

**Recenzja rozprawy doktorskiej mgr Marcina Nowaka
„Sztuczne sieci neuronowe i algorytm MARS w regresyjnym modelowaniu odpływu
ze zlewni zlodowaconych na przykładzie Rzeki Waldemara (Svalbard)”**

Recenzowana rozprawa doktorska jest pracą z zakresu geografii fizycznej i hydrologii glacialnej. Dotyczy matematycznego modelowania dobowego odpływu wód z częściowo zlodowaconej zlewni górskiej w warunkach klimatu polarnego w atlantyckim sektorze Arktyki. Podjęty temat wpisuje się w główny nurt współczesnych badań nad problemem globalnych zmian klimatu i środowiska, w których zasoby wodne kriosfery odgrywają zarówno kluczową rolę w mechanizmach zmian jak również są ważne w skali regionalnej w szacowaniu zasobów wodnych dla celów gospodarczych i utrzymania jakości życia. Uzyskane wyniki stanowią propozycję optymalnego wykorzystania wybranych algorytmów regresyjnych (jako metod prostych) w modelowaniu odpływu.

Rozprawa liczy 107 stron, w tym 88 stron tekstu zasadniczego, zawartego w dziewięciu rozdziałach. Problem badawczy omawiany i dyskutowany jest w oparciu o 166 pozycji literatury, gdzie zaznacza się światowa literatura najnowsza lecz należy wyróżnić innych sześć prac, w których Kandydat jest autorem lub współautorem, m.in. w czasopiśmie *Boreas* i *Geografiska Annaler*. Dokumentację badań stanowi 19 tabel i 33 ryciny.

Głównym celem pracy jest ocena możliwości wykorzystania sztucznych sieci neuronowych (SSN) oraz funkcji sklepanych (MARS) jako technik regresyjnych w matematycznym modelowaniu dobowych wartości odpływu z częściowo zlodowaconej, arktycznej zlewni górskiej na przykładzie Rzeki Waldemara. Dla celów porównawczych Autor wybrał stosowaną w zagadnieniach hydrologicznych klasyczną metodę regresji wielokrotnej, podkreślając, że bardziej złożone techniki regresyjne, do których należą SSN i MARS, posiadają „*zdolność do wydobywania wiedzy na temat przebiegu procesów kształtujących dane zjawisko w określonym systemie jedynie na podstawie informacji zawartych w zmierzonych lub oszacowanych sygnałach wejściowych do tego systemu i odpowiadających im reakcji*”.

Autor postawił sobie cztery zasadnicze zadania:

- konstrukcja modeli ma być możliwie prosta;

- modele mają być oparte na standardowych danych meteorologicznych, rutynowo badanych jako tło badanych procesów fizycznych oraz możliwych do pozyskania z jednego systemu pomiarowego (np. automatycznej stacji pogodowej);
- konstrukcja, wykorzystanie i ocena modeli muszą być oparte na popularnym i szeroko wykorzystywanym oprogramowaniu;
- od użytkownika opracowanych modeli nie powinna być wymagana znajomość warsztatu programistycznego.

Obiektem badań jest zlewnia Lodowca Waldemara usytuowana na Ziemi Oskara II w północno-zachodniej części Spitsbergenu Zachodniego. Lodowiec Waldemara (norw. *Waldemarbreen*), stanowiący główny element krajobrazu zlewni jest jednym z kilku lodowców regionu Kaffiøyry, tamtejszej nadmorskiej równiny.

Jest to dobrze rozpoznany obszar, posiadający bogatą literaturę (meteorologia, hydrologia i hydrochemia, bilans masy, geomorfologia strefy marginalnej, geologia). Początki badań polskich sięgają tam roku 1938 a regularne prace badawcze rozpoczęto w połowie lat 70-tych. Materiał badawczy wykorzystany w niniejszej pracy zebrano w trakcie sezonów letnich 2010, 2011 i 2012 roku.

W pierwszych trzech rozdziałach Autor przedstawia podstawy modelowania odpływu ze zlewni zlodowaconych, dokonuje przeglądu głównych podejść obliczeniowych z uwzględnieniem metod regresyjnych oraz określa cel i zakres pracy. Najwięcej uwagi poświęca przeglądowi głównych podejść metodycznych z naciskiem na metody regresyjne w modelowaniu odpływu proglacjalnego. Autor omawia ich główne cechy a w podsumowaniu przedstawia szereg opracowań odpływu dla wybranych basenów glacialnych, uwzględniając te prace, w których standardowe elementy meteorologiczne w jawny sposób określają zależności w empirycznych równaniach regresji i odznaczają się wysokimi współczynnikami determinacji. Ten fragment pracy tworzy właściwy wstęp merytoryczny do zasadniczego celu niniejszej dysertacji. Z trzynastu wybranych prac, w 11 z nich zastosowano regresję wielokrotną a dwóch przypadkach użyto metody sztucznych sieci neuronowych. W omówionych pracach, znajduje się jedno opracowanie współautorstwa Kandydata, dotyczące modelowania odpływu proglacjalnego, właśnie z zastosowaniem sztucznych sieci neuronowych.

Rozdział czwarty zawiera charakterystykę fizyczno-geograficzną obszaru badań. Przedstawione są ogólne cechy rzeźby terenu i geologii regionu Kaffiøyry, główne cechy klimatu i zarys sieci hydrograficznej obszaru. W kolejnych podrozdziałach przedstawione są ogólne cechy glacialnej części zlewni Rzeki Waldemara (4.4) i sam lodowiec (4.5). W ostatnim podrozdziale (4.6) zawarto przegląd dotychczasowych badań hydrologicznych w tym regionie. Pod względem redakcyjnym ale też pod względem przekazu wiedzy o

badanym obszarze cały rozdział czwarty zatytułowany „Obszar badań” jest dla czytelnika mało interesujący i momentami lakoniczny, pomimo sporego wsparcia cytowaną literaturą. Niniejsza opinia dotyczy części klimatologicznej (4.2) ale szczególnie części hydrologicznej, na którą składają się ostatnie trzy podrozdziały (4.4 – 4.6). Skoro Autor zajmuje się regresyjnymi modelami odpływu, w których istotną rolę odgrywają wielkości meteorologiczne a sam w swojej cytowanej pracy bada zależności pomiędzy odpływem a temperaturą, opadem, wilgotnością powietrza, usłonecznieniem i promieniowaniem słonecznym (Nowak i Sobota, 2015) to czytelnik dysertacji spodziewa się nawet ogólnej informacji o tychże elementach klimatu, jako dobrego wstępu do charakterystyki ustroju hydrologicznego badanej zlewni, czego zabrakło w ostatnich podrozdziałach. Dane takie odnajdują się dopiero w następnych rozdziałach (5.1).

Ważną i obszerną częścią dysertacji są rozdziały poświęcone przyjętemu przez Autora podejściu metodycznemu. W rozdziale piątym zatytułowanym „Metody badań” Kandydat przedstawia ogólny przebieg swoich badań i etapy procedur metodycznych, zastosowanych w części terenowej jak i na etapie modelowania. Materiał badawczy (pomiarów hydrometryczne i meteorologiczne) zebrano w trakcie kolejnych trzech sezonów letnich w 2010, 2011 i 2012 roku. W pierwszym etapie postępowania zebrany materiał terenowy oceniono pod względem adekwatności do przeprowadzenia modelowania za pomocą regresji wielorakiej. Ocenie poddano związek odpływu z temperaturą powietrza, opadem atmosferycznym, wilgotnością powietrza, prędkością wiatru, promieniowaniem słonecznym i ciśnieniem atmosferycznym. Uzyskane wyniki potwierdzają znaną naturę ustroju hydrologicznego zlewni proglacjalnej, na którą pierwszorzędny wpływ ma temperatura powietrza. Przeprowadzona ocena wykazała, że uzyskane dane pomiarowe nie spełniają oczekiwań do poprawnej metodycznie konstrukcji parametrycznych modeli regresji liniowej.

Wniosek z poprzedniego rozdziału uzasadnia zastosowanie przez Autora innych metod tj. sztucznych sieci neuronowych (SSN) oraz funkcji sklepanych (MARS) w modelowaniu dobowych wartości odpływu, co jest obiektem dalszych badań, przedstawionych w kolejnych rozdziałach dysertacji.

W obszernym rozdziale szóstym zatytułowanym „Charakterystyka wykorzystanych metod modelowania” Kandydat szczegółowo objaśnia algorytmy zastosowanych sztucznych sieci neuronowych oraz funkcji sklepanych w modelach regresji wielokrotnej opisujących dobowy odpływ, następnie tłumaczy zastosowaną metodykę weryfikacji wyników modelowania oraz przeprowadza analizę wrażliwości modeli regresyjnych traktując procedurę jako metodę identyfikacji zależności pomiędzy odpływem a warunkami meteorologicznymi. W konstrukcji modeli odpływu Autor zaproponował trzy metody, z dwoma wariantami w każdej. Pierwsza z nich to dwa warianty sieci neuronowych, nazywane modelami MLP1 i MLP2, uzyskane w procedurach „samouczenia się” sieci i ocenionych jako posiadające najbardziej optymalne

algorytmy uczenia się. Kolejna metoda nazwana MARS to dwa warianty algorytmu regresji z wykorzystaniem funkcji sklejaných, różniące się jakością dopasowania funkcji z użyciem wielkości promieniowania słonecznego (MARS1) oraz natężeniem przepływu z dnia poprzedniego (MARS2). Trzecia z zaproponowanych metod stanowi tło analizy porównawczej, na podstawie której dokonano oceny wydajności modeli sieci neuronowych oraz modeli MARS. W metodzie tej opartej na liniowej regresji wielokrotnej zaproponowano dwa warianty, opracowane metodą regresji krokowej wstecznej, która zaproponowała dwa najlepsze przybliżenia nazwane tu modelami MVR1 oraz MVR2.

W mojej ocenie jest to niezwykle cenny rozdział dysertacji, w której Autor przedstawia procedury i sam dba o szczegóły analityczne prowadzące do wiarygodności zastosowanych metod i uzyskanych wyników. Świadczy to o dużej odpowiedzialności badawczej i rzetelności naukowej Autora. Opinię tę należy przenieść na rozdział następny, zatytułowany „Wyniki”, który zawiera rezultaty weryfikacji wyników modelowania przeprowadzonego na podstawie trzech metod i sześciu różnych wariantów.

W rozdziale „Wyniki” przedstawionych jest sześć wariantów wyników, które są dwuczęściowe. Pierwsze z nich są efektem weryfikacji wewnętrznej, przeprowadzonej na serii pomiarów z 2010 i 2011 roku, które posłużyły do budowy modelu i jego kalibracji. Materiał pomiarowy z roku 2012 potraktowano jako dane niezależne w procedurze weryfikacji zewnętrznej. W rozdziale tym analizowane są wyniki oceny statystycznej, przedstawione są hydrogramy wartości mierzonych i modelowanych, dokonana jest ocena wrażliwości modeli na zmiany wybranych parametrów meteorologicznych.

Uzyskane wyniki poddane są rozważaniom w rozdziale kolejnym „Dyskusja”, gdzie Autor dokonuje oceny porównawczej poprawności zastosowanych metod oraz konfrontuje możliwości i ograniczenia metody sztucznych sieci neuronowych (algorytm MLP) w modelowaniu odpływu ze zlewni zlodowaconych. Autor wykazuje, że spośród sześciu ocenionych, jeden z wariantów algorytmu sztucznych sieci neuronowych (MLP2) jest modelem o największej poprawności, przewyższającej jakość powszechnie stosowanych modeli budowanych na liniowej regresji wielokrotnej. Kandydat szczególnie podkreśla, że głównym ograniczeniem metod przedstawionych w pracy jest ich rozdzielczość czasowa. Skonstruowane modele bazują na wartościach średniego dobowego natężenia przepływu. Można je zatem stosować do rekonstrukcji odpływu w skali sezonowej, z pominięciem jego zmienności w ciągu doby. Podkreślić należy jeszcze raz ostrożność Kandydata w podejściu metodycznym i w interpretacji wyników. Autor świadomie zwraca uwagę, że przeprowadzona ocena zastosowanych modeli ma na celu *„...podkreślenie faktu, że bezkrytyczne podejście użytkownika do wybranej przez niego metody modelowania (nawet najbardziej kompleksowej) jest błędem i może prowadzić do otrzymania mało wartościowych rezultatów. Każda próba ujęcia rzeczywistego zjawiska za pomocą modelu matematycznego jest*

pewnym jego uproszczeniem. Wybór konkretnej metody modelowania powinien być zatem świadomym kompromisem między jej możliwościami i ograniczeniami. "

Ostatni, dziewiąty rozdział zawiera „Podsumowanie i wnioski”. Autor formułuje szereg konkluzji, a najważniejsze wnioski dedykowane geografii fizycznej / hydrologii glacialnej są następujące:

- wybrane techniki regresyjne pozwalają osiągnąć zadowalające wyniki w zakresie szacowania wielkości i zmienności średniego dobowego odpływu z niewielkich zlewni zlodowaconych;
- niskie wymagania dotyczące danych wejściowych opisywanych modeli mogą stanowić o ich wysokiej aplikacyjności;
- spośród wybranych trzech metod (6 wariantów) największą poprawność wykazują nieparametryczne modele regresji realizowanej przez sztuczne sieci neuronowe (modele MLP);
- zastosowanie metod analizy wrażliwości modeli regresyjnych może stanowić dodatkowe źródło wiedzy na temat ilościowej charakterystyki procesów kształtujących wielkość i zmienność odpływu proglaclalnego;
- modele MLP mogą być cennym uzupełnieniem badań z zakresu bilansowania zasobów wodnych zlewni zlodowaconych oraz stanowić wydajną metodę pomocniczą w prowadzonych współcześnie wieloaspektowych badaniach geograficznych, np. z zakresu dynamiki procesów geomorfologicznych, gdzie wielkość i zmienność natężenia przepływu rzecznoego jest jedną z ważniejszych zmiennych środowiskowych.

Przedstawiona dysertacja jest kompetentna i rzeczowa, napisana zrozumiałym, przejrzystym językiem. Praca jest momentami niełatwa w charakterze jej odbioru, co po części wynika ze specyfiki postawionego problemu badawczego a po części ze sposobu podziału treści i odpowiedniego akcentowania w nich głównych wątków i konkluzji. Opinia ta nie rzutuje na ogólną wartość merytoryczną pracy. Autor wykazuje się nie tylko odpowiednią wiedzą i umiejętnościami ale również odpowiedzialnością naukową i głębokim zaangażowaniem poświadczonym autorstwem i współautorstwem prac w podjętej dziedzinie badań.

Wyniki postawione w pracy mają dużą wartość poznawczą i naukową i można mieć nadzieję, że znajdą się w przyszłych publikacjach Autora. Temperatura a potem opad atmosferyczny okazują się dwoma najważniejszymi zmiennymi w modelowaniu odpływu. Tu można postawić pytanie lub sformułować kolejny problem badawczy, którym byłaby realizacja scenariusza zmian odpływu w przyszłości lecz także odtworzenie ustroju

hydrologicznego zlewni na podstawie danych historycznych i próba interpretacji skutków środowiskowych.

Wydawać się może, że od pracy powstałej w środowisku przyrodników realizujących badania terenowe w obszarach trudno dostępnych i takich, które dysponują wieloletnimi seriami danych oczekuje się rozwiązania problemu dotyczącego cech obiektu/zjawiska ujętego w przestrzeni geograficznej i w czasie. Ta praca nie może spełniać takich oczekiwań jednakże jest to ważne dzieło naukowe, praca metodyczna oparta o nowoczesny warsztat naukowy i techniki badawcze, które powinny częściej być obecne w naukach o ziemi uprawianych w naszym kraju.

Jestem przekonany, że przedstawiona rozprawa spełnia warunki określone w Ustawie o Tytule i Stopniach Naukowych dla rozpraw doktorskich i wnioskuje o dopuszczenie mgra Marcina Nowaka do dalszego etapu przewodu doktorskiego.

Wrocław, 18 lipca 2016

