

**Maurycy Ciupak**

**Instytut Meteorologii i Gospodarki Wodnej**

---

**Krzysztof Rokiciński**

**Akademia Marynarki Wojennej**

**ZASTOSOWANIE METODY KOPULI  
DO DWUWYMIAROWEJ ANALIZY  
WEZBRAŃ SZTORMOWYCH  
W PROFILACH WODOWSKAZOWYCH  
ŚWINOUJŚCIE I KOŁOBRZEG**

**STRESZCZENIE**

W artykule zastosowano metodę kopuli do dwuwymiarowej analizy wezbrań sztormowych. Wezbrania sztormowe scharakteryzowano dwuwymiarową zmienną losową: maksymalnym poziomem morza ponad napelnienie w profilach wodowskazowych Świnoujście  $PM_{ŚWI}$  i Kołobrzeg  $PM_{KOL}$ . Próba losowa została wyznaczona na podstawie zidentyfikowanych wezbrań sztormowych wzdłuż południowych wybrzeży Bałtyku w latach 1976–2000. Celem artykułu było znalezienie najlepszego dwuwymiarowego rozkładu prawdopodobieństwa badanej zmiennej losowej ( $PM_{ŚWI}$ ,  $PM_{KOL}$ ). Badaniu poddano eliptyczną kopulę gaussowską i archimedesowską jednoparametrową Gumela-Hougaard. Do badania zgodności kopuli z zaobserwowanymi realizacjami zmiennej losowej ( $PM_{ŚWI}$ ,  $PM_{KOL}$ ) użyto współczynnika korelacji rang Spearmana. Najlepszą zgodność uzyskano dla kopuli Gumbela-Hougaard  $\rho_s = 0,9871$ .

Słowa kluczowe:

metoda kopuli, wezbrania sztormowe, profile wodowskazowe, Świnoujście, Kołobrzeg.

**WSTĘP**

Najczęstszymi przyczynami okresowych zmian poziomu morza są pływy morskie oraz wezbrania sztormowe. W wielu morzach śródładowych o małych masach

wody pływy mają znaczenie marginalne. Wynika to ze znikomego wpływu grawitacyjnego przyciągania Księżyca i reakcji siły odśrodkowej mas wody wywołanej ruchem okrężnym Ziemi wokół środka mas Ziemi i Księżyca na spiętrzanie wody i w konsekwencji na jej przepływ. Przykładem morza śródlądowego o nieistotnym znaczeniu zjawiska pływów jest Morze Bałtyckie. W obszarze południowego Bałtyku, gdzie głębokości nie przekraczają 100 m, poziom wody podczas przyprływu podnosi się zaledwie o kilka centymetrów.

Podstawową przyczyną dużych zmian poziomu morza w takiej sytuacji jest oddziaływanie niskiego ciśnienia atmosferycznego oraz wiatru. Ryzyko pojawienia się ekstremalnych poziomów morza wzdłuż południowych wybrzeży Bałtyku zwiększa wysokie napelnienie Bałtyku w okresie poprzedzającym wezbranie sztormowe.

Powodzie sztormowe związane są z przemieszczającymi się sztormowymi układami barycznymi z zachodu na wschód. Największe wezbrania sztormowe na południowym wybrzeżu Bałtyku związane są z przemieszczaniem się układu niskiego ciśnienia wraz z towarzyszącymi frontami atmosferycznymi od Morza Norweskiego, przez Skandynawię i Morze Bałtyckie. W miarę przemieszczania się niżu na południe i południowy-wschód, Bałtyk wraz z południowym wybrzeżem po krótkim okresie wiatrów odlądowych przedostaje się w sektor sztormowych wiatrów z kierunków północno-zachodnich do północno-wschodnich. Tego typu zdarzenie może trwać od kilku do kilkudziesięciu godzin.

Różne tory przemieszczania się układów niskiego ciśnienia oraz różna ekspozycja linii brzegowej poszczególnych akwenów Bałtyku wywołują wzrosty poziomu morza w wielu częściach wybrzeża. W takiej sytuacji wezbranie sztormowe powinno zostać scharakteryzowane więcej niż jedną zmienną losową. W artykule wezbranie sztormowe scharakteryzowano dwuwymiarową zmienną losową: maksymalnym poziomem morza ponad napelnienie w Świnoujściu  $PM_{\text{SWI}}$  i w Kołobrzegu  $PM_{\text{KOL}}$ .

Celem artykułu jest znalezienie najlepszego dwuwymiarowego rozkładu prawdopodobieństwa zmiennej losowej ( $PM_{\text{SWI}}$ ,  $PM_{\text{KOL}}$ ). Funkcję gęstości prawdopodobieństwa dwuwymiarowej zmiennej losowej zbudowano z zastosowaniem kopuli (*copula*). W badaniu użyto eliptyczną kopulę gaussowską i archimedesowską jednoparametrową Gumbela-Hougaard<sup>1</sup>.

---

<sup>1</sup> Zob. E. J. Gumbel, *Distributions des valeurs extremes en plusieurs dimensions*, Publ. Inst. Statist. Univ., Paris 1960, No 9, pp. 171–173; P. Hougaard, *A class of multivariate failure time distributions*, 'Biometrika', 1986, No 73, pp. 671–678; S. Śmiech, *Analiza zależności ekstremalnych*, „Zeszyty Naukowe” Akademii Ekonomicznej, Kraków 2006, nr 726, s. 91–101.

## MATERIAŁY I METODY

Do dwuwymiarowej analizy wezbrań sztormowych przygotowano ciągi pomiarowe maksymalnych poziomów morza zaobserwowanych na posterunkach pomiarowych w Świnoujściu i Kołobrzegu (rys. 1.) w latach 1976–2000. Wejściowe dane do opracowania uzyskano z katalogu wezbrań i obniżeń sztormowych poziomów morza oraz ekstremalnych poziomów wód na polskim wybrzeżu<sup>2</sup> oraz z monografii Sztobryn i Stigge (2005). Jej autorzy przyjęli definicję wezbrania sztormowego (wg Hydrological Aspects of Combined Effects of Storm Surges and Heavy Rainfall on River Flow, WMO, No 704, Geneva 1988) jako gwałtowny wzrost poziomu morza powyżej poziomu, który byłby zaobserwowany w tym samym miejscu i czasie, gdyby nie wystąpiły silne wiatry dolądowe. Do identyfikacji wezbrań sztormowych przyjęto wartość progową poziomu morza<sup>3</sup> odpowiadającą poziomowi 570 cm N.N. (Normal Null), tj. w odniesieniu do poziomu morza wyznaczonego dla Amsterdamu.



Rys. 1. Polska część południowego wybrzeża Bałtyku, porty morskie oznaczone są symbolem kotwicy

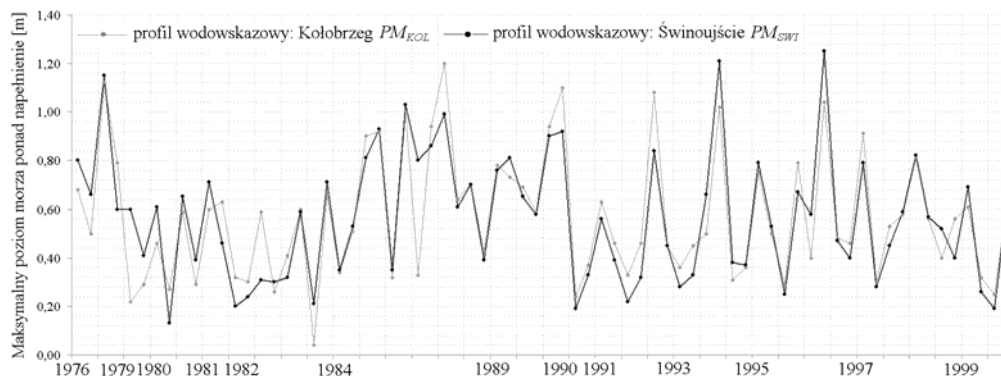
Źródło: opracowanie własne.

<sup>2</sup> B. Wiśniewski, T. Wolski, *Katalogi wezbrań i obniżeń sztormowych poziomów morza oraz ekstremalne poziomy wód na polskim wybrzeżu*, Wyd. Akademii Morskiej, Szczecin 2009, s. 156.

<sup>3</sup> A. Majewski, Z. Dziadziuszko, A. Wiśniewska, *Monografia powodzi sztormowych 1951–1975. Ogólna charakterystyka powodzi sztormowych u polskiego wybrzeża Bałtyku*, WKiŁ, Warszawa 1983.

Za początek wezbrania sztormowego przyjęto moment, w którym poziom morza przekracza wartość dziennego poziomu odniesienia, tzw. napełnienia<sup>4</sup>. Powyższy parametr określa hipotetyczny poziom morza, który zostałby odnotowany, gdyby nad morzem nie występowały czynniki meteorologiczne mające wpływ na spiętrzanie wody.

W artykule analizie częstotliwości poddawana jest dwuwymiarowa zmienna losowa (rys. 2.), którą tworzy maksymalny poziom morza ponad napełnienie w Świnoujściu  $PM_{\text{ŚWI}}$  [m] i w Kołobrzegu  $PM_{\text{KOL}}$  [m] zanotowany przy danym zdarzeniu sztormowym.



Rys. 2. Maksymalny poziom morza ponad napełnienie zaobserwowany na posterunkach pomiarowych Świnoujście i Kołobrzeg w latach 1976–2000

Źródło: opracowanie własne.

### Ogólna postać kopuli dla rozkładu $m$ -wymiarowego

Założmy, że dysponujemy  $m$  zmiennymi losowymi:  $X_1, X_2, \dots, X_m$  z funkcjami rozkładów odpowiednio:  $F_1(X_1), F_2(X_2), \dots, F_m(X_m)$ . Celem jest otrzymanie funkcji  $m$ -wymiarowego rozkładu  $F(X_1, X_2, \dots, X_m)$ . Sklar (1959) udowodnił, że istnieje funkcja  $C$  (kopula) taka, że:

$$F(X_1 = x_1, X_2 = x_2, \dots, X_m = x_m) = C(F_1(x_1), F_2(x_2), \dots, F_m(x_m)) \quad (1)$$

jest dystrybuantą łączną posiadającą rozkłady brzegowe  $F_1(X_1), F_2(X_2), \dots, F_m(X_m)$ .

Dla  $m$ -wymiarowego przypadku dla każdego  $u \in [0, 1]^m$ ,  $C(u) = 0$ , jeżeli przynajmniej jeden element  $u$  jest równy 0. Natomiast  $C(u) = u_j$  dla  $j = 1, 2, \dots, m$ , jeżeli

<sup>4</sup> *Wezbrania sztormowe wzdłuż południowego Bałtyku (zachodnia i środkowa część)*, red. M. Sztobryn, H.-J. Stigge, IMGW, Warszawa 2005, s. 9.

wszystkie elementy  $u$  są równe 1, z wyjątkiem  $u_j$ . Funkcja łącząca  $C$  powinna być  $m$ -niemalejąca, czyli niemalejąca we wszystkich  $m$  argumentach.

Rolą funkcji łączącej  $C$  jest określenie zależności pomiędzy zmiennymi  $X_1, X_2, \dots, X_m$ . W przypadku rozkładów brzegowych ciągłych zmiennych losowych, różniczkując równanie (1), otrzymujemy  $m$ -wymiarową gęstość:

$$f(X_1 = x_1, X_2 = x_2, X_m = x_m) = c(F_1(x_1), F_2(x_2), \dots, F_m(x_m))f_1(x_1)f_2(x_2) \dots f_m(x_m), \quad (2)$$

gdzie:

wyrażenie  $c(F_1(x_1), F_2(x_2), \dots, F_m(x_m))$  nazywane jest gęstością kopuli, której wzór otrzymujemy, obliczając  $m$ -tą pochodną cząstkową kopuli względem  $m$  zmiennych:

$$C(F_1(x_1), \dots, F_m(x_m)) = \frac{\partial^m C}{\partial u_1 \partial u_2 \dots \partial u_m} \text{ w } u_1 = F_1(x_1), \dots, u_m = F_m(x_m). \quad (3)$$

O wyborze kopuli jako metody estymacji parametrów wielowymiarowych rozkładów prawdopodobieństwa zadecydowały następujące jej własności:

- kopula pozwala łączyć dowolne rozkłady brzegowe;
- teoria kopuli umożliwia budowanie statystyk wielowymiarowych, dysponując statystykami brzegowymi, zwłaszcza wtedy kiedy rozważane zmienne losowe są od siebie statystycznie zależne.

#### **Postać eliptycznej kopuli gaussowskiej dwuwymiarowej zmiennej losowej ( $PM_{\dot{S}WI}$ , $PM_{KOL}$ )**

Kopulę gaussowską definiuje się<sup>5</sup> za pomocą całki podwójnej:

$$C_\rho^G(v, z) = \Phi_\rho(\Phi^{-1}(v), \Phi^{-1}(z)), \quad (4)$$

gdzie:  $\Phi_\rho(\Phi^{-1}(v), \Phi^{-1}(z)) = \int_{-\infty}^{\Phi^{-1}(v)} \int_{-\infty}^{\Phi^{-1}(z)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(\frac{2\rho su - s^2 - u^2}{2(1-\rho^2)}\right) ds du$ ,

$v = F_1(PM_{\dot{S}WI})$  i  $z = F_2(PM_{KOL})$ , zmienne losowe  $PM_{\dot{S}WI}$  i  $PM_{KOL}$  mogą podlegać dowolnym rozkładom prawdopodobieństwa.

<sup>5</sup> Zob. U. Cherubini, E. Luciano, W. Vecchiato, *Copula Method in Finance*, John Wiley & Sons, Ltd, 2004; Garcia A., Gençay R., *Managing Adverse Dependence for Portfolios of Collateral In Financial Infrastructures*, Bank of Canada, 2007, s. 25.

### Postać archimedesowskiej kopuli Gumbela-Hougaard dla rozkładu dwuwymiarowej zmiennej losowej ( $PM_{SWI}$ , $PM_{KOL}$ )

Kopula Gumbela-Hougaard<sup>6</sup> służąca do wyznaczania ekstremalnych wartości zmiennej losowej w pewnym okresie czasu wyrażona jest wzorem:

$$C_{\theta}^{GH}(v, z) = \exp\left\{-\left[(-\log v)^{\theta} + (-\log z)^{\theta}\right]^{\frac{1}{\theta}}\right\}, \quad (5)$$

gdzie:  $\theta \in [1; +\infty)$ .

### WYNIKI BADAŃ

Wejściem do analizy była dwuwymiarowa zmienna losowa ( $PM_{SWI}$ ,  $PM_{KOL}$ ). W badaniu estymację parametrów kopuli wykonano w dwóch etapach. W pierwszym etapie wykonano estymację parametrów brzegowych rozkładów prawdopodobieństwa metodą największej wiarygodności. W drugim etapie, w przypadku eliptycznej kopuli gaussowskiej, estymację współczynnika korelacji  $\rho$  wykonano metodą symulacji bayesowskiej<sup>7</sup>, natomiast parametry archimedesowskiej kopuli Gumbela-Hougaard wykonano metodą IFM (*Inference Functions for the Margins*)<sup>8</sup>.

W opracowaniu do obliczenia miary zależności między zaobserwowanymi realizacjami zmiennej losowej ( $PM_{SWI}$ ,  $PM_{KOL}$ ) zastosowano współczynnik korelacji liniowej oraz korelację nieparametryczną w postaci współczynnika rang Spearmana.

Do obliczenia miary zależności między wygenerowanymi realizacjami zmiennej losowej ( $PM_{SWI}$ ,  $PM_{KOL}$ ) zastosowano nieparametryczny współczynnik korelacji rang Spearmana  $\rho_s$ , który można wyrazić za pomocą kopuli<sup>9</sup> w następujący sposób:

$$\rho_s \stackrel{def.}{=} 12 \int_0^1 \int_0^1 C(v, z) dv dz - 3. \quad (6)$$

### Estymacja parametrów rozkładów brzegowych metodą największej wiarygodności (MNW)

W pierwszym etapie procedury zestawiono próbę losową  $Z = \{PM_{SWI}, PM_{KOL}\}$  o liczebności  $n = 71$ . Dla każdej zmiennej losowej określono miary rozkładu: asymetrii, koncentracji, rozproszenia i położenia (tabela 1.).

<sup>6</sup> Zob. E. J. Gumbel, wyd. cyt., pp. 171–173; P. Hougaard, wyd. cyt., pp. 671–678; tenże, *Analysis of Multivariate Survival Data*, Springer Verlag, New York — Berlin — Heidelberg 2000.

<sup>7</sup> P. J. Danaher, M. S. Smith, *Modeling Multivariate Distribution Using Copulas: Applications in Marketing*, Melbourne Business School, University of Melbourne, 2009, s. 47.

<sup>8</sup> U. Cherubini, E. Luciano, W. Vecchiato, wyd. cyt.

<sup>9</sup> P. J. Danaher, M. S. Smith, wyd. cyt.

Tabela 1. Miary rozkładu zmiennej losowej: maksymalnego poziomu morza ponad napełnienie w Świnoujściu  $PM_{\text{SWI}}$  i w Kołobrzegu  $PM_{\text{KOL}}$  w latach 1976–2000

	$PM_{\text{SWI}}$ [m]	$PM_{\text{KOL}}$ [m]
Miara asymetrii rozkładu		
Współczynnik skośności	0,509	0,591
Miara koncentracji rozkładu		
Współczynnik skupienia (kurtoza)	-0,322	-0,334
Miary rozproszenia rozkładu		
Odchylenie standardowe	0,212 [m]	0,258 [m]
Wariancja	0,067 [m]	0,066 [m]
Miary położenia rozkładu		
Średnia arytmetyczna	0,565 [m]	0,568 [m]
Mediana	0,575 [m]	0,520 [m]

Źródło: opracowanie własne.

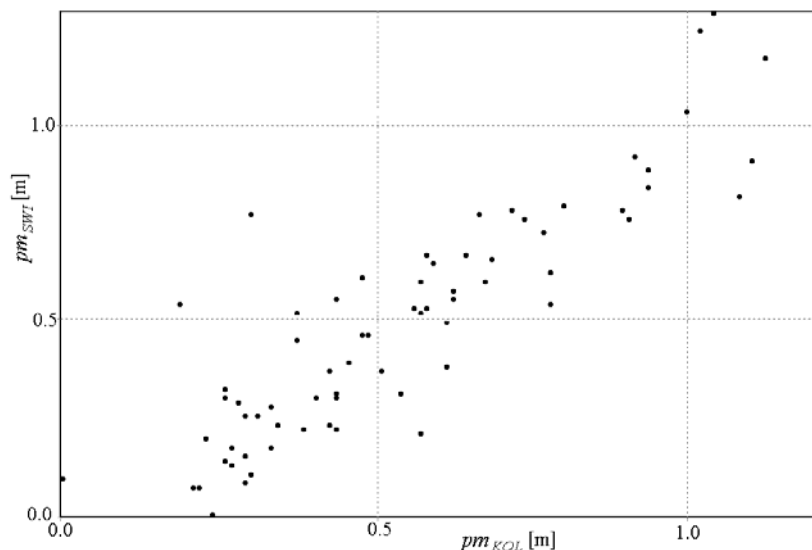
Ujemna wartość kurtozy obu zmiennych świadczy o spłaszczeniu rozkładu i tym samym o mniejszej koncentracji wartości cechy. Obie analizowane zmienne mają skośność dodatnią, co wskazuje na niewielką prawostronną asymetrię (wydłużone prawe ramie rozkładu).

Zbadano zależność między analizowanymi zmiennymi losowymi. Wartość współczynnika korelacji liniowej  $\rho_p = 0,878$ , natomiast współczynnik korelacji rang Spearmana  $\rho_s = 0,866$ . Diagram korelacyjny zaobserwowanych realizacji zmiennej losowej ( $PM_{\text{SWI}}$ ,  $PM_{\text{KOL}}$ ) w okresie od 1976 do 2000 roku przedstawiono na rysunku 3.

Współczynnik korelacji rang Spearmana jest w niewielkim stopniu wrażliwy na obserwacje odstające. W związku z tym wykonano badanie jednorodności obu zmiennych losowych<sup>10</sup>. W badaniu jednorodności zakłada się, że nie jest znany rozkład badanej zmiennej losowej. Do weryfikacji charakterystyk statystycznych najwygodniejszymi i jednocześnie najprostszymi w użyciu są nieparametryczne testy statystyczne. Najczęstszymi charakterystykami statystycznymi ciągów pomiarowych ulegającymi zmianom są wartość średnia i wariancja. Zmiany tych wartości mają charakter skokowy lub występują w postaci zmiennych trendów. Najskuteczniejszymi testami w wykrywaniu zaburzeń zjawisk hydrologicznych są test sumy rang Kruskala-Wallisa oraz test współczynnika Spearmana korelacji rangowej na trend wartości średniej i trend wariancji zmiennej losowej. Ciągi pomiarowe zostały sprawdzone pod kątem występowania w nich elementów odstających przy użyciu testu Grubbsa-Becka.

<sup>10</sup> B. Ozga-Zieliński, *Metody analizy niejednorodności ciągów pomiarowych zjawisk hydrologicznych*, „Wiadomości IMGW”, 1999, z. 2, s. 3.

Na podstawie testów statystycznych przyjęto, że ciągi pomiarowe zmien-  
nych losowych  $PM_{SWI}$  i  $PM_{KOL}$  są statystycznie jednorodne.



Rys. 3. Diagram korelacyjny zaobserwowanych realizacji zmiennej losowej ( $PM_{SWI}$ ,  $PM_{KOL}$ ):  
maksymalnego poziomu morza powyżej poziomu napełnienia  
w profilach wodowskazowych: Świnoujście i Kołobrzeg w latach 1976–2000

Źródło: opracowanie własne.

Estymację parametrów rozpoczęto od doboru odpowiednich teoretycznych  
funkcji rozkładu prawdopodobieństwa  $f_1(PM_{SWI}, g_1, \dots, g_k)$  i  $f_2(PM_{KOL}, g_1, \dots, g_k)$  ma-  
jących stanowić matematyczny model własności losowych zmiennych  $PM_{SWI}$  i  $PM_{KOL}$   
oddzielnie dla każdego rozkładu brzegowego<sup>11</sup>.

Rozpatrywane były następujące rozkłady prawdopodobieństwa: trójparame-  
trowy logarytmiczno-normalny, dwuparametrowy gamma, Weibulla, Gumbela maksi-  
mum, Gumbela minimum, logistyczny, Cauchy'ego, Laplace'a i jednoparametrowa  
funkcja błędów<sup>12</sup>, której funkcja gęstości wyrażona jest wzorem:

$$f(x) = h\pi^{-2} e^{-(hx)^2}, \quad (7)$$

gdzie:

$h$  — parametr skali  $h > 0$ ;

$e$  — podstawa logarytmu naturalnego.

<sup>11</sup> M. Ozga-Zielińska, J. Brzeziński, *Hydrologia stosowana*, PWN, Warszawa, 1994, 1997.

<sup>12</sup> M. Abramowitz, I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, New York 1972, s. 297–309.



Do estymacji parametrów rozkładów prawdopodobieństwa zastosowano metodę największej wiarygodności. Dobrano odpowiednią teoretyczną funkcję gęstości rozkładu prawdopodobieństwa  $f(Z, g_1, \dots, g_k)$  z parametrami  $g_1, \dots, g_k$ , gdzie  $k = 2$  — liczba parametrów. Rozpatrywano trójparametrowy rozkład logarytmiczno-normalny z dolnym ograniczeniem rozkładu prawdopodobieństwa zmiennej losowej równym zero. Wynikiem zastosowanej metody estymacji nieznanego parametru  $g_j$  było znalezienie odpowiedniej funkcji zależnej od elementów próby losowej  $\hat{g}_j = g_j(Z_1, \dots, Z_n)$  dla  $j = 1, \dots, k$  i  $n = 71$  — liczebność próby losowej.

Sprawdzenie zgodności przyjętego rozkładu teoretycznego z rozkładem empirycznym wykonano przez zastosowanie testu zgodności  $\lambda$  Kołmogorowa z obniżoną wartością krytyczną<sup>13</sup> oraz testu zgodności  $\chi^2$  Pearsona przy poziomie istotności  $\alpha = 0,05$ . Wyniki sprawdzenia zawarto w tabeli 2. Najlepszą zgodność przyjętego rozkładu teoretycznego z rozkładem empirycznym dla obu zmiennych losowych  $PM_{\dot{S}WI}$  i  $PM_{KOL}$  uzyskano za pomocą funkcji rozkładu Gumbela maksimum.

Tabela 2. Wyniki testów zgodności empirycznych rozkładów z rozkładami teoretycznymi dla zmiennej losowej maksymalnego poziomu morza ponad napelnienie w Świnoujściu  $PM_{\dot{S}WI}$  i w Kołobrzegu  $PM_{KOL}$  w latach 1976–2000

Teoretyczne rozkłady prawdopodobieństwa	Testy zgodności			
	$\lambda$ Kołmogorowa		$\chi^2$ Pearsona	
	$PM_{\dot{S}WI}$ [m]	$PM_{KOL}$ [m]	$PM_{\dot{S}WI}$ [m]	$PM_{KOL}$ [m]
log-normalny (3P) dla $\varepsilon = 0$	1,515	0,766	7,789	9,273
Gamma (2P)	0,918	1,604	7,763	16,038
Weibull (2P)	0,987	0,828	7,321	8,894
Gumbel maksimum (2P)	<b>0,748</b>	<b>0,723</b>	<b>7,205</b>	<b>8,887</b>
Gumbel minimum (2P)	1,305	1,382	115,399	90,816
logistyczny (2P)	1,108	1,009	11,478	21,560
Laplace (2P)	1,429	1,366	22,829	32,678
Cauchy (2P)	8,303	8,368	-----	-----
funkcja błędu (1P)	6,222	6,572	4037,695	6090,522

Źródło: opracowanie własne.

Na podstawie powyższych wyników do dalszej analizy wezbrań sztormowych przyjęto zmienne losowe:  $PM_{\dot{S}WI}$  i  $PM_{KOL}$  podlegające dwuparametrowemu rozkładowi Gumbela maksimum. Uzyskane wartości parametrów rozkładów brzegowych dla obu zmiennych losowych przedstawiono w tabeli 3.

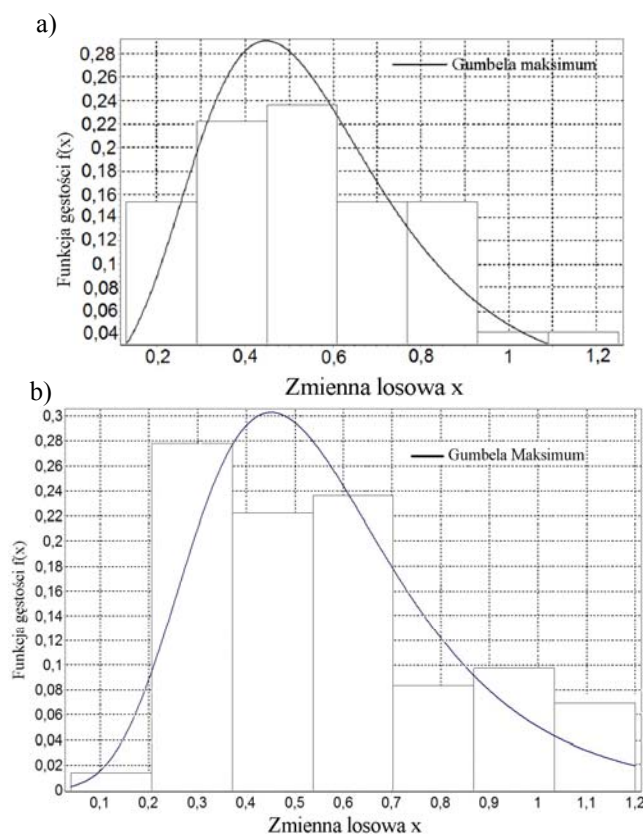
<sup>13</sup> S. Węglarczyk, *Test zgodności testowania rozkładów z parametrami estymowanymi*, „Przegląd Geofizyczny”, 1993, z. 3–4.

Tabela 3. Parametry brzegowych rozkładów teoretycznych estymowanych metodą największej wiarygodności (MNW)

Parametry rozkładów brzegowych	
$PM_{\text{ŚWI}}$ — rozkład Gumbela maksimum	$PM_{\text{KOL}}$ — rozkład Gumbela maksimum
$\mu = 0,4481$ — parametr położenia [m]	$\mu = 0,4514$ [m]
$\beta = 0,2024$ — parametr skali [m]	$\beta = 0,2014$ [m]

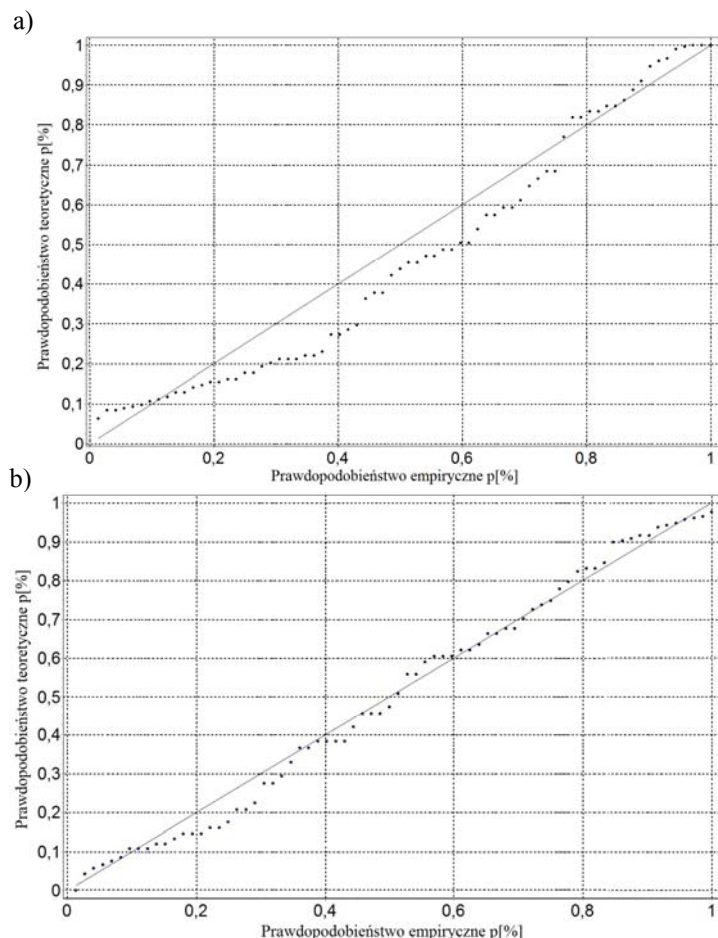
Źródło: opracowanie własne.

Na rysunkach 4. i 5. przedstawiono proste metody graficzne służące do weryfikacji hipotezy o zgodności empirycznego rozkładu prawdopodobieństwa z badanym rozkładem hipotetycznym.



Rys. 4. Histogramy z dopasowanymi funkcjami gęstości prawdopodobieństwa maksymalnego poziomu morza ponad poziom napełnienia w Świnoujściu  $PM_{\text{ŚWI}}$  (a) i w Kołobrzegu  $PM_{\text{KOL}}$  (b) w latach 1976–2000

Źródło: opracowanie własne.



Rys. 5. Wykresy prawdopodobieństwo-prawdopodobieństwo ( $P-P$ ) porównujące zgodność rozkładu badanej zmiennej (kropki):  $PM_{SWI}$  (a) i  $PM_{KOL}$  (b) z rozkładem teoretycznym Gumbela maksimum (linia prosta)

Źródło: opracowanie własne.

Rysunek 4. prezentuje wyniki badań w postaci histogramów liczebności zmiennej losowej:  $PM_{SWI}$  (a) i  $PM_{KOL}$  (b) z zaznaczoną funkcją gęstości rozkładu Gumbela maksimum. Na rysunku 5. przedstawiono wykresy prawdopodobieństwo-prawdopodobieństwo ( $P-P$ ) badanych zmiennych losowych. Prawdziwość postawionej hipotezy ocenia ułożenie punktów wzdłuż prostej. W przypadku zmiennej losowej  $PM_{KOL}$  (b) punkty układają się dość wyraźnie wzdłuż prostej, a dla zmiennej  $PM_{SWI}$  (a) znacznie gorzej.

**Estymacja parametrów kopuli oraz generowanie realizacji zmiennej losowej  $(v, z) = (F_1(p_{m_{SWI}}), F_2(p_{m_{KOL}}))$**

Estymacja eliptycznej kopuli gaussowskiej

W przypadku dwuwymiarowej eliptycznej kopuli gaussowskiej należy estymować jeden parametr — współczynnik korelacji  $\rho$ . Estymację wartości parametru eliptycznej kopuli wykonano metodą bayesowską z zastosowaniem symulacji komputerowej, stosując metodę Monte Carlo<sup>14</sup> opartą na łańcuchu Markowa (*Markov chain Monte Carlo*, MCMC). Wykonano następującą procedurę obliczeniową:

1. Podwójna transformacja zmiennych losowych:

$PM_{SWI}$  i  $PM_{KOL}$  wg wzoru  $x_{1,i}^* = \Phi^{-1}(F_1(p_{m,i}))$  i  $x_{2,i}^* = \Phi^{-1}(F_2(t_{w,i}))$  dla  $i = 1, \dots, n$ , gdzie  $F_1$  i  $F_2$  są funkcjami brzegowych rozkładów prawdopodobieństwa, a  $\Phi^{-1}$  jest funkcją odwrotnego rozkładu normalnego.

Wynikiem procedury jest zbiór realizacji zmiennych losowych  $x^* = \{x_{1,i}^*, x_{2,i}^*\}$  dla  $i = 1, \dots, n$ , wygenerowanych za pomocą eliptycznej kopuli transformowanych do  $R^2 = R^+ \times R^+$ . Zbiór zaobserwowanych realizacji zmiennych losowych  $PM_{SWI}$  i  $PM_{KOL}$  oznaczono  $x = \{p_{m,i}, t_{w,i}\}$  dla  $i = 1, \dots, n$  o liczebności  $n = 71$ .

2. Estymacja współczynnika korelacji  $r$  uwarunkowanego zbiorami  $x^*$  i  $x^{15}$  za pomocą metody Metropolisa-Hastingsa<sup>16</sup>.

Rozkład Choleskiego macierzy  $\Sigma$  o wymiarach  $2 \times 2$  według wzoru  $\Sigma = (R^T R)^{-1}$ , gdzie  $R^T$  jest transponowaną macierzą trójkątną górną o wymiarach  $2 \times 2$ , której elementami jest estymowany współczynnik korelacji  $r$  otrzymany w drugim kroku procedury.

Obliczenie macierzy korelacji  $\Gamma$  o wymiarach  $2 \times 2$   $\Gamma = \text{diag}(\Sigma)^{-\frac{1}{2}} \Sigma \text{diag}(\Sigma)^{-\frac{1}{2}}$ , gdzie  $\text{diag}(\Sigma)$  jest macierzą diagonalną o elementach będących wartościami własnymi  $\Sigma$ .

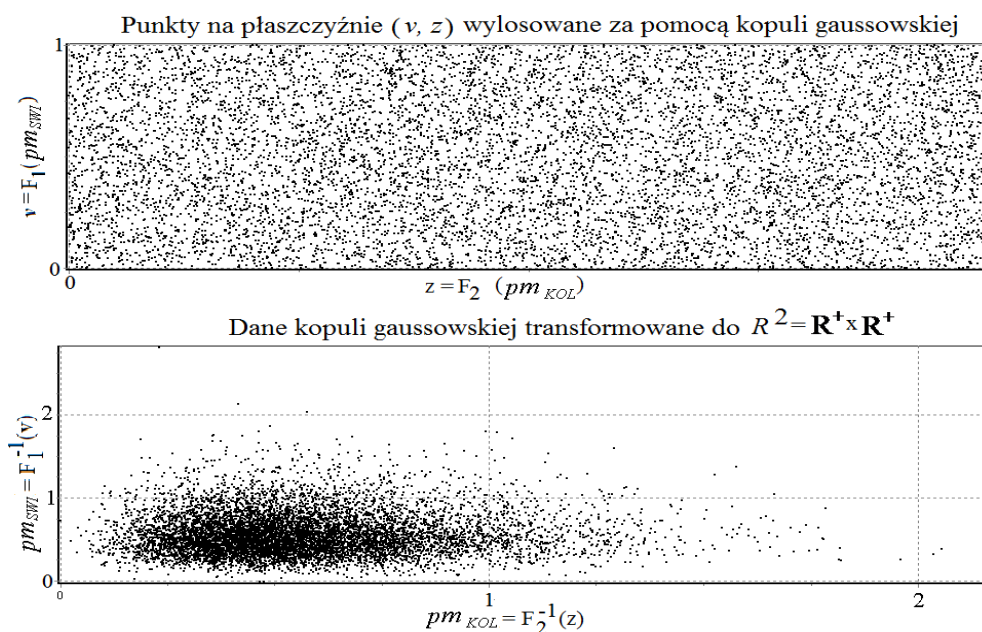
<sup>14</sup> P. J. Danaher, M. S. Smith, wyd. cyt., s. 47.

<sup>15</sup> L. Dalla Valle, *Bayesian Copulae Distributions with Application to Operational Risk Management*, 'Methodology and Computing in Applied Probability', 2009, Vol. 11, No 1, pp. 95–115; P. J. Danaher, M. S. Smith, wyd. cyt., s. 47.

<sup>16</sup> S. Chib, E. Greenberg, *Understanding the Metropolis-Hastings Algorithm*, 'The American Statistician', 1995, Vol. 49, No 4, pp. 327–335.

Po wykonaniu powyższych przekształceń, wartość współczynnika korelacji  $\rho$  otrzymywana jest na podstawie wzoru  $\rho = \frac{-r}{r^2 + 1}$ .

3. Losowanie realizacji zmiennych losowych o normalnym rozkładzie prawdopodobieństwa  $\mathbf{Z} \sim N(0, \rho)$ .
4. Generowanie realizacji dwuwymiarowej zmiennej losowej za pomocą kopuli gaussowskiej według wzoru  $\mathbf{U} = (\Phi(Z_1), \Phi(Z_2))$  (górną część rys. 6.).
5. Transformacja zmiennych losowych  $\mathbf{U}$  otrzymanych za pomocą kopuli gaussowskiej do przestrzeni  $R^2 = R^+ \times R^+$ , czyli  $\mathbf{X} = (F_1^{-1}(U_1), F_2^{-1}(U_2))$  (dolną część rys. 6.).



Rys. 6. Górny rysunek — punkty na płaszczyźnie  $(v, z)$  wylosowane za pomocą kopuli gaussowskiej; dolny rysunek — dane kopuli gaussowskiej transformowane do  $R^2 = R^+ \times R^+$  za pomocą odwrotnego rozkładu brzegowego

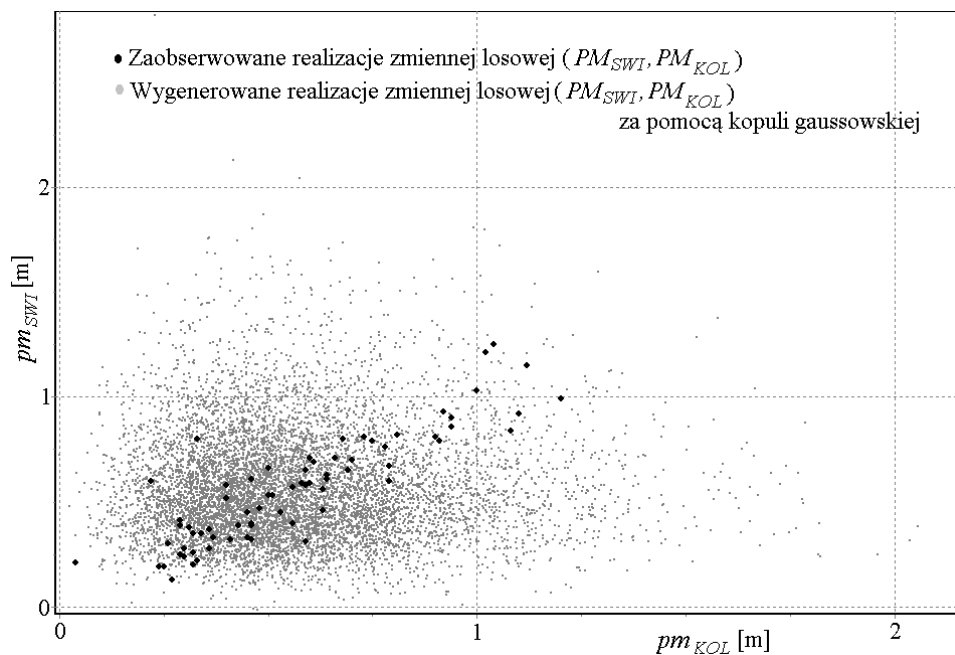
Źródło: opracowanie własne.

W przypadku gdy brzegowe rozkłady prawdopodobieństwa są rozkładami normalnymi, kopuła gaussowska generuje łączny standardowy rozkład normalny (równomierne zagęszczenie punktów), w przeciwnym wypadku wylosowane za pomocą kopuli gaussowskiej punkty na płaszczyźnie  $(v, z)$  nie są równomiernie rozłożone (górną część rys. 6.).

Wykonano  $k = 10000$  powtórzeń opisanej procedury. Uzyskano następujący zbiór zmiennych losowych  $\{(P^{(1)}, \mathbf{X}^{*(1)}, \mathbf{U}^{(1)}, \mathbf{X}^{(1)}), \dots, (P^{(k)}, \mathbf{X}^{*(k)}, \mathbf{U}^{(k)}, \mathbf{X}^{(k)})\}$ . Estymatorem wartości oczekiwanej wykorzystującym realizację łańcucha generowanego przez algorytm M-H jest średnia wartość współczynnika korelacji  $\rho$  obliczona według wzoru

$$\bar{\rho} = \frac{1}{k} \sum_{i=1}^k \rho^{(i)}. \quad (8)$$

Na rysunku 7. przedstawiono diagram korelacyjny zaobserwowanych i wygenerowanych realizacji zmiennej losowej  $(PM_{SWI}, PM_{KOL})$  za pomocą kopuli gaussowskiej.



Rys. 7. Diagram korelacyjny zaobserwowanych i wygenerowanych realizacji zmiennej losowej  $(PM_{SWI}, PM_{KOL})$ . za pomocą kopuli gaussowskiej

Źródło: opracowanie własne.

Estymacja archimedesowskiej kopuli Gumbela-Hougaard

Estymację parametrów archimedesowskiej kopuli Gumbela-Hougaard wykonano metodą IFM (*Inference Functions for the Margins*)<sup>17</sup> złożoną z dwóch etapów.

<sup>17</sup> U. Cherubini, E. Luciano, W. Vecchiato, wyd. cyt.

W pierwszym etapie procedury oszacowano parametry brzegowych rozkładów prawdopodobieństwa  $\theta_j$ :

$$\hat{\theta}_j = \arg \max_{\theta_j} \sum_{i=1}^n \log f_j(x_{ji}; \theta_j) \text{ dla } j = 1, 2, \quad (9)$$

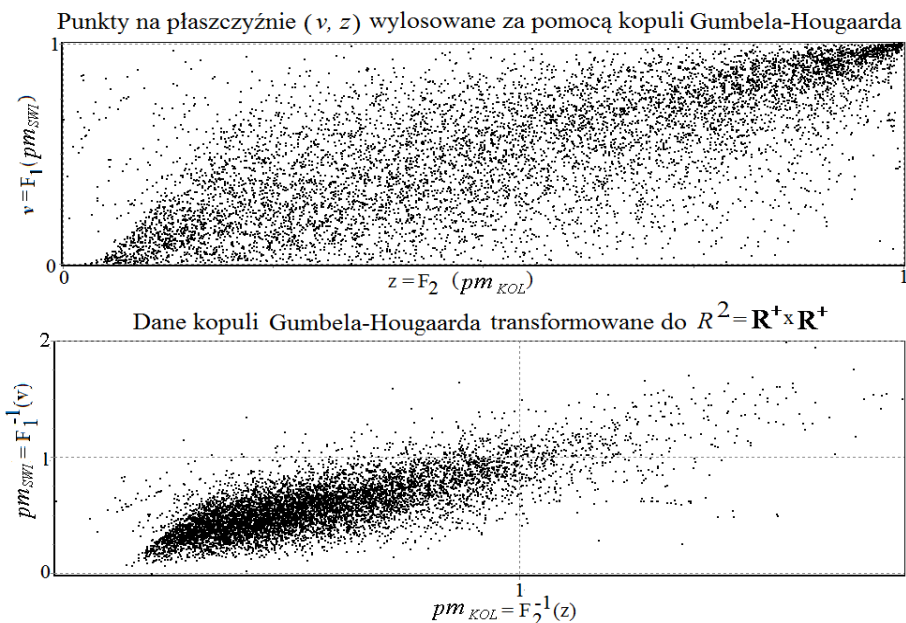
gdzie:

$n = 71$  — liczebność próby losowej.

W drugim etapie, na podstawie zbioru estymowanych parametrów brzegowych rozkładów prawdopodobieństwa  $\hat{\theta}_j$ , szacowane są parametry kopuli  $\Phi$ :

$$\hat{\Phi} = \arg \max_{\Phi} \sum_{i=1}^n \log c(F_1(x_{1i}, \hat{\theta}_1), F_2(x_{2i}, \hat{\theta}_2)). \quad (10)$$

Do generowania realizacji dwuwymiarowej zmiennej losowej ( $PM_{SWI}$ ,  $PM_{KOL}$ ) (rys. 8.) za pomocą kopuli Gumbela-Hougaard użyto jej generatora  $\varphi(u) = (-\log u)^\theta$  z funkcją odwrotną  $\varphi^{-1}(t) = \exp(-t^{\frac{1}{\theta}})$ .

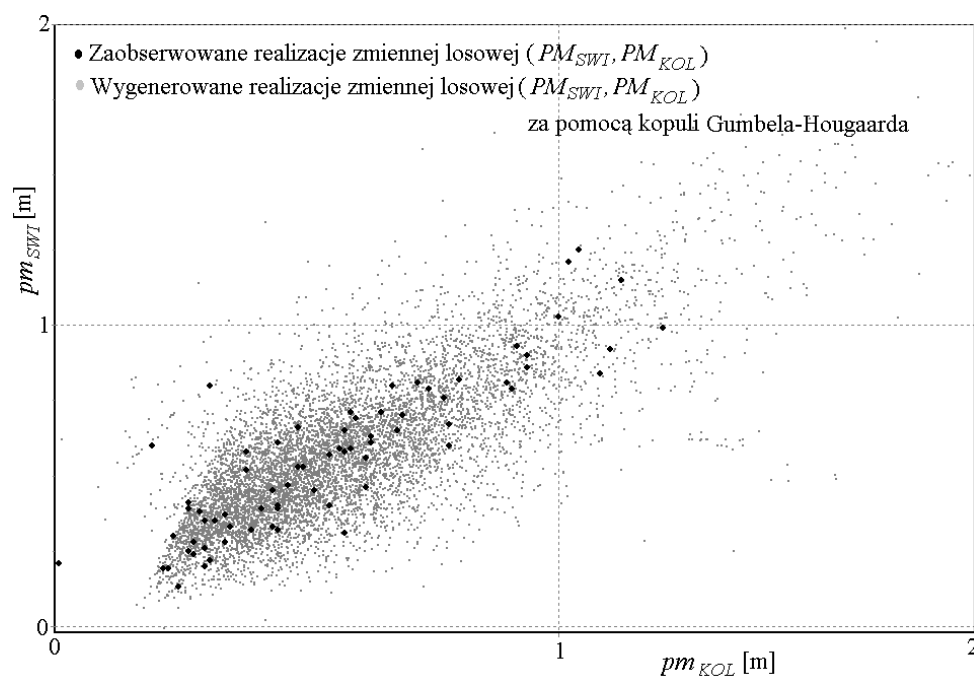


Rys. 8. Górny rysunek — punkty na płaszczyźnie  $(v, z)$  wylosowane za pomocą kopuli Gumbela-Hougaard; dolny rysunek — dane kopuli Gumbela-Hougaard transformowane do  $R^2 = R^+ \times R^+$  za pomocą odwrotnego rozkładu brzegowego

Źródło: opracowanie własne.

Teoria kopuli oraz wykorzystujące ją symulacje Monte Carlo stosowane są do określania stopnia pogrubienia ogonów dystrybuant wielowymiarowych jak też do wzajemnej zależności ogonów dystrybuant brzegowych. Skoncentrowane punkty na płaszczyźnie  $(v, z)$  w obszarze o współrzędnych  $(0, 0)$  i  $(1, 1)$  (górną część rys. 8.) mogą świadczyć o pogrubionym lewym i prawym ogonie dystrybuanty dwuwymiarowej.

Na rysunku 9. przedstawiono diagram korelacyjny zaobserwowanych i wygenerowanych realizacji zmiennej losowej  $(PM_{SWI}, PM_{KOL})$  za pomocą kopuli Gumbela-Hougaard.



Rys. 9. Diagram korelacyjny zaobserwowanych i wygenerowanych realizacji zmiennej losowej  $(PM_{SWI}, PM_{KOL})$  za pomocą kopuli Gumbela-Hougaard

Źródło: opracowanie własne.

W tabeli 4. przedstawiono wartości współczynnika korelacji rang Spearmana  $\rho_s$  oraz parametry estymowanych rozkładów prawdopodobieństwa. Współczynnik  $\rho_s$  przyjmuje wartość zero dla zmiennych niezależnych oraz 1, -1 odpowiednio dla idealnej zależności rosnącej, malejącej. Najlepszą zgodność  $\rho_s = 0,9871$  uzyskano w przypadku kopuli Gumbela-Hougaard, natomiast zdecydowanie najgorszą dla kopuli gaussowskiej  $\rho_s = 0,02$ .



Tabela 4. Wartości uzyskanych parametrów statystycznych

Parametry estymowanych rozkładów		Współczynnik korelacji rang Spearmana $\rho_s$
Kopula gaussowska		0,0200
$\bar{\rho}$	-0,647	
Kopula Gumbela-Hougaard		<b>0,9871</b>
$\theta$	2,252	

Wartość parametru  $\theta$  kopuli Gumbela-Hougaard może zmieniać się w zakresie  $[1, +\infty)$ .  $\theta = 1$  oznacza brak zależności, natomiast uzyskana wartość  $\theta = 2,252$  oznacza zależność rosnącą. W przypadku kopuli gaussowskiej  $\bar{\rho} = -0,647$  wskazuje na zależność malejącą.

### PODSUMOWANIE

1. Wezbrania sztormowe wzdłuż południowych wybrzeży Bałtyku scharakteryzowano dwuwymiarową zmienną losową ( $PM_{\dot{S}WI}$ ,  $PM_{KOL}$ ): maksymalnym poziomem morza ponad napętnienie w profilach wodowskazowych Świnoujście i Kołobrzeg.
2. Funkcję gęstości prawdopodobieństwa dwuwymiarowej zmiennej losowej ( $PM_{\dot{S}WI}$ ,  $PM_{KOL}$ ) zbudowano z zastosowaniem kopuli, gdyż metoda ta pozwala łączyć dowolne rozkłady brzegowe oraz umożliwia budowanie statystyk wielowymiarowych, kiedy rozważane są zmienne losowe statystycznie od siebie zależne.
3. Wartość współczynnika korelacji liniowej między  $PM_{\dot{S}WI}$  i  $PM_{KOL}$   $\rho_p = 0,878$ , natomiast współczynnik korelacji rang Spearmana  $\rho_s = 0,866$ .
4. W badaniu użyto eliptyczną kopulę gaussowską i archimedesowską jednoparametrową Gumbela-Hougaard. Zastosowany algorytm estymacji parametrów obu kopuli złożony był z dwóch etapów.
5. W pierwszym etapie wykonano estymację parametrów brzegowych rozkładów prawdopodobieństwa metodą największej wiarygodności. Na podstawie wyników testów zgodności  $\lambda$  Kołmogorowa i  $\chi^2$  Pearsona ustalono, że najlepszą zgodność rozkładu teoretycznego z rozkładem empirycznym dla obu zmiennych losowych ( $PM_{\dot{S}WI}$  i  $PM_{KOL}$ ) uzyskano za pomocą dwuparametrowego rozkładu Gumbela maksimum.
6. W drugim etapie do estymacji eliptycznej kopuli gaussowskiej zastosowano podejście bayesowskie, natomiast parametry archimedesowskiej kopuli Gumbela-Hougaard estymowano metodą IFM. Do obliczenia miary zależności między

wygenerowanymi realizacjami zmiennej losowej ( $PM_{SWI}$ ,  $PM_{KOL}$ ) zastosowano nieparametryczny współczynnik korelacji rang Spearmana. Uzyskane wyniki pokazują, że zależności między zmiennymi najlepiej uwzględnić kopuła Gumbela-Hougaard  $\rho_s = 0,9871$ , najgorzej kopuła gaussowska  $\rho_s = 0,02$ . Wartość parametru kopuli Gumbela-Hougaard  $\theta = 2,252$  oznacza zależność rosnącą.

7. Należy stwierdzić, że dwuwymiarowa archimedesowska kopuła Gumbela-Hougaard dobrze estymuje dwuwymiarowy rozkład maksymalnego poziomu morza ponad napełnienie w profilach wodowskazowych Świnoujście i Kołobrzeg.
8. W celu zastosowania opisanej w artykule metody kopuli do estymacji dwuwymiarowego rozkładu prawdopodobieństwa zmiennej losowej ( $PM_{SWI}$ ,  $PM_{KOL}$ ) należy opracować metodykę jej interpretacji probabilistycznej.

### BIBLIOGRAFIA

- [1] Abramowitz M., Stegun I. A., *Handbook of Mathematical Functions with Formulas. Graphs and Mathematical Tables*, New York 1972.
- [2] Cherubini U., Luciano E., Vecchiato W., *Copula Method In Finance*, John Wiley & Sons Ltd., 2004.
- [3] Chib S., Greenberg E., *Understanding the Metropolis-Hastings Algorithm*, 'The America Statistician', 1995, Vol. 49, No 4, pp. 327–335.
- [4] Dalla Valle L., *Bayesian Copulae Distributions with Application to Operational Risk Management*, 'Methodology and Computing in Applied Probability', 2009, Vol. 11, No 1, pp. 95–115.
- [5] Danaher P. J., Smith M. S., *Modeling Multivariate Distribution Using Copulas: Applications in Marketing*, Melbourne Business School, University of Melbourn, 2009.
- [6] Garcia A., Gençay R., *Managing Adverse Dependence for Portofolios of Collateral In Financial Infrastructures*, Bank of Canada, 2007, 25.
- [7] Gumbel E. J., *Distributions des valeurs extremes en plusieurs dimensions*, 'Publ. Inst. Statist. Univ.', Paris 1960, 9.
- [8] Hougaard P., *A class of multivariate failure time distributions*, 'Biometrika', 1986, No 73, pp. 671–678.
- [9] Hougaard P., *Analysis of Multivariate Survival Data*, Springer Verlag, New York — Berlin — Heidelberg 2000.

- [10] Majewski A., Dziadziuszko Z., Wiśniewska A., *Monografia powodzi sztormowych 1951–1975. Ogólna charakterystyka powodzi sztormowych u polskiego wybrzeża Bałtyku*, WKiŁ, Warszawa 1983.
- [11] Ozga-Zielińska M., Brzeziński J., *Hydrologia stosowana*, PWN, Warszawa 1994, 1997.
- [12] Ozga-Zieliński B., *Metody analizy niejednorodności ciągów pomiarowych zjawisk hydrologicznych*, „Wiadomości IMGW”, 1999, z. 2.
- [13] Sklar A., *Fonctions de répartition à dimensions et leurs marges*, ‘Publ. Inst. Statist. Univ.’, Paris 1959, 8.
- [14] *Wezbrania sztormowe wzdłuż południowego Bałtyku (zachodnia i środkowa część)*, red. M. Sztobryn, H.-J. Stigge, IMGW, Warszawa 2005.
- [15] Śmiech S., *Analiza zależności ekstremalnych*, „Zeszyty Naukowe” Akademii Ekonomicznej, Kraków 2006, nr 726.
- [16] Węglarczyk S., *Test zgodności testowania rozkładów z parametrami estymowanymi*, „Przegląd Geofizyczny”, 1993, z. 3–4.
- [17] Wiśniewski B., Wolski T., *Katalogi wzebrań i obniżeń sztormowych poziomów morza oraz ekstremalne poziomy wód na polskim wybrzeżu*, Wyd. Akademii Morskiej, Szczecin 2009.

## **APPLICATION OF COPULA METHOD IN TWODIMENSIONAL STORM SURGES ANALYSIS AT ŚWINOUJŚCIE AND AT KOŁOBRZEG**

### **ABSTRACT**

The copula method was applied to analyze 2D storm surges. Storm surges were described with 2D variable: the maximum water level above the filling level at Świnoujście  $PM_{SWI}$  and at Kołobrzeg  $PM_{KOL}$ . The identified storm surges along of southern Baltic coasts in the period 1976–2000 were used to set the random sample. The aim of this paper is to find the best fit 2D probability distribution of the variable  $(PM_{SWI}, PM_{KOL})$ . Two copula functions: elliptical Gaussian and Archimedean Gumbel-Hougaard copula were investigated. To verify copula conformity with

the data observed the non-parametric dependence measure such as Spearman's  $\rho_s$  was used. The best conformity was recorded for Gumbel-Hougaard copula  $\rho_s = 0,9871$ .

Keywords:

storm surge, the maximum water level above the filling level, bayesian method, Southern Baltic, Gaussian copula, Gumbel-Hougaard copula.