Artur Hołda
Uniwersytet Ekonomiczny w Krakowie

Gabriela Malik
Wyższa Szkoła Ekonomii i Informatyki w Krakowie

MODELOWANIE ZALEŻNOŚCI CEN KONTRAKTÓW TERMINOWYCH NA PRODUKTY ROLNE NOTOWANYCH NA GIEŁDZIE TOWAROWEJ W CHICAGO Z WYKORZYSTANIEM FUNKCJI KOPULI

Wprowadzenie

Jednym z najważniejszych pojęć na rynkach finansowych, budzącym szerokie zainteresowanie zarówno praktyków, jak i teoretyków zajmujących się badaniem tychże rynków, jest pojęcie struktury zależności. Jej określenie ma istotne znaczenie w zarządzaniu instrumentami finansowymi, a błędna interpretacja siły powiązań może prowadzić do niewłaściwych decyzji inwestycyjnych. Najpowszechniejszą i nadal szeroko stosowaną miarą zależności, ze względu na jej prostotę i łatwość interpretacji, jest współczynnik korelacji liniowej Pearsona. Należy jednak pamiętać, że z powodu liniowości, jest on odpowiednim narzędziem do mierzenia zależności tylko pomiędzy zmiennymi rozkładów eliptycznych. W sytuacjach, kiedy dane empiryczne pochodzą na przykład z rozkładów asymetrycznych i nieeliptycznych, charakteryzują się wysoką kurtozą lub skośnością, stosowanie współczynnika korelacji liniowej nie jest wskazane. W celu zbadania struktury zależności należy wówczas zastosować inną miarę.

Narzędziem, które doskonale nadaje się do modelowania niemal wszystkich typów zależności między szeregami czasowymi są kopule. Pojęcie kopuli, jako nazwy obiektu matematycznego, zostało wprowadzone po raz pierwszy w artykule Sklara [1959]. W ogólnym i zarazem najprostszym ujęciu kopula jest funkcją, która pozwala wyróżnić z dystrybuanty łącznego rozkładu wektora losowego składową opisującą tylko strukturę zależności.

Funkcje kopuli stały się powszechnie stosowanym w wielu dziedzinach wielowymiarowym narzędziem do modelowania, ze względu na ich łatwą implementację, niezmienniczość względem silnie rosnących przekształceń zmiennych losowych oraz fakt, że wybór konkretnej postaci może być dokonany spośród szerokiego wachlarza znanych rodzin kopul. Jedną z najważniejszych zalet kopuli jest jednak możliwość osobnego modelowania rozkładów brzegowych i rozkładu łącznego. Z tego powodu kopule mają bardzo duże zastosowanie w finansach. Umożliwiają one ścisłą analizę zależności, które występują, na przykład między cenami poszczególnych akcji, indeksów giełdowych lub instrumentów dłużnych.

Zagadnienie zależności pomiędzy cenami lub stopami zwrotu różnych instrumentów finansowych doczekało się szerokiego omówienia w literaturze przedmiotu. Na szczególną uwagę zasługuje w tym kontekście praca Gurgul i in. [2008], gdzie autorzy rozważają zależności równoczesne i przyczynowe pomiędzy stopami zwrotu, ich zmiennością i wielkością obrotów dla wybranych największych spółek giełdowych, tworzących indeks DAX. Równie interesujący jest artykuł Gurgula i Syreka [2007], w którym porównano klasyczną metodę Markowitza konstrukcji portfela z metodą opartą na kopule i teorii wartości ekstremalnych. Inne przykłady zastosowań kopuli w finansach to: alokacja zasobów, ocena zdolności kredytowej, modelowanie ryzyka niewypłacalności, zarządzanie ryzykiem [zob. Embrechts et al. 2003; Cherubini et al. 2004].

W kontekście badania współzależności za pomocą kopuli zaznaczyła się również popularność modelu warunkowej wariancji, w szczególności modelu GARCH w jego najczęstszej parametryzacji GARCH(1,1), stosowanego do modelowania stóp zwrotu instrumentów finansowych. W pracy Rocha i Alegre [2005] model tej klasy, uzupełniony o ARMA(1,1) w równaniu średniej, został użyty do wyestymowania warunkowych rozkładów brzegowy. Przefiltrowane w ten sposób szeregi zostały następnie użyte w procesie estymacji parametrów kopuli. Jak można zauważyć, wybór właściwego rozkładu warunkowego modelu GARCH odgrywa w praktyce kluczową rolę i decyduje o poprawności całego podejścia. W związku z tym w literaturze przedmiotu są prezentowane liczne modyfikacje podstawowej parametryzacji GARCH. Przykładem jest praca Jondeau i Rockingera [2006], w której autorzy wykorzystują skośny rozkład Hansena ze zmiennymi w czasie parametrami skośności i kurtozy.

Kopule mają również szerokie zastosowanie w modelowaniu ryzyka ubezpieczeniowego. Przykładem jest rachunek aktuarialny, gdzie kopule są używane do modelowania zależności pomiędzy śmiertelnością a stratami ubezpieczyciela [zob. Frees et al. 1996; Frees et al. 2005]. Trivedi i Zimmer [2006] wykorzystują z kolei kopule do modelowania ubezpieczenia kosztów leczenia, popytu na usługi medyczne i wyborów dokonywanych przez zamężne pary. Pełny model wspomnianych wyżej autorów zawiera dychotomiczne równanie wyboru dla rodzinnego ubezpieczenia kosztów leczenia i osobne równanie dla wykorzystania usług medycznych przez parę. W inżynierii kopule są stosowane w procesach wielowymiarowego sterowania i modelowania hydrologicznego [zob. Yan 2006; Genest i Faure 2007].

Niniejsza praca została poświęcona zbadaniu struktury współzależności cen kontraktów terminowych na kukurydzę, soję oraz pszenicę. Aby zrealizować ten cel, dopasowano rożne rozkłady teoretyczne do szeregów przetransformowanych cen. Następnie zaś wyznaczono zmienną, co do której przyjęto założenie o jednostajności na podstawie dystrybuanty najlepszego rozkładu opisującego proces przetransformowanych cen oraz wskazania testów dobroci dopasowania rozkładów teoretycznych. Dwuetapowa metoda największej wiarygodności została wykorzystana na etapie estymacji parametrów rozważanych trójwymiarowych kopuli, normalnej oraz *t*-studenta. Oprócz testowania dobroci dopasowania alternatywnych kopuli, zweryfikowano również zespół hipotez odnoszących się do postaci macierzy korelacji. Otrzymane wyniki empiryczne wskazują jednoznacznie na przewagę kopuli *t*-studenta jako narzędzia modelowania współzależności cen kontraktów terminowych na produkty rolne, o niestrukturyzowanej macierzy korelacji.

W pierwszym podrozdziale zaprezentowano elementarne podstawy teorii kopul. Kolejny został podzielony na dwie części. Pierwsza z nich zawiera charakterystykę próby badawczej oraz metodykę dopasowania rozkładów teoretycznych do danych empirycznych, wraz z omówieniem wyników dopasowania rozważanych rozkładów. Druga natomiast przedstawia metodologię estymacji parametrów funkcji kopuli oraz procedury testowania. Prezentacja wyników badań empirycznych dopasowania kopul znajduje się w podrozdziale trzecim. Podsumowanie stanowi przegląd najważniejszych wniosków sformułowanych w toku badania, z uwzględnieniem dywagacji na temat celowości i kierunków przyszłych wysiłków badawczych.

1. Elementy teorii kopul

Funkcje kopula, nazywane również funkcjami powiązań, to funkcje, które łączą wielowymiarową dystrybuantę zmiennej losowej z jej jednowymiarowymi dystrybuantami brzegowymi. Ogólnie można powiedzieć, że kopula jest wielowymiarowym rozkładem prawdopodobieństwa określonym na d-wymiarowej kostce $[0,1]^d$, z jednostajnymi rozkładami brzegowymi na odcinku $[0, 1]$. W podrozdziale tym przypomnimy krótko podstawowe informacje dotyczące elementów teorii kopul. Czytelnik zainteresowany szerzej tematem może znaleźć pełen przegląd zastosowań i własności funkcji kopuli w pracach: Joego [1997] oraz Nelsena [1999]. Rozpoczniemy od prezentacji definicji kopuli trójwymiarowej, która jest przedstawiona następująco:

Definicja 1

Kopulą trójwymiarową nazywamy każdą funkcję $C:[0,\,1]^3 \to [0,\,1]$, która spełnia następujące warunki:

1. Dla każdego $u_1, u_2, u_3 \in [0,\,1]$, mamy: $C(u_1, u_2, u_3) = 0$, jeżeli co najmniej jedna ze współrzędnych wynosi zero, czyli: $u_i = 0$, dla $i = 1,\,2,3$.

Inaczej mówimy po prostu, że funkcja C jest przytwierdzona lub uziemiona.

2. Dla każdego $u_1, u_2, u_3 \in [0,\,1]$, mamy: $C(u_1, 1, 1) = u_1$ oraz $C(1, 1, u_3) = u_3$.
3. C jest 3-rosnąca, czyli C-miara na każdej kostce, której wierzchołki leżą w $[0,\,1]^3$ jest nieujemna.

Jednym z podstawowych i niezwykle ważnych ze względu na analizę zależności wielowymiarowych twierdzeń dotyczącym teorii kopul jest twierdzenie Sklara [1959][1].

Podstawową konkluzją z twierdzenia Sklara jest fakt, że nieznany trójwymiarowy, jak i również uogólniając, wielowymiarowy, łączny rozkład, może być przybliżony najlepiej dopasowaną funkcją kopula oraz odpowiednimi rozkładami brzegowymi. Przydatne wydaje się również twierdzenie odwrotne, które przy podanym rozkładzie łącznym oraz dystrybuantach rozkładów brzegowych daje jawny wzór na funkcję kopuli. Należy jednak pamiętać, że twierdzenie odwrotne będzie działać natychmiastowo w przypadku zmiennych losowych ciągłych. W przeciwnym razie może się okazać, że odwrócenie funkcji nie zawsze jest możliwe.

Ważną własnością kopul jest ich niezmienniczość względem ściśle monotonicznych przekształceń. Własność ta oznacza, że struktura zależności między zmiennymi jest ujęta za pomocą kopuli, niezależnie od rozkładów brzegowych.

W części trzeciej artykułu, wykorzystując metodę największej wiarygodności, będziemy zajmować się estymacją parametrów funkcji kopuli, dlatego warto w tym miejscu wprowadzić pojęcie gęstości kopuli oraz jej reprezentację kanoniczną.

[1] W pracy Sklara z 1959 r. oraz w późniejszym artykule z 1974 r., którego autorami są Schweizer i Sklar można znaleźć dowody tego twierdzenia, przy czym dowód zamieszczony w publikacji z 1974 r. ma nieco mniej skomplikowaną postać niż dowód, który powstał jako pierwszy.

Definicja 2
Gęstość trójwymiarowej kopuli *C*, o ile istnieje, oznaczamy symbolem *c* i określamy wzorem:

$$c(u_1,u_2,u_3)=\frac{\partial^3 C(u_1,u_2,u_3)}{\partial u_1 \partial u_2 \partial u_3}. \tag{1}$$

Definicja 3
Dla ciągłych zmiennych losowych z trójwymiarową dystrybuantą łączną *H* jest związana gęstość rozkładu *h*, którą nazywamy reprezentacją kanoniczną *H*. Dana jest ona wzorem:

$$h(x_1,x_2,x_3)=c\left(F_1(x_1),F_2(x_2),F_3(x_3)\right)\cdot f_1(x_1)\cdot f_2(x_2)\cdot f_3(x_3), \tag{2}$$

gdzie f_1, f_2, f_3 – to gęstości rozkładów brzegowych odpowiednio F_1, F_2, F_3.

Wprowadźmy teraz definicje kopul, które będziemy stosować w naszych badaniach empirycznych. Podstawową rodzinę kopul stanowią tzw. kopule eliptyczne, wśród których do najczęściej stosowanych należą kopule gaussowskie oraz kopule t-studenta.

Kopula Gaussa
Definicja wielowymiarowej kopuli Gaussa, czy też ujmując inaczej kopuli wielowymiarowego rozkładu normalnego jest opisana następująco:

$$C_R^{Ga}(u_1,...,u_d)=\Phi_R\left(\Phi^{-1}(u_1),...,\Phi^{-1}(u_d)\right), \tag{3}$$

gdzie Φ_R oznacza dystrybuantę wielowymiarowego standaryzowanego rozkładu normalnego z macierzą korelacji *R*, natomiast Φ to standardowy, jednowymiarowy rozkład normalny. Wzór (3) można napisać również następująco:

$$C_R^{Ga}(u_1,\dots,u_d)=\int_{-\infty}^{\Phi^{-1}(u_1)}\dots\int_{-\infty}^{\Phi^{-1}(u_d)}\frac{1}{\sqrt{(2\pi)^d|R|}}\exp\left(-\frac{1}{2}x^T R^{-1}x\right)dx_1\dots dx_d. \tag{4}$$

Gęstość kopuli Gaussa wyraża się wzorem:

$$c_R^{Ga}(u_1,\dots,u_d)=\frac{1}{\sqrt{|R|}}\exp\left(-\frac{1}{2}\varsigma^T\left(R^{-1}-I\right)\varsigma\right), \tag{5}$$

gdzie, $\varsigma = \left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \ldots, \Phi^{-1}(u_d)\right)^T$.

Kopula t-studenta

Definicja wielowymiarowej kopuli t-studenta wygląda następująco:

$$C_{R,v}^{St}(u_1,...,u_d) = t_{R,v}\left(t_v^{-1}(u_1),...,t_v^{-1}(u_d)\right), \tag{6}$$

gdzie $t_{R,v}$ jest dystrybuantą wielowymiarową rozkładu t-studenta o v stopniach swobody, z macierzą korelacji R, natomiast t_v oznacza standardowy jednowymiarowy rozkład t-studenta o v stopniach swobody.
Wzór (6) definiujący kopulę t-studenta można również przedstawić następująco:

$$C_{R,v}^{St}(u_1,...,u_d) = \int_{-\infty}^{t_v^{-1}(u_1)} \cdots \int_{-\infty}^{t_v^{-1}(u_d)} \frac{\Gamma\left(\frac{v+d}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\sqrt{(v\pi)^d |R|}} \left(1+\frac{1}{v}x^T R^{-1} x\right)^{-\frac{v+d}{2}} dx_1 \ldots dx_d. \tag{7}$$

Gęstość kopuli t-studenta można opisać za pomocą wzoru:

$$c_{R,v}^{St}(u_1,...,u_d) = \frac{\Gamma\left(\frac{v+d}{2}\right)}{\sqrt{|R|}\Gamma\left(\frac{v}{2}\right)} \cdot \left(\frac{\Gamma\left(\frac{v}{2}\right)}{\Gamma\left(\frac{v+1}{2}\right)}\right)^d \frac{\left(1+\frac{1}{v}\varsigma^T R^{-1} \varsigma\right)^{-\frac{v+d}{2}}}{\prod_{j=1}^{d}\left(1+\frac{\varsigma_j^2}{v}\right)^{-\frac{v+1}{2}}}, \tag{8}$$

gdzie, $\varsigma_j = t_v^{-1}(u_j)$.

2. Opis próby badawczej, metodologia badania oraz procedura testowania

2.1. Opis próby badawczej i metodyka ustalenia najlepszego rozkładu teoretycznego

Próbę badawczą stanowiły dzienne notowania kontraktów terminowych dla trzech produktów, mianowicie kukurydzy, soi oraz pszenicy, notowanych na giełdzie towarowej w Chicago. Dobór produktów uwzględniał zarówno znaczenie dla obrotów na rynku terminowym, jak i dostępność odpowiednio długich szeregów czasowych. Dane pochodziły z lat 1975-2010 i obejmowały cenę zamknięcia kontraktu o najkrótszym terminie wygaśnięcia. W ten sposób szereg notowań można uważać za cenę terminową o najkrótszym możliwym terminie realizacji. Dane empiryczne zostały sprawdzone pod względem ewentualnych nieciągłości i błędów. Nie stosowano żadnych metod korygowania lub uzupełniania danych empirycznych, aby zminimalizować wpływ arbitralnych ingerencji na otrzymywane wyniki. Pojedynczy szereg danych liczył przeciętnie ponad dziesięć tysięcy obserwacji.

Szeregi cen instrumentów finansowych należą najczęściej do grupy procesów niestacjonarnych, których stopień integracji nie przekracza jednak jedności. Zazwyczaj też występuje autokorelacja, choć szybko zanikająca dla wyższych opóźnień [zob. Brzeszczyński i Kelm 2002; Weron i Weron 1999]. Aby wyeliminować zarówno niestacjonarność, jak i możliwą autokorelację badanych szeregów, wykorzystano model ARIMA [zob. Box i Jenkins 1983], który jest odpowiedni dla szeregów o pewnym skończonym i całkowitym stopniu integracji d oraz strukturze zależności zawierającej obok parametrów autoregresyjnych również parametry średniej ruchomej błędów. Ogólnie postać ARIMA wyraża się wzorem:

$$\left(1-\sum_{i=1}^{p}\phi_i L^i\right)(1-L)^d y_t=\left(1-\sum_{i=0}^{q}\theta_i L^i\right)\varepsilon_t, \tag{9}$$

gdzie L oznacza operator opóźnienia.

Estymację modelu ARIMA powinno poprzedzać wyznaczenie stopnia integracji szeregu d oraz identyfikacja właściwej postaci modelu, tzn. określenie liczby parametrów autoregresyjnych (p) i średniej ruchomej (q). Dla zbadania stopnia integracji szeregów cen rozważanych produktów rolnych wykorzystano uogólniony test DF [zob. Charemza i Deadman 1997; Dickey i Fuller 1979, 1981]. Liczba opóźnień w teście została dobrana w sposób empiryczny, tzn. testowanie rozpoczęto dla maksymalnego opóźnienia, a następnie w kolejnej rundzie testu opóźnienie zmniejszano o jeden, aż do odrzucenia hipotezy zerowej o istnieniu pierwiastka jednostkowego. Testowanie pozwoliło na ustalenie stopnia integracji szeregów cen produktów rolnych na $d=1$. Wybór konkretnej parametryzacji modelu przebiegał dwuetapowo. Ogląd przebiegu funkcji autokorelacji i funkcji autokorelacji cząstkowej dostarczył ogólnych wskazówek co do liczby parametrów autoregresyjnych i średniej ruchomej [zob. Gurgul i Majdosz 2003]. Ostatecznym sprawdzianem była wartość kryterium informacyjnego AIC. Estymację modelu ARIMA przeprowadzono metodą największej wiarygodności, zbadano statystyczną istotność parametrów oraz wyznaczono szeregi reszt, które posłużyły za podstawę dalszego badania. Czytelnik zainteresowany przeglądem wyników estymacji dla poszczególnych modeli znajdzie go w pracy Malik [2011].

W celu jak najlepszego scharakteryzowania statycznych własności rozkładu cen produktów rolnych na rynku terminowym zbadano ich zgodność z różnymi rozkładami teoretycznymi. Wybrano najpopularniejsze rozkłady, których przydatność do opisu procesu cen znajduje potwierdzenie w literaturze przedmiotu [zob. Suder et al. 2004; Aparicio i Estrada 2001], a mianowicie: rozkład logistyczny, skalowany rozkład t-studenta, rozkład potęgowo-wykładniczy, rozkład hiperboliczny, mieszanka rozkładów normalnych, rozkład normalny odwrotny gaussowski oraz rozkład α-stabilny. W celu dopasowania rozkładów teoretycznych w większości przypadków wykorzystano metodę największej wiarygodności. Wyjątek stanowiła mieszanka rozkładów normalnych, gdzie ze względu na problem nieokreśloności parametru prawdopodobieństwa dla zwyczajnej metody największej wiarygodności, wykorzystano maksymalizację wartości oczekiwanej. Ze względu na ograniczoną obszerność niniejszego artykułu nie prezentujemy wyników estymacji parametrów rozkładów dla rozważanych szeregów danych, ani wyników testowania dobroci dopasowania, przeprowadzonych za pomocą trzech testów najczęściej rekomendowanych w literaturze [zob. Weron i Weron 1999], a mianowicie testu χ^2, testu Kołmogorowa i testu Andersona-Darlinga. Czytelnik zainteresowany wynikami znajdzie pełen ich przegląd w pracy Malik [2011].

Spośród siedmiu rozważanych rozkładów teoretycznych najlepiej do opisu rozkładu cen produktów rolnych notowanych na giełdzie towarowej w Chicago nadają się rozkłady, które są w stanie modelować zjawisko grubych ogonów, w tym w szczególności skalowany rozkład t-studenta oraz rodzina rozkładów α-stabilnych. Wniosek taki sugerują wyniki testowania dobroci dopasowania na podstawie trzech wspomnianych powyżej testów, dla których wartości krytyczne otrzymano metodami symulacyjnymi [Malik 2011].

2.2. Metodologia estymacji parametrów funkcji kopuli i procedury testowania

Jedną z podstawowych i zarazem najpopularniejszą metodą estymacji parametrów jest metoda największej wiarygodności. Teoria kopul pokazuje, że każdą wielowymiarową łączną dystrybuantę zmiennej losowej możemy rozdzielić na jednowymiarowe dystrybuanty brzegowe i funkcję kopula, która opisuje zależność pomiędzy zmiennymi. Niech $X = (X_1, X_2, ..., X_d)$ będzie wektorem zmiennych losowych, gdzie X_i posiada jednowymiarową funkcję dystrybuanty F_i z parametrem α_i, oznaczaną $F_i(x_i;\alpha_i)$ i odpowiadającą jej gęstość $f_i(x_i;\alpha_i)$. Parametry odpowiedzialne za rozkłady brzegowe oznaczmy przez $\theta_1 = (\alpha_1^T, ..., \alpha_d^T)^T$, natomiast parametry związane z kopulą przez θ_2 i niech $\theta = (\theta_1^T, \theta_2^T)^T$. Rozważmy model postaci:

$$F\left(x_1,\ldots,x_d;\theta\right)=C\left(F_1\left(x_1;\alpha_1\right),\ldots,F_d\left(x_d;\alpha_d\right);\theta_2\right), \tag{10}$$

gdzie C jest kopulą z wektorem parametrów θ_2, przy czym pomiędzy parametrami θ_1,θ_2 nie występują żadne wiążące je ograniczenia. Dla danych rozkładów brzegowych można sformułować różne wielowymiarowe dystrybuanty wykorzystując różne funkcje kopuli. Zakładając, że kopula C ma gęstość c, dla zadanej próbki $x_t=\left(x_{t1},x_{t2},\ldots,x_{td}\right)$, gdzie $t=1,\ldots,T$ logarytmiczną funkcję wiarygodności modelu (10) można, wykorzystując zależność (2), zapisać za pomocą wzoru poniżej, dekomponując ją na dwie składowe:

$$l(\theta)=\sum_{t=1}^{T}\ln c\left(F_1\left(x_{t1};\alpha_1\right),\ldots,F_d\left(x_{td};\alpha_d\right);\theta_2\right)+\sum_{t=1}^{T}\sum_{j=1}^{d}\ln f_j\left(x_{tj};\alpha_j\right). \tag{11}$$

Estymacja metodą największej wiarygodności zakłada maksymalizację logarytmu funkcji wiarygodności w odniesieniu do wszystkich parametrów jednocześnie. W praktyce jednak, ze względu na skomplikowaną postać funkcji wiarygodności, co w konsekwencji wpływa na nadmierną złożoność obliczeniową, trudno jest osiągnąć ten cel. Po za tym, estymując jednocześnie wszystkie parametry nie wykorzystalibyśmy zasadniczej własności teorii kopul, a mianowicie możliwości rozdzielenia wielowymiarowej dystrybuanty na rozkłady brzegowe i funkcję kopula. Z tych właśnie powodów wprowadzono dwustopniową metodę największej wiarygodności. Metoda ta jest nazywana metodą funkcji wnioskowania dla rozkładów brzegowych, w skrócie IFM (ang. *the method of Inference Functions for Margins*), a zaproponowano ją w pracach Joego i Xu w 1996 r. oraz Joego w 1997 r. Procedura przebiega dwuetapowo. W pierwszym kroku za pomocą metody największej wiarygodności są wyestymowane parametry rozkładów brzegowych. W tym celu maksymalizujemy funkcję:

$$L_j\left(\alpha_j\right)=\sum_{t=1}^{T}\ln f_j\left(x_{tj};\alpha_j\right),\quad j=1,\ldots d. \tag{12}$$

W kroku drugim z kolei, wykorzystując już wyestymowane parametry $\hat{\theta}_1$, przeprowadzamy estymację parametrów kopuli jako:

$$\hat{\theta}_2=\underset{\theta_2}{ArgMax}\sum_{t=1}^{T}\ln c\left(F_1\left(x_{t1};\hat{\alpha}_1\right),\ldots,F_d\left(x_{td};\hat{\alpha}_d\right);\theta_2\right). \tag{13}$$

Patton w swojej pracy [2006] wykazał, że pod pewnymi warunkami regularności metoda IFM daje zgodne i asymptotycznie normalne estymatory. Należy jednak pamiętać, że istnieje pewna strata efektywności estymacji, ponieważ krok pierwszy nie bierze pod uwagę zależności między rozkładami brzegowymi.

Aby sprawdzić czy struktura zależności wyestymowanego modelu kopuli jest odpowiednio dopasowana zastosujemy testy zgodności. W literaturze można znaleźć kilka propozycji dotyczących testów zgodności dla modeli kopuli. W niniejszym artykule test dobroci dopasowania kopuli do danych empirycznych opiera się na statystyce Cramer-Mises zdefiniowanej w pracy Genesta i in. [2009], dla której wartości krytyczne zostały wyznaczone za pomocą boostrapu parametrycznego [zob. też Kojadinovic i Yan 2011; Kojadinovic et al. 2011; Nikoloulopoulos i Karlis 2008].

3. Wyniki empiryczne dopasowania kopul

Kopule eliptyczne, do których należą zarówno kopula Gaussa, jak i kopula t-studenta posiadają macierze korelacji, wywiedzione jako konsekwencja rozkładu eliptycznego, przy czym kopula t-studenta ma jeszcze dodatkowy parametr odpowiedzialny za liczbę stopni swobody. Jak wynika z teorii zawartej w podrozdziale 1, kopule są niezmiennicze względem ściśle monotonicznych przekształceń rozkładów brzegowych, dlatego strukturę zależności między różnymi obserwacjami empirycznymi możemy określić za pomocą macierzy korelacji. Wprowadzone do powszechnego stosowania struktury zależności to między innymi struktura wymienna, gdzie wszystkie współczynniki korelacji są sobie równe oraz macierz niestrukturyzowana o wszystkich parametrach różnych [zob. Yan 2007]. Odpowiednie macierze korelacji w przypadku wymiaru równego 3 dane są za pomocą wzorów:

$$ex := \begin{bmatrix} 1 & p & p \\ p & 1 & p \\ p & p & 1 \end{bmatrix}, \qquad un := \begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_3 \\ \rho_2 & \rho_3 & 1 \end{bmatrix}, \tag{14}$$

gdzie, p oraz $\rho_j,\ j = 1,2,3$ oznaczają współczynniki korelacji. Pierwsza macierz (*ex*) zakłada, że wszystkie współczynniki korelacji są jednakowe, podczas gdy druga (*un*) dopuszcza różne współczynniki dla każdej pary zmiennych. Macierze korelacji (14) pozwalają zbadać zależności, jakie mogą występować pomiędzy trzema badanymi produktami rolnymi notowanymi na giełdzie towarowej w Chicago, mianowicie: kukurydzą, soją i pszenicą. W przypadku macierzy *ex* wyestymowane współczynniki korelacji są takie same, co możemy zinterpretować nie tylko jako dowód silnej lub słabej zależności pomiędzy cenami każdej z par produktów, lecz także jako przejaw wysokiego stopnia jednorodności grupy produktów rolnych pod względem struktury wewnętrznych zależności. W macierzy *un* współczynniki korelacji są z kolei różne, co oznacza, że mimo przynależności do tej samej kategorii produktów rolnych, każda para produktów powinna być rozpatrywana oddzielnie z punktu widzenia jej struktury zależności.

Wnioski otrzymane na podstawie badania dobroci dopasowania różnych rozkładów teoretycznych do szeregu przefiltrowanych cen produktów rolnych dowodzą, że rozkłady, które najlepiej nadają się do opisu to skalowany rozkład t-studenta (dla kukurydzy) oraz rodzina rozkładów α-stabilnych (dla soi i pszenicy). Użycie dystrybuant odpowiednich rozkładów z parametrami wyestymowanymi na podstawie danych empirycznych daje zatem wystarczającą podstawę do twierdzenia, że przetransformowany za pomocą dystrybuanty szereg ma rozkład jednostajny. Na podstawie skonstruowanych w ten sposób szeregów danych przeprowadzono następnie estymację parametrów rozważanych kopuli (drugi etap estymacji). Wyniki zostały zebrane w tabeli 1.

Tabela 1

Wyniki estymacji parametrów kopuli

	p	**kukurydza – soja** ρ_1	**kukurydza – pszenica** ρ_2	**soja – pszenica** ρ_3
Kopula Gaussa	0,5025 (4,453E-03)	0,6005 (5,060E-03)	0,4926 (5,915E-03)	0,4077 (7,248E-03)
Kopula t-studenta	0,5078 (2,139E-04)	0,6107 (1,579E-04)	0,4966 (1,562E-04)	0,4100 (8,188E-05)

* W nawiasach okrągłych podano wartości błędów estymacji odpowiednich parametrów.

Kolejnym etapem badania było sprawdzenie, czy struktura zależności opisana za pomocą wyestymowanych kopuli jest dostatecznym przybliżeniem rzeczywistości i nadaje się do modelowania danych empirycznych. W tym celu zastosowano test dobroci dopasowania, opierający się na statystyce Cramer-Mises, dla której wartości krytyczne zostały wyznaczone za pomocą boostrapu parametrycznego. Wyniki testowania dobroci dopasowania znajdują się w tabeli 2.

Tabela 2

Wyniki testowania dobroci dopasowania kopul

Kopula	**Macierz**	**Statystyka Cramer-Mises**
Kopula Gaussa	***ex***	0,4759***
	un	0,2222***
Kopula t-studenta	***ex***	0,3012***
	un	0,1214

* – istotność na poziomie 10%, ** – istotność na poziomie 5%, *** – istotność na poziomie 1%.

Obraz wyłaniający się z wyników testów dobroci dopasowania wskazuje jednoznacznie na odrzucenie kopuli normalnej jako przydatnej do opisu zależności między badanymi produktami rolnymi zarówno dla macierzy wskazującej na jednakowe współczynniki korelacji, jak i dla macierzy o różnych współczynnikach korelacji. W przypadku kopuli *t*-studenta wskazania testów dobroci dopasowania są różne dla hipotezy o równych wszystkich współczynnikach korelacji (macierz *ex*) oraz hipotezy zakładającej różne współczynniki korelacji dla każdej pary badanych produktów rolnych (macierz *un*). O ile bowiem w pierwszym przypadku statystyka testowa jest istotna na wysokim (1%) poziomie istotności, w drugim – brak podstaw do odrzucenia hipotezy na zwyczajowo przyjmowanym poziomie istotności (statystyka testowa pozostaje nieistotna nawet dla 10% poziomu istotności). Otrzymane wyniki empiryczne wskazują zatem jednoznacznie na kopule *t*-studenta jako właściwą do opisu współzależności cen produktów rolnych. Każda para produktów powinna jednak być rozpatrywana oddzielnie, co wyraża niestrukturyzowana macierz korelacji.

Podsumowanie

W artykule tym autorzy zbadali zależności, jakie mogą występować pomiędzy produktami rolnymi notowanymi na giełdzie towarowej w Chicago. Próbę badawczą stanowiły trzy podstawowe produkty, a mianowicie: kukurydza, soja oraz pszenica. Określenie struktury zależności ma istotne znaczenie w zarządzaniu instrumentami finansowymi, a błędna interpretacja siły powiązań może prowadzić do niewłaściwych decyzji inwestycyjnych. Wyniki empiryczne dostarczyły przekonujących dowodów, że kopula t-studenta zdecydowanie lepiej modeluje zależności pomiędzy badanymi produktami rolnymi niż kopula Gaussa. Wyestymowane współczynniki korelacji okazały się natomiast różne dla każdej pary zmiennych, co może sugerować znaczne zróżnicowanie własności cen poszczególnych produktów rolnych między sobą.

Niniejsze badanie jest pierwszym zamiarem autorów w powyższym obszarze. Jako takie nie stanowi ono kompleksowego ujęcia badanego problemu, lecz raczej pierwszy krok na drodze wiodącej ku temu celowi. Prezentowane wyniki należy zatem oceniać przez pryzmat ograniczeń przyjętego podejścia. Autorzy są ich w pełni świadomi i mają nadzieję na kontynuowanie badań w powyższym zakresie w przyszłości, w celu wyeliminowania przynajmniej niektórych z nich. Wysiłek badawczy powinien się skoncentrować w szczególności na dwóch obszarach. Pierwszym z nich jest pełniejsze zweryfikowanie założeń dotyczących rozkładów brzegowych modelowanych kopuli. Wydaje się celowe rozważenie alternatywnych modeli cen, w tym modeli klasy GARCH i ARIMA-GARCH oraz pogłębione testowanie hipotezy przetransformowanych szeregów reszt. Drugim obszarem, w którym autorzy dostrzegają możliwość ulepszeń jest dobór wyjściowego zbioru funkcji kopuli do przetestowania. Oprócz kopuli normalnej oraz *t*-studenta, warto również uwzględnić kopule, które dopuszczają asymetryczną relację w ogonach rozkładu.

Literatura

Aparicio F.M., Estrada J. (2001): *Empirical Distribution of Stock Returns: European Securities Markets 1990-1995*. „The European Journal of Finance", No. 7.

Box G., Jenkins G. (1983): *Analiza szeregów czasowych*. PWN, Warszawa.

Brzeszczyński J., Kelm R. (2002): *Ekonometryczne modele rynków finansowych*. WIG-Press, Warszawa.

Charemza W.W., Deadman D.F. (1997): *Nowa ekonometria*. PWE, Warszawa.

Cherubini U., Luciano E., Vecchiato W. (2004): *Copula Methods in Finance*. John Wiley & Sons, New York.

Dickey D.A., Fuller W.A. (1979): *Distribution of the Estimators for Autoregressive Time Series with a Unit Root*. „Journal of the American Statistical Association", No. 74.

Dickey D.A., Fuller W.A. (1981): *Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root*. „Econometrica", No. 49.

Embrechts P., Lindskog F., McNei L.A. (2003): *Modelling Dependence with Copulas and Applications to Risk Management*. In: *Handbook of Heavy Tailed Distribution in Finance*. Ed. S. Rachev.

Frees E.W., Wang P. (2005): *Credibility Using Copulas*. „North American Actuarial Journal", No. 9(2).

Frees E.W., Carriere J., Valdez E.A. (1996): *Annuity Valuation with Dependent Mortality*. „Journal of Risk and Insurance", No. 63.

Genest C., Favre A.C. (2007): *Everything You Always Wanted to Know about Copula Modeling abut Were Afraid to Ask*. „Journal of Hydrologic Engineering", No. 12.

Genest C., Remillard B., Beaudoin D. (2009): *Goodness-of-fit Tests for Copulas: A Review and a Power Study*. „Mathematics and Economics", No. 44.

Gurgul H., Majdosz P. (2003): *Analiza empiryczna efektu polepszania wyników w sektorze otwartych funduszy emerytalnych w Polsce*. Folia Oeconomics Cracoviensia.

Gurgul H., Syrek R. (2007): *Wykorzystanie kopul do konstrukcji portfeli inwestycyjnych*. „Ekonomia Menedżerska", nr 2.

Gurgul H., Mestel R., Syrek R. (2008): *Kopule i przyczynowość w badaniach związków pomiędzy zmiennymi finansowymi wybranych spółek z DAX*. „Ekonomia Menedżerska", nr 3.

Joe H. (1997): *Multivariate Models and Dependence Concepts*. Chapman and Hall, London.

Joe H., Xu J.J. (1996): *The Estimation Method of Inference Function for Margins for Multivariate Models*. Department of Statistics, University of British Columbia, Technical Report.

Jondeau E., Rockinger M. (2006): *The Copula-GARCH Model of Conditional Dependencies: An International Stock Market Application*. „Journal of International Money and Finance", No. 25.

Kojadinovic I., Yan J. (2011): *A Goodness-of-fit Test for Multivariate Multiparameter Copulas Based on Multiplier Central Limit Theorems*. „Statistics and Computing", 21.

Kojadinovic I., Yan J., Holmes M. (2011): *Fast Large-sample Goodness-of-fit Tests for Copulas*. „Statistica Sinica", 21.

Malik G. (2011): *Statystyczne własności rozkładu cen produktów rolnych na rynku terminowym na przykładzie giełdy towarowej w Chicago* W: *Nauka i gospodarka w dobie destabilizacji*. Red. J. Teczke, J. Czekaj, B. Mikuła, R. Oczkowska. Biuro Projektu Nauka i Gospodarka, Kraków.

Nelsen R.B. (1999): *An Introduction to Copulas*. Springer Verlag, New York.

Nikoloulopoulos A.K., Karlis D. (2008): *Copula Model Evaluation Based on Parametric Bootstrap*. „Computational Statistics & Data Analysis", 52.

Patton A.J. (2006): *Estimation of Multivariate Models for Time Series of Possibly Different Lengths*. „Journal of Applied Econometrics", Vol. 21 (2).

Roch O., Alegre A. (2005): *Testing the Bivariate Distribution of Daily Equity Returns Using Copulas. An Application to the Spanish Stock Market*. „Computational Statistics & Data Analysis", 51.

Schweizer B., Sklar A. (1974): *Operations on Distributions Functions not Derivable from Operations on Random Variables*. „Studia Mathematica", 52.

Sklar A. (1959): *Fonctions de répartition á n dimensions et leurs marges*. Publications de l'Institut Statistique de l'Université de Paris 8.

Suder M., Wolak J., Wójtowicz T. (2004): *Własności dziennych stóp zwrotu na przykładzie indeksów giełd europejskich* W: *Zarządzanie przedsiębiorstwem w warunkach integracji europejskiej*. Uczelniane Wydawnictwo Naukowo-Dydaktyczne Akademia Górniczo-Hutnicza.

Trivedi P.K., Zimmer D.M. (2006): *Using Trivariate Copulas to Model Sample Selection and Treatment Effects: Application to Family Health Care Demand*. „Journal of Business and Economic Statistics", 24.

Weron A., Weron R. (1999): *Inżyneria finansowa*. Wydawnictwa Naukowo-Techniczne, Warszawa.

Yan J. (2006): *Multivariate Modeling with Copulas and Engineering Applications*. W: *Handbook in Engineering Statistics*. Red. H. Pham.

Yan J. (2007): *Enjoy the Joy of Copulas: With a Package copula*. „Journal of Statistical", 21.

COPULA-BASED MODELING THE CORRELATIONS BETWEEN COMMODITY FUTURES PRICES QUOTED ON THE CME

Summary

The goal of this article was to investigate the correlations between futures prices of commodities quoted on the CME. The sample includes corn, soybeans and wheat. Using ARIMA model for which best parameterization was identified based upon the AIC value, the raw time series of the prices for the contract with the shortest time left to expiration were subject to the process of removing a stochastic trend as well as autocorrelation. The transformed time series were then used as an input in fitting various theoretical distributions whose practical importance in describing the process of prices had been proven in the literature. The unknown parameters were estimated by means of the ML. Three different tests, namely χ^2, Kolomogorov and AD, were employed in order to investigate/verify the goodness-of-fit of these distributions. Finally, the parameters of normal as well as t copulas were estimated by means of the two-step ML method, with different hypotheses concerning the form of a correlation matrix. The goodness-of-fit test based on Cramer-Mises statistic was used to choose between the alternative copulas, with the critical values being obtained via non-parametric boostrap.