

Analiza porównawcza cen ropy i gazu

Mackiewicz-Kubiak Aleksander, Pagielska Marta

2025-?-?

Pakiety

Do analizy użyto następujących pakietów:

```
library(tidyr)
library(gamlss)
library(dplyr)
library(fitdistrplus)
library(usefun)
library(quantmod)
library(scales)
library(copula)
library(psych)
library(MVN)
library(readr)
library(xts)
options(scipen = 999) # celem wyłączenia notacji naukowej
pdf.options(encoding='ISOLatin2.enc')
```

KROK 1

Do analizy wybrano ceny paliw: ropy oraz gazu ziemnego pochodzące z okresu ponad 20 lat (począwszy od 4 stycznia 2000 r. do 17 czerwca 2022 r. włącznie). Ropa oraz gaz są surowcami energetycznymi, których ceny często reagują na te same czynniki, np. napięcia geopolityczne, decyzje OPEC czy kryzysy energetyczne. Wybrana ropa to ropa typu Brent, która wykorzystywana jest jako jeden z najważniejszych światowych benchmarków cen ropy naftowej. Wydobywana jest z Morza Północnego i służy jako punkt odniesienia dla wyceny około dwóch trzecich globalnych dostaw ropy. Jej cena wpływa zatem na ceny paliw i energii na całym świecie.

Wczytanie danych z pliku CSV:

##	Symbol	Date	Open	High	Low	Close	Volume	Currency
## 1	Brent Oil	2000-01-04	23.90	24.70	23.89	24.39	32509	USD
## 2	Brent Oil	2000-01-05	24.25	24.37	23.70	23.73	30310	USD
## 3	Brent Oil	2000-01-06	23.55	24.22	23.35	23.62	44662	USD
## 4	Brent Oil	2000-01-07	23.57	23.98	23.05	23.09	34826	USD
## 5	Brent Oil	2000-01-10	23.04	23.78	23.04	23.73	26388	USD
## 6	Brent Oil	2000-01-11	23.90	24.67	23.83	24.62	17410	USD

Podział danych na ceny ropy i gazu:

```
oil <- subset(data, Symbol == "Brent Oil", select = c(Date, Close))
gas <- subset(data, Symbol == "Natural Gas", select = c(Date, Close))

oil$Date <- as.Date(oil$Date)
gas$Date <- as.Date(gas$Date)

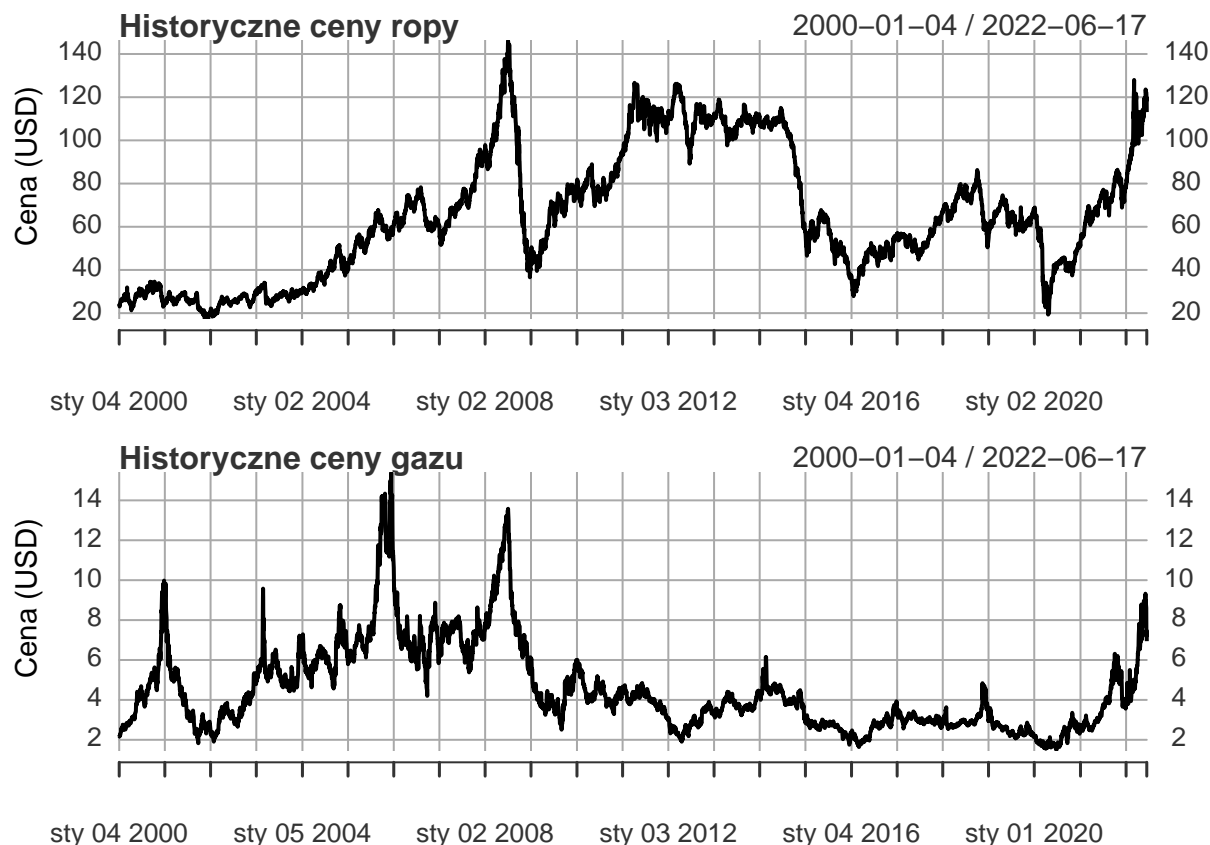
oil <- xts(oil$Close, order.by = oil$Date)
colnames(oil) <- "oil.Close"
gas <- xts(gas$Close, order.by = gas$Date)
colnames(gas) <- "gas.Close"

head(oil); head(gas)
```

```
##          oil.Close
## 2000-01-04      24.39
## 2000-01-05      23.73
## 2000-01-06      23.62
## 2000-01-07      23.09
## 2000-01-10      23.73
## 2000-01-11      24.62
```

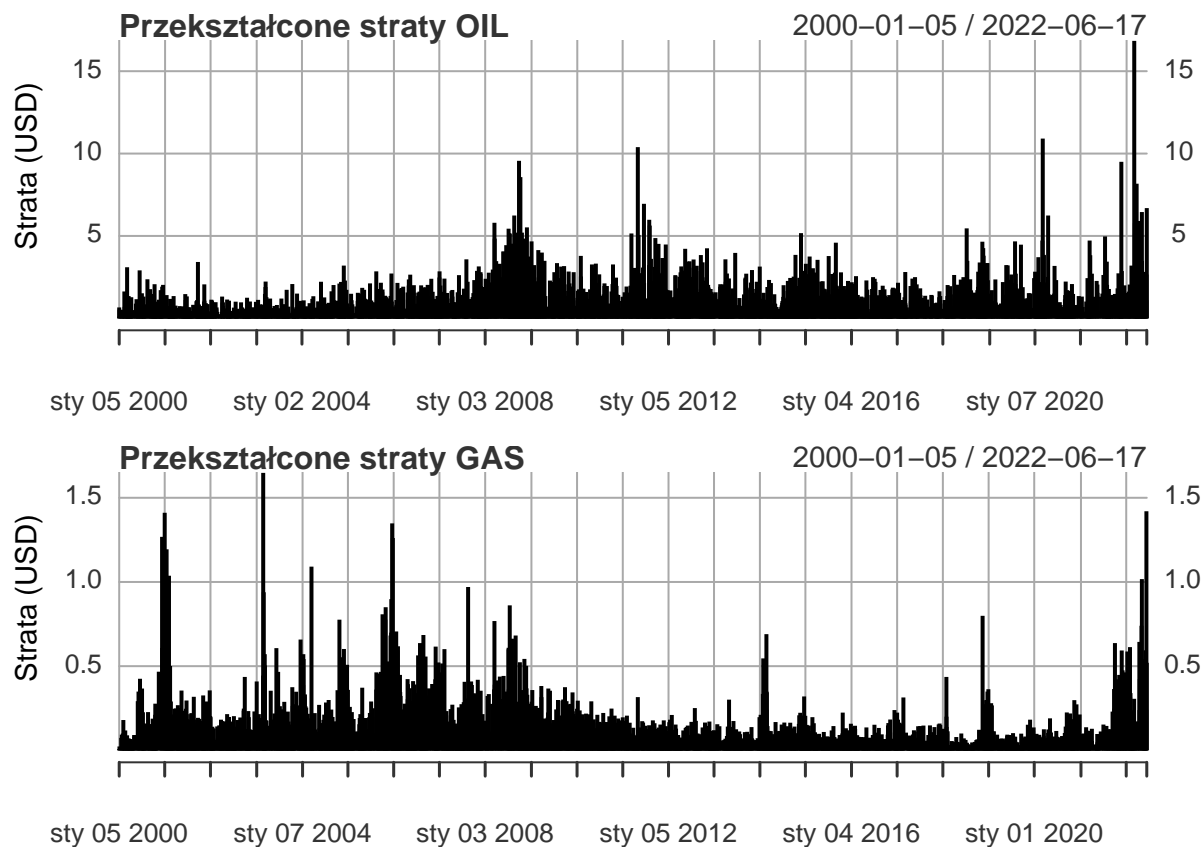
```
##          gas.Close
## 2000-01-04       2.176
## 2000-01-05       2.168
## 2000-01-06       2.196
## 2000-01-07       2.173
## 2000-01-10       2.216
## 2000-01-11       2.260
```

Wizualizacja historycznych cen ropy i gazu.



Ceny ropy wykazują duże wahania, w tym gwałtowny wzrost przed rokiem 2008, kiedy to nastąpił kryzys finansowy. Po roku 2014 zaobserwować można spadek cen ropy, zapewne związany z rewolucją łupkową. Wyraźne zaburzenia można odnotować w trakcie pandemii COVID-19, a zatem w okolicach roku 2020. Ogólny trend cen ropy można opisać jako zmienny, w tym silnie reagujący na czynniki globalne i geopolityczne. Ceny gazu ziemnego również cechują się zmiennością, ale w nieco innych momentach niż ropa, co może wskazywać na częściową niezależność rynków. Gwałtowne wzrosty cen gazu pojawiają się w momentach wysokiego zapotrzebowania energetycznego, takich jak kryzysy zimowe czy wojna na Ukrainie.

Straty obliczamy jako różnice wartości cen zamknięcia między kolejnymi dniami. Dodatkowo, w przypadku gdy mamy do czynienia z zyskami a nie stratami, to usuwamy te wartości, gdyż nie one są celem analizy. Dodatkowo, przekształcamy ujemne wartości start na wartości dodanie, by ułatwić dalsze analizy i obliczenia.



Wykresy strat również pokazują nieregularność między cenami gazu i ropy, okresy dużych lub małych strat zupełnie się nie pokrywają.

KROK 2

Obliczenie podstawowych statystyk:

```
##      Index      oil.Close
##  Min.   :2000-01-05  Min.    : 0.010
## 1st Qu.:2005-10-29  1st Qu.: 0.310
## Median :2011-05-16  Median : 0.680
## Mean   :2011-03-29  Mean    : 1.009
## 3rd Qu.:2016-06-14  3rd Qu.: 1.343
## Max.   :2022-06-17  Max.    :16.840
```

```
##      Index      gas.Close
##  Min.   :2000-01-05  Min.    :0.0010
## 1st Qu.:2005-11-28  1st Qu.:0.0320
## Median :2011-06-13  Median :0.0700
## Mean   :2011-05-19  Mean    :0.1142
## 3rd Qu.:2016-12-06  3rd Qu.:0.1410
## Max.   :2022-06-17  Max.    :1.6480
```

Widać znacząco różnice między maksymalnymi wartościami, co może wynikać z faktu, że ceny ropy są o wiele większe od cen gazu i dlatego też straty w jej cenach są większe. Mediany obu szeregów strat są

bardzo podobne, mimo dużej różnicy w średniej, co może oznaczać, że oba rozkłady są prawoskośne (czego oczekujemy po tych rozkładach strat), a większe możliwe spadki ceny ropy mocniej wpływają na średnią. Oba minima są powyżej 0, co oznacza, że poprzednie transformacje danych zadziałały tak jak chcieliśmy.

Problemem powyższych przekształceń jest fakt, że szeregi strat dla obu firm niekonieczne muszą być równej długości. Ponieważ naszym celem jest badanie zależności tych szeregów, należy zadbać by były one równej długości. W tym celu wyselekcjonujemy dane z dni, w których obie firmy zanotowały stratę, by te szeregi były jak najbardziej porównywalne, jednocześnie sprawdzając, czy nie usuniemy zbyt dużo danych.

```
## [1] "Długość nowych szeregów strat:"
```

```
## [1] 1557
```

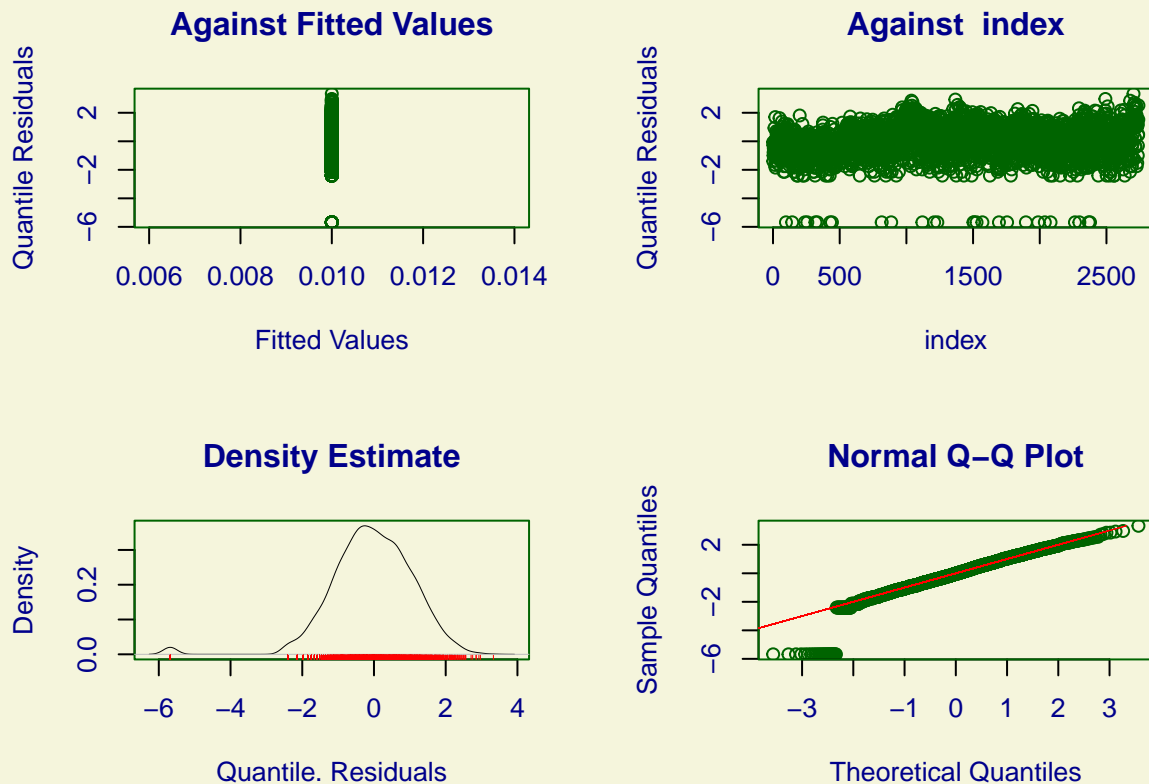
```
## [1] "Ilość straconych wartości dla dłuższego szeregu"
```

```
## [1] 1304
```

Mimo, że usuneliśmy prawie 50% danych to nowe, równe szeregi mają wciąż wystarczającą liczbę obserwacji, by móc przeprowadzać dla nich wiarygodne analizy.

Dopasowanie rozkładów:

```
##
## Family: c("ST3", "skew t type 3")
## Fitting method: "nlminb"
##
## Call: gamlssML(formula = y, family = DIST[i])
##
## Mu Coefficients:
## [1] 0.01
## Sigma Coefficients:
## [1] -9.512
## Nu Coefficients:
## [1] 9.413
## Tau Coefficients:
## [1] 1.057
##
## Degrees of Freedom for the fit: 4 Residual Deg. of Freedom 2736
## Global Deviance: 5460.73
## AIC: 5468.73
## SBC: 5492.39
```

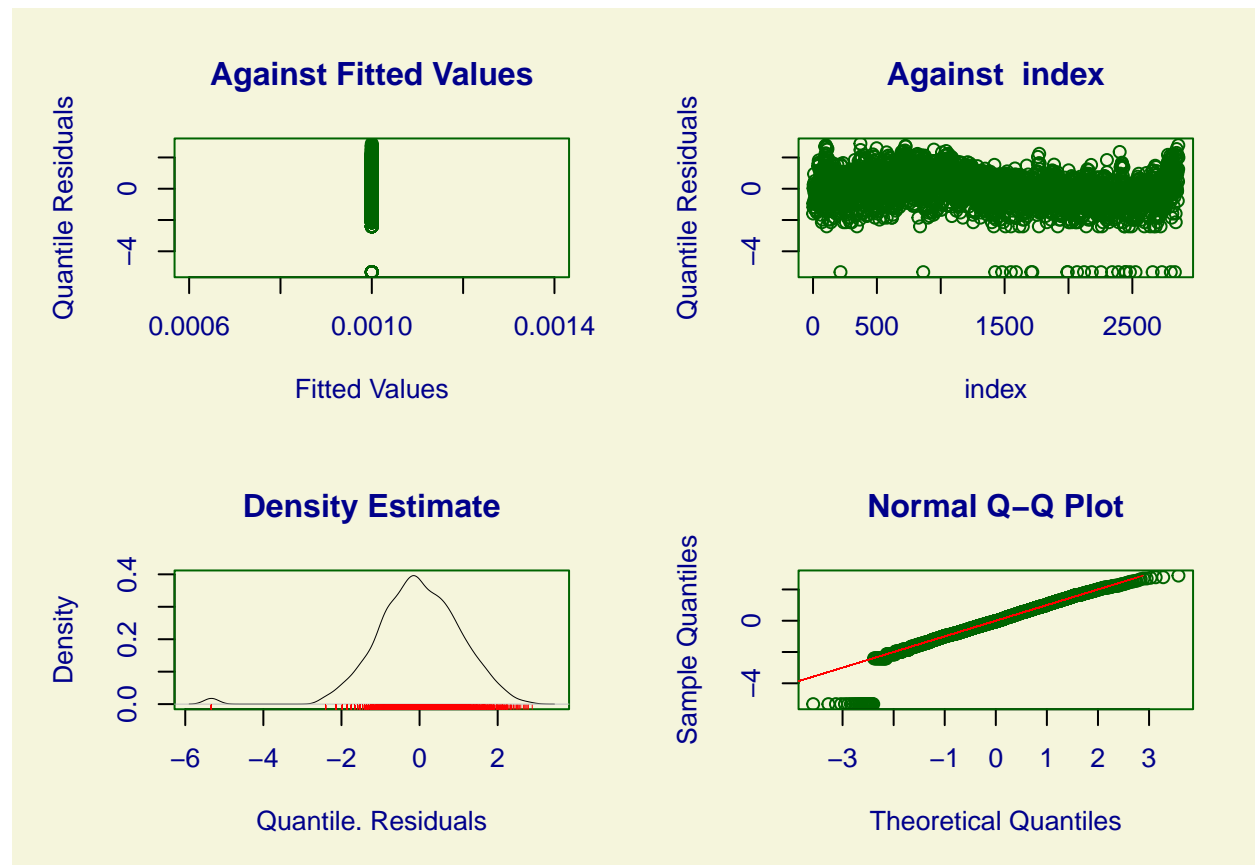


```
## *****
##      Summary of the Quantile Residuals
##              mean   = -0.0549827
##              variance = 1.311312
##              coef. of skewness = -1.029483
##              coef. of kurtosis = 7.304407
## Filliben correlation coefficient = 0.9696299
## *****
```

Dla wartości strat dla cen ropy został dopasowany rozkład ST3. Zwarte skupienie punktów na pierwszym wykresie sugeruje brak istotnej zależności reszt od wartości dopasowanych. Rozkład reszt względem indeksu wskazuje na losowość błędów, bez systematyczności czy struktury czasowej. Estymacja gęstości rozkładu reszt pokazuje, że reszty mają rozkład normalny. Wykres kwantyl-kwantyl również sugeruje normalność rozkładu reszt, poza widocznym ciężkim ogonem z lewej strony wykresu.

```
##
## Family: c("ST3", "skew t type 3")
## Fitting method: "nlminb"
##
## Call:  gamlssML(formula = y, family = DIST[i])
##
## Mu Coefficients:
## [1] 0.001
## Sigma Coefficients:
## [1] -10.84
```

```
## Nu Coefficients:
## [1] 8.418
## Tau Coefficients:
## [1] 0.7775
##
## Degrees of Freedom for the fit: 4 Residual Deg. of Freedom 2857
## Global Deviance: -6856.64
## AIC: -6848.64
## SBC: -6824.81
```



```
## *****
## Summary of the Quantile Residuals
## mean = -0.03258705
## variance = 1.205787
## coef. of skewness = -0.8075908
## coef. of kurtosis = 6.29924
## Filliben correlation coefficient = 0.9777378
## *****
```

Dla wartości strat dla cen gazu ponownie został dopasowany rozkład ST3, tym razem z lekko innymi parametrami. Tak jak dla dopasowania dla strat cen ropy, zwarte skupienie punktów na pierwszym wykresie sugeruje brak zależności reszt a rozkład reszt względem indeksu sugeruje losowość błędów. Estymacja gęstości rozkładu reszt ponownie pokazuje, że reszty mają rozkład normalny, co również pokazuje wykres kwantyl-kwantyl, z identycznym ciężkim ogonem z lewej strony wykresu.

Obliczenie korelacji dla przekształconych strat i pełnych danych cenowych:

```
## [1] "Współczynnik Pearsona"
```

```
##          gas.Close  
## oil.Close 0.07096298
```

```
## [1] "Współczynnik Spearmana"
```

```
##          gas.Close  
## oil.Close 0.08026153
```

```
## [1] "Współczynnik Kendalla"
```

```
##          gas.Close  
## oil.Close 0.05395803
```

Korelacja dla przekształconych strat jest niska, co sugeruje, że spadki cen nie mają ze sobą znaczącej zależności liniowej.

```
## [1] "Współczynnik Pearsona"
```

```
##          gas.Close  
## oil.Close 0.1181318
```

```
## [1] "Współczynnik Spearmana"
```

```
##          gas.Close  
## oil.Close 0.1326297
```

```
## [1] "Współczynnik Kendalla"
```

```
##          gas.Close  
## oil.Close 0.09183231
```

Korelacja dla cen zamknięcia również jest niska, co sugeruje, że ceny nie mają ze sobą znaczącej zależności liniowej. O ile dla strat, które poddaliśmy wielu przekształceniom można się spodziewać niskich wartości korelacji, tak dla cen zamknięcia jest to trochę dziwniejsze, jednak porównując oryginalne wykresy cen, można mieć wrażenie jakby te wartości nie były od siebie zbyt zależne.

KROK 3

Dopasowanie modeli kopuły:

```
##          oil.Close  gas.Close  
## [1,] 0.50545571 0.03883184  
## [2,] 0.09403081 0.12002567  
## [3,] 0.42811297 0.08087291  
## [4,] 0.18324775 0.39152760  
## [5,] 0.25770218 0.46245186  
## [6,] 0.34595635 0.40949936
```



```
##      oil.Close      gas.Close
## Min.   :0.0009628   Min.   :0.001284
## 1st Qu.:0.2500000   1st Qu.:0.250321
## Median :0.5003209   Median :0.501284
## Mean   :0.5000000   Mean   :0.500000
## 3rd Qu.:0.7496791   3rd Qu.:0.749679
## Max.   :0.9993582   Max.   :0.999358
```

1. Gumbela:

```
## Call: fitCopula(gumbel_copula, data = data_matrix)
## Fit based on "maximum pseudo-likelihood" and 1557 2-dimensional observations.
## Gumbel copula, dim. d = 2
##      Estimate Std. Error
## alpha    1.044      0.017
## The maximized loglikelihood is 4.348
## Optimization converged
## Number of loglikelihood evaluations:
## function gradient
##           6           6
```

2. Franka:

```
## Call: fitCopula(frank_copula, data = data_matrix)
## Fit based on "maximum pseudo-likelihood" and 1557 2-dimensional observations.
## Frank copula, dim. d = 2
##      Estimate Std. Error
## alpha    0.4848      0.153
## The maximized loglikelihood is 5.043
## Optimization converged
## Number of loglikelihood evaluations:
## function gradient
##           4           4
```

3. Claytona:

```
## Call: fitCopula(clayton_copula, data = data_matrix)
## Fit based on "maximum pseudo-likelihood" and 1557 2-dimensional observations.
## Clayton copula, dim. d = 2
##      Estimate Std. Error
## alpha    0.1141      0.032
## The maximized loglikelihood is 5.137
## Optimization converged
## Number of loglikelihood evaluations:
## function gradient
##           3           3
```

4. Normalna:

```
## Call: fitCopula(normal_copula, data = data_matrix)
## Fit based on "maximum pseudo-likelihood" and 1557 2-dimensional observations.
## Normal copula, dim. d = 2
```

```
##           Estimate Std. Error
## rho.1    0.0949      0.026
## The maximized loglikelihood is 6.92
## Optimization converged
## Number of loglikelihood evaluations:
## function gradient
##           6           6
```

5. T-studenta:

```
## Warning in var.mpl(copula, u): the covariance matrix of the parameter estimates
## is computed as if 'df.fixed = TRUE' with df = 163.167789543001
```

```
## Call: fitCopula(t_copula, data = data_matrix)
## Fit based on "maximum pseudo-likelihood" and 1557 2-dimensional observations.
## t-copula, dim. d = 2
##           Estimate Std. Error
## rho.1    0.09447      0.026
## df      163.16779      NA
## The maximized loglikelihood is 6.893
## Optimization converged
## Number of loglikelihood evaluations:
## function gradient
##          913          453
```

```
## [1] "Kryterium loglikelihood"
```

```
## fit_gumbel  fit_frank fit_clayton  fit_normal  fit_t
##    4.347928    5.043053    5.137189    6.919763    6.892620
```

```
## [1] "Kryterium AIC"
```

```
## fit_gumbel  fit_frank fit_clayton  fit_normal  fit_t
##   -6.695857   -8.086105   -8.274378  -11.839525   -9.785241
```

```
## [1] "Kryterium BIC"
```

```
## fit_gumbel  fit_frank fit_clayton  fit_normal  fit_t
##  -1.3453405  -2.7355890  -2.9238615  -6.4890093   0.9157917
```

Wyznaczenie najlepiej dopasowanej kopuły względem:

1. Kryterium loglikelihood(im większa wartość, tym lepiej dopasowany model):

```
## [1] "fit_normal"
```

```
## [1] 6.919763
```

2. Kryterium AIC(im mniejsza wartość, tym lepiej dopasowany model):

```
## [1] "fit_normal"
```

```
## [1] -11.83953
```

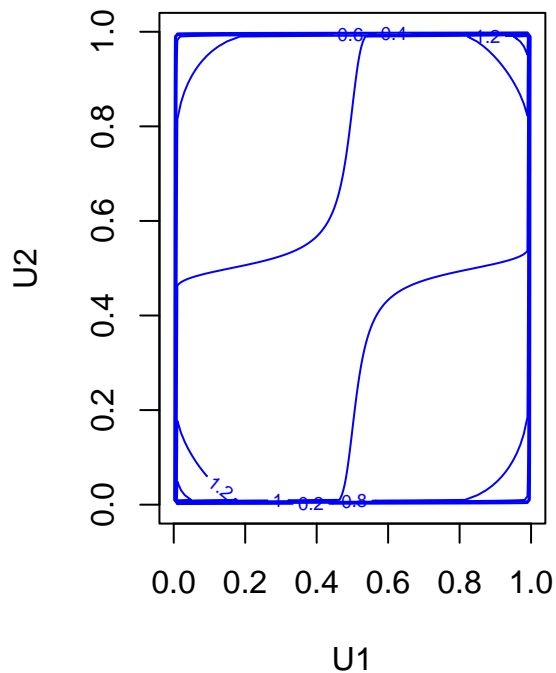
3. Kryterium BIC(im mniejsza wartość, tym lepiej dopasowany model):

```
## [1] "fit_normal"
```

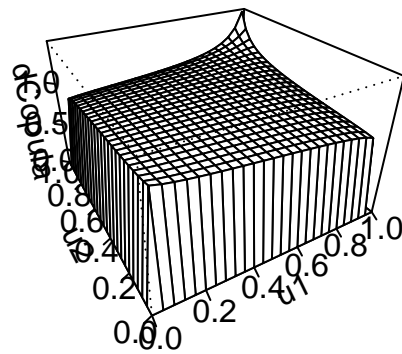
```
## [1] -6.489009
```

Wszystkie ww. kryteria wybrały kopułę normalną jako tą najlepiej dopasowaną. Wizualizacja najlepiej dopasowanej kopuły:

Wykres konturowy kopuły



Dopasowana kopuła normalna



Wykonanie testu Mardia:

```
## $multivariateNormality
##           Test      Statistic p value Result
## 1 Mardia Skewness 4492.63618540665      0    NO
## 2 Mardia Kurtosis 138.079796047741      0    NO
## 3           MVN          <NA>    <NA>    NO
##
## $univariateNormality
##           Test Variable Statistic  p value Normality
## 1 Anderson-Darling oil.Close  93.8806 <0.001      NO
## 2 Anderson-Darling gas.Close  98.9668 <0.001      NO
```

```
##
## $Descriptives
##           n      Mean   Std.Dev Median   Min    Max  25th  75th    Skew
## oil.Close 1557 0.9881952 1.0692720  0.650 0.010 10.390 0.280 1.290 2.642203
## gas.Close 1557 0.1475594 0.1618081  0.104 0.001  1.648 0.048 0.185 3.197267
##           Kurtosis
## oil.Close 11.38847
## gas.Close 16.28629
```

Test Mardia wskazuje, że dane nie pochodzą z dwuwymiarowego rozkładu normalnego.

KROK 3

Obliczamy wartość zagrożoną (VaR) portfela składającego się z dwóch składników, korzystając zarówno z danych rzeczywistych, jak i danych wygenerowanych na podstawie najlepiej dopasowanej kopuły. Najpierw generowane są dane z dopasowanych rozkładów brzegowych, a później dane z kopuły. Następnie dla każdej próbki i poziomu istotności (α) obliczane są optymalne wagi (β), które minimalizują ryzyko portfela. Obliczenia są wykonywane na dwa sposoby: za pomocą funkcji optymalizującej (optimize) oraz ręcznie, za pomocą iteracji po wszystkich możliwych wartościach β z określoną odległością przedziału. Wszystko to, aby upewnić się, że wyniki są spójne i aby móc porównać ich dokładność w różnych sytuacjach.

```
## Wartości Beta (z funkcji optimize) z danych empirycznych dla N = 500 i Alpha = 0.95 wynosi 0.000077
## Wartości Beta (liczona ręcznie) z danych empirycznych dla N = 500 i Alpha = 0.95 wynosi 0
## Wartości Beta (z funkcji optimize) z danych z kopuły dla N = 500 i Alpha = 0.95 wynosi 0.011582
## Wartości Beta (liczona ręcznie) z danych z kopuły dla N = 500 i Alpha = 0.95 wynosi 0
##
## Wartości Beta (z funkcji optimize) z danych empirycznych dla N = 500 i Alpha = 0.99 wynosi 0.098535
## Wartości Beta (liczona ręcznie) z danych empirycznych dla N = 500 i Alpha = 0.99 wynosi 0.0985
## Wartości Beta (z funkcji optimize) z danych z kopuły dla N = 500 i Alpha = 0.99 wynosi 0.030921
## Wartości Beta (liczona ręcznie) z danych z kopuły dla N = 500 i Alpha = 0.99 wynosi 0.0309
##
## Wartości Beta (z funkcji optimize) z danych empirycznych dla N = 1000 i Alpha = 0.95 wynosi 0.000066
## Wartości Beta (liczona ręcznie) z danych empirycznych dla N = 1000 i Alpha = 0.95 wynosi 0
## Wartości Beta (z funkcji optimize) z danych z kopuły dla N = 1000 i Alpha = 0.95 wynosi 0.00949
## Wartości Beta (liczona ręcznie) z danych z kopuły dla N = 1000 i Alpha = 0.95 wynosi 0.0008
##
## Wartości Beta (z funkcji optimize) z danych empirycznych dla N = 1000 i Alpha = 0.99 wynosi 0.064072
## Wartości Beta (liczona ręcznie) z danych empirycznych dla N = 1000 i Alpha = 0.99 wynosi 0.0011
## Wartości Beta (z funkcji optimize) z danych z kopuły dla N = 1000 i Alpha = 0.99 wynosi 0.064948
## Wartości Beta (liczona ręcznie) z danych z kopuły dla N = 1000 i Alpha = 0.99 wynosi 0.0228

## Wartości VaR (z funkcji optimize) z danych empirycznych dla N = 500 i Alpha = 0.95 wynosi 0.422655
## Wartości VaR (liczona ręcznie) z danych empirycznych dla N = 500 i Alpha = 0.95 wynosi 0.422642
## Wartości VaR (z funkcji optimize) z danych z kopuły dla N = 500 i Alpha = 0.95 wynosi 0.320041
## Wartości VaR (liczona ręcznie) z danych z kopuły dla N = 500 i Alpha = 0.95 wynosi 0.312425
##
## Wartości VaR (z funkcji optimize) z danych empirycznych dla N = 500 i Alpha = 0.99 wynosi 0.938719
## Wartości VaR (liczona ręcznie) z danych empirycznych dla N = 500 i Alpha = 0.99 wynosi 0.938753
## Wartości VaR (z funkcji optimize) z danych z kopuły dla N = 500 i Alpha = 0.99 wynosi 0.816875
## Wartości VaR (liczona ręcznie) z danych z kopuły dla N = 500 i Alpha = 0.99 wynosi 0.816882
##
## Wartości VaR (z funkcji optimize) z danych empirycznych dla N = 1000 i Alpha = 0.95 wynosi 0.328267
```

```
## Wartości VaR (liczona ręcznie) z danych empirycznych dla N = 1000 i Alpha = 0.95 wynosi 0.328221
## Wartości VaR (z funkcji optimize) z danych z kopuły dla N = 1000 i Alpha = 0.95 wynosi 0.359476
## Wartości VaR (liczona ręcznie) z danych z kopuły dla N = 1000 i Alpha = 0.95 wynosi 0.354975
##
## Wartości VaR (z funkcji optimize) z danych empirycznych dla N = 1000 i Alpha = 0.99 wynosi 0.813129
## Wartości VaR (liczona ręcznie) z danych empirycznych dla N = 1000 i Alpha = 0.99 wynosi 0.686386
## Wartości VaR (z funkcji optimize) z danych z kopuły dla N = 1000 i Alpha = 0.99 wynosi 0.890412
## Wartości VaR (liczona ręcznie) z danych z kopuły dla N = 1000 i Alpha = 0.99 wynosi 0.889527
```