

KRYSTIAN SZCZĘSNY

<https://doi.org/10.33995/wu2023.4.7>

data wpływu: 18.05.2023

data akceptacji: 16.01.2024

Głębokie sieci neuronowe w identyfikacji rozkładów brzegowych i wielowymiarowej kopuli w kontekście agregacji ryzyka w *Solvency II*

Jednym z podstawowych aspektów dyrektywy Solvency II jest ochrona ubezpieczonych przed niewypłacalnością zakładów ubezpieczeń. W tym celu poprzez agregację kapitałowych wymogów wypłacalności dla poszczególnych rodzajów ryzyka wyznaczany jest kapitałowy wymóg wypłacalności (SCR) oraz efekt dywersyfikacji (DE). SCR można wyznaczyć za pomocą formuły standardowej (SF), podanej przez autorów dyrektywy w oparciu o metodę wariancji-kowariancji (V-C), lub za pomocą modeli wewnętrznych, opracowywanych niezależnie (przez zakłady). W niniejszym artykule przedstawiamy propozycję metody, która może być wykorzystana podczas opracowywania modeli wewnętrznych. Istotnym etapem tego procesu jest ustalanie najprecyzyjniejszej zależności pomiędzy czynnikami ryzyka. Proponujemy wykorzystanie kopuli estymowanych za pomocą sieci neuronowych w modelowaniu zależności w podmodule ryzyka składki i rezerw. Konstruujemy dwie sieci neuronowe: pierwszą do rozkładów brzegowych, a drugą do estymacji kopuli. W badaniu analizujemy wskaźniki dla segmentów ubezpieczycieli majątkowych, wyznaczone na podstawie raportów o wypłacalności i kondycji finansowej polskich zakładów ubezpieczeń. Porównano DE uzyskane metodą V-C, metodą kaskad kopuli i estymacją kopuli przy użyciu sieci neuronowych. Przeprowadzone badania wskazują na znaczne różnice w wartościach DE uzyskanych dla kopuli oszacowanej z wykorzystaniem sieci neuronowych, kopuli parametrycznej oraz podejścia zaproponowanego przez autorów dyrektywy w oparciu o metodę V-C. Uzyskane wyniki mogą być wykorzystane w modelach wewnętrznych.

Słowa kluczowe: *Solvency II*, kapitałowy wymóg wypłacalności, efekt dywersyfikacji, kaskady kopuli, kopule neuronowe.

Publikacja prezentuje wyniki badań naukowych przeprowadzonych w ramach projektu nr 88/SD/2023/PRO finansowanego ze środków subwencji przyznanej Uniwersytetowi Ekonomicznemu w Krakowie.

Wprowadzenie

W styczniu 2016 roku weszła w życie dyrektywa *Solvency II*, będąca największą zmianą regulacyjną unijnego rynku ubezpieczeń. Dyrektywa ma strukturę trzech filarów: filar I określa wymogi kapitałowe uwzględniające wszystkie mierzalne rodzaje ryzyka działalności zakładu ubezpieczeń; filar II reguluje mechanizmy wewnętrznego zarządzania zakładem ubezpieczeń oraz standardowych procedur nadzoru; filar III reguluje obowiązek zachowania transparentności rynku i ujawniania sytuacji finansowej zakładu ubezpieczeń. W badaniach referowanych w niniejszym artykule skupiliśmy się na filarze I.

Jednym z najistotniejszych celów *Solvency II* jest ochrona właścicieli polis. Dlatego zgodnie z dyrektywą każdy zakład ubezpieczeń zobowiązany jest do spełnienia dwóch wymogów wypłacalności: kapitałowego wymogu wypłacalności (SCR – *Solvency Capital Requirement*) oraz minimalnego wymogu wypłacalności (MCR – *Minimal Capital Requirement*)¹. Z punktu widzenia nadzoru poziom MCR odzwierciedla poziom kapitału, poniżej którego nadzór może wprowadzić radykalne działania. SCR natomiast reprezentuje poziom kapitału umożliwiający danej instytucji pokrycie znaczących, nieprzewidzianych strat, co zapewnia wypłacalność zobowiązań i daje gwarancje właścicielom polis. Firmy, które są w stanie pokryć SCR, zajmują silną pozycję i powinny mieć swobodę zarządzania własnym biznesem bez działań nadzorczych wykraczających poza rutynowy proces nadzoru.

W artykule koncentrujemy się na procesie ustalania SCR, który zgodnie z dyrektywą *Solvency II* bazuje na agregacji czynników ryzyka, na jakie jest narażony zakład ubezpieczeń. Twórcy dyrektywy podają formułę standardową (SF – *Standard Formula*), zgodnie z którą do agregacji ryzyka wykorzystywana jest metoda wariancji-kowariancji, a macierz korelacji między czynnikami ryzyka jest z góry ustalona i podana w dyrektywie. W piątym badaniu ilościowym (QIS5 – *Fifth Quantitative Impact Study*), przeprowadzonym przez Europejski Urząd Nadzoru Ubezpieczeń i Pracowniczych Programów Emerytalnych (EIOPA – *European Insurance and Occupational Pensions Authority*), europejskie zakłady ubezpieczeń zgłaszały, że wartości korelacji między czynnikami ryzyka w ich zakładach są inne niż te dane w dyrektywie. Z niedoskonałości SF zdają sobie sprawę sami jej twórcy, zachęcając zakłady ubezpieczeń do budowania modeli wewnętrznych, które w lepszym stopniu oddadzą profil działalności ubezpieczyciela. Doświadczenie praktyków zarówno w stosowaniu SF, jak i w budowaniu modeli wewnętrznych, skłoniły EIOPA do rozpoczęcia w październiku 2020 roku ogólnoeuropejskiego badania porównawczego dotyczącego dywersyfikacji w modelach wewnętrznych. Jednym z celów badania było lepsze zrozumienie związku między sposobami modelowania zależności i agregacją ryzyka a wynikającymi z nich korzyściami z dywersyfikacji. Tematyka niniejszego artykułu wpisuje się w najnowsze badania zainicjowane przez EIOPA.

Analiza wpływu informacji o strukturze zależności pomiędzy różnymi rodzajami ryzyka na SCR już od jakiegoś czasu jest poruszana w literaturze. Krystian Szczęsny analizuje wrażliwość SCR na zmiany

1. Art. 129 Dyrektywy *Solvency II*.

wartości w macierzy korelacji². Lluís Bermúdez i współautorzy stwierdzają, że modyfikacje założeń dotyczących korelacji i zależności mają istotny wpływ na szacowanie SCR³. Stefan Mittnik⁴ oraz Arturo Cifuentes i Ventura Charlin⁵ zwracają uwagę, iż korelacja liniowa nie jest dobrym narzędziem do prawidłowego opisu zależności między rozkładami innymi niż rozkłady normalne, gdyż nie jest uwzględniana zależność nieliniowa i w ogonach. Ponieważ ubezpieczyciele potrzebują elastycznych narzędzi, które opisują rzeczywistą strukturę zależności, zastosowanie kopuli w tym kontekście staje się coraz bardziej popularne. Kopule to wielowymiarowe dystrybuanty o jednostajnych rozkładach brzegowych. Dzięki twierdzeniu Sklara, które mówi o związku kopuli z rozkładami wielowymiarowymi, rozkład ten jesteśmy w stanie przedstawić za pomocą wielowymiarowej funkcji kopuli oraz rozkładów brzegowych agregowanych czynników ryzyka. Martin Eling i Kwangmin Jung⁶ (2020) oraz Krystian Szczęsny⁷ modelują strukturę zależności za pomocą wprowadzonych przez Tima Bedforda i Rogera M. Cooke'a⁸ kaskad kopuli. Wielowymiarowy rozkład dekomponowany jest na parametryczne warunkowe kopule dwuwymiarowe. W praktyce często zakłada się, że dwuwymiarowe kopule zależą od zmiennych warunkujących tylko pośrednio, poprzez warunkowe rozkłady brzegowe. Jest to założenie upraszczające. Elif F. Acar i współautorzy⁹ oraz Fabian Spanhel i Malte S. Kurz¹⁰ wskazują, iż założenie to może prowadzić do błędnych wyników. Dodatkowo wybór kopuli dwuwymiarowej na danym poziomie dekompozycji zależy od wyboru kopuli na poprzednich poziomach. Ingrid H. Haff (2013)¹¹ zwraca uwagę, że taka strategia wyboru kopuli wyraźnie kumuluje niepewność i model musi być dokładnie oceniony. Powyższe przykłady potwierdzają potrzebę zastosowania bardziej elastycznych narzędzi, dzięki którym będziemy w stanie zidentyfikować kopulę opisującą właściwą strukturę zależności.

W ostatnich latach można zauważyć silne zainteresowanie technikami uczenia maszynowego (ML – *Machine Learning*) w kontekście modelowania zależności. Przykładowe zastosowanie głębokich sieci

2. K. Szczęsny, *Wpływ błędnej specyfikacji struktury zależności w procesie agregacji ryzyka na efekt dywersyfikacji w Solvency II*, [w:] *Problemy gospodarcze i społeczne Polski i Europy*, [red.] E. Sojka, J. Acedański, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice 2022, s. 98–112.
3. L. Bermúdez, A. Ferri, M. Guillén, *A correlation sensitivity analysis of non-life underwriting risk in solvency capital requirement estimation*, "ASTIN Bulletin" 29.04.2013, s. 21–37; doi: 10.1017/asb.2012.1.
4. S. Mittnik, *Solvency II Calibrations : Where Curiosity Meets Spuriousity*, "Preprints" July 2016; doi: 10.20944/preprints201607.0021.v1.
5. A. Cifuentes, V. Charlin, *Operational risk and the Solvency II capital aggregation formula: Implications of the hidden correlation assumptions*, "Journal of Operational Risk" 2016, No. 11(4), s. 23–33; doi: 10.21314/JOP.2016.181.
6. M. Eling, K. Jung, *Risk aggregation in non-life insurance: Standard models vs. internal models*, "Insurance: Mathematics and Economics" 2020, No. 95, s. 183–198; doi: 10.1016/j.insmatheco.2020.09.003.
7. K. Szczęsny, *Wykorzystanie kaskad kopuli w agregacji ryzyka w procesie wyznaczania kapitałowych wymogów wypłacalności w Solvency II*, [w:] *Sektor ubezpieczeń w obliczu wyzwań współczesności*, [red.] M. Lemkowska, M. Wojtkowiak, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań 2022, s. 98–117; <https://doi.org/10.18559/978-83-8211-131-6/7>.
8. T. Bedford, R.M. Cooke, *Vines – A new graphical model for dependent random variables*, "Annals of Statistics" 2002, No. 30(4), s. 1031–1068; doi: 10.1214/aos/1031689016.
9. E.F. Acar, C. Genest, J. Nešlehová, *Beyond simplified pair-copula constructions*, "Journal of Multivariate Analysis" 2012, No. 110, s. 74–90; doi: 10.1016/j.jmva.2012.02.001.
10. F. Spanhel, M.S. Kurz, *Simplified vine copula models: Approximations based on the simplifying assumption*, "Electronic Journal of Statistics" 2019, No. 13(1), s. 1254–1291; doi: 10.1214/19-EJS1547.
11. I.H. Haff, *Parameter estimation for pair-copula constructions*, "Bernoulli" 2013, No. 19(2), s. 462–491; doi: 10.3150/12-BEJ413.

neuronowych (DNN – *Deep Neural Networks*) do modelowania zależności odnajdujemy w uczeniu się funkcji generatora kopuli Archimedeseskiej¹² czy też przy ustalaniu najlepszej reprezentacji modelu kaskady kopuli¹³. Algorytmy ML stosowane są w wielu obszarach działalności ubezpieczeniowej: do wykrywania oszustw ubezpieczeniowych¹⁴ czy modelowania szkód komunikacyjnych¹⁵; do przewidywania wypłacalności^{16,17} oraz do wyznaczenia rozkładu prawdopodobieństwa wartości aktywów netto¹⁸.

Niniejsza praca jest kontynuacją badań¹⁹, w których do identyfikacji struktury zależności, niezbędnej w celu wyznaczenia SCR, zastosowaliśmy kaskady kopuli. W aktualnych badaniach idziemy o krok dalej, wykorzystując sieci neuronowe najpierw do identyfikacji rozkładów brzegowych, a następnie do identyfikacji kopuli. Zhi Zeng i Ting Wang²⁰ proponują wykorzystanie sieci neuronowych do wyznaczenia rozkładu wielowymiarowego. W tym celu konstruuje się dwie sieci: jedną do dopasowania rozkładów brzegowych, a drugą do wyznaczenia kopuli. Podejście to autorzy stosują do wyznaczenia rozkładu dwuwymiarowego. W naszych badaniach uogólniamy tę metodę do wyznaczenia rozkładu czterowymiarowego. Według naszej najlepszej wiedzy w zakresie wyznaczenia SCR takie podejście nie było przedstawiane w dotychczasowej literaturze. Badania prowadzimy dla danych rzeczywistych pozyskanych z raportów o wypłacalności i kondycji finansowej (SFCRs – *Solvency and Financial Condition Reports*) polskich ubezpieczycieli majątkowych. Modelujemy zależność w oparciu o C-kaskady kopuli oraz kopulę dopasowaną do danych rzeczywistych z wykorzystaniem sieci neuronowych. Dokładność dopasowanych modeli do danych rzeczywistych oceniamy za pomocą odległości energetycznej²¹. Ostatecznie porównujemy efekty dywersyfikacji dla kapitałowego wymogu wypłacalności dla ryzyka składki i rezerw w ubezpieczeniach innych niż na życie wyznaczone z wykorzystaniem trzech różnych metod modelowania zależności. Pierwszy sposób zaprezentowany w niniejszym artykule to sposób ustalania DE, bazujący na zastosowaniu zalecanej w dyrektywie SF, w której zależności określone są macierzą korelacji. Drugi bazuje na metodzie opisanej przez Krystiana Szczęsnego, który strukturę zależności wyznacza za pomocą

12. C.K. Ling, F. Fang, J.Z. Kolter, *Deep Archimedean Copulas*, [w:] *Advances in Neural Information Processing Systems*, [ed.] H. Larochelle et al., Curran Associates Inc., New York 2020, s. 1535–1545; <https://proceedings.neurips.cc/paper/2020/file/10eb6500bd1e4a3704818012a1593cc3-Paper.pdf> [dostęp: 8.02.2024].
13. Y. Sun, A. Cuesta-Infante, K. Veeramachaneni, *Learning vine copula models for synthetic data generation*, 33rd AAAI Conference on Artificial Intelligence, AAAI 2019, 31st Innovative Applications of Artificial Intelligence Conference, IAAI 2019 and the 9th AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, s. 5049–5057; doi: 10.1609/aaai.v33i01.33015049.
14. A.K.I. Hassan, A. Abraham, *Modeling insurance fraud detection using imbalanced data classification*, “Advances in Intelligent Systems and Computing” 2016, No. 419, s. 117–127; doi: 10.1007/978-3-319-27400-3_11.
15. Z.M. Yunos et al. *Predictive modelling for motor insurance claims using artificial neural networks*, “International Journal of Advances in Soft Computing and its Applications” 2016, No. 8(3), s. 160–172.
16. A. Ibiwoye, O.O. Ajibola, A. Sogunro, *Artificial Neural Network Model for Predicting Insurance Insolvency*, “International Journal of Management and Business Research” 2012, No. 2(1), s. 59–68.
17. G. Castellani et al., *Accuracy and Robustness of Machine Learning Methods in SCR Estimation. Nested Monte Carlo Simulation*, Insurance Data Science Conference June 2019, s. 1–7.
18. M. Lintvelt, *Least-Squares Monte Carlo Within Solvency II*, Amsterdam School of Economics, Amsterdam 2016.
19. K. Szczęśny, *Wykorzystanie kaskad kopuli...*
20. Z. Zeng, T. Wang, *Neural Copula: A unified framework for estimating generic high-dimensional Copula functions*, Research Gate 2022; doi: <https://doi.org/10.48550/arXiv.2205.15031>.
21. T. Gneiting, A.E. Raftery, *Strictly proper scoring rules, prediction, and estimation*, “Journal of the American Statistical Association” 2007, No. 102(477), s. 359–378; doi: 10.1198/016214506000001437.

kaskad kopuli²². Trzeci sposób, który jest nowością w niniejszych badaniach, polega na ustaleniu DE w oparciu o nieparametryczną metodę identyfikacji zależności, bazującą na sieci neuronowej.

W pierwszym rozdziale omawiamy standardową procedurę szacowania wymogu kapitałowego dla podmodułu ryzyka składki i rezerw w ubezpieczeniach innych niż na życie. W rozdziale drugim przedstawiamy zastosowaną metodologię: nakreśliłyśmy ideę oceny efektu dywersyfikacji oraz istotę analizowanego problemu, która leży zarówno w metodzie wyznaczenia wymogu kapitałowego dla poszczególnych segmentów, jak i w poprawnym zidentyfikowaniu struktury zależności między segmentami. Trzeci rozdział to opis strategii empirycznej. Natomiast w rozdziale czwartym prezentujemy opis danych. W ostatnim, piątym rozdziale przedstawiamy wyniki badań empirycznych dla danych z SFCRs oraz konkluzje.

1. Wymóg wypłacalności dla ryzyka składki i rezerw w ubezpieczeniach innych niż na życie

Kapitałowy wymóg wypłacalności jest poziomem dopuszczonych środków własnych, dzięki którym w razie wystąpienia nieprzewidzianych strat zakład ubezpieczeń wywiąże się ze swych zobowiązań i wypłaci ubezpieczonym odszkodowania. Ustalany jest na drodze agregacji wcześniej wyznaczonych wymogów kapitałowych dla wszystkich mierzalnych rodzajów ryzyka, na które narażony jest ubezpieczyciel. Sposób wyznaczania agregowanych wymogów kapitałowych oraz sama procedura agregacji prowadząca do ustalenia SCR dla zakładu ubezpieczeń zostały szczegółowo opisane w Rozporządzeniu Delegowanym Komisji UE²³ (dalej: RDK).

W niniejszym artykule skupiamy się na ryzyku składki i rezerw w ubezpieczeniach innych niż na życie. Początkowo twórcy dyrektywy proponowali, aby wymóg kapitałowy dla tego ryzyka w sposób standardowy był wyznaczany zgodnie z formułą:

$$SCR_{pr,res} = \rho(\sigma) \cdot V, \quad (1)$$

gdzie $\rho(\sigma)$ jest tzw. funkcją odchylenia standardowego związaną z rozkładem zmiennej modelującej ryzyko składki i rezerw, natomiast V oznacza miarę działalności ubezpieczyciela. Zakładając rozkład logarytmiczno-normalny dla zmiennej modelującej ryzyko składki i rezerw, funkcję $\rho(\sigma)$ określono w następujący sposób:

$$\rho(\sigma) = \frac{\exp(N_{0,995} \cdot \sqrt{\ln(\sigma^2 + 1)})}{\sqrt{\sigma^2 + 1}} - 1, \quad (2)$$

gdzie $N_{0,995}$ oznacza kwantyl rzędu 0.995 rozkładu normalnego. Ostatecznie jednak, po konsultacji z praktykami, autorzy dyrektywy zmienili postać tej funkcji, zakładając rozkład normalny. W RDK przyjęto $\rho(\sigma) = 3 \cdot \sigma$, co oznacza, że zatwierdzono uproszczoną następującą formułę na wymóg kapitałowy dla ryzyka składki i rezerw:

22. K. Szczęsny, *Wykorzystanie kaskad kopuli...*

23. Rozporządzenie delegowane Komisji [UE] 2015/35 z dnia 10 października 2014 r. uzupełniające dyrektywę Parlamentu Europejskiego i Rady 2009/138/WE w sprawie podejmowania i prowadzenia działalności ubezpieczeniowej i reasekuracyjnej (Wypłacalność II).

$$SCR_{pr,res} = 3 \cdot \sigma \cdot V. \quad (3)$$

Z powyższego wzoru wynika, że wyznaczenie $SCR_{pr,res}$ wymaga ustalenia miary działalności ubezpieczyciela V oraz odchylenia standardowego σ , mierzącego ryzyko składki i rezerw w ubezpieczeniach innych niż na życie. Wartości V i σ wyznacza się, uwzględniając podział na dwanaście segmentów portfela ubezpieczeń innych niż na życie (tabela 1.). Miara działalności V jest sumą miar działalności segmentów V_i , przy czym V_i jest sumą miary ekspozycji na ryzyko składki $V_{(prem,i)}$ oraz miary ekspozycji na ryzyko rezerw $V_{(res,i)}$, skorygowaną o współczynnik dywersyfikacji geograficznej składki i rezerw w podziale na 18 regionów świata dla każdego segmentu (oznaczany dalej przez DIV_i). Zatem miara działalności V ustalana jest z wykorzystaniem następującego wzoru:

$$V = \sum_{i=1}^{12} V_i = \sum_{i=1}^{12} (V_{(prem,i)} + V_{(res,i)}) \cdot (0,75 + 0,25 \cdot DIV_i) \quad (4)$$

Z kolei odchylenie standardowe σ wyznacza się na drodze agregacji:

$$\sigma = \sqrt{\frac{1}{V^2} \cdot \left(\sum_{i=1}^{12} \sum_{j=1}^{12} R_{i,j} \cdot \sigma_i \cdot \sigma_j \cdot V_i \cdot V_j \right)} \quad (5)$$

gdzie: V – miara działalności ubezpieczyciela,

V_i, V_j – miary działalności w ramach segmentów $i, j = 1, \dots, 12$,

σ_i, σ_j – odchylenia standardowe dla poszczególnych segmentów (RDK, art. 117; Załącznik II),

$R_{i,j}$ – współczynniki korelacji między i -tym i j -tym segmentem.

Wartości korelacji między segmentami zostały ustalone i są podane w Annex IV RDK.

2. Ocena efektu dywersyfikacji – metodyka badań

Wyznaczony w opisanym powyżej sposób wymóg wypłacalności dla ryzyka składki i rezerw w ubezpieczeniach innych niż na życie jest nie większy od sumy kapitałów potrzebnych do zabezpieczenia się przed ryzykiem każdego segmentu z osobna. Otrzymana różnica nazywana jest efektem dywersyfikacji. Wielkość tego efektu najczęściej oceniana jest za pomocą współczynnika dywersyfikacji:

$$DE = 1 - \frac{\kappa(L)}{\sum_{i=1}^d \kappa(L_i)}, \quad (6)$$

gdzie: $\kappa(L_i)$ oznacza wymóg kapitałowy dla segmentu i , a $\kappa(L)$ – wymóg wypłacalności dla ryzyka składki i rezerw. Przy czym L_i jest zmienną losową modelującą ryzyko i -tego segmentu, a L – zmienną losową modelującą zagregowane ryzyko segmentów:

$$L = \psi(L_1, \dots, L_d), \quad (7)$$

czyli ryzyko składki i rezerw w ubezpieczeniach innych niż na życie.

Kluczową kwestią w ocenie efektu dywersyfikacji na właściwym poziomie jest dobór metod zarówno do wyznaczenia wymogu kapitałowego dla poszczególnych segmentów L_i , jak i dla zagregowanej zmiennej ryzyka L , która zależy w głównej mierze od wyboru funkcji agregującej ψ . W formule standardowej (3) zakłada się, że L jest sumą zależnych zmiennych losowych L_i o rozkładach normalnych

z parametrami określonymi za pomocą odchylenia standardowego dla poszczególnych segmentów. Przy czym zależność między L_i opisywana jest macierzą korelacji [stosuje się agregację metodą wariancji-kowariancji]. Zatem przyjmuje się, że zmienna ryzyka L , będąca podstawą wyznaczenia $SCR_{pr,res}$, ma rozkład normalny. Jednak rodzi się zasadnicze pytanie, czy to założenie jest zasadne? Podana przez autorów dyrektywy metoda wariancji-kowariancji z metodycznego punktu widzenia jest poprawna wówczas, gdy agregowane czynniki ryzyka mają wielowymiarowy rozkład normalny. Jak potwierdzają badania przedstawione np. przez Christiana Genesta²⁴ oraz Richarda A. Shawa i Grigory'ego Spivaka²⁵, tylko wówczas struktura zależności między nimi jest opisana macierzą korelacji w sposób jednoznaczny. Stosowanie FS z góry ustalonymi w dyrektywie *Solvency II* wartościami macierzy jest bardzo ograniczające. Heterogeniczność ubezpieczeń tworzących poszczególne segmenty sprawia, że rozkłady zmiennych L_i często należą do różnych rodzin. Zatem zależności między nimi mają złożoną strukturę i nie podlegają wielowymiarowemu rozkładowi normalnemu.

Oszacowanie rozkładu L jest zadaniem na ogół trudnym, gdyż zależy od rozkładu wektora losowego (L_1, \dots, L_d) , czyli wielowymiarowego rozkładu zmiennych L_i . W praktyce najczęściej znane są rozkłady brzegowe L_i lub można je w miarę prosto szacować. Natomiast kluczowe znaczenie odgrywają identyfikacja i opis struktury zależności między zmiennymi L_i . Zastosowanie w tym celu macierzy korelacji, jak ma to miejsce w standardowej formule na $SCR_{pr,res}$, jest niewystarczające, ponieważ w sposób jednoznaczny określają one jedynie wielowymiarowe rozkłady eliptyczne.

Elastycznym narzędziem, za pomocą którego można modelować szerszą klasę struktur zależności między agregowanymi modułami ryzyka, jest kopula. Może ona być stosowana do opisu zależności między rozkładami asymetrycznymi oraz do modelowania zależności w ogonach. Kopula $C(u_1, \dots, u_d)$ to dystrybuenta d -wymiarowego rozkładu z jednostajnymi rozkładami brzegowymi. Za jej pomocą można połączyć dowolne jednowymiarowe rozkłady brzegowe w rozkład wielowymiarowy. Otrzymujemy to z fundamentalnego w teorii kopuli twierdzenia Sklara, z którego wynika, że dla d -wymiarowej łącznej dystrybuanty $F_{1\dots d}$ z rozkładami brzegowymi F_1, \dots, F_d istnieje kopula $C: I^d \rightarrow I$, ($I = [0,1]$) taka, że dla każdego $x_1, \dots, x_d \in R$

$$F_{1\dots d}(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)). \quad (8)$$

Także odwrotnie, jeżeli C jest kopulą a F_1, \dots, F_d są dystrybuantami rozkładów jednowymiarowych, wówczas funkcja $F_{1\dots d}$ określona powyższym wzorem (8) jest dystrybuantą łączną rozkładu d -wymiarowego o rozkładach brzegowych F_1, \dots, F_d . Dodatkowo, jeśli F_1, \dots, F_d są ciągłe, to kopula C wyznaczona jest jednoznacznie. Ponadto, jeśli F_1, \dots, F_d są rosnące i iniektywne, to dla każdej F_i ($i = 1, \dots, d$) istnieje funkcja odwrotna. Stąd dla $F_{1\dots d}$ oraz F_1, \dots, F_d , określonych jak w twierdzeniu Sklara i ich funkcji odwrotnych $F_1^{-1}, F_2^{-1}, \dots, F_d^{-1}$, kopula C ma następującą postać:

$$C(u_1, u_2, \dots, u_d) = F_{1\dots d}\left(F_1^{-1}(u_1), F_2^{-1}(u_2), \dots, F_d^{-1}(u_d)\right), \quad (9)$$

dla $(u_1, \dots, u_d) \in [0,1]^d$. Zatem, jeśli znane są rozkłady brzegowe L_i to wszystkie informacje na temat zależności między zmiennymi losowymi L_i zawarte są w kopuli.

24. C. Genest, *Correlation and Dependence*, "Journal of the American Statistical Association" 2002, No. 97 [458], s. 653–654; doi: 10.1198/jasa.2002.s472.

25. R.A. Shaw, G. Spivak, *Correlations and Dependencies in Economic Capital Models*, "Discussion Paper" September 2009.

Zakładając, że F_i są ciągłe i ściśle rosnące, funkcję gęstości prawdopodobieństwa dla rozkładu $F_{1\dots d}$ można zapisać następująco:

$$f(x_1, x_2, \dots, x_d) = \frac{\partial F_{1\dots d}(x_1, x_2, \dots, x_d)}{\partial x_1 \dots \partial x_d} = \frac{\partial C(u_1, u_2, \dots, u_d)}{\partial u_1 \dots \partial u_d} \prod_{i=1}^d \frac{\partial F_i(x_i)}{\partial x_i} = \quad [10]$$

$$c(u_1, u_2, \dots, u_d) \prod_{i=1}^d f_i(x_i)$$

Z powyższych rozważań wynika, że chcąc oszacować gęstość wielowymiarowego rozkładu $f(x_1, x_2, \dots, x_d)$, można do danych, niezależnie od siebie, dopasować rozkłady brzegowe $f_i(x_i)$ oraz kopulę. W celu dopasowania kopuli najczęściej korzysta się z podejścia parametrycznego, wybierając konkretną jej postać z rodziny eliptycznej lub Archimedesesa. Kopule eliptyczne są elastyczne w wielowymiarowym modelowaniu zależności, ponieważ pozwalają uwzględnić różne rodzaje zależności między poszczególnymi parami zmiennych losowych. Jednak, jako że są to kopule symetryczne, nie można za ich pomocą modelować dowolnej struktury zależności. Z kolei kopule Archimedesesa wykorzystuje się w celu oddania różnorodnych dwuwymiarowych zależności, w szczególności asymetrycznych lub w ogonach rozkładów. O ile zatem stosunkowo łatwo można opisać strukturę zależności między dwoma czynnikami ryzyka, o tyle trudność pojawia się, gdy należy wybrać kopulę dla wymiaru większego niż dwa. Wówczas wspomniane kopule dwuwymiarowe wykazują szereg ograniczeń związanych ze swoją strukturą oraz parametrami. Jeśli modeluje się zależności między większą liczbą zmiennych losowych, to ustalony parametr zależności określa tę samą siłę zależności między wszystkimi zmiennymi losowymi, co sprawia, że model kopuli staje się sztywny i traci wiele informacji na temat zależności. Dlatego do modelowania zależności pomiędzy liczbą zmiennych losowych większą niż dwa stosuje się kaskady kopuli. Dzięki nim wielowymiarowy rozkład dekomponuje się na kopule dwuwymiarowe, uwzględniając zależności między każdą zmienną losową. Niestety, przedstawiona procedura wyznaczania rozkładu wielowymiarowego nadal niesie ze sobą pewne ograniczenia. Dopasowuje się rozkłady brzegowe, które dla danych parametrów opisują określony kształt rozkładu, oraz kopule parametryczne, które dla określonych parametrów narzucają dany kształt zależności. W niniejszym artykule w celu wyznaczenia wielowymiarowego rozkładu zmiennych modelujących ryzyko segmentów ubezpieczeń innych niż na życie wykorzystujemy innowacyjne podejście oparte na sieciach neuronowych, dzięki czemu nie jesteśmy narażeni na ograniczenia związane z dopasowaniem parametrycznym. Sieci neuronowe są wyrafinowaną techniką modelowania, zdolną do odwzorowywania nadzwyczaj złożonych zależności. Najbardziej znaną ich cechą jest generalizacja, czyli zdolność uczenia się na podstawie przykładów i możliwość automatycznego uogólniania zdobytej wiedzy. W oszacowaniu wielowymiarowej funkcji gęstości skorzystamy z podejścia zaprezentowanego w pracy Zhi Zenga i Ting Wang²⁶, w którym konstruuje się dwie sieci neuronowe.

26. Z. Zeng, T. Wang, *Neural Copula*

3. Budowa i uczenie się sieci neuronowej

Modele sieci neuronowych można postrzegać jako zdefiniowaną funkcję, która pobiera dane wejściowe (argumenty funkcji) i generuje wynik (wartości funkcji). Postać funkcji jest ściśle związana z określoną regułą uczenia się sieci neuronowej. W badaniach wykorzystywane są sieci typu perceptronu wielowarstwowego. Na sieć neuronową patrzymy jak na funkcję rozkładu prawdopodobieństwa. Innymi słowy: uczymy naszą sieć na danych, do których chcemy dopasować rozkład prawdopodobieństwa, i na tę sieć nakładamy warunki ograniczające, wynikające z własności rozkładu prawdopodobieństwa. Konstruujemy dwie sieci: pierwszą dla jednowymiarowego rozkładu brzegowego, a drugą dla rozkładu wielowymiarowego. W badaniu uwzględniamy cztery segmenty ubezpieczyciela, dlatego budujemy cztery sieci neuronowe dla poszczególnych segmentów ubezpieczycieli, a następnie jedną sieć neuronową dla rozkładu wielowymiarowego. Konstrukcja tych sieci została opisana w dalszej części rozdziału.

Rozkłady brzegowe

W celu identyfikacji jednowymiarowego rozkładu prawdopodobieństwa przez DNN budujemy sieć neuronową składającą się z $l_m - 1$ warstw ukrytych. Każda warstwa ukryta zbudowana jest z pewnej liczby neuronów. Warstwa wejściowa definiowana jest jako:

$$\mathbf{h}_m^0 = x \in \Omega \subset \mathbb{R} \quad (11)$$

W warstwach ukrytych jako funkcję aktywacji wykorzystujemy tangens hiperboliczny:

$$\mathbf{h}_m^{j+1} = \tanh(\mathbf{w}_m^j \mathbf{h}_m^j + \mathbf{b}_m^j), j \in \{0, 1, \dots, l_m - 1\} \quad (12)$$

natomiast w warstwie wyjściowej funkcję sigmoidalną:

$$\hat{F}_m = \text{sigmoid}(\mathbf{w}^{l_m} \mathbf{u}^{l_m} + \mathbf{b}^{l_m}) \in [0, 1] \quad (13)$$

Po wyznaczeniu dystrybuanty rozkładu prawdopodobieństwa $\hat{F}_m(x, \boldsymbol{\theta}_m)$, gdzie $\boldsymbol{\theta}_m$ oznacza wagi sieci neuronowej $\boldsymbol{\theta}_m = \{\mathbf{w}_m^j\} \cup \{\mathbf{b}_m^j\}$, wyznaczamy gęstość prawdopodobieństwa:

$$\hat{f}_m(x, \boldsymbol{\theta}_m) = \frac{d\hat{F}_m(x, \boldsymbol{\theta}_m)}{dx} \quad (14)$$

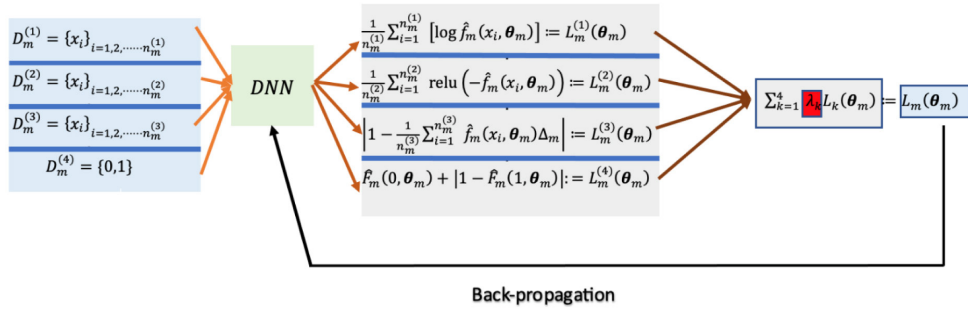
Na rysunku 1. przedstawiamy procedurę uczenia DNN. Podczas uczenia na sieć nałożono ograniczenia wynikające z własności rozkładu prawdopodobieństwa, a sieć uczono na czterech zestawach treningowych. Pierwszym z nich były dane rzeczywiste $D_m^{(1)}$, których rozkład ustalamy z wykorzystaniem DNN. Jako kryterium oceny dopasowania funkcji sieci do danych wykorzystujemy funkcję straty w postaci logarytmu największej wiarygodności $L_m^{(1)}(\boldsymbol{\theta}_m)$.

Następnie na funkcję sieci neuronowej nakładamy ograniczenia wynikające z własności rozkładu prawdopodobieństwa. Zbudowano dwa zbiory treningowe $D_m^{(2)}, D_m^{(3)}$ o punktach równomiernie rozmieszczonych, dla których zakres zmienności jest taki, jak dla $D_m^{(1)}$. Dla $D_m^{(2)}$ i $D_m^{(3)}$ wyznaczamy odpowiednio funkcje straty $L_m^{(2)}(\boldsymbol{\theta}_m), L_m^{(3)}(\boldsymbol{\theta}_m)$ odpowiadające za nieujemność funkcji gęstości

$\hat{f}_m(x, \theta_m)$ oraz unormowanie $\hat{f}_m(x, \theta_m)$. Ostatni zbiór treningowy $D_m^{(4)} = \{0,1\}$ odpowiada funkcji straty $L_m^{(4)}(\theta_m)$ i opisuje własność funkcji dystrybucyjnej:

$$\begin{cases} \hat{F}_m(0, \theta_m) = 0 \\ \hat{F}_m(1, \theta_m) = 1 \end{cases} \quad [15]$$

Rysunek 1. Procedura nauki sieci neuronowej dla rozkładów brzegowych



Źródło: opracowanie własne.

Na końcu wyznaczamy funkcję straty dla modelu jako kombinację liniową powyższych czterech funkcji strat:

$$L_m(\theta_m) = \sum_{k=1}^4 \lambda_k L_k(\theta_m) \quad [16]$$

i poprzez propagację wsteczną uczymy model. W celu uniknięcia przetrenowania modelu wykorzystujemy regularyzację l_2 , a wagi aktualizujemy z wykorzystaniem optymalizatora Adam (*Adaptive Moment Estimation*).

Funkcja kopuli

Sieć neuronowa dla kopuli ma podobną architekturę do sieci przedstawionej powyżej architektury sieci dla rozkładów brzegowych. Zbudowana jest z $l_c - 1$ warstw ukrytych, a główna różnica tkwi w liczbie neuronów w warstwie wejściowej:

$$\mathbf{h}_c^0 = \mathbf{u} = [u_1, u_2, \dots, u_4]^T \in [0,1]^4 \quad [17]$$

W warstwach ukrytych jako funkcję aktywacji wykorzystujemy tangens hiperboliczny:

$$\mathbf{h}_c^{j+1} = \tanh(\mathbf{w}_c^j \mathbf{h}_c^j + \mathbf{b}_c^j) \quad j \in \{0,1, \dots, l_c - 1\} \quad [18]$$

natomiast w warstwie wyjściowej funkcję sigmoidalną:

$$\hat{C} = \text{sigmoid}(\mathbf{w}^{l_c} \mathbf{u}^{l_c} + \mathbf{b}^{l_c}) \in [0,1] \quad [19]$$

Wartości, które otrzymujemy z sieci neuronowej, to dopasowany rozkład $\hat{C}(\mathbf{u}, \boldsymbol{\theta}_c)$ gdzie $\boldsymbol{\theta}_c$ oznacza wszystkie wagi modelu $\boldsymbol{\theta}_c = \{\mathbf{w}_c^j\} \cup \{\mathbf{b}_c^j\}$. Z twierdzenia Skłara wynika, że wielowymiarowy rozkład można przedstawić za pomocą wielowymiarowej funkcji kopuli oraz rozkładów brzegowych. Stąd przybliżony wielowymiarowy rozkład ma postać:

$$\hat{F}(\mathbf{x}, \{\boldsymbol{\theta}_m^{(i)}\}, \boldsymbol{\theta}_c) = \hat{C}(\hat{F}_1(x_1, \boldsymbol{\theta}_m^{(1)}), \hat{F}_2(x_2, \boldsymbol{\theta}_m^{(2)}), \hat{F}_3(x_3, \boldsymbol{\theta}_m^{(3)}), \hat{F}_4(x_4, \boldsymbol{\theta}_m^{(4)}), \{\boldsymbol{\theta}_m^{(i)}\}, \boldsymbol{\theta}_c) \quad [20]$$

a gęstość rozkładu:

$$\hat{f}(\mathbf{x}, \{\boldsymbol{\theta}_m^{(i)}\}, \boldsymbol{\theta}_c) = \hat{c}(\hat{F}_1(x_1, \boldsymbol{\theta}_m^{(1)}), \dots, \hat{F}_4(x_4, \boldsymbol{\theta}_m^{(4)}), \boldsymbol{\theta}_c) \prod_{i=1}^4 f_i(x_i, \boldsymbol{\theta}_m^{(i)}) \quad [21]$$

Na rysunku 2. przedstawiamy procedurę uczenia DNN. Podczas uczenia na sieć nałożono ograniczenia wynikające z własności wielowymiarowego rozkładu oraz kopuli. Rozkłady brzegowe wyznaczamy z wykorzystaniem sieci DNN dla rozkładów brzegowych, stąd w procedurze uczenia DNN dla kopuli wartości parametrów $\boldsymbol{\theta}_m$ są ustalone. Pierwszym zbiorem treningowym sieci jest zbiór danych rzeczywistych $D_c^{(1)}$, dla którego wyznaczamy strukturę zależności oraz rozkład wielowymiarowy. Dla $D_c^{(1)}$ wyliczamy stratę w postaci logarytmu największej wiarygodności $L_c^{(1)}(\{\boldsymbol{\theta}_m^{(i)}\}, \boldsymbol{\theta}_c)$. Następnie wyznaczamy dwa zbiory treningowe $D_c^{(2)}, D_c^{(3)}$ składające się z macierzy danych o czterech kolumnach. Bez straty ogólności przyjmujemy, że każda kolumna zawiera losowo wybrane punkty z dziedziny. Dla macierzy $D_c^{(2)}, D_c^{(3)}$ wyznaczamy odpowiednio funkcje straty $L_c^{(2)}(\{\boldsymbol{\theta}_m^{(i)}\}, \boldsymbol{\theta}_c), L_c^{(3)}(\{\boldsymbol{\theta}_m^{(i)}\}, \boldsymbol{\theta}_c)$ które odpowiadają za nieujemność funkcji gęstości $\hat{f}_m(x, \{\boldsymbol{\theta}_m^{(i)}\}, \boldsymbol{\theta}_c)$ oraz unormowanie $\hat{f}_m(x, \{\boldsymbol{\theta}_m^{(i)}\}, \boldsymbol{\theta}_c)$. Z własności funkcji kopuli wynika:

$$\begin{cases} \hat{C}(\underline{\mathbf{u}}_1, \boldsymbol{\theta}_c) = \hat{C}(0, u_2, u_3, u_4, \boldsymbol{\theta}_c) = 0 \\ \hat{C}(\underline{\mathbf{u}}_2, \boldsymbol{\theta}_c) = \hat{C}(u_1, 0, u_3, u_4, \boldsymbol{\theta}_c) = 0 \\ \hat{C}(\underline{\mathbf{u}}_d, \boldsymbol{\theta}_c) = \hat{C}(u_1, u_2, 0, u_4, \boldsymbol{\theta}_c) = 0 \\ \hat{C}(\underline{\mathbf{u}}_d, \boldsymbol{\theta}_c) = \hat{C}(u_1, u_2, u_3, 0, \boldsymbol{\theta}_c) = 0 \end{cases} \quad \begin{cases} \hat{C}(\bar{\mathbf{u}}_1, \boldsymbol{\theta}_c) = \hat{C}(u_1, 1, 1, 1, \boldsymbol{\theta}_c) = u_1 \\ \hat{C}(\bar{\mathbf{u}}_2, \boldsymbol{\theta}_c) = \hat{C}(1, u_2, 1, 1, \boldsymbol{\theta}_c) = u_2 \\ \hat{C}(\bar{\mathbf{u}}_d, \boldsymbol{\theta}_c) = \hat{C}(1, 1, u_4, 1, \boldsymbol{\theta}_c) = u_3 \\ \hat{C}(\bar{\mathbf{u}}_d, \boldsymbol{\theta}_c) = \hat{C}(1, 1, 1, u_4, \boldsymbol{\theta}_c) = u_4 \end{cases} \quad [22]$$

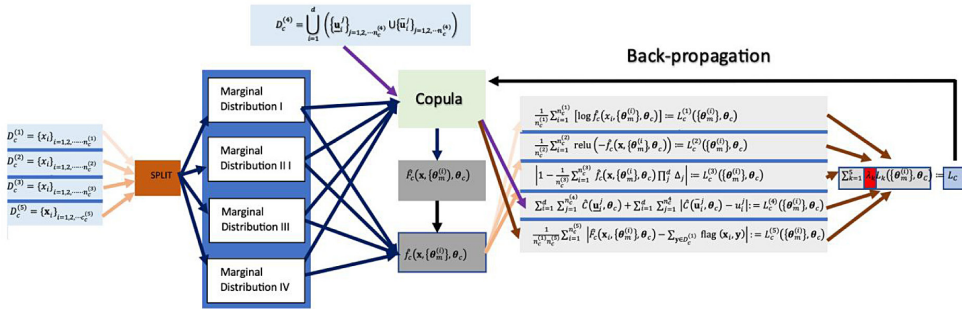
stąd wyznaczamy zbiór treningowy $D_c^{(4)}$ zawierający argumenty powyższych funkcji \hat{C} , a funkcja straty $L_c^4(\{\boldsymbol{\theta}_m^{(i)}\}, \boldsymbol{\theta}_c)$ odpowiada $D_c^{(4)}$. Należy zauważyć, że pobieranie próbek z CDF jest zwykle czasochłonne. Dla zbioru treningowego $D_c^{(5)}$ danego macierzą złożoną z liczb losowych z dziedziny określamy funkcję straty $L_c^{(5)}(\{\boldsymbol{\theta}_m^{(i)}\}, \boldsymbol{\theta}_c)$ i funkcję $\hat{F}_c(\mathbf{x}, \{\boldsymbol{\theta}_m^{(i)}\}, \boldsymbol{\theta}_c)$ aproksymowano:

$$\hat{F}_c(\mathbf{x}, \{\boldsymbol{\theta}_m^{(i)}\}, \boldsymbol{\theta}_c) \approx \frac{1}{n_c^{(1)}} \sum_{\mathbf{y} \in D_c^{(1)}} \text{flag}(\mathbf{x}, \mathbf{y}) \quad [23]$$

gdzie:

$$\text{flag}(\mathbf{x}, \mathbf{y}) = \begin{cases} 1 & \forall j \leq d, y_j < x_j \\ 0 & p.p \end{cases} \quad (24)$$

Rysunek 2. Procedura nauki sieci neuronowej dla kopuli



Źródło: opracowanie własne.

Ostatecznie wyznaczamy funkcję straty dla modelu jako kombinacją liniową powyższych czterech funkcji strat:

$$L_m(\{\theta_m^i\}, \theta_c) = \sum_{k=1}^5 \lambda_k L_k(\{\theta_m^i\}, \theta_c) \quad (25)$$

i poprzez propagację wsteczną uczymy model. Podczas nauki w celu uniknięcia przetrenowania modelu wykorzystujemy regularyzację l_2 , a wagi aktualizujemy z wykorzystaniem optymalizatora Adam.

4. Strategia empiryczna

Z uwagi na brak ogólnodostępnych danych na potrzeby niniejszego badania ryzyko składki i rezerw dla poszczególnych segmentów modelujemy za pomocą współczynnika zespolonego, danego poniższym wzorem²⁷:

$$L_i = \frac{NCE_i + OE_i}{NP_i} \quad (26)$$

gdzie dla i -tego segmentu NCE_i oznacza odszkodowania i świadczenia netto (bezpośrednia działalność ubezpieczeniowa), OE_i to koszty poniesione, natomiast NP_i to składki zarobione netto (bezpośrednia działalność ubezpieczeniowa). Zmienną losową L , modelującą zagregowane ryzyko segmentów, otrzymujemy, agregując współczynniki zespolone L_i za pomocą następującej formuły:

$$L = \sum_i w_i L_i \quad (27)$$

27. T. Schubert, G. Griebmann, *German Proposal for a Standard Approach for Solvency II*, "The International Association for the Study of Insurance Economic" 2007, No. 32(1), s. 133–150.

gdzie w_i jest udziałem poszczególnych segmentów w portfelu ubezpieczyciela wyliczonym jako:

$$w_i = \frac{NP_i}{\sum_i NP_i}. \quad [28]$$

Wymogi kapitałowe dla poszczególnych segmentów $\kappa(L_i)$ oraz zagregowanego ryzyka $\kappa(L)$ wyznaczamy odpowiednio:

$$\kappa(L_i) = VaR_{0,995}(L_i) - E(L_i) \quad [29]$$

$$\kappa(L) = VaR_{0,995}(L) - E(L) \quad [30]$$

Analizę rozpoczynamy od oszacowania rozkładów brzegowych L_i . W tym celu stosujemy dwa podejścia: parametryczne oraz nieparametryczne z zastosowaniem sieci neuronowych. Natomiast kluczową z punktu widzenia oceny efektu dywersyfikacji strukturę zależności między zmiennymi L_i modelujemy z wykorzystaniem trzech różnych metod:

1. Stosujemy macierz korelacji podaną w Rozporządzeniu Delegowanym (tabela 1.). W praktyce oznacza to wykorzystanie wielowymiarowego rozkładu normalnego i agregację ryzyka standardową metodą wariancji-kowariancji.
2. Podobnie jak przy szacowaniu rozkładów brzegowych, stosujemy podejście parametryczne, przy czym strukturę zależności modelujemy za pomocą C-kaskady kopuli. Sposób ten został szczegółowo opisany w przywoływanej już pracy²⁸.
3. Stosujemy podejście nieparametryczne, bazujące na sieci neuronowej.

Dokładność dopasowania do danych rzeczywistych, oszacowanych w ten sposób rozkładów wielowymiarowych, oceniamy, odwołując się do metod stosowanych w analizie jakości prognoz probabilistycznych, w której bazujemy na odległości energetycznej. Tilman Gneiting i Adrian E. Raftery²⁹ rozwijają w oparciu o tę odległość nową metodę oceny jakości prognozowania, stawiając tezę, że jest ona odpowiednią odległością dla rozkładów wielowymiarowych. Dla H -wymiarowej zmiennej losowej X i wektora obserwacji y pochodzącego z rozkładu Y odległość energetyczną definiujemy jako:

$$ES_{\beta}(F_X, y) = \mathbb{E}\left(\|X - y\|_2^{\beta}\right) - \frac{1}{2}\mathbb{E}\left(\|X - \tilde{X}\|_2^{\beta}\right) \quad [31]$$

gdzie \tilde{X} jest podzbiorem X . Ponadto $\beta \in (0,2)$ a $\|\cdot\|_2$ jest normą Euklidesową. $ES_{\beta}(F_X, y)$ jest poprawnie określona dla wszystkich β , natomiast w niniejszej pracy przyjmujemy $\beta = 1$.

Florian Ziel i Kevin Berg³⁰ na podstawie badań symulacyjnych stwierdzają, że odległość energetyczna powinna być preferowaną miarą oceny prognoz wieloczynnikowych.

28. K. Szczęsny, *Wykorzystanie kaskad kopuli...*

29. T. Gneiting, A.E. Raftery, *Strictly proper scoring rules, prediction, and estimation*, "Journal of the American Statistical Association" 2007, No. 102[477], s. 359–378; doi: 10.1198/016214506000001437.

30. F. Ziel, K. Berk, *Multivariate Forecasting Evaluation: On Sensitive and Strictly Proper Scoring Rules*, arXiv 2019.

5. Charakterystyka danych

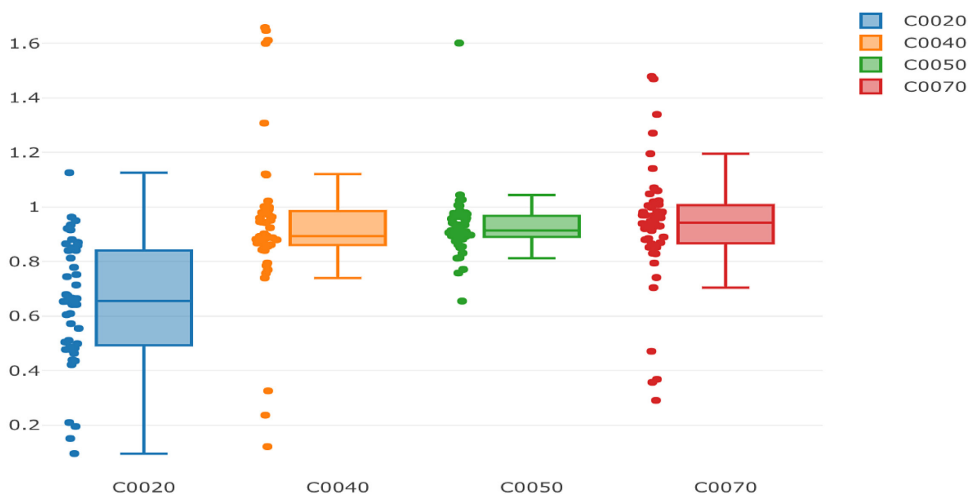
Z uwagi na fakt, że ubezpieczyciele w Polsce nie prowadzą działalności w ramach wszystkich 12 segmentów, które wymienione są w dyrektywie, w badaniach empirycznych analizujemy wskaźniki zespolone L_i dla czterech segmentów polskich ubezpieczycieli majątkowych, wyznaczonych na podstawie danych pozyskanych z raportów o wypłacalności i kondycji finansowej z lat 2017–2022. Z uwagi na małą liczebność danych w badaniach zastosowano technikę bootstrapową, aby uzyskać dane zawierające 10 000 obserwacji.

Wybrane segmenty to:

- C0020 – ubezpieczenia na wypadek utraty dochodów;
- C0040 – ubezpieczenia odpowiedzialności cywilnej z tytułu użytkowania pojazdów mechanicznych;
- C0050 – pozostałe ubezpieczenia pojazdów;
- C0070 – ubezpieczenia od ognia i innych szkód rzeczowych.

Na rysunku 3. za pomocą wykresów pudełkowych przedstawiamy rozkład wskaźników zespolonych dla analizowanych segmentów ubezpieczycieli, natomiast w tabeli 1. prezentujemy ich podstawowe statystyki opisowe.

Rysunek 3. Rozkład wskaźników zespolonych dla segmentów ubezpieczycieli



Źródło: opracowanie własne.

Tabela 1. Podstawowe statystyki dla segmentów ubezpieczycieli

	C0020	C0040	C0050	C0070
Minimum	0,096	0,121	0,655	0,291
Maximum	1,126	1,658	1,601	1,479
Kwartyl 1	0,496	0,861	0,892	0,868
Kwartyl 3	0,841	0,983	0,967	1,006
Średnia	0,647	0,936	0,930	0,932
Mediana	0,655	0,893	0,914	0,943

LCL Średnia	0,582	0,853	0,894	0,865
UCL Średnia	0,712	1,020	0,965	0,999
Wariancja	0,050	0,083	0,015	0,053
Odchylenia st.	0,223	0,288	0,123	0,231
Skośność	-0,410	0,227	3,042	-0,494
Kurtoza	-0,090	2,385	16,210	1,765

Źródło: opracowanie własne.

LCL Średnia i UCL Średnia oznaczają odpowiednio dolną i górną granice 95 % przedziału ufności dla średniej. Wskaźniki C0040, C0050, C0070 mają obserwacje odstające. Na podstawie mediany (poziomej linii w pudełku wykresu) zauważamy skośność tych rozkładów. Dalej liczymy współczynnik korelacji liniowej pomiędzy każdym segmentem ubezpieczyciela. Wyniki przedstawiamy w tabeli 2.

Tabela 2. Macierz korelacji między segmentami ubezpieczyciela

	C0020	C0040	C0050	C0070
C0020	1	0,2286	0,0034	0,358
C0040	0,2286	1	0,3028	0,8171
C0050	0,0034	0,3028	1	0,0378
C0070	0,3580	0,8171	0,0378	1

Źródło: opracowanie własne.

Największą wartość korelacji otrzymujemy między segmentami C0040 i C0070, natomiast najmniejszą między segmentami C0020 i C0050. Warto zwrócić uwagę, że wyznaczona macierz korelacji liniowej różni się od macierzy podanej przez autorów dyrektywy.

6. Wyniki badań empirycznych

W niniejszym rozdziale prezentujemy wyniki przeprowadzonej na dwa sposoby estymacji rozkładów brzegowych oraz struktury zależności za pomocą kopuli. Następnie omawiamy sposób, w jaki wyznaczamy rozkład wielowymiarowy i oceniamy jego dopasowanie do danych rzeczywistych. Na końcu ustalamy efekty dywersyfikacji.

6.1 Estymacja rozkładów brzegowych

Do wskaźników zespolonych L_i dopasowujemy parametryczne rozkłady za pomocą metody największej wiarygodności i najlepiej dopasowany rozkład wybieramy w oparciu o Bayesowskie kryterium informacyjne. Wybieramy spośród rodzin rozkładów normalnych, logarytmiczno-normalnych, logistycznych, Weibulla i gamma. Dopasowane rozkłady prawdopodobieństwa i ich parametry dla poszczególnych segmentów prezentuje tabela 3.

Tabela 3. Rozkłady brzegowe dla czterech segmentów (estymacja parametryczna)

Segment	C0020	C0040	C0050	C0070
Rozkład	Norm (0,65; 0,22)	Logis (0,92; 0,13)	Logis (0,92; 0,05)	Logis (0,94; 0,11)

Źródło: opracowanie własne.

Następnie dla każdego wskaźnika zespolonego L_i trenujemy sieć neuronową tak, aby nauczyć się rozkładu prawdopodobieństwa. Każda sieć ma tę samą budowę: cztery warstwy i w każdej warstwie 10 neuronów.

Dalej korzystamy z faktu, że jeśli F jest ciągłą zmienną losową $X \sim F$, to $F(X)$ ma rozkład jednostajny $U(0,1)$. Dla wszystkich segmentów oceniamy dopasowanie rozkładu prawdopodobieństwa, sprawdzając, czy wartości $\hat{F}_s^m(x_{i,s})$ mają rozkład $U(0,1)$, gdzie \hat{F}_s^m oznacza oszacowaną dystrybuantę dla segmentu s (C0020, C0040, C0050, C0070) metodą m (sieć neuronowa, podejście parametryczne), a $x_{i,s}$ – zaobserwowane wartości wskaźnika [25] dla poszczególnych segmentów s .

W ocenie korzystamy z dwóch testów: Kołmogorova-Smirnova oraz testu Chi-kwadrat. Wyniki testów prezentujemy w tabeli 4.

Tabela 4. Dopasowanie rozkładu prawdopodobieństwa dla poszczególnych segmentów z wykorzystaniem testów Kołmogorova-Smirnova i Chi-kwadrat

Segment	Opis rozkładu	Kołmogorov–Smirnov		Chi-kwadrat	
		Statystyka	p-value	Statystyka	p-value
C0020	Sieć neuronowa	0,055	0,997	0,229	0,892
	Parametryczne	0,096	0,730	0,235	0,235
C0040	Sieć neuronowa	0,093	0,763	0,133	0,715
	Parametryczne	0,187	0,061	0,193	0,634
C0050	Sieć neuronowa	0,095	0,749	1,167	0,558
	Parametryczne	0,103	0,655	1,580	0,209
C0070	Sieć neuronowa	0,054	0,998	0,229	0,892
	Parametryczne	0,137	0,301	6,081	0,014

Źródło: opracowanie własne.

Otrzymane p-value, zarówno dla sieci neuronowych, jak i dla rozkładów brzegowych, pozwala na potwierdzenie, że $\hat{F}_s^m(x_{i,s})$ mają rozkłady jednostajne. Natomiast wartości statystyk wskazują, że rozkłady wyznaczone przez sieci neuronowe są lepiej dopasowane do danych rzeczywistych.

6.2 Estymacja struktury zależności za pomocą kopuli

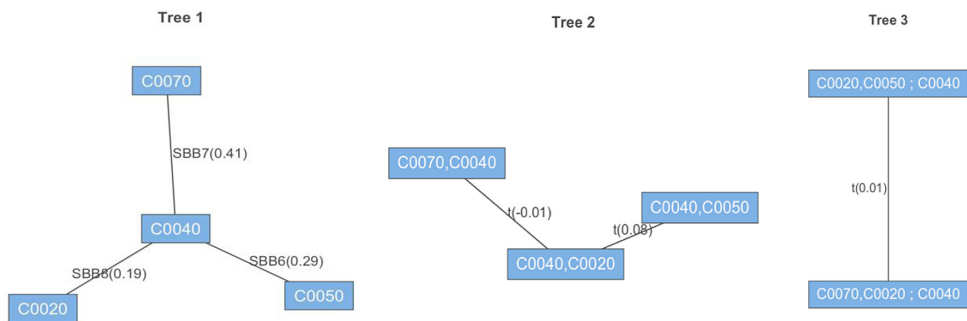
W kolejnym etapie za pomocą kopuli opisujemy strukturę zależności między segmentami C0020, C0040, C0050, C0070 ubezpieczyciela. Podobnie jak w przypadku dopasowania rozkładów

brzegowych, stosujemy dwa podejścia: C-kaskady kopuli oraz kopuli identyfikowanej za pomocą głębokich sieci neuronowych.

Identyfikacja kopuli za pomocą C-kaskady kopuli

Parametry dla kopuli dwuwymiarowych wybieramy przy użyciu metody największej wiarygodności, a najlepsza kopula dwuwymiarowa jest wybrana z wykorzystaniem kryterium logarytmu wiarygodności. Stosowana tutaj metoda została dokładnie opisana³¹. Rysunek 4. prezentuje uzyskaną dekompozycję czterowymiarowego rozkładu na kopule dwuwymiarowe.

Rysunek 4. Drzewo C-kaskady kopuli



Źródło: opracowanie własne.

Identyfikacja kopuli za pomocą sieci neuronowej

Postępujemy zgodnie z metodologią opisaną w artykule Zhi Zenga i Ting Wang³². Budujemy głęboką sieć neuronową składającą się z 5 warstw ukrytych, każda złożona z 10 neuronów. Jako funkcję aktywacji w warstwach ukrytych, wykorzystujemy tangens hiperboliczny. Dla każdej warstwy stosujemy regularyzację L2 (regularyzacja Tichonowa) ze współczynnikiem regularyzacji 0,001. Do trenowania sieci używamy optymalizatora Adam ze współczynnikiem uczenia się 0,0001. Na sieć nakładamy ograniczenia w postaci kombinacji własności funkcji kopuli.

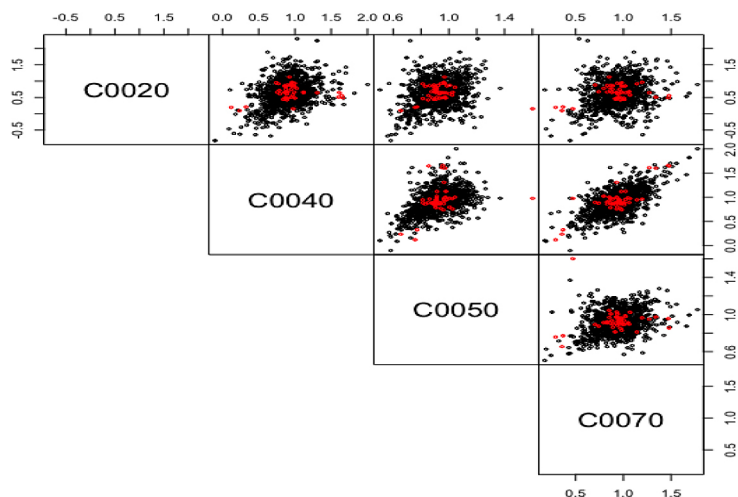
6.3 Wyznaczenie rozkładu wielowymiarowego i ocena dopasowania

Mając wyznaczone rozkłady gęstości dla poszczególnych segmentów oraz rozkład gęstości kopuli, wyznaczamy gęstość rozkładu czterowymiarowego. Na rysunku 5. przedstawiamy wykres par realizacji wylosowanych dla rozkładu czterowymiarowego wyznaczonego za pomocą podejścia parametrycznego, natomiast na rysunku 6. – z wykorzystaniem głębokich sieci neuronowych. Punkty czerwone oznaczają pary obserwacji rzeczywistych, a punkty czarne są to pary realizacji, wylosowane z czterowymiarowego rozkładu łącznego.

31. Zob. K. Szczęśny, *Wykorzystanie kaskad kopuli...*

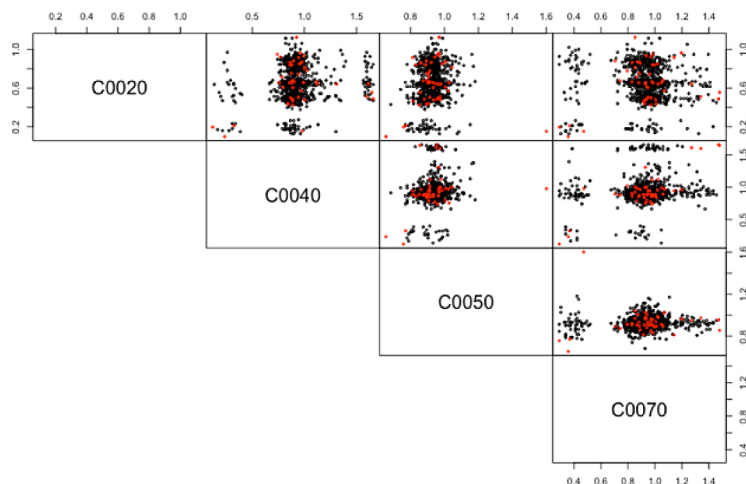
32. Z. Zeng, T. Wang, *Neural Copula...*

Rysunek 5. Wykres par realizacji wylosowanych dla rozkładu czterowymiarowego wyznaczonego za pomocą podejścia parametrycznego



Źródło: opracowanie własne.

Rysunek 6. Wykres par realizacji wylosowanych dla rozkładu czterowymiarowego wyznaczonego z wykorzystaniem głębokich sieci neuronowych



Źródło: opracowanie własne.

W kolejnym etapie, po oszacowaniu na dwa sposoby czterowymiarowych rozkładów, oceniamy, który z nich lepiej opisuje dane rzeczywiste. Zarówno dla rozkładu wyznaczonego za pomocą głębokich sieci neuronowych, jak i dla rozkładu wyznaczonego z wykorzystaniem podejścia parametrycznego wyznaczamy wartość odległości energetycznej między każdym wektorem obserwacji rzeczywistych a wektorami realizacji wylosowanych z oszacowanych rozkładów. Następnie wyznaczamy średnią wartość odległości energetycznej dla każdego rozkładu. Uzyskane wyniki prezentujemy w tabeli 5. Poprawa wyniku odległości energetycznej, kiedy zastosowana jest kopuła

neuronowa, wskazuje na fakt, że modyfikacja i doprecyzowanie konstrukcji stosowanych sieci powinno wpłynąć na lepsze dopasowanie do danych rzeczywistych.

Tabela 5. Wartości błędu dopasowania

	C – kaskady kopuli	kopula neuronowa
Błąd	0,2589487	0,256471

Źródło: opracowanie własne.

6.4. Wyznaczenie efektu dywersyfikacji

Mając wyznaczone rozkłady czterowymiarowe, porównujemy efekty dywersyfikacji otrzymane za pomocą podejść, w których kolejno:

- strukturę zależności modelujemy za pomocą ustalonej w Rozporządzeniu Delegowanym macierzy korelacji (wybieramy z niej wartości współczynników korelacji między analizowanymi segmentami), czyli stosujemy metodę wariancji-kowariancji;
- strukturę zależności modelujemy za pomocą C-kaskady kopuli oszacowanej w sposób parametryczny;
- strukturę zależności modelujemy za pomocą kopuli oszacowanej w sposób nieparametryczny z wykorzystaniem głębokich sieci neuronowych.

W tabeli 6. przedstawiamy otrzymane efekty dywersyfikacji.

Tabela 6. Efekt dywersyfikacji

	SF	C-kaskady kopuli	Sieci neuronowe
DE	0,31	0,35	0,53

Źródło: opracowanie własne.

Dla wyznaczonych rozkładów z tabeli 4. oraz zależności między nimi opisanych za pomocą macierzy korelacji danej przez twórców dyrektywy *Solvency II* (tabela 1.) efekt dywersyfikacji wynosi 0,31. Wykorzystując podejście parametryczne do wyznaczenia rozkładów brzegowych dla segmentów (patrz tabela 4.) oraz modelując strukturę zależności z wykorzystaniem C-kaskady kopuli, otrzymujemy wartość efektu dywersyfikacji 0,35. Jest to wartość o 4 punkty procentowe wyższa niż uzyskana za pomocą SF. Natomiast efekt dywersyfikacji otrzymany w wyniku zastosowania podejścia nieparametrycznego do szacowania rozkładów brzegowych i kopuli, wykorzystującego głębokie sieci neuronowe, jest wyższy o 22 punkty procentowe niż efekt dywersyfikacji uzyskany za pomocą SF. Zatem w każdym przypadku wykorzystanie kopuli do modelowania struktury zależności daje wynik wyższy niż proponowana metoda wariancji-kowariancji w dyrektywie *Solvency II*.

Podsumowanie

Autorzy dyrektywy *Solvency II* zachęcają zakłady ubezpieczeń do tworzenia modeli wewnętrznych, które lepiej odzwierciedlą profil działalności ubezpieczyciela. Celem szeroko prowadzonych przez

nas badań jest opracowanie metody, która może być wykorzystana w modelach wewnętrznych do szacowania zależności między segmentami w podmodule ryzyka składki i rezerw. Badania prowadzimy dla danych rzeczywistych pozyskanych z raportów o wypłacalności i kondycji finansowej polskich ubezpieczycieli majątkowych.

Rozkłady brzegowe wybranych segmentów oraz kopule szacujemy na dwa sposoby: wykorzystując podejście parametryczne z zastosowaniem C-kaskady kopuli³³ oraz innowacyjnie, wykorzystując głębokie sieci neuronowe. Mając jednowymiarowe rozkłady brzegowe oraz funkcję kopuli, powołujemy się na twierdzenie Sklara i łączymy wyznaczone rozkłady brzegowe w rozkład czterowymiarowy w każdym z podejść. Następnie z wyznaczonych rozkładów czterowymiarowych losujemy realizacje i za pomocą odległości energetycznej sprawdzamy dopasowania rozkładów do danych rzeczywistych. Prowadzone badania wyłaniają dwa istotne aspekty, na które musi zwrócić uwagę zakład ubezpieczeń. Pierwszym jest fakt, że DE ubezpieczyciela zależy od właściwie zidentyfikowanej struktury zależności, a zaproponowane podejście w SF oparte na korelacji liniowej jest niewystarczające. Drugą kwestią jest wybór zarówno metod estymacji rozkładów brzegowych, jak i metody identyfikacji kopuli. Z badań wynika, że model oparty na głębokich sieciach neuronowych lepiej dopasowuje się do danych rzeczywistych, a wykorzystany do estymacji rozkładów brzegowych oraz funkcji kopuli daje dokładniejszy DE.

Uzyskane wyniki wskazują, że zaproponowana architektura głębokiej sieci neuronowej jest dobrym kandydatem do estymacji rozkładów brzegowych oraz identyfikacji kopuli wykorzystywanej do opisu struktury zależności między rodzajami ryzyka w modelach wewnętrznych.

Wykaz źródeł

- Acar E.F., Genest C., Nešlehová J., *Beyond simplified pair-copula constructions*, "Journal of Multivariate Analysis" 2012, No. 110; doi: 10.1016/j.jmva.2012.02.001.
- Bedford T., Cooke R.M., *Vines – A new graphical model for dependent random variables*, "Annals of Statistics" 2002, No. 30(4); doi: 10.1214/aos/1031689016.
- Bermúdez L., Ferri A., Guillén M., *A correlation sensitivity analysis of non-life underwriting risk in solvency capital requirement estimation*, "ASTIN Bulletin" 29.04.2013; doi: 10.1017/asb.2012.1.
- Castellani G. et al., *Accuracy and Robustness of Machine Learning Methods in SCR Estimation Nested Monte Carlo Simulation*, Insurance Data Science Conference June 2019.
- Cifuentes A., Charlin V., *Operational risk and the Solvency II capital aggregation formula: Implications of the hidden correlation assumptions*, "Journal of Operational Risk" 2016, No. 11(4); doi: 10.21314/JOP.2016.181.
- Eling M., Jung K., *Risk aggregation in non-life insurance: Standard models vs. internal models*, "Insurance: Mathematics and Economics" 2020, No. 95; doi: 10.1016/j.insmatheco.2020.09.003.
- Genest C., *Correlation and Dependence*, "Journal of the American Statistical Association" 2002, 97(458); doi: 10.1198/jasa.2002.s472.
- Gneiting T., Raftery A.E., *Strictly proper scoring rules, prediction, and estimation*, "Journal of the American Statistical Association" 2007, No. 102(477); doi: 10.1198/016214506000001437.

33. K. Szczęsny, *Wykorzystanie kaskad kopuli ...*

- Haff I.H., *Parameter estimation for pair-copula constructions*, "Bernoulli" 2013, No. 19(2); doi: 10.3150/12-BEJ413.
- Hassan A.K.I., Abraham A., *Modeling insurance fraud detection using imbalanced data classification*, "Advances in Intelligent Systems and Computing" 2016, No. 419; doi: 10.1007/978-3-319-27400-3_11.
- Ibiwoye A., Ajibola O.O., Sogunro A., *Artificial Neural Network Model for Predicting Insurance Insolvency*, "International Journal of Management and Business Research" 2012, No. 2(1).
- Inanoglu H., Jacobs M., *Models for Risk Aggregation and Sensitivity Analysis: An Application to Bank Economic Capital*, "Journal of Risk and Financial Management" 2009; doi: 10.3390/jrfm2010118.
- Kingma D.P., Ba J.L., *Adam: A method for stochastic optimization*, 3rd International Conference on Learning Representations, ICLR 2015 – Conference Track Proceedings.
- Ling C.K., Fang F., Kolter J.Z., *Deep Archimedean Copulas*, [w:] *Advances in Neural Information Processing Systems*, [ed.] Larochelle H. et al., Curran Associates, Inc. New York 2020.
- Lintvelt M., *Least-Squares Monte Carlo Within Solvency II*, Amsterdam School of Economics, Amsterdam 2016.
- Mittnik S., *Solvency II Calibrations: Where Curiosity Meets Spuriousity*, "Preprints" 2016; doi: 10.20944/preprints201607.0021.v1.
- Rozporządzenie delegowane Komisji (UE) 2015/35 z dnia 10 października 2014 r. uzupełniające dyrektywę Parlamentu Europejskiego i Rady 2009/138/WE w sprawie podejmowania i prowadzenia działalności ubezpieczeniowej i reasekuracyjnej (Wyłącalność II), Dz.U. Unia Europejska, 17.1.2015.
- Shaw R.A., Spivak G., *Correlations and Dependencies in Economic Capital Models*, "Discussion Paper" September 2009.
- Schubert T., Griebmann G., *German Proposal for a Standard Approach for Solvency II*, "The International Association for the Study of Insurance Economic" 2007, No. 32(1).
- Spanhel F., Kurz M.S., *Simplified vine copula models: Approximations based on the simplifying assumption*, "Electronic Journal of Statistics" 2019, No. 13(1); doi: 10.1214/19-EJS1547.
- Sun Y., Cuesta-Infante A., Veeramachaneni K., *Learning vine copula models for synthetic data generation*, 33rd AAAI Conference on Artificial Intelligence, AAAI 2019, 31st Innovative Applications of Artificial Intelligence Conference, IAAI 2019 and the 9th AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019; doi: 10.1609/aaai.v33i01.33015049.
- Szczęsny K., *Wpływ błędnej specyfikacji struktury zależności w procesie agregacji ryzyka na efekt dywersyfikacji w Solvency II*, [w:] *Problemy gospodarcze i społeczne Polski i Europy*, Sojka E., Acedański J. [red.], Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice 2022.
- Szczęsny K., *Wykorzystanie kaskad kopuli w agregacji ryzyka w procesie wyznaczania kapitałowych wymogów wyłącalności w Solvency II*, [w:] *Sektor ubezpieczeń w obliczu wyzwań współczesności*, Lemkowska M., Wojtkowiak M. [red.], Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań 2022; <https://doi.org/10.18559/978-83-8211-131-6/7>.
- Yunos Z.M. et al., *Predictive modelling for motor insurance claims using artificial neural networks*, "International Journal of Advances in Soft Computing and its Applications" 2016, No. 8(3).
- Zeng Z., Wang T., *Neural Copula: A unified framework for estimating generic high-dimensional Copula functions*, Research Gate 2022; doi: <https://doi.org/10.48550/arXiv.2205.15031>.
- Ziel F., Berk K., *Multivariate Forecasting Evaluation: On Sensitive and Strictly Proper Scoring Rules*, arXiv 2019.

Deep neural networks in identifying marginal distributions and multidimensional copula in the context of risk aggregation in *Solvency II*

One of the basic aspects of the Solvency II Directive is the protection of the insured against the insolvency of insurance companies. For this purpose, by aggregating the Solvency Capital Requirements for individual types of risk, the Solvency Capital Requirement (SCR) and the diversification effect (DE) are determined. The SCR may be determined using the Standard Formula (SF) given by the authors of the Directive is based on the variance-covariance (V-C) method or using internal models developed independently (by the companies). The aim of our present research is to use copulas estimated with the use of neural networks in modeling dependencies in the premium and reserve risk submodule. Two neural networks are constructed: the first to boundary distributions and the second to estimate the copula. In the research, we analyze indicators for the segments of non-life insurers, determined from reports on the solvency and financial condition of Polish insurance companies. We compare the DE obtained by the V-C method, Vine copula and copula estimation approach using neural networks. The conducted research indicates significant differences in the DE obtained for the copula estimated with the use of neural networks, parametric copula and the approach proposed by the authors of the Directive based on the V-C method. The obtained results can be used in internal models.

Keywords: *Solvency II, Solvency Capital Requirement, diversification effect, Vine copula, neural copulas.*

KRYSTIAN SZCZĘSNY – Katedra Matematyki Uniwersytetu Ekonomicznego w Krakowie.

ORCID: 0000-0001-6557-3836

e-mail: szczesnk@uek.krakow.pl