

Prof. dr hab. Stanisław Lasocki
Instytut Geofizyki PAN
ul. Księcia Janusza 64, 01-452 Warszawa

Recenzja rozprawy doktorskiej mgr inż. Justyny Sowińskiej-Botor pt. "Estymacja i weryfikacja parametrów strefy małych prędkości z wykorzystaniem metod statystycznych"

Celem recenzowanej rozprawy doktorskiej, zdefiniowanym we Wstępie był „...dobór metod statystycznych, które w trudnych warunkach danych, mocno odbiegających od rozkładu normalnego, pozwoliły by uzyskać akceptowalny wynik estymacji parametrów prędkości sejsmicznej i miąższości SMP...” (cyt. Wstęp, str. 9; Nb. nie ma prędkości sejsmicznej - chodzi o prędkość fali sejsmicznej, prędkość nie ma parametrów – chodzi o poziomy rozkład prędkości fali w pierwszej i drugiej strefie małych prędkości).

Przyjęto, że cel ten będzie osiągnięty, jeśli udowodnione zostaną wymienione dalej tezy, cyt. Wstęp, str. 9:

„Zastosowanie w interpolacji danych z mikroprofilowań prędkości w płytkich otworach, zmienności warunków hydrogeologicznych, istotnie poprawia estymację parametrów strefy małych prędkości”.

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.

Interpolacja danych mikroprofilowań prędkości w warunkach rozkładów odbiegających od normalnego jest możliwa i wiarygodna, ale wymaga zastosowania specjalnych procedur statystycznych”.

Tezy pierwszej nie zrozumiałem. Teza druga jest mocno wątpliwa, gdyż jeśli estymuje się parametry to znaczy, że ma się w ręku dane pochodzące z punktów pomiarowych, np. otworów. A więc otwory już powstały. Informacja, że jest ich za dużo nie zmienia stanu rzeczy. Rozkład punktów pomiarowych musi się optymalizować przed jego ustaleniem, a nie po. Ogólne sformułowanie tezy trzeciej (każda interpolacja?, wiarygodna według jakiego kryterium? według kogo?) czyni ją nie do udowodnienia.

Zarówno cel jak i tezy są ogólne. Przy takim ich określeniu w rozprawie powinna zostać przedstawiona ogólna metodyka estymacji parametrów strefy małych prędkości poza punktami obserwacyjnymi i wyniki weryfikacji tej metodyki poprzez dobraną jakościowo, pewną liczbę przypadków polowych. Ale rozprawa przedstawia wyniki przetwarzania danych pomiarowych tylko z jednego obszaru.

W stanie prawnym obowiązującym recenzowaną rozprawę nie żąda się by w rozprawie doktorskiej została postawiona i udowodniona nowa teza (lub tezy). W Ustawie z dnia 14 marca 2003 r. „o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki”, art. 13 pkt 1 jest:

„Rozprawa doktorska ... powinna stanowić oryginalne rozwiązanie problemu naukowego, oraz wykazywać ogólną wiedzę teoretyczną kandydata w danej dyscyplinie naukowej lub artystycznej oraz umiejętność samodzielnego prowadzenia pracy naukowej lub artystycznej”.

Dlatego w dalszej analizie rozprawy skupiłem się nad tym czy zostało przedstawione poprawne rozwiązanie jakiegoś problemu naukowego w badanym *case study*, w oderwaniu od ogólnych zdań na temat celu i tezy.

Podstawowe analizowane dane to „wyniki mikroprofilowań sejsmicznych wykonanych na 310 punktach pomiarowych rozmieszczonych na obszarze badań” (str. 57) – miąższości pierwszej i drugiej warstwy strefy małych prędkości i prędkości fali sejsmicznej w tych warstwach. Ze względu na pochodzenie danych obserwacyjnych, można się spodziewać, że są one silnie heterogeniczne tak w przestrzeni jak i czasie poboru. I niestety jest to skokowa heterogeniczność klastrowa wynikająca po pierwsze z faktu, że obszar badań składa się z dwóch różnych części (wysoczyzna polodowcowa i Pradolina Toruńsko-Eberswaldzka). Po drugie, z tego, że pomiary wykonano w latach 70-80 XX wieku i od 2007 roku, a więc można się spodziewać znacznych różnic w dokładności. Jeśli dane tworzą rozdzielone klastry, tak jak powinno się spodziewać w przypadku danych wykorzystanych w rozprawie, to z wyjątkiem bardzo specjalnych technik służących bardzo specjalnym celom, analizę statystyczną prowadzi się odrębnie w każdym klastrze. Jeśli nie jest to możliwe to, o ile to możliwe, przed rozpoczęciem analizy dane trzeba zhomogenizować. Zebranie danych z rozdzielonych klastrów w jeden zbiór badany daje niepoprawne i nieprzydatne wyniki analizy statystycznej.

Heterogeniczność danych została w pracy zauważona, wskazana na rysunkach 5.3 – 5.6 (Rozdział 5.1) i zignorowana. Nie mogę zrozumieć jak przy tak wyraźnym zróżnicowaniu, jak na wspomnianych rysunkach można było zignorować niebezpieczeństwo jaki ten typ heterogeniczności niesie z sobą. Połączenie grup danych o całkowicie różnych charakterystykach statystycznych było kardynalnym błędem o bardzo dużych konsekwencjach.

Kriging jest metodą interpolacji wykorzystującą własności pewnego typu procesu gaussowskiego. W uproszczeniu, proces gaussowski jest procesem stochastycznym, indeksowanym w przypadku krigingu położeniem, takim, w którym łączny rozkład wszystkich zmiennych losowych (parametrów) jest wielowymiarowym rozkładem Gaussa. Czyli parametry, dla których będzie prowadzona interpolacja - w rozprawie miąższości i prędkości w pierwszej i drugiej warstwie strefy małych prędkości, mają mieć rozkłady normalne.

Przedstawione na Rys. 5.7 (Rozdział 5.2.1) histogramy przetwarzanych parametrów pokazują wyraźnie, że rozkłady parametrów zdecydowanie odbiegają od rozkładów normalnych. Aby do takich parametrów można było zastosować kriging trzeba je najpierw znormalizować czyli tak przekształcić by przekształcone miały rozkład normalny. Aby można było dokonać takiego przekształcenia, trzeba wiedzieć jaki jest rozkład parametrów przed normalizacją. W rozprawie udało się znaleźć model rozkładu tylko dla jednego parametru – prędkości w warstwie pierwszej (log-normalny). Dla pozostałych parametrów modeli nie znaleziono i nie można ich było poprawnie znormalizować. W związku z tym, na przykład, wykorzystano w kringingu surowe dane miąższości warstwy pierwszej mimo, że miały rozkład całkowicie odbiegający od rozkładu normalnego (prawdopodobnie rozkład wykładniczy), a więc nie spełniały warunku stosowalności kringingu.

Co więcej, miąższości i prędkości drugiej warstwy są wyraźnie dwumodalne. Zostało to zauważone w tylko przypadku miąższości oraz bezrefleksyjnie, bo zignorowane w dalszej pracy. Tymczasem, więcej niż jedna moda w gęstości prawdopodobieństwa (wyraźna modalność histogramu, jak tutaj, to pewna modalność gęstości) to cecha opisowa wskazująca na wymieszanie różnych populacji (Cox, Brit. J. Math. Statist. Psychol. 19, 39-47. 1966). I wbrew temu co znajduje się w R6.2, cyt. „dwumodalność rozkładów danych zdecydowanie przemawia za przeprowadzeniem ich normalizacji”, żadna normalizacja temu nie zaradzi.

Wykorzystanie na wejściu czy to danych surowych (miąższość warstwy pierwszej), czy to danych niepoprawnie normalizowanych (miąższość i prędkość warstwy drugiej) wyklucza poprawność wyników dalszej analizy.

Beztroskiego podejścia do danych wejściowych nie zachwiała przeprowadzona analiza anizotropii (Rozdział 6.1) mimo, że mapy wariogramów kierunkowych (Rys. 6.1-6.4) wykazują bardzo duże zróżnicowania kierunkowe i niepokojąco zmienne wartości wzdłuż części kierunków. Wariogramy kierunkowe robi się po to by wybrać takie kierunki, wzdłuż których długozasięgowa zmienność przestrzenna (drift) jest mała. Do takich wariogramów dopasowuje się modele teoretyczne. Po co zostały zrobione w recenzowanej rozprawie – nie wiem, bo ani ich nie zinterpretowano ani z nich dalej nie korzystano. Może dlatego, że software umożliwił?

W Rozdziale 6.3 do wariogramów empirycznych, obliczanych dwoma technikami z danych surowych i przekształczonych dopasowuje się modele teoretyczne. Modele teoretyczne wariogramów są kluczowe w krigingu gdyż na ich podstawie określa się wagi obserwacji wykorzystanych w estymacji parametru poza punktami obserwacji. Niestety modele uznane w rozprawie za najlepsze są złe, gdyż nie modelują wariogramów empirycznych. Wynika to z kształtów wariogramów empirycznych – odwrócona parabola wariogramu miąższości pierwszej warstwy (Rys. 6.5-1b), długookresowa oscylacja wariogramu miąższości drugiej warstwy (Rys. 6.7-2b). Jeszcze jako tako prezentuje się wariogram prędkości drugiej warstwy (Rys. 6.8-2b), w którym oscylacje krótkookresowe można by ostatecznie potraktować jako rozproszenie statystyczne. Jednak nie widać by określał położenie wariancji progowej (*sill*)¹. Te niemodelowalne kształty wariogramów empirycznych wynikają ze złego wyboru danych wejściowych, o czym było powyżej. W efekcie różnice pomiędzy danymi eksperymentalnymi, a modelami są systematyczne. Dla miąższości pierwszej warstwy mniej więcej połowa wartości wariogramu empirycznego (do $h=1500$) jest poniżej wartości z modelu, a druga połowa powyżej. Dla miąższości drugiej warstwy wariogram empiryczny wykonuje jedną pełną oscylację wokół linii wymodelowanej.

Jak złe są opracowane modele, z tzw. najlepszymi włącznie, pokazuje Rys. 6.9 zestawiający obserwacje z estymatami krigingowymi. Obserwacje i estymaty pochodzące z akceptowalnych modeli kompletnie się z sobą nie zgadzają. Nie jest bowiem najważniejsze to jaka jest korelacja (która zresztą jest bardzo mała) tylko to jakie są odchylenia punktów od linii zaczynającej się w punkcie (0, 0) i współczynnika nachylenia 1.0. W modelowaniu statystycznym są dwa źródła niepewności – aleatoryczne, wynikające z czysto losowej składowej obserwacji i epistemiczne, wynikające z błędności modelu. Gdyby w jakimkolwiek przypadku z przedstawionych na Rys. 6.9 punkty były rozrzucone losowo wokół zaczynającej się od zera linii o nachyleniu 45 stopni, to, nawet przy znacznym rozrzucie, świadczyłyby o

¹ Wariogramu prędkości w warstwie pierwszej nie komentuję, bo chyba rysunek najlepszego dopasowania (Fig. 6.6-3a) został pomyłony.

poprawności modelowania. Taki obraz jak jest na Rys. 6.9 świadczy o tym, że modelowanie jest katastrofalnie niepoprawne. Przy okazji: o tym, że współczynnik korelacji dla danych składających się z dwóch chmur punktów, jak na Rys. 6.9 na wykresach rozrzutu dla parametrów drugiej warstwy, nie ma żadnej wartości interpretacyjnej i nie powinien być liczony, uczyłem studentów geologii bodajże na drugim roku.

Nie pomoże tu żadne zaklinanie rzeczywistości podawaniem różnych miar statystycznych. Nb. pierwszy raz widzę, by jakiś średni błąd estymacji był zero (w rozprawie MSE). Pewnie zabrakło wartości bezwzględnej we wzorze 4.52.

Poza zakresem zmienności modeli znalazło się od 20 (miąższość warstwy pierwszej) do blisko 40 procent (prędkość w warstwie drugiej) obserwacji. Nie ma to większego znaczenia i wspominam tylko dla porządku, bo to co pozostało w zakresie jest i tak źle modelowane. A tymczasem na str. 78 czytamy „*Wszystkie testowane modele spełniały kryteria bardzo dobrego dopasowanie według przyjętych miar dobroci*”.

Każda metoda estymacji daje najlepsze wyniki w punktach, z których pochodzą dane na których metoda bazuje. Jeśli estymacje krigingowe w rozprawie tak fatalnie nie radzą sobie w punktach obserwacyjnych to analizowanie wyników poza punktami obserwacyjnymi (Rozdział 7) nie ma sensu. Zgodność uzyskanych wyników (Rys. 7.2, 7.5, 7.7, 7.8) z rzeczywistością w użytecznym stopniu jest niemożliwa.

W Rozdziale 8 próbuje się poprawić estymację miąższości warstwy pierwszej strefy małych prędkości integrując dane o miąższości z danymi o głębokości zwierciadła wód gruntowych, traktowanej jako zmienna poboczna pomimo małej wartości współczynnika korelacji. I istotnie, zgodność estymat z obserwacjami (Rys. 8.12) jest znacznie lepsza niż przedstawiona uprzednio w pierwszej kolumnie i drugim wierszu Rys. 6.9. Ale wbrew temu co konkluduje się w rozprawie to nie integracja danych była główną przyczyną poprawy modelowania. W tej części pracy obszar badań musiał zostać zmniejszony gdyż dane hydrologiczne pochodziły z mniejszego obszaru. Zmniejszony obszar badań pokazany jest na Rys. 8.3-5. Zmniejszenie spowodowało wykluczenie z analizy zdecydowaną większość danych z pradoliny oraz ponad połowę danych z wysoczyzny zebranych od 2007 roku. W ten sposób dane zostały znacząco zhomogenizowane. Być może również ich rozkład przybliżył się do rozkładu normalnego. Niestety nie ma histogramów pozwalających to chociaż jakościowo sprawdzić.

Przy próbie uzasadnienia dlaczego nastąpiła poprawa estymacji, nie zwrócono uwagi na to, że dzięki szczęśliwej, przypadkowej, bo niezamierzonej homogenizacji danych wynikającej z konieczności usunięcia części z nich, wyniki krigingu bez integracji w tej części pracy (Rys. 8.12 z lewej strony) są dużo lepsze niż poprzednie, przedstawione na Rys. 6.9 w Rozdziale 6 (także bez integracji). Punkty na Rys. 8.12 są rozmieszczone znacznie bardziej liniowo, chociaż także nie wzdłuż linii nachylonej 45 stopni co pokazuje obciążenie estymacji. Zawierając podanym wartościom liczbowym, współczynnik korelacji między estymatami bez integracji danych zwiększył się o 0.23 (z 0.64 na 0.87). Integracja danych zwiększyła go dalej o 0.12 (na 0.99) i usunęła obciążenie wynikające z dryftu. Gdyby zwrócono uwagę na te zmiany bez integracji, być może zorientowano by się o katastrofalnych skutkach łącznego analizowania rozdzielonych klastrów danych.

Do badania jakości estymacji w funkcji zmniejszanej liczby danych (Rozdział 9) wzięto jeszcze bardziej zhomogenizowane dane, a mianowicie dane z wysoczyzny zebrane w latach

1976-1981. Niestety nie zobrazowano ani danych wejściowych ani nie przeprowadzono analizy wybranych modeli. Jeśli jednak estymaty o korelacji z obserwacjami na poziomie od 0.54, 0.53, 0.35. 0.10 (Tab. 9.2, współczynniki determinacji od 30% do 1%) uważa się za dobre, to już nie dziwi, że zmniejszenie tej korelacji z 0.53 na 0.23 (prędkość w warstwie pierwszej) określa się jako „predykcja wartości ... w wyniku rozrzedzenia nie uległa znacznemu pogorszeniu”.

Końcowa ocena rozprawy

Odnosząc się do wymagań stawianych rozprawie doktorskiej w Ustawie z dnia 14 marca 2003 r. „o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki”, art. 13 pkt 1 stwierdzam, że:

1. recenzowana rozprawa zawiera kardynalne błędy i nie przedstawia żadnego poprawnego rozwiązania podjętych problemów;
2. rozprawa nie wykazała, by Doktorantka posiadała umiejętność samodzielnego prowadzenia pracy naukowej.

W związku z powyższym stwierdzam, że przedstawiona do recenzji rozprawa mgr inż. Justyny Sowińskiej-Botor nie spełnia wymagań stawianych rozprawom doktorskim w/w Ustawie i wnioskuję o niedopuszczenie Doktorantki do dalszego postępowania w przewodzie doktorskim.

Kraków, 12/09/2023

