącznikami obejmuje 106 stron. Układ rozdziałów i treści w nich zawartych jest przejrzysty, stanowi logiczną całość. Na wstępie rozprawy znajduje się streszczenie pracy w języku polskim i angielskim, które w sposób jasny i precyzyjny wprowadza czytelnika w tematykę rozprawy. Na początku rozprawy znajduje się również wykaz akronimów stosowanych w dalszej części rozprawy. Następnie autor w sposób przejrzysty określił hipotezy i cele nie tylko opisując je, ale obrazując na schemacie logiczną strukturę powiązań postawionych w rozprawie hipotez, które znalazły swoje odbicie w poszczególnych publikacjach, z ujęciem następstwa celów szczegółowych osiąganych w toku podjętych przez siebie badań, ukazując ich spójność i kompleksowość. Kolejnym oryginalnym elementem rozprawy jest fragment zatytułowany „Czym rozprawa ta nie jest”, w którym Doktorant jasno precyzuje cel podjętych prac wskazując na możliwość ich praktycznego zastosowania w procesie oceny stanu ekologicznego ekosystemów jeziornych. Rozprawa zawiera pełne teksty 3 artykułów poprzedzone szerokim wprowadzeniem wskazującym na potrzebę realizacji podjętego zagadnienia, omówieniem problematyki oceny ekologicznej oraz modelowania stanu ekologicznego jezior, jakości i dostępności danych wykorzystywanych do oceny stanu ekologicznego ekosystemów jeziornych oraz technik uczenia maszynowego wykorzystywanych w omawianym zagadnieniu. W następnej kolejności Doktorant wskazał na źródło danych wykorzystanych do badań i omówił zastosowane narzędzia i metody oraz dokonał podsumowania rezultatów podjętych badań w kontekście zdefiniowanych celów szczegółowych i zidentyfikowanych luk w wiedzy. W podsumowaniu tej części rozprawy Doktorant określił wyzwania i dalsze kierunki badań i przedstawił spis literatury wykorzystanej w rozprawie. Część merytoryczną rozprawy zamyka zestaw 3 załączników zawierający: charakterystykę danych wejściowych, wyniki analizy

wrażliwości stanu ekologicznego na zmiany w parametrach oraz szczegóły obliczeniowe modelowania wykorzystanych w jednej z publikacji, stanowiącej podstawę rozprawy. Doktorant dołączył również oświadczenia współautorów publikacji naukowych zaliczonych do cyklu wraz z informacją o charakterze ich udziału w publikacji.

W rozprawie Doktorant zastosował numerację rozdziałów tylko w części obejmującej publikacje naukowe cyklu, stosując jedynie pogrubienie liter w pozostałych częściach rozprawy jako zaznaczenie ich odrębności, co w nieznacznym stopniu może utrudniać jej odbiór.

Przedłożona do recenzji rozprawa jest w mojej ocenie kompletna, a jej układ oceniam jako prawidłowy.

Zastosowane piśmiennictwo

W rozprawie, w części poprzedzającej cykl publikacji, dotyczącej omówienia podjętej problematyki oraz omówieniu danych, metod i narzędzi Doktorant powołuje się na 122 pozycje literaturowe (prawie wszystkie w języku angielskim). Publikacje są wymienione w spisie według kolejności cytowania w tekście. Bibliografia zawiera publikacje z okresu 1999 do 2022, przy czym publikacje naukowe opublikowane w ciągu ostatnich 10 lat stanowią ponad 68%. Przytoczone prace są w przeważającej części publikacjami o zasięgu międzynarodowym. Dobór źródeł literaturowych jest trafny, a przegląd wiedzy zarówno w części wprowadzającej, jak i dołączonych publikacjach został przeprowadzony wnikliwie, trafnie wskazując zarówno na osiągnięcia w tematyce poruszonej w rozprawie, jak i luki w dotychczasowej wiedzy i praktyce. W części opartej na przeglądzie dotychczasowej wiedzy Doktorant zwrócił uwagę na różnorodność i niespójność metod stosowanych przy ocenie stanu ekologicznego wód w poszczególnych krajach, a nawet w obrębie jednego kraju. Szczegółowo omówił rozwój metod tych ocen, rolę modelowania stanu ekologicznego jezior oraz problemów z nim związanych takich jak jakość i dostępność danych. Omówił stosowane metody uzupełniania brakujących danych oraz wskazał na problemy z tym związane.

Cel pracy

Autor rozprawy w sposób jasny określił główny cel pracy jako „wsparcie oceny stanu ekologicznego jezior algorytmami uczenia maszynowego”. Wyznaczył również 5 celów szczegółowych, adekwatnych do przyjętego celu głównego, będących jednocześnie głównymi celami cyklu publikacji.

W pierwszej z cyklu publikacji Doktorant postawił sobie zadania zaproponowania metod wsparcia procesu uzupełniania danych pomiarowych wykorzystywanych do oceny stanu ekologicznego jezior C_1 oraz określenie sposobu przetwarzania danych na linii ekspert-

analitik w procesie takiej oceny C₂. Osiągnięcie tych celów pozwoliło na przejście do kolejnego etapu realizacji celu głównego, to znaczy do redukcji liczby wymiarów, w których prowadzona jest ocena w celu uproszczenia dalszej analizy przy jednoczesnej minimalizacji utraty zdolności objaśniającej użytych zmiennych C₃ i stworzenie modelu pozwalającego na powtarzalną i weryfikowalną ocenę stanu ekologicznego jezior C₅. Ostatnim etapem realizacji celu głównego było stworzenie modelu klastrującego, wspierającego proces priorytetyzacji jezior wskazanych do wprowadzenia działań naprawczych i/lub ochronnych C₆. Połączenie realizacji kolejnych celów szczegółowych w przedstawionych publikacjach, stanowi osiągnięcie celu głównego rozprawy. Określone cele są spójne, stanowią swego rodzaju etapy niezbędne do stworzenia narzędzia do usprawnienia oceny stanu ekologicznego jezior poprzez wykorzystanie algorytmów uczenia maszynowego.

Metody badawcze

W części zatytułowanej „Dane, metody i narzędzia” Doktorant wskazał na źródło danych wykorzystanych w badaniach (otwarte zasoby Głównego Inspektoratu Ochrony Środowiska). Przedstawił metody użyte w badaniach w odniesieniu do osiągniętych dzięki nim celów szczegółowych wraz ze wskazaniem miejsc wykorzystania w dołączonych publikacjach. Przyjęte metody i narzędzia z zakresu identyfikacji i imputacji danych, opracowania modelu wykorzystywanego do stworzenia narzędzia wspierającego priorytetyzację ocenianych jezior oparte o algorytmy uczenia maszynowego oraz bogaty materiał badawczy (w ocenianej rozprawie analizie poddano zbiór pomiarów środowiskowych 499 jezior w Polsce) pozwolił Doktorantowi na osiągnięcie zakładanych celów.

Zarówno metodyka, jak i wyniki zawarte w przedłożonych publikacjach zostały już ocenione przez recenzentów tych czasopism, w których się ukazały i merytorycznie nie budzą żadnych zastrzeżeń.

Omówienie wyników badań

Ostatnim podrozdziałem pierwszej części rozprawy jest omówienie najważniejszych osiągnięć związanych bezpośrednio z realizacją założonych celów. Doktorant omówił rezultaty swoich badań porządkując je względem założonych celów oraz wskazując, w której publikacji zostały one przedstawione.

Jednym z problemów, na które zwrócił uwagę Doktorant w swojej pracy jest problem braku danych niezbędnych do prawidłowego przeprowadzania oceny stanu ekologicznego wód ekosystemów wód jeziornych oraz stosowanych dotychczas metod ich imputacji. Jako rozwiązanie dotychczasowych problemów w ewaluacji tych wód Autor proponuje nowy

algorytm uzupełniania danych z wykorzystaniem kombinacji metod imputacji wielokrotnej i grupowania hierarchicznego możliwych wyników, zwracając uwagę na etap przygotowania danych do analiz we wczesnym stadium procesu oceny oraz na rolę ekspertów (operatora modelu) jako osoby odpowiedzialnej za prawidłowość imputacji danych. Zwrócił też uwagę na niezwykle istotny aspekt komunikacji i interakcji na linii ekspert-analityk-odbiorca. Jako rezultat drugiego celu wskazał schemat przepływu danych i informacji pomiędzy ekspertem ds. wdrażania zapisów RDW, a analitykiem danych, kładąc nacisk na wczesny etap ewaluacji ekologicznej, w którym imputacja danych i dobór prawidłowy metod analizy zmiennych minimalizują szum decyzyjny. Autor zaproponował redukcję liczby wymiarów do dwóch wiodących z użyciem metod: analizy głównych składników (PCA) oraz liniowej analizy dyskryminacyjnej (LDA). Kolejnym krokiem w badaniach było odtworzenie wyników oceny stanu ekologicznego dla kampanii pomiarowej 2011-2015. W konstrukcji klasyfikatora użyto liniowej kombinacji zmiennych pozyskanych za pomocą LCA. Uzyskana średnia skuteczność klasyfikacji stanu ekologicznego 499 jezior wyniosła 94%. Kolejnym etapem prac Doktoranta było rozpoznanie możliwości wykorzystania algorytmu uczenia maszynowego w postaci sieci Kohonena do wsparcia procesu priorytetyzacji ekosystemów jeziornych do działań ochronnych i zaradczych.

Doktorant w tej części ograniczył się tylko do omówienia najważniejszych osiągnięć, szczegółowe omówienie wyników badań przeprowadzonych przez Doktoranta jest zamieszczone w publikacjach stanowiących integralną część rozprawy oraz rozszerzone o informacje zawarte w załącznikach 1-3.

Możliwości praktycznego zastosowania uzyskanych wyników badań

Po zapoznaniu się z przedłożoną do oceny rozprawą doktorską Pana mgr inż. Grzegorza Chrobaka uważam, że istnieje duża możliwość praktycznego zastosowania zawartych w niej wyników badań. Dotychczasowy system oceny stanu ekologicznego wód oparty na ocenie wybranych parametrów napotyka na wiele problemów. Odpowiedzią na te potrzeby w zakresie ocen wód jeziornych, mogą być m.in. osiągnięcia zaprezentowane w ocenianej rozprawie doktorskiej. Oparty na trzech komponentach (imputacji danych brakujących, modelowaniu procesu i priorytetyzacji ekosystemów wskazanych do działań naprawczych i/lub ochronnych) system wspomagany algorytmami uczenia maszynowego może być przydatnym narzędziem w rękach ekspertów zajmujących się taką oceną.

Oryginalność rozwiązania problemu badawczego

W swojej rozprawie Doktorant łączy wiele aspektów nowoczesnej nauki w celu opracowania nowego, kompleksowego, opartego o trzy współzależne komponenty systemu

wspierającego proces decyzyjny w trakcie oceny, podczas priorytetyzacji działań ochronnych, a także w prowadzeniu monitoringu osiągnięcia celów środowiskowych. Podejście doktoranta do rozwiązywanego problemu badawczego jakim jest opracowanie skutecznego narzędzia wspierającego pracę ekspertów charakteryzuje się oryginalnością i nowoczesnym podejściem do wykonywanych ocen. Wprowadzenie do systemu oceny stanu ekologicznego elementów szeroko rozumianej sztucznej inteligencji zastosowanych przez autora jest podejściem nowatorskim i przyszłościowym.

Poziom ogólnej wiedzy teoretycznej kandydata w dyscyplinie ochrona i kształtowanie środowiska

Kandydat wykazał się dużą wiedzą w zakresie podzbioru sztucznej inteligencji jakim jest uczenie maszynowe. Chociaż w samej rozprawie teoretyczna wiedza kandydata w dyscyplinie ochrona i kształtowanie środowiska jest tłem i środowiskiem stosowania metod z zakresu sztucznej inteligencji, to jej właściwe stosowanie niewątpliwie wymaga znajomości i zrozumienia zjawisk, zależności i procesów zachodzących w środowisku, występujących relacji pomiędzy elementami badanych ekosystemów wodnych oraz wiedzy na temat technik prowadzenia pomiarów terenowych, czy metod monitoringowych. Wykorzystanie w badaniach danych rzeczywistych (dla 499 jezior) pochodzących z drugiego cyklu planistycznego aktualizacji Planów Gospodarowania Wodami wymaga znajomości istniejących metod opartych na badaniu i ocenie określonych parametrów objętych oceną, charakteryzujących elementy jakości: biologiczne, fizykochemiczne i hydromorfologiczne, jak również podstaw ochrony, utrzymania i poprawy stanu ekologicznego ekosystemów wodnych. O odpowiednim poziomie wiedzy Doktoranta świadczy dobór i umiejętność wykorzystania dostępnej bibliografii. Swoją ogólną wiedzę z zakresu dyscypliny ochrona i kształtowanie środowiska, a zwłaszcza w zakresie badania stanu ekologicznego i ochrony jakości wód Doktorant wykazał również w analizie wyzwań i dalszych kierunków badań.

Nieprawidłowości

- W części literatura brak jednolitości w opisie pozycji literaturowych. Doktorant czasami podaje pełne nazwy czasopism a czasami stosuje skróty (np. pozycja 7, 11, 71, 72 Environmental Modelling and Software, 36 i 39 Ecological Modelling, 38, 41 Ecological Indicators)
- Brak w spisie literatury: Dyrektywa 2000/60/WE Parlamentu Europejskiego i Rady z dnia 23 października 2000 r. przytoczonej w części „Wprowadzenie i luki w wiedzy”
- Autor nie ustrzegł się popełnienia drobnych błędów stylistycznych, interpunkcyjnych i językowych co jednak w żaden sposób nie umniejsza wartości merytorycznej

rozprawy i nie wpływa negatywnie na jej odbiór. Z obowiązku recenzenta poniżej załączam wykaz zauważonych błędów:

-Streszczenie pracy w języku polskim:

– w zdaniu „Jeziora w każdym z nowo powstałych... pozycji w zarówno w odniesieniu do oryginalnej...”

-Czym rozprawa ta nie jest:

– w zdaniu „Opracowany system nie jest także stworzony...”

Wprowadzenie i luki wiedzy:

– w zdaniu „Jednym z kluczowych elementów w zakresie ochrony i zarządzania ekosystemami wodnymi...”

Problematyka oceny ekologicznej – szeroka perspektywa:

– w zdaniu „W związku ze specyfiką oceny..., w roku 2012 rozpoznano po raz pierwszy zwrócono uwagę na problem...-styl – propozycje poprawy: „rozpoznano i po raz pierwszy zwrócono uwagę na problem...” lub „rozpoznano po raz pierwszy i zwrócono uwagę na problem...”

Modelowanie stanu ekologicznego jezior:

– w zdaniu „Zbiory wykorzystanych danych obejmują... procesy zachodzące w ekosystemach”

Jakość danych z perspektywy oceny ekologicznej:

– w zdaniu „W przypadku zestawów pomiarowych służących do... spotykają się z brakami definiowanymi...”

- w zdaniu „Optymalne rozwiązanie tego problemu..., a także od znajomości kontekstu...”

- w zdaniu „Najpopularniejsze konsekwencje niewłaściwego...”

- w zdaniu „Dane wykorzystywane do oceny stanu... braki mogą wystąpić na wczesnym etapie...”

- w zdaniu „Stosowana jest zwykle w sytuacjach... uznanych za analogiczne dla rozpatrywanych.”

Uczenie maszynowe w ocenie ekologicznej jezior:

- w zdaniu „Wśród metod stosowanych w modelowaniu... oparte o..., algorytmy genetyczne...”

- w zdaniu „Jest to szczególnie ważne w kontekście..., na których podejmowane są decyzje na podstawie dokonanej już oceny...”

Rezultaty:

- w zdaniu „Praca odpowiada na zidentyfikowane... algorytmu uzupełniania danych...”

- w zdaniu „Oznacza to, że zanim jezioro osiągnie docelowy stan ekologiczny, logicznym celem jest eel jego awans jakościowy do podgrupy...”

Podsumowanie recenzji

Recenzowana rozprawa doktorska Pana mgr inż. Grzegorza Chrobaka przygotowana w formie zwarteo opracowania z włączonym, spójnym tematycznie cyklem 3 publikacji prezentuje wysoki poziom naukowy, zawiera oryginalne osiągnięcia naukowe o dużym znaczeniu poznawczym i szansą na praktyczne wykorzystanie uzyskanych wyników badań. Sposób realizacji postawionych założeń, zastosowana metodyka, interpretacja uzyskanych wyników, charakteryzują się wysokim poziomem merytorycznym co pozwoliło na publikację osiągniętych rezultatów w wysoko punktowanych czasopismach naukowych. Tematyka rozprawy mieści się w dyscyplinie ochrona i kształtowanie środowiska.

Rozprawa ta spełnia wszystkie wymagania stawiane rozprawom doktorskim określone w ustawie o stopniach i tytule naukowym oraz o stopniach i tytule w zakresie sztuki z dnia 14 marca 2003 roku (tekst jedn. Dz. U. 2017 poz. 1789), w związku z czym wnoszę o dopuszczenie Pana mgr inż. Grzegorza Chrobaka do dalszych etapów przewodu doktorskiego.


dr hab. inż. Maria Strzelczyk prof. ITP-PIB