

Warszawa, 22.08.2023

Prof. dr hab. Mariusz Majdański
Instytut Geofizyki
Polskiej Akademii Nauk
ul. Księcia Janusza 64
01-452 Warszawa
email: mmajd@igf.edu.pl

Recenzja rozprawy doktorskiej

Pani mgr Justyny Sowińskiej-Botor zatytułowanej

„Estymacja i weryfikacja parametrów strefy małych prędkości z wykorzystaniem metod statystycznych”

Recenzowana rozprawa doktorska wykonana została pod kierunkiem dr hab. inż. Michała Stefaniuka, prof. AGH. Poniższa recenzja została przygotowana na zlecenie Rady Dyscypliny Naukowej „Nauki o Ziemi i Środowisku” Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie z dnia 5 czerwca 2023r. numer dokumentu RDN-NoZiŚ-dz.510-7/2023.

1. Informacje ogólne

Recenzowana rozprawa ma klasyczny układ rozprawy doktorskiej w formie monografii. Dokument spełnia wszystkie formalne wymagania, jest spójny i krótki, opisany w sposób prosty i czytelny z dobrze dobranymi figurami ilustrującymi poszczególne etapy przeprowadzonych analiz. Rozprawa czytelnie określa problem badawczy i jasno stawia tezy pracy. Forma rozprawy jest bardzo dobra, mam jednak istotne uwagi merytoryczne dotyczące samego celu pracy, a przede wszystkim danych użytych w prezentowanych analizach i zasadności całej analizy.

2. Problem badawczy

Tematem rozprawy jest określenie strefy małych prędkości badanego obszaru w oparciu o punktowe dane sejsmiczne. Sama rozprawa określa główny cel pracy jako „...dobór optymalnych metod statystycznych pozwalający uzyskać akceptowalny wynik estymacji ...”. W przeciwieństwie do tematu rozprawy jej cel jasno określa, że prezentowane wyniki mocno skupiają się na samych metodach geostatystycznych, a nie na wynikach rozkładu parametrów sejsmicznych ośrodka.

Tezy rozprawy zdefiniowane są jako:

1. Zastosowanie w interpolacji danych z mikroprofilowań prędkości w płytkich otworach, zmienności warunków hydrogeologicznych, istotnie poprawia estymację parametrów strefy małych prędkości.
2. Estymacja parametrów strefy małych prędkości z użyciem metod geostatystycznych umożliwia ocenę wpływu gęstości opróbowania i zmienności przestrzennej na błędy interpolacji, co pozwala na określenie minimalnej ilości otworów.
3. Interpolacja danych mikroprofilowań prędkości w warunkach rozkładów odbiegających od normalnego jest możliwa i wiarygodna, ale wymaga zastosowania specjalnych procedur statystycznych.

Tezy pracy wyraźnie wskazują na statystyczną część badanego problemu. Jako geofizyk pracujący z danymi sejsmicznymi liczyłem na większy komponent przetwarzania danych, albo przynajmniej na przykłady danych (w pracy nie ma ani jednego zapisu sejsmicznego). Doceniam natomiast duży nacisk i staranność przy detalicznej analizie błędów w zastosowanych metodach interpolacji.

3. Struktura rozprawy

Rozprawa jest klasyczną monografią o standardowym układzie zgodnym z wymogami pracy naukowej. Całość jest w języku polskim i włącznie z referencjami zamyka się w 128 stronach. Napisana jest bardzo starannie, prostym i zrozumiałym językiem. W całym tekście nie znalazłem nawet literówek. Są za to drobne błędy interpunkcyjne. Figury i tabele w rozprawie wykonane są bardzo starannie i spełniają wymogi publikacji naukowych.

Wstęp bardzo zwięźle określa cel pracy oraz definiuje tezy. Zawiera również odpowiednie referencje do prac poruszających tematykę badawczą związaną z geostatystyką.

Rozdział drugi, charakterystyka obszaru badań, definiuje miejsce analiz w okolicach Poznania oraz dokładnie prezentuje budowę utworów czwartorzędowych. Rozdział dokładnie opisuje budowę geologiczną, skład petrograficzny oraz zawiera liczne odniesienia do prac naukowych badających obszar badań. Całość uzupełniona jest wieloma przekrojami hydrogeologicznymi. Po tym rozdziale spodziewałem się dokładnych analiz sejsmicznych i interpretacji wyników w nawiązaniu do budowy geologicznej.

Trzeci rozdział do bardzo pobieżne przedstawienie metod geofizycznych używanych do rozpoznania strefy przypowierzchniowej. Rozdział ten jest bardzo ogólny, a jego przydatność dyskusyjna. Jeśli miałby to być przegląd metod to brakuje podstawowych rysunków i referencji do głównych angielskojęzycznych źródeł. Nie rozumiem też przedstawiania tak skrótowej informacji o metodach MASW, tomografii czy sejsmice refleksyjnej skoro i tak metody te nie są dyskutowane w rozprawie. Rozdział ten ma jeszcze dwa niedociągnięcia. Na str. 26 ponownie definiuje cel pracy jako „rozpoznanie strefy przypowierzchniowej oraz wykorzystanie uzyskanych wyników w przetwarzaniu i interpretacji danych metodą sejsmiki refleksyjnej” co, po pierwsze, nie jest prawdą, a po drugie mijają się z celem pracy ze str.9. Dodatkowo opis na str.29 definiuje technikę mikro-sejsmocarotazu jako „uphole w otworach wiertniczych oraz płytkich pomiarów refrakcyjnych”. Mogę się jedynie domyślać, że chodzi o tzw. *vertical seismic profiling* i pomiar wyłącznie w kierunku pionowym. Opis jest niejednoznaczny i w danych mogą być wykorzystane pomiary refrakcyjne rozpoznające prędkości poziome albo tzw.

walkaway VSP rozpoznający kombinację prędkości pionowych i poziomych. Ma to naprawdę duże znaczenie jeśli weźmiemy pod uwagę istotną anizotropię pomiędzy pionem i poziomem typowo występującą w płytkich obszarach. W dalszej części rozdziału znajdujemy jedyny w całej rozprawie przykład interpretacji pomiaru strefy małych prędkości (Fig.3.2). Figura jest jasna i wykonana zgodnie z praktyką pomiarów komercyjnych. Niestety nawet w tym jedynym przykładzie nie ma zapisów sejsmicznych. Co warto podkreślić wynik ten jest zawsze obarczona błędem subiektywnej interpretacji osoby analizującej odnośnie wyznaczenia czasów przebiegu i ugięć hodografu, a tym samym wynik ten zawsze obarczony jest błędem. Niestety w całej rozprawie dane te, które są danymi wejściowymi do analiz geostatystycznych, traktowane są jako dane bezbłędne. Co więcej nawet ten jedyny przedstawiony przykład przedstawia interpretację trójwarstwową w pierwszych 40 metrach ośrodka, a w dalszych analizach wszystkie dane traktowane są jako ośrodek dwu warstwowy. W jaki sposób dokonano reinterpretacji wszystkich danych na dwie warstwy?

Rozdział czwarty opisuje metodykę badań i jest znacznie lepiej napisany w stosunku do rozdziału poprzedniego. Opisy metod i definicje są jasne i czytelne, a także dobrze ilustrowane i zawierają właściwe odnośniki literaturowe. Jedyna drobna uwaga dotyczy opisu metody Krigingu z zewnętrznym trendem zdefiniowanym jako „trend o powolnych zmianach wartości zmiennej”, a uważam że powinno być dodane że chodzi o zmienność przestrzenną.

Piąty rozdział opisuje analizę danych pomiarowych. W tym miejscu widzę największy problem pracy i błędne założenia przedstawionej analizy. Zbiór danych zawiera cztery podzbiory mikroprofilowań sejsmicznych wykonanych na 310 punktach. Cytując „172 pomiary wykonano w latach 70tych wzdłuż linii pomiarowych”. Tu domyślam się, że chodziło o płytką refrakcję. 119 pomiarów pochodzi z lat 2007-2008. Tu brak informacji jak były wykonane. Następnie jest jasna informacja o analizach w 7 otworach z 2011 roku oraz 12 otworach z 2014 roku. Tu jest jasna informacja, że chodzi o pionowy profil sejsmiczny w otworze, czyli wysokiej jakości dane o spodziewanych małych błędach pomiarowych i interpretacyjnych. Jeżeli te dane uzyskane były różnymi technikami to przez istnienie wyraźnej anizotropii pomiędzy prędkościami pionowymi a poziomymi (to może być nawet 20% w nieskonsolidowanych ośrodkach) dostajemy podzbiory danych, których nie powinno się interpretować jednocześnie. Anizotropia taka jest typowym zjawiskiem i wpływa zarówno na określenie prędkości jak i miąższości warstw. Kolejnym istotnym czynnikiem jest zmienność prędkości sejsmicznych w czasie. Z doświadczenia w realiach Polski widzimy zmiany dochodzące do 10% prędkości fal P w najpłytszych warstwach przy zmianach sezonowych, czyli w odstępie kilku tygodni. Wynikają one z zawartości wody opadowej w ośrodku, jego przenikalności, a w szczególnych przypadkach nawet zamarzania ośrodka. Oczywiście zmiany sezonowe dotyczą jedynie pierwszych kilku metrów, ale ma to wpływ na cały profil prędkości. Porównywanie prędkości ośrodka na przestrzeni 40 lat z pewnością może generować takie różnice. Dlatego w przypadku wysokorozdzielczych pomiarów refleksyjnych pomiary przypowierzchniowe do poprawek statycznych wykonuje się w trakcie prac. Ostatnie zastrzeżenie dotyczy ilości warstwy w danych wejściowych. Z pewnością dane pochodzące z 40 lat były interpretowane przez różne osoby (błąd subiektywny), i z pewnością zawierały rozpoznanie różnej ilości warstw. Jak zostało to przełożone na dwie warstwy użyte w dalszych analizach (błąd systematyczny)? W rzeczywistości na profilu do głębokości 40 metrów z pewnością prędkości sejsmiczne są różne na różnych głębokościach i opisanie ich jako 2-3 warstwy o stałym prędkości, co zwyczajowo się stosuje, jest uproszczeniem. Dalej w rozdziale 5.1 rozprawa dzieli podzbiory danych na stare (OLD) i nowsze (NEW), oraz rozdziela uwzględniając lokalne warunki morfologiczne. Wg mnie tak istotna różnica w prędkościach i miąższościach warstw OLD i NEW

powinna wzbudzić zastrzeżenia i być analizowana oddzielnie, albo powinna być lepiej zweryfikowana pod kontem metod pomiaru i podstaw fizycznych. Zamiast tego doktorantka przeprowadza szczegółowe analizy statystyczne danych wejściowych traktując je jako spójny zbiór wiarygodnych danych bez błędów pomiarowych, co w najlepszym przypadku spowoduje sztuczne uśrednienie różnych wyników. W rozdziale 5.2 już na fig.5.7 widać rozkład dwumodalny, którego podstawy fizyczne zostały pominięte. Tu trzeba przyznać, że analizy statystyczne wykonane są prawidłowo, a wszystkie wyniki przedstawione czytelnie i zrozumiale.

Rozdział szósty to już zaawansowane analizy geostatystyczne, którymi doktorantka posługuje się biegle. Analizy anizotropii w formie wariogramów kierunkowych są czytelne i wykonane zgodnie ze sztuką. Zmiany na krótkich odległościach interpretowane jako efekty samorodkowe są wg mnie niespójności danych wejściowych. Fig.6.5 przedstawiająca dopasowanie wariogramów eksperymentalnych ponownie pokazuje dużą niespójność danych, która została sztucznie uśredniona procesem normalizacji, czyli w tym przypadku silnym i nieliniowym przekształceniem funkcją anamorfozy. Wg mnie ponownie jest to niewłaściwy krok przetwarzania pomijający podstawy fizyczne w danych wejściowych. Z drobnych uwag w tekście jak i w tytule rozdziału pojawia się „dane odbiegające od normalności”. To jest oczywiście kolokwializm, który powinien być zastąpiony formalnym wyrażeniem „dane nie posiadające cech rozkładu normalnego”, ale są to pojedyncze przypadki w tej świetnie napisanej rozprawie.

Rozdział siódmy prezentuje główne wyniki rozprawy czyli wyinterpolowane parametry warstwy przypowierzchniowej. Wyniki uzyskane metodą Krigingu (OK) pokazują średni błąd 37% z górnym kwartylem 72%. Tak duże błędy w wyniku podważają samą ideę interpolacji tych danych w tak zróżnicowanym obszarze i tak różnych danych wejściowych. Doktorantka przedstawia jeszcze analizy blokowe, które powoduje dalsze uśrednianie wyników i tym samym istotną poprawę dokładności. To oczywisty wynik zgodny z regułą Ockhama. Nadal jednak brakuje krytycznej analizy i zastanowienia się nad sensem tak skomplikowanych analiz geostatystycznych przy pominięciu podstaw fizycznych w danych. Na str.85 znalazłem kolejne kolokwializmy w postaci „rozkłady błędu spadły o ponad 100%”, albo „nadal jest to poziom zawyżony ... jednak wydaje się być bardziej realistyczny”. Cały opis jest jasny i poprawny, ale język w pracy naukowej powinien być jednak bardziej precyzyjny. Mam także zastrzeżenie do fig.7.6, która przedstawia błędy estymacji V1 w skali barw przedstawiającej wartości jedynie poza obszarem badań.

Rozdział ósmy prezentuje integracje danych sejsmicznych z danymi hydrogeologicznymi i skuteczną próbę zmniejszenia błędów interpolacji przez uwzględnienie dodatkowych danych i zjawisk fizycznych jakimi są efekty zwierciadła wód gruntowych. Doktorantka uwzględnia w analizach dodatkowe dane z niepublikowanych prac Zajac i in. 2014, w których jest współautorem. Wykorzystuje także dodatkowe dane, głębokość zwierciadła wód gruntowych, z mapy hydrograficznej (Adamski 2002). Uwzględnienie dodatkowych danych i ograniczenie do podzbioru najnowszych danych sejsmicznych znacznie poprawia wyniki uzyskane kokrigingiem kolokacyjnym. Wg mnie wynika to z ograniczenia danych co spowodowało ich spójność. Dodatkowy wiąż wynikający z danych o wodach gruntowych z pewnością był pomocny, a błędy w wynikach są istotnie mniejsze.

Rozdział dziewiąty to bardzo krótka analiza zależności wyniku od zagęszczenia punktów pomiarowych. Analiza jest statystycznie poprawna i wykazuje optymalne zagęszczenie punktów pomiarowych. Taką analizę łatwo jest przeprowadzić *post factum* kiedy znamy już strukturę. W rzeczywistości często

niewiele wiemy o badanym ośrodku więc zagęszczenie punktów pomiarowych wynika głównie z ograniczeń budżetu projektu. W fig.9.2 i 9.3 dziwią mnie wartości 0 dla minimalnych wartości prędkości i miąższości. Prędkości 0 są niefizyczne.

Podsumowanie jest bardzo krótkim streszczeniem wcześniejszych analiz. Kolejny raz pojawia się „rozkład danych znacznie odbiegających od normalności”, ale opis jest dobry i zawiera ilościowe wyniki. Nie zgadzam się jednak z tezą (str.120), że „prawidłowo dobrane model geostatystyczne mogą zastąpić gęstą siatkę pomiarów, dając możliwość obniżenia kosztów”. Analizy takie możemy przeprowadzić już po uzyskaniu danych, a optymalne dobranie siatki pomiarowej wymaga wstępnej wiedzy o badanej strukturze. Wg mnie bez tej wiedzy najlepsze modele geostatystyczne nie pomogą w optymalizacji kosztów.

4. Krytyczne i dyskusyjne uwagi dotyczące rozprawy

Praca przygotowana jest bardzo starannie. Język jest poprawny, a kilka kolokwializmów które odnalazłem opisałem powyżej. Figury również przygotowane są bardzo starannie. Struktura rozprawy spełnia formalne wymogi pracy naukowej. Analizy geostatystyczne przedstawione w rozprawie wykonane są z niezwykłą starannością i dowodzą opanowania używanych metod. Szczególnie cenne wg mnie są bardzo precyzyjne ilościowe analizy dokładności przedstawione dla każdego z wyników oraz prawidłowy dobór metod nakierowany na dalsze poprawienie dokładności.

Jak już opisałem wcześniej mam jednak istotne uwagi merytoryczne które postaram się tu podsumować:

- 1) O ile doktorantka dobrze zna się na metodach geostatystycznych to podstawy geofizyczne pracy, które są celem omawianej rozprawy, są zmarginalizowane. Opisy metod geofizycznych z rozdziału trzeciego są szczątkowe, a ich przydatność w tej rozprawie dyskusyjna. Pominięcie podstaw geofizycznych uważam za największy problem tej pracy. Brak krytycznego podejścia do danych wejściowych stworzył trudny problem statystyczny, który doktorantka starała się rozwiązać. Wg mnie połączenie danych z mikroprofilowań sejsmicznych uzyskanych na przestrzeni 40 lat, na dużym zróżnicowanym geologicznie obszarze, prawdopodobnie różnymi technikami pomiarów, doprowadziło do utworzenia zbioru z rozkładem dwumodalnym i zmusiło do użycia silnych, nieliniowych funkcji normujących, co doprowadziło do bardzo dużych błędów w uzyskanych wynikach.
- 2) Praca ma bardzo dobry komponent analizy geostatystycznej, ale brakuje informacji o danych wejściowych. W jakiś sposób były one zebrane, jak interpretowane, jakimi błędami obarczone. Dużym niedociągnięciem jest brak informacji o tym jak profile prędkościowe zostały przekształcone w jednorodny model dwuwarstwowy. Dane wejściowe do analiz geostatystycznych traktowane są jak pewniki bez błędów pomiaru, co dla mnie jako fizyka jest niedopuszczalne.
- 3) Głównym celem pracy jest optymalny dobór metod geostatystycznych i taka analiza została przedstawiona. Z tytułu rozprawy spodziewałbym się, że celem rozprawy będzie wyznaczenie parametrów badanej struktury i wtedy krytyczne podejście do danych wejściowych dałoby dokładniejszy wynik przy użyciu najprostszej metody krigingu. Zresztą jak pokazują wyniki z

rozdziału 8 wyniki uzyskane najprostszą i najbardziej zaawansowaną metodą są bardzo bliskie i w rozwiązaniach praktycznych równoważne.

- 4) Nie zgadzam się z drugą tezą rozprawy, która twierdzi że metody geostatystyczne pozwolą na określenie optymalnej ilości pomiarów do określenia parametrów struktury. Do tego musimy znać strukturę, a zazwyczaj nie mamy takiej wiedzy projektując pomiary geofizyczne. Zgadzam się, że *post factum* geostatystyka może określić co zrobiliśmy nieoptymalnie, ale nie jesteśmy w stanie pokazać tego przed pomiarami.
- 5) Dodatkowo mam uwagę formalną. Do uzyskania doktoratu w Instytucie Geofizyki, w którym pracuje niezbędne jest opublikowanie pracy naukowej. W spisie literatury znalazłem referencję do pracy Zajac i in. 2014, w której doktorantka jest współautorem. Niestety praca ta nie jest opublikowana i znajduje się jedynie w archiwum AGH. Drugi artykuł do praca pierwszoautorska doktorantki, która jest w druku. Niestety nie miałem dostępu do tej pracy.

5. Wniosek końcowy

Przedstawiona rozprawa przedstawiają oryginalne i metodycznie poprawne rozwiązanie problemu naukowego jakim jest estymacja parametrów strefy małych prędkości w badanym obszarze. Rozprawa jest dowodem opanowania przez doktorantkę warsztatu badawczego oraz dogłębnej wiedzy w zakresie geostatystyki. Potwierdzam, że recenzowana praca spełnia warunki określone w art.13 ust 1 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki (tekst jednolity w Dz. U. z 2017 r. poz. 1789) w związku z art. 179 ust. 1 i ust. 2 ustawy z dnia 3 lipca 2018 r. przepisy wprowadzające. Na tej podstawie wnioskuję do Wysockiej Rady Dyscypliny Naukowej AGH o dopuszczenie mgr. inż. Justyny Sowińskiej-Botor do publicznej obrony przedstawionych w jej rozprawie tez.

Warszawa, 22.08.2023

