

|   |           |
|---|-----------|
| SEKRETARIAT NAUKOWY<br>INSTYTUT GEOFIZYKI PAN |           |
| WPŁYNEŁO                                      |           |
| la ..... 04.12.2014                           |           |
| Nr dz. ....                                   | zał. .... |
| Ref. ....                                     |           |

dr hab. Wojciech Jakubowski  
Katedra Matematyki  
Uniwersytet Przyrodniczy we Wrocławiu  
ul. Grunwaldzka 53, 50-357 Wrocław  
e-mail: wojciech.jakubowski@upwr.edu.pl

Wrocław, 3 grudnia 2017 r.

## Recenzja rozprawy doktorskiej

mgr inż. Sisay'a Eshetu Debele

### „Frequency analysis of extreme river flows: selected methods and their application”

#### 1. Przedmiot recenzji

Przedmiotem recenzji jest rozprawa doktorska pod tytułem „Frequency analysis of extreme river flows: selected methods and their application”. Autorem rozprawy jest mgr inż. Sisay Eshetu Debele, ubiegający się o stopień naukowy doktora, asystent w Zakładzie Hydrologii i Hydrodynamiki w Instytucie Geofizyki Polskiej Akademii Nauk.

Promotorem pracy jest prof. dr hab. inż. Renata J. Romanowicz.

#### 2. Cel pracy, uwagi wstępne

Rozprawa dotyczy istotnego w hydrologii i gospodarce wodnej zagadnienia wyboru modelu probabilistycznego przepływów maksymalnych rocznych.

W recenzowanej pracy rozważane są dwie metody ich wyznaczania:

- w oparciu o analizę częstościową – z każdego roku (sezonu) pobierana jest tylko jedna obserwacja przepływu maksymalnego – FFA (*Flood Frequency Analysis*);
- na podstawie analizy wezbrań – z każdego wezbrania o przepływach przekraczających pewien próg odcięcia wybierany jest najwyższy przepływ – POT (*Peak Over Threshold*).

Przedstawione są cztery zagadnienia związane bezpośrednio lub pośrednio z wyznaczaniem rozkładów prawdopodobieństwa przepływów maksymalnych rocznych. Są nimi:

1. estymacja rozkładu przepływów maksymalnych rocznych w oparciu o zależne od siebie przepływy sezonowe (zastosowanie funkcji kopuli) – rozdział 3;
2. analiza trendów przepływów maksymalnych rocznych wyznaczonych metodą FFA – rozdział 4, (rozszerzona wersja artykułu [2] wydrukowanego w *Acta Geophysica*);
3. zastosowanie pakietu GAMLSS do analizy trendów przepływów maksymalnych rocznych wyznaczonych metodą FFA – rozdział 5, (rozszerzona wersja artykułu [1] wydrukowanego w *Acta Geophysica*);
4. analiza trendów przepływów maksymalnych rocznych wyznaczonych metodą POT – rozdział 6.

Powyższe rozdziały poprzedzone są obszernym wstępem (rozdziały 1, 2) oraz podsumowaniem (rozdział 7). Do rozprawy dołączone zostały trzy dodatki (A, B, C).

Każdy z rozdziałów został przedstawiony jako osobne opracowanie składające się ze wstępu, przedstawienia problemu i metodyki badań, jeśli to było możliwe – przykładowej aplikacji modelu, końcowych wniosków oraz spisu literatury. Taki układ pracy po części został wymuszony poprzez dołączenie do rozprawy rozszerzonej wersji publikowanych w *Acta Geophysica* artykułów.

### 3. Charakterystyka rozprawy

Recenzowana praca jest podzielona na siedem rozdziałów, łącznie 135 stron. Poniżej scharakteryzowano poszczególne rozdziały rozprawy.

1. W pierwszym rozdziale oprócz ogólnego wstępu jest sformułowany zasadniczy cel pracy. Autor pisze, że w dysertacji przedstawione zostały wybrane metody estymacji rozkładów przepływów maksymalnych rocznych. Analizuje je przy założeniu ich stacjonarności lub ich niestacjonarności. Takie podejście spowodowało, że zasadnicza część dysertacji dzieli się na dwie rozłączne części (rozdział trzeci – przepływy stacjonarne; rozdziały od czwartego do szóstego włącznie – przepływy niestacjonarne).
2. Rozdział drugi zawiera przegląd literatury opisujący stosowane w praktyce hydrologicznej standardowe rozkłady prawdopodobieństwa oraz metody ich estymacji. Niestety rozdział ten nie jest napisany porządnie. Autor
  - definiując rozkład GEV myli jego funkcję gęstości z dystrybuantą;
  - definiując rozkład Gumbela nie podaje wzoru na współczynnik skośności;
  - przedstawiając współczynnik skośności rozkładu GEV nie zaznacza, że współczynnik rozkładu  $k$  musi być mniejszy niż  $\frac{1}{3}$ ;
  - definiując rozkłady GG, GIG oraz RGE pomylił odnośnik do literatury. Powinien był zacytować pracę [3], tym bardziej, że niektóre definicje są stamtąd bezpośrednio cytowane.
3. W rozdziale 3, w pierwszej zasadniczej części rozprawy, autor zaproponował aplikację dwuwymiarowych kopuł do estymacji łącznego rozkładu maksymalnych przepływów sezonowych. Zastosowano kopuły typu Gumbel-Hougaard'a, Franka oraz gaussowską do estymacji łącznych sezonowych przepływów maksymalnych w dorzeczu Sanu. Opracowywano pary przepływów maksymalnych w układzie zima – lato. Postępowanie takie ma sens jedynie w przypadku gdy przepływy sezonowe są od siebie zależne. Przeprowadzono zatem rangowe testy niezależności oraz test  $t$  badający istotność współczynnika korelacji Pearsona. Do dalszej analizy wybrano tylko te stacje wodowskazowe w których hipotezę o niezależności należało odrzucić. W zestawie badanych 20 stacji aż w 14 nie stwierdzono zależności pomiędzy przepływami maksymalnymi zimowymi, a następującymi po nich przepływami maksymalnymi letnimi. (Czy badano drugi z układów: lato – zima? Przy dominancie maksymalnych przepływów letnich taki wybór kolejności danych mógłby dać lepsze rezultaty.) Estymację parametru kopuł przeprowadzono trzema metodami, następnie przeprowadzono testy ich zgodności z danymi empirycznymi. Wyniki przedstawiono w postaci odpowiednich tabel i wykresów. W rozprawie podczas analizy zależności pomiędzy przepływami maksymalnymi zimowymi oraz następującymi po nich przepływami maksymalnymi letnimi zastosowano test  $t$  badający istotność współczynnika korelacji. Jest to poważny błąd, wyraźnie obniżającym wartość przedstawionej dysertacji. Test  $t$  wolno stosować tylko wtedy, gdy badane pomiary pochodzą z populacji o rozkładzie normalnym, czy zbliżonym do normalnego. Rozkłady prawdopodobieństwa sezonowych przepływów maksymalnych są asymetryczne i na pewno nie pochodzą z próby o rozkładzie normalnym.
4. Rozdziały 4 i 5 przedstawiają pakiet GAMLSS (Generalized Additive Model for Location, Scale and Shape parameters) oraz jego zastosowanie do estymacji sezonowego modelu przepływów maksymalnych rocznych w ośmiu zlewniach rozsianych po Polsce i dwóch zlewniach norweskich. Poprzez model sezonowy rozumiana jest estymacja rozkładu  $\max(W, S)$  podana równaniem 4.1, gdzie  $W$  oznacza przepływ maksymalny zimowy, zaś  $S$  – przepływ maksymalny letni. W tej części dysertacji (dla badanych stacji wodowskazowych) wyznaczono:

- najlepsze, w sensie kryterium AIC, dopasowane rozkłady prawdopodobieństwa dla sezonowych przepływów maksymalnych przy założonych czterech możliwościach ich liniowej zmienności w czasie;
- estymatory parametrów modelu liniowego (opcja III – średnia i odchylenie standardowe zmieniają się w sposób liniowy);
- estymowany rozkład przepływów maksymalnych rocznych otrzymany z modelu sezonowego.

Na podstawie otrzymanych wyników estymacji, w drugiej części rozdziału czwartego, autor proponuje model rozkładu prawdopodobieństwa, który mógłby określać prognozowane przepływy maksymalne w dwóch stacjach wodowskazowych, a także przedstawia prognozę wody stuletniej dla dwóch posterunków wodowskazowych w Kłodzku i w Fustvatn (Norwegia).

Wyniki przedstawiane w rozdziałach 4 i 5 zostały opublikowane w czasopiśmie *Acta Geophysica* [2, 1] przez troje autorów. Zgodnie z oświadczeniem doktoranta, ocenił on swój udział pracy w obu opracowaniach od 60 do 70%.

5. W rozdziale 6 dysertacji przedstawiono metodologię modelowania niestacjonarnych przepływów maksymalnych rocznych wykorzystującą model Poissona oraz metodę POT. Ilustracją było badanie sezonowych przepływów maksymalnych w dwóch wodowskazach Kłodzko i Fustvatn. Ta część dysertacji jest dość zwięźle napisana, chyba nie wszystkie obliczenia, przejścia logiczne zostały tu dookreślone. Z analizy przedstawionych tabel i wykresów wynika że:

- poziomy odcięcia w sezonie zimowym i letnim znacznie się od siebie różnią (tab. 6.1), jak to zostało uwzględnione w modelu sezonowym;
- niestacjonarność w przedstawianym modelu została uwzględniona poprzez uzmiennienie parametru  $\Lambda$  w modelu Poissona, czy w tej części dysertacji postępowano podobnie, jak w metodzie GAMLSS?

Uogólniony rozkład Pareto (GP) ma następującą własność. Jeśli zmienna losowa  $X$  ma rozkład GP to warunkowa zmienna losowa  $X - b | X > b$  ma także rozkład GP ze zmienionym parametrem  $\alpha$ , a niezmiennym parametrem  $k$ . Progi odcięcia  $60 \frac{\text{m}^3}{\text{s}}$  – zima oraz  $100 \frac{\text{m}^3}{\text{s}}$  – lato są położone bardzo wysoko. Uważam, że należało przyjąć dużo niższe wartości rzędu  $Q_{30\%} - Q_{20\%}$ , a następnie odpowiednio dobierając parametr  $b$  dokonać estymacji metodą największej wiarygodności.

Podsumowujący rozprawę rozdział 7 zawiera 6 ogólnych oraz 30 szczegółowych wniosków wynikających z całej dysertacji. Szczególnie ważnym wydaje się być wniosek 4 o niestacjonarności hydrologicznych procesów wynikających ze zmian klimatu, zmian antropogenicznych w skali lokalnej i globalnej. Ma ona dalekie konsekwencje szczególnie w jakości uzyskiwanych prognoz przepływów maksymalnych sezonowych i rocznych.

#### 4. Uwagi szczegółowe

- Praca jest napisana jasnym i w miarę prostym językiem, jednakże autor ma tendencję do nadużywania akronimów (np. na stronie 9 rozprawy jest ich aż dwadzieścia). Znaczenie części z nich jest oczywiste, jednakże nie wszystkich. Na początku rozprawy przydałby się tabela ze wszystkimi występującymi w rozprawie akronimami.
- Na stronie 20 w podrozdziale „*The GAMLSS fitting algorithms*” jest odwołanie do wzoru 2.20. Wzoru o takim numerze nie ma w rozprawie.
- Tabela 3.5 ze strony 42 jest powtórzeniem części tabeli 3.3 ze strony 30.
- Na stronie 38, przy opisie własności funkcji kopuli popełniono błąd pisząc  $C(1, v) = 1$ , powinno być  $C(1, v) = v$ .



- Skala na wykresach przedstawiających dopasowanie rozkładów do danych empirycznych jest opartą o rozkład Gumbela. Nie rozumiem dlaczego na jednych opis osi poziomej pozostawiono w skali Gumbela (rys. 3.12), a na innych w skali prawdopodobieństwa (rys. 6.3). Uważam, że wszędzie powinna być stosowana skala prawdopodobieństwa.

## 5. Oceny rozprawy

Recenzowana praca dotyczy ważnego zagadnienia wyznaczania, na podstawie niejednorodnych obserwacji przepływów maksymalnych, kwantyli odpowiadającym założonemu prawdopodobieństwu ich przekroczenia. Autor stawia następujące tezy:

1. Dwuwymiarowe rozkłady typu kopuła opisujące łączny rozkład przepływów maksymalnych zimowych i letnich dokładniej niż standardowe metody przedstawiają losowość zjawiska;
2. Zastosowanie metod pakietu GAMLSS do estymacji rozkładu prawdopodobieństwa niejednorodnych przepływów maksymalnych rocznych, w modelu sezonowym, wiarygodniej przedstawia ich losowość, niż podczas wykorzystania modelu ogólnego.
3. Zastosowanie niejednorodnego procesu Poissona pozwala na wyznaczenie trendu zmienności wysokich wartości, uzyskanych metodą POT, przepływów maksymalnych rocznych lub sezonowych.

Autor w dysertacji przedstawił aparat statystyczny umożliwiający potwierdzenie stawianych tez. Otrzymane wyniki stanowią oryginalne rozwiązanie postawionego problemu naukowego. Doktorant również wykazał się wysoką umiejętnością połączenia postawionych na początku rozprawy celów w jedną całość.

Uważam także, że zastosowanie pakietu GAMLSS środowiska R do estymacji rocznych przepływów maksymalnych wymagało od doktoranta bardzo dobrej znajomości wielu dziedzin nauki. Już samo zastosowanie tego pakietu do określenia trendów zmienności przepływów sezonowych maksymalnych, z punktu widzenia zastosowań w polskiej praktyce hydrologicznej, stanowi znaczący wynik. Uważam, że same rozdziały 4 i 5 wystarczałyby na dobrą rozprawę doktorską. W takim przypadku powinna być tylko zwiększona liczba badanych posterunków, szczególnie w regionach górskich i podgórskich, gdzie w ostatnich 20 latach obserwowana jest wysoka liczba gwałtownych letnich powodzi.

## 6. Wnioski końcowe

Stwierdzam, że przedstawiona praca doktorska pt. „Frequency analysis of extreme river flows: selected methods and their application” spełnia wymagania stawiane pracom doktorskim w ustawie o tytule naukowym oraz o stopniach i tytule w zakresie sztuki z dnia 14 marca 2003 roku (Dziennik Ustaw nr 65, poz. 595, z późniejszymi zmianami) i wnoszę o dopuszczenie jej do publicznej obrony.



## Literatura

- [1] S.E. Debele, E. Bogdanowicz, and W.G. Strupczewski. Around and about an application of the gamlss package to non-stationary flood frequency analysis. *Acta Geophys.*, 65:885–892, 2017.
- [2] S.E. Debele, W.G. Strupczewski, and E. Bogdanowicz. A comparison of three approaches to non-stationary flood frequency analysis. *Acta Geophys.*, 65:863–883, 2017.
- [3] B. Rigby, M. Stasinopoulos, G Heller, and V. Voudouris. *The Distribution Toolbox of GAMLSS*, 2014.