International Journal of Applied Technology & Leadership
ISSN 2720-5215
Volume 1, Issue 6, July 2022
ijatl@org

EXTREME VALUE DISTRIBUTION OF PRICES OF CHOSEN AGRICULTURAL PRODUCTS LISTED ON FUTURES MARKET

Gabriela Malik, PhD

College of Economics and Computer Studies, Cracow, Poland

Abstract

The aim of this article is to present statistical properties of price distribution of agricultural products listed on the Commodities Exchange in Chicago, in particular corn, soy and wheat, as well as an analysis of behaviour of bordering extreme values of prices of futures contracts, moreover defining a generalized extreme value distribution together with evaluation of accuracy of model choice to empirical data. To be able to reach this aim descriptive statistics as well as normality of return rate of examined agricultural products were presented along with their graphic representations. Parameters of distribution of tails of random variables, describing values of given futures contracts, are estimated with the maximum likelihood method using the block method. On the other hand, evaluation of adjustment of generalized extreme value distribution function for the tails of empirical distribution is conducted on the basis of quantile plots of appropriate distributions.

Introduction

In decision processes, as well as in other scientific disciplines, research of extreme value dependencies has very significant meaning. In numerous situations correct identification of dependencies allows avoiding making a wrong investment choice. Researching dependencies between extreme prices of agricultural products is crucial to the analysis process of the structure of dependencies, especially when modelling of dependency between maximum and minimum values of analysed observations. Dependencies for extremely small values or large values may be characterised by different features than those indicated on the basis of the whole test. Reasons for fluctuation of prices of agricultural products may include: financial crises, wars, speculation on markets, as well as droughts, fires, rains or natural disasters. For these reasons, variability of prices in times of extreme changes, i.e. unusual and rarely observed events on markets calls for close monitoring, because in the process of investing on futures markets, it may lead to above average gains or losses.

EVT – (*Extreme Value Theory)* is used for describing behaviour of extreme value property limitations. It consists in dealing with stochastic behaviour of extreme events found in the tails of probability, which are only a small fraction of the whole analysed distribution. Its goal is not describing usual behaviour of schoastic phenomena but the unusual and rarely observed events. Extreme Value Theory has broad usage everywhere where modelling of dependency between analysed extreme values (maximal and minimal) is of importance. Scientific fields where Extreme Value Theory plays an important role are: hydrology, [cf. A. C Davison, R. L Smith, 1990; R. W. Katz, M. B Parlange, P. Naveau, 2002], insurance [compare. A. J. McNeil, 1999] and finances [cf. J. Danielsson, C. G. de Vries, 1997; A. J. McNeil, 1997; P. Embrechts, S. Resnick, G. Samorodnitsky, 1999; R. Gencay, F. Selcuk, 2004].

The aim of this article is to present statistical stochastic properties of distribution of prices of agricultural products listed on the Commodity Exchange in Chicago, analysis of behaviours of extreme price values of futures contracts and evaluation of accuracy of the model choice to the empirical data.

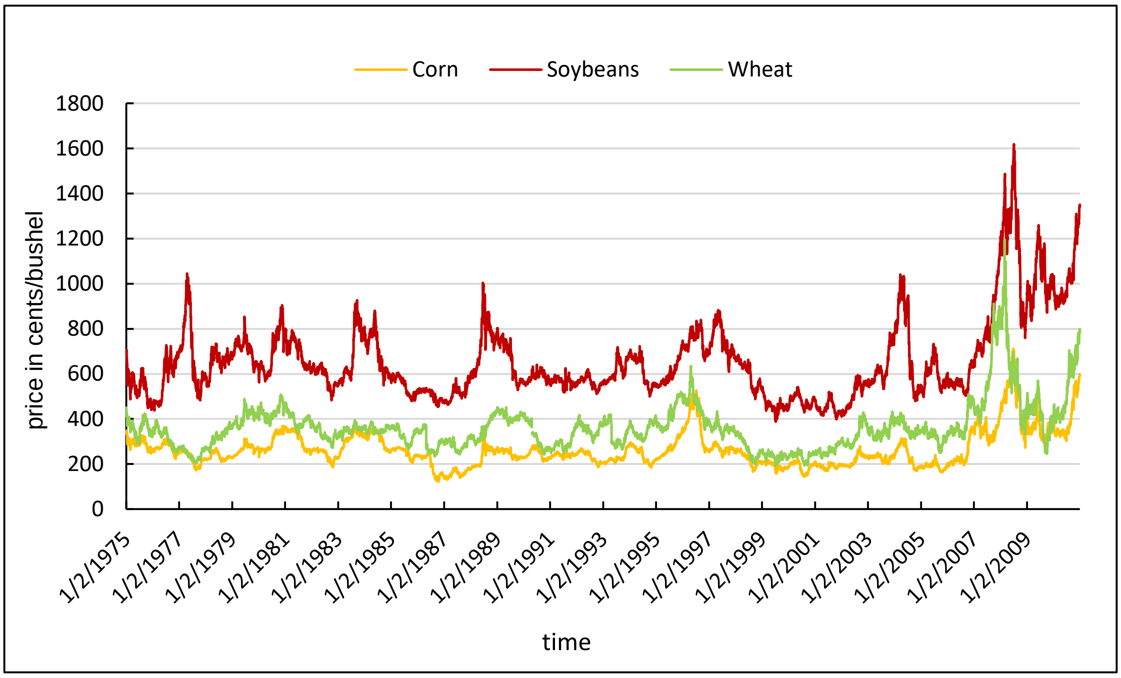
The article is divided into six parts. The second part encompasses the characteristics of a probe sample. Third part consists of statistical properties of price distribution of agricultural products on the futures market. It presents descriptive statistics, results of normality tests for the rate of return of agricultural products and their graphic illustration. Methodology of researching extreme values with various types of distribution of extreme values is presented in part four of the article. The fifth chapter is composed of a report and discussion of the results of empirical research. The final sixth part is an overview of the most important proposals formulated within this research.

Description of the research sample

The research sample consists of daily stock quotes of nominal prices of futures contracts for three agricultural products i.e. corn, wheat and soy, enlisted on the Commodity Exchange in Chicago. Choice of this particular Commodity Exchange is determined by the meaning of this exchange on the global market of agricultural products. Value of a contract is expressed in the price of square bushel – transaction unit for a given product in US dollars. Data consists of closing price of the contract of shortest expiry date, so that the chain of quotations can be treated as forward price of the shortest possible completion date. The choice of products is dictated by availability of adequately long time series. Empirical data encompasses years 1975-2010, and the single time series comprises 9070 observations. They are checked for possible discontinuity and errors. To minimize impact of arbitrary interference in the obtained results no correction procedures nor completing data has been used.

Figure 2.1 represents behaviour of prices of researched agricultural products in the analysed time period.

Figure 2.1. Daily changes of prices in future contracts – corn, wheat and soy enlisted on the Commodity Exchange in Chicago in the years 1975-2010.



Source: [own study]

Analysing behaviour of prices of futures contract we can clearly observe that in the middle of 2008 there was a rapid growth of prices of each of the agricultural products, particularly soy, for which the minimal price reaches the level of 1618,5 cents per bushel, whereas the maximal prices, respectively of wheat and corn, were at the level of 1195 and 711 cents per bushel. Direct cause of such a significant growth of the prices of agricultural products was the financial crisis in the USA, with September 2008 considered to be the beginning of the crisis, with the bankruptcy of the American investment bank- Lehman Brothers. Increase of analysed prices of agricultural products can be observed also in 2010. It was not as distinct as two years before, and its main causes were attributed to droughts and fires in Russia.

Statistical properties of agricultural product price distribution on futures market

Since chains of prices of financial instruments belong to the stationary processes, in order to make statistical analysis between prices of agricultural products, on the basis of price chains` daily constant growth rate. Usage of logarithmic growth rate does not remain without importance for the properties of the examined data chain. It is known that logarithmising, as one of Box-Cox transformations stabilized chain variance.

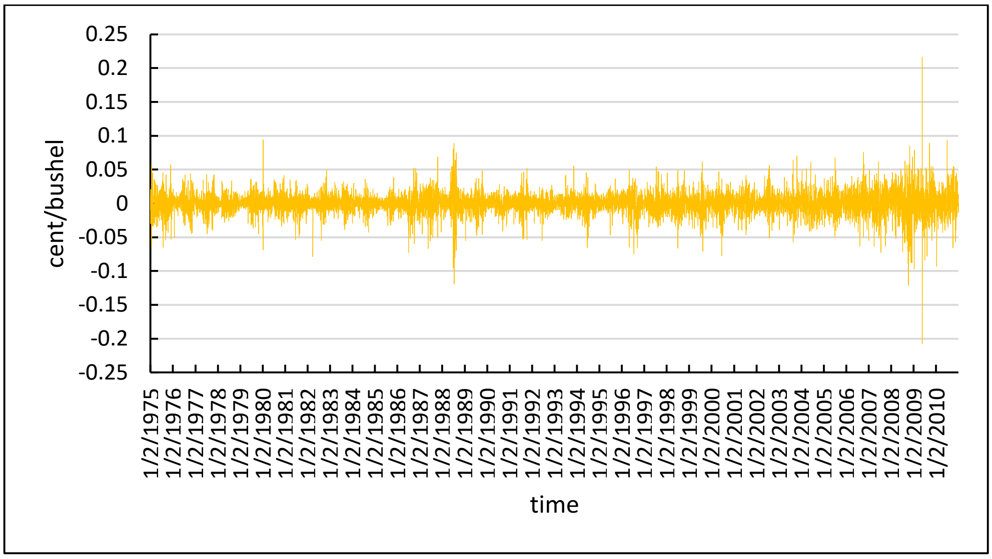
Therefore, on the basis of daily closing price logarithmic rate of return was calculated for particular agricultural product, using the following formula:

$$R_t = ln\left(\frac{X_t}{X_{t-1}}\right), \tag{3.1}$$

where X_t stands for the value of the futures contract of day t.

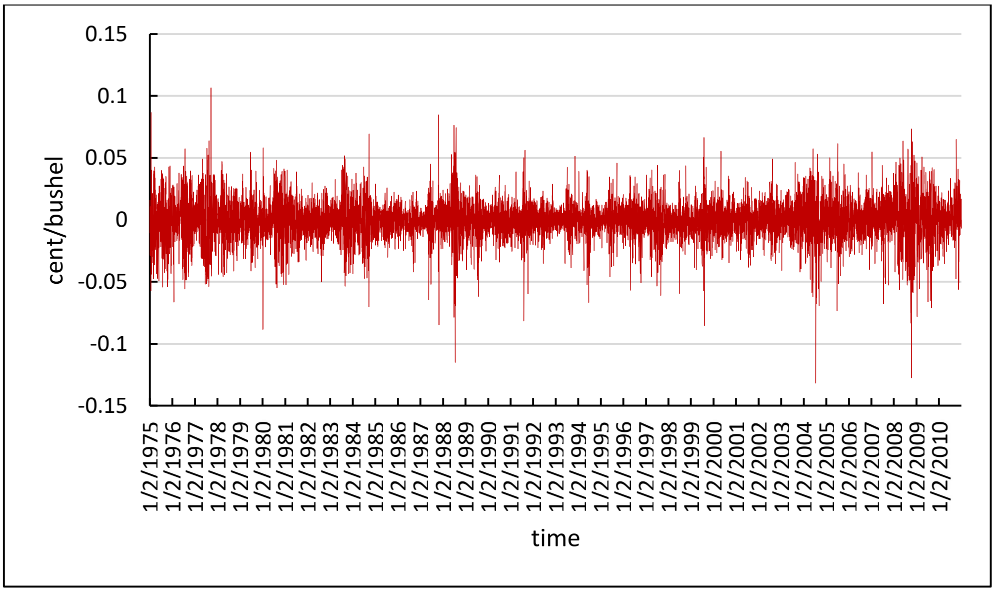
Figures 3.1.-3.4. represent graphical illustrations of daily growth rate for futures contracts on corn, wheat and soy listed on Commodity Exchange in Chicago in the analysed period in the years 1975-2010.

Figure 3.1. Daily rate of return of futures contracts for corn listed on Commodity Exchange in Chicago in the years 1975-2010



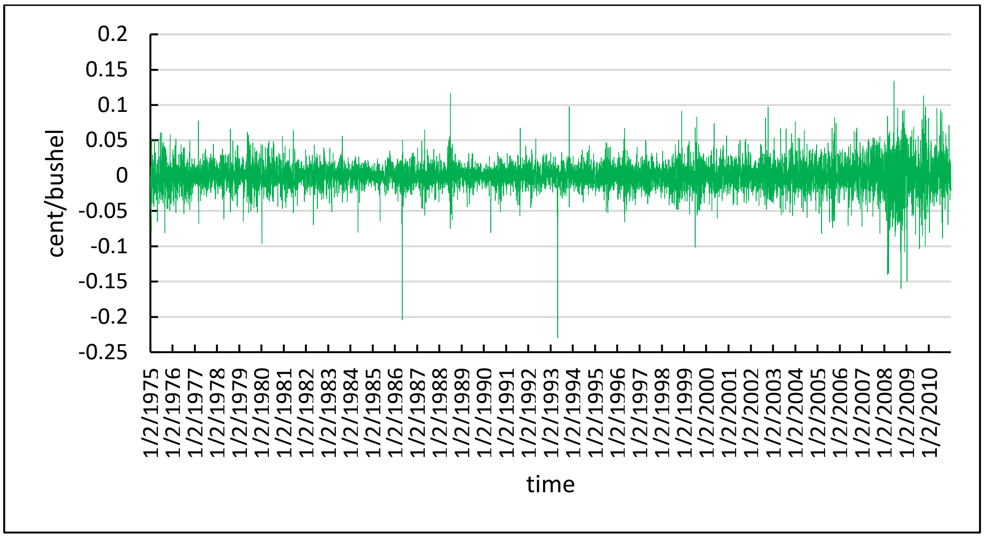
Source: [own study]

Figure 3.2.Daily rate of return of futures contract for soy listed on Commodity Exchange in Chicago in the years of 1975-2010



Source: [own study]

Figure 3.3. Daily rate of return of futures contract for wheat listed on Commodity Exchange in Chicago in the years of 1975-2010



Source: [own study]

Analysing the diagrams (figures 3.1.-3.3.) one can observe very significant variability of rates of return of futures contract for given agricultural products and occurrence of single observations which are characteristic for their size that significantly varies from the average level. Exhibitions of slight asymmetry are also perceived in the case of rate of return for futures contracts for wheat and soy, where more observances deviate toward minus. Moreover, the figure clearly shows periods of concentration of higher values of rates of return and periods when next to each other relatively low values are concentrated which indicates occurrence of groups of variances. To make a more thorough analysis of empirical data in table 3.1. basic descriptive statistics is presented, depicting the rate of return for analysed agricultural products. Last two rows comprise results of normality tests with the usage of Shapiro-Wilk and Jarque-Bera tests.

Table 3.1. Descriptive statistics and normality test results for daily rate of return for chosen agricultural products.

	Corn	**Soy**	**Wheat**
Minimum	-0,208	-0,132	-0,230
Quartile 1.	-0,008	-0,008	-0,010
Median	0,000	0,000	0,000
Average	0,000	0,000	0,000
Standard deviation	0,017	0,016	0,020
Quartile 3.	0,009	0,009	0,010
Maximum	0,216	0,107	0,134
Range	0,424	0,238	0,363

Skewness	-0,252	-0,418	-0,428
Kurtosis	9,544	3,821	7,502
Shapiro-Wilk Test	0,942 (0,000)	0,958 (0,000)	0,946 (0,000)
Jarque-Ber Test $\left(\times 10^{-3}\right)$	34,453 (0,000)	5,774 (0,000)	21,561 (0,000)

* In the round brackets values of critical probability (*p*-value) were given.

Source: [own study]

As the results presented in table 2.1. indicate, for each of the analysed agricultural products average daily rate of return is close to zero, as might be expected. Comparison of minimal and maximal value adequately with quartile one and three clearly show a strong tendency for extreme value occurrence, particularly on the negative side of rate of return distribution. Minimal daily rate of return ranges from -23% (wheat) to -13,2% (soy). On the other hand, the maximum rate of return was reached for corn and it is 21,6%, lowest for soy being on the level of 10,7%. Standard deviation assumes low values and varies from number 0,016 in case of corn to number 0,020 for soy. Since the value of deviation is higher than average, one can conclude that the daily rate of return for chosen agricultural products is characterized by high variability. However, insignificantly low value of skewness and relatively high kurtosis confirm phenomenon well known in the literature on the subject, concerning minor usefulness of normal distribution for description of rate of return. Moreover, they show minor left sided asymmetry of analysed series. For each of the agricultural products empirical distribution of rates of return show features largely deviating from normality, which is exemplified in the rejection of the null hypothesis about conformity of empirical distribution with normal distribution in case of both used normality tests with particularly stringent significance level.

It is worth emphasising that when it comes to the usefulness of normal distribution for description of empirical distribution of prices and rate of return on futures market, it had been proved already in the 1960s`, that it is characterized by low usefulness in this area. Both E. Fama [cf. E. Fama, 1965], as well as P. Clark [cf. P. Clark, 1973] presented in their research relevant evidence showing the dissonance between what is observable and what the normal distribution postulated. Similarly, B. Mandelbrot [cf. B. Mandelbrot, 1963] in respect to the futures market of agricultural products, analysing prices of cotton confirmed that prices cannot be described with the usage of normal distribution.

Reader who is particularly interested in research of statistical properties of distribution of prices of agricultural products on futures markets may find more information on the subject in the publication of G. Malik [cf. G. Malik, 2011]. The author provided numerous evidence against the hypothesis of normality of distribution of prices of agricultural products in her article, what is more the author shows that calibrated distribution t-student and a-stable family are most appropriate for description. In the course of empirical analysis, the author determines that price series are a non-stationary process, which through single calculation of growth may easily lead to a stationary process.

Similar conclusions were drawn in the scientific articles of S. Deng in. [S. Deng, W. Jiang, Z. Xia, 2002, as well as H. J. Jin [H. J. Jin, 2007] emphasising the significance of the stable distribution family in the context of their usage in description of prices of products listed on commodity exchanges. In world literature broad application of t-Student scaled distribution was shown by [cf. a.o. P. D. Praetz, 1972; R. Blattberg, N. Gonedes, 1974; J. B. Gray, D. W. French, 1990; A. Peiro, 1994; F. M. Aparicio, J. Estrada, 2001; D. S. Broca, 2002].

Extreme Value Theory – types of distribution

Classical approach to usage for modelling of extreme values is based on the block maxima model. This method is used for a broad number of observances which were chosen from a large probe. Modelling of behaviour of extreme values of independent random variables of identical probability distribution in practice is based on usage of maxima and minima observances in set time periods. Blocks are designated by distinct ranges of time of equal length, mostly months, quarters or years. [cf. A. J. McNeil, 1999].

Let $X_1, X_2, \dots, X_n$ be independent random variables of the same distribution as distribution *F*. Let us recall the following designation: $M_n = max\{X_1, \dots, X_n\}$. Considered random variables $X_1, X_2, \dots, X_n$ may represent various meteorological and hydrological characteristics (e.g. state of water in rivers, daily concentration of ozone, average daily temperature etc.), but they might also be financial variables (e.g. prices or rate of return of different financial instruments). Extreme Value Theory will be presented in case of reaching maximal values for set time periods. To receive analogical theory for minimal values it is necessary to use a simple relation, i.e.: $max\{X_1, \dots, X_n\} = -min\{-X_1, \dots, -X_n\}$.

Let us have progression: $a_n > 0$ and b_n such that for certain distribution G a following relation is taking place:

$$\lim_{n \to \infty} P\left(\frac{M_n - b_n}{a_n} \le z\right) = G(z), \tag{4.1}$$

Then distribution *G* takes one of the three forms below:

- Gumbel

$$G(z) = \exp\left\{-\exp\left[-\left(\frac{z-b}{a}\right)\right]\right\}, \quad -\infty < z < \infty, \tag{4.2}$$

- Frecht

$$G(z) = \begin{cases} 0, & z \le b \\ \exp\left\{-\left(\frac{z-b}{a}\right)^{-\alpha}\right\}, & z > b \end{cases} \tag{4.3}$$

- Weibull

$$G(z)=\begin{cases}\exp\left\{-\left[-\left(\frac{z-b}{a}\right)^{\alpha}\right]\right\}, & z<b\\ 1, & z\geq b\end{cases}, \quad (4.4)$$

where a, b and $\alpha > 0$ are parameters.

Therefore, three different types of distributions of extreme values exist, although distribution described with Gumbel distribution function is described as type I, which is concentrated on the whole straight line and characterized by light distribution. On the other hand Frechta distribution represents type II, with heavy distribution and concentration on a certain half –line (c, ∞), where $c \in IR$. In the Weibull distribution, which is called distribution type III, specified on half line $(-\infty, c)$ where $c \in IR$, and containing limitations for the upper tail.

In literature of the subject distributions defined by formulas $(4.1) - (4.4)$ are called extreme value distributions. We may write them in a cumulative group, with one family of distributions, which is called generalized extreme value distribution. (GEV – *generalized extreme value distribution*) [cf. S. Coles, 2001, s. 45-58]. We receive the following generalized formula:

$$G(z)=\begin{cases}\exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]_{+}^{-1/\xi}\right\}, & \xi\neq 0\\ \exp\left\{-\exp\left[-\left(\frac{z-b}{a}\right)\right]\right\}, & \xi=0\end{cases}, \quad (4.5)$$

where:

$x_{+} = max\{x, 0\}$,

μ – location parameter, $\mu \in IR$,

σ – scale parameter, $\sigma > 0$,

ξ – shape parameter, $\xi \in IR$.

Obtained value of shape parameter ξ decides which of the extreme value distributions should be taken into consideration at the next step of our analysis. And therefore:

- $\xi > 0$ corresponds to Frecht distribution, where $\xi = \frac{1}{\alpha}$,
- $\xi < 0$ corresponds to Weibull distribution, where $\xi = -\frac{1}{\alpha}$,
- $\xi = 0$ corresponds to Gumbel distribution.

Extreme value distribution of prices of agricultural products of futures market

In order to analyse behaviour of bordering extreme values of futures contracts and defining type of generalized extreme value distribution of the daily rate of return of chosen agricultural products i.e. corn, soy and wheat were examined, obtained on the basis of the formula (3.1) presented in chapter III of this article. Parameters of distribution of tails of random variables, describing values of chosen futures contract, were estimated with maximum likelihood method with the usage of block method. Observations were divided into 303 sub periods, with length corresponding to the next months. Results of estimation of parameters of generalized extreme value distribution of chosen agricultural products with average estimation errors given in brackets are presented in the table below.

Table 5.1. Results of estimation of generalized extreme value parameter for corn, soy and wheat

Agricultural products	Left tail			Right tail		
	μ	σ	ξ	μ	σ	ξ
Corn	0,024	0,012	0,165	0,022	0,010	0,242
	(0,001)	(0,001)	(0,048)	(0,001)	(0,001)	(0,052)
Soy	0,024	0,011	0,176	0,022	0,010	0,138
	(0,001)	(0,001)	(0,049)	(0,001)	(0,001)	(0,054)
Wheat	0,042	0,017	0,258	0,029	0,012	0,214
	(0,002)	(0,001)	(0,093)	(0,001)	(0,001)	(0,055)

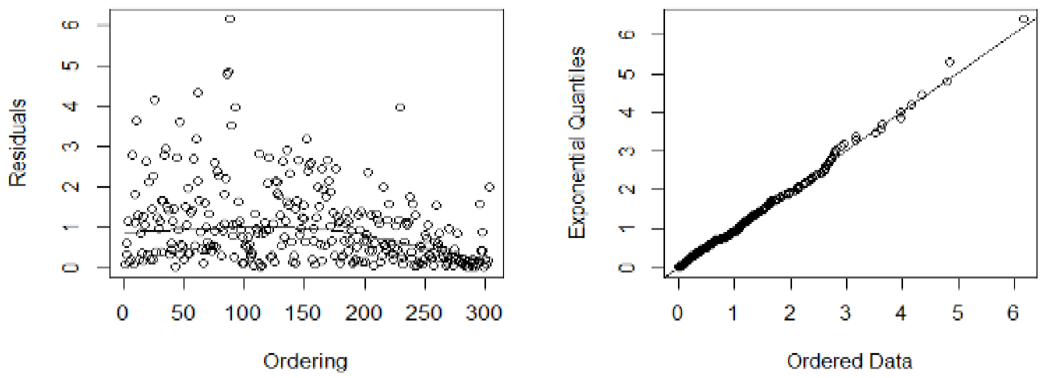
Source: [own study]

Analysing results presented in table 5.1.we might notice that shape parameter ξ of generalized extreme value distribution of examined daily rates of return of futures contract is positive , which corresponds to Frecht distribution and signifies fat tail of distribution of chosen agricultural products. All obtained parameters are statistically important for the adopted significance level 0,05.

Next step of the conducted research aims at evaluating accuracy of model choice for empirical data. Since the conducted analysis above generalized extreme distribution value is adjusted to the distribution tail, the choice of such a popular matching test as Kolmogorov test or Anderson-Darling test is not recommended [cf. A. Weron, R. Weron, 2021]. Similarly , a very popular and at the same oldest test used in the literature on the subject , namely matching test χ^2, does not apply to the analysis of goodness of matching of extreme value distributions. Main cause of lacking applicability of this non-compliance with the main assumption that research should come from a large, independently chosen probe [cf. A. Zeliaś, 2000, s. 280]. Therefore, to evaluate matching of distribution in the tail section, graphic test will be conducted. Relatively popular and recommended in the literature of the subject, proposed in 1984 by A.C. Davison matching test based on residua [cf. A. C. Davison, 1984; A. C. Davison and R.L. Smith, 1990].

Evaluation of matching of generalized extreme value distribution for tails of empirical distribution is conducted on the basis of quantile graphs of adequate distributions. Alignment of points on a straight line signifies goodness of matching of the examined distribution, whereas concave deviation from straight line is a sign of heavy tails. To be able to verify the goodness of matching of the generalized extreme value distribution for each of the examined agricultural products on figures 5.1.-5.6. two types of residua graphs are presented i.e. diagram of dispersion and quantile diagram.

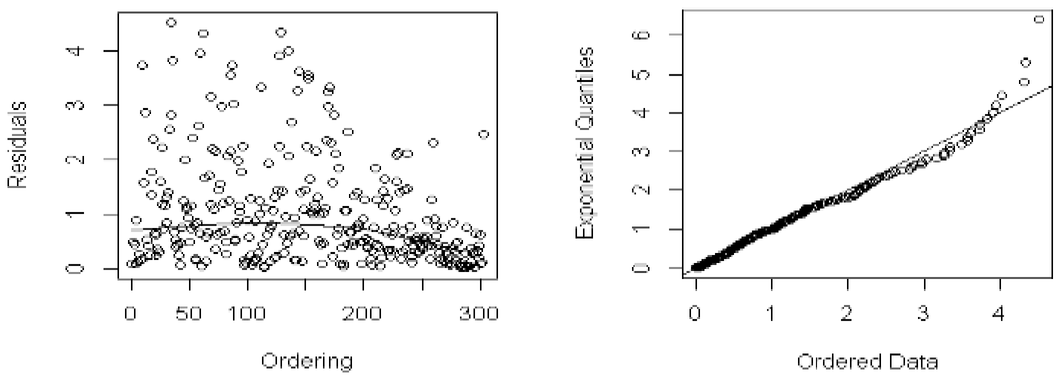
Figure 5.1. Diagram of dispersion of residue [fig. a)] and quantile diagram [fig. b)] for corn (right tail)



a) b)

Source: [own study with usage of R programme]

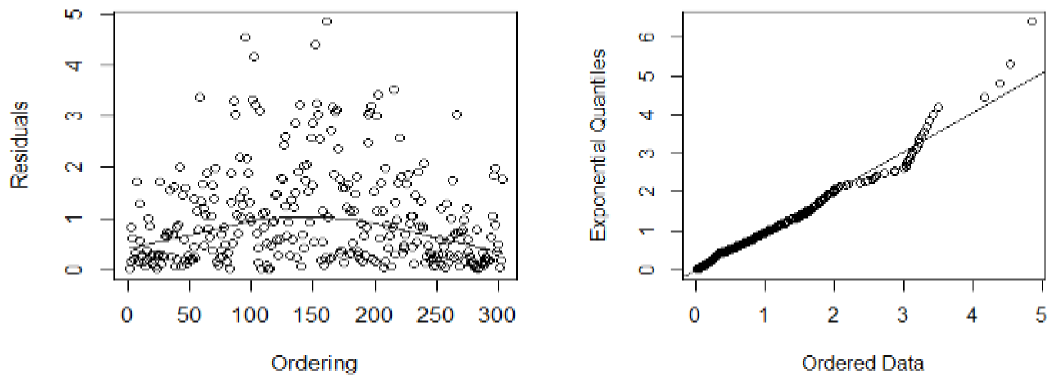
Figure 5.2. Diagram of dispersion of residue [fig. a)] quantile diagram [fig. b)] for corn (left tail)



a) b)

Source: [own study with usage of R programme]

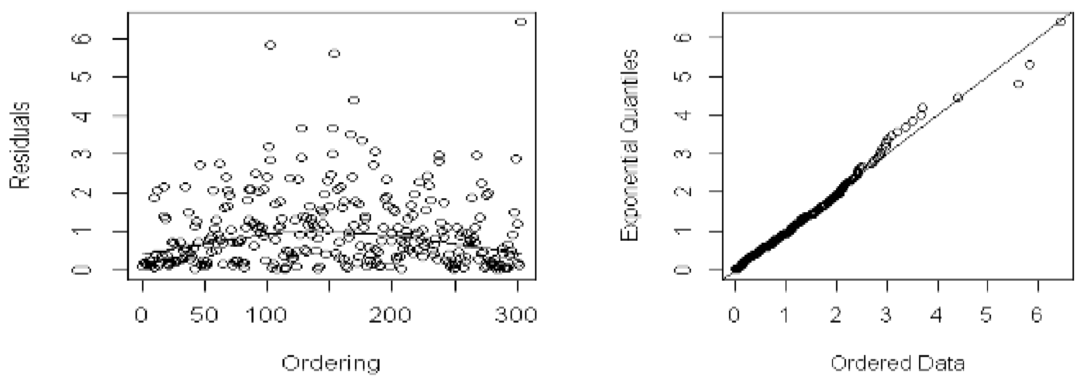
Figure 5.3. Diagram of dispersion of residue [fig. a)] quantile diagram [fig. b)] for soy (right tail)



a) b)

Source: [own study with usage of R programme]

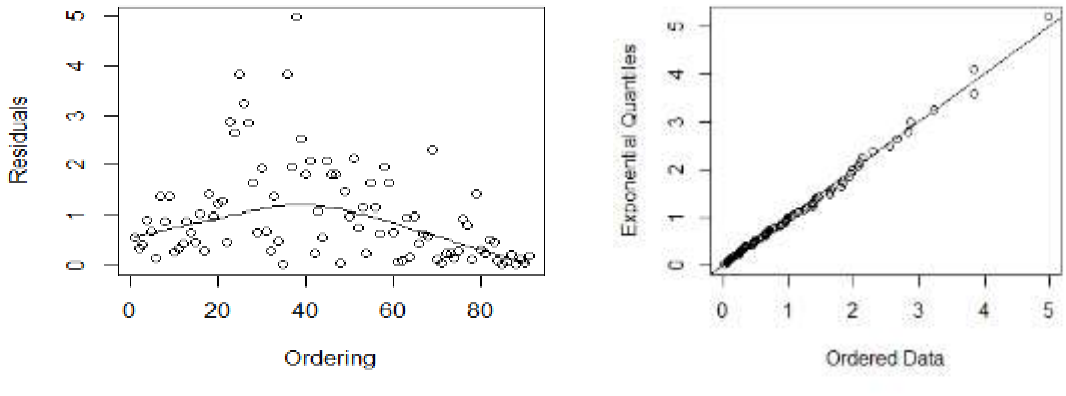
Figure 5.4. Diagram of dispersion of residue [fig. a)] quantile diagram [fig. b)] for soy (left tail)



a) b)

Source: [own study with usage of R programme]

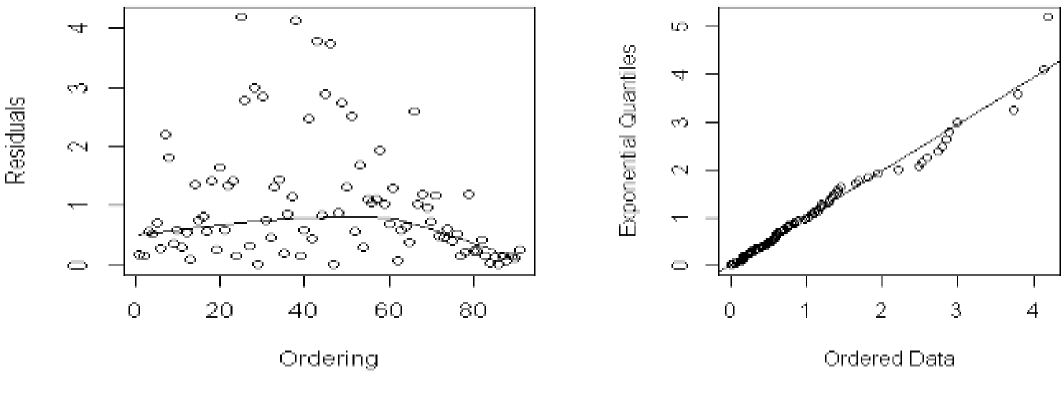
Figure 5.5. Diagram of dispersion of residue [fig. a)] quantile diagram [fig. b)] for wheat (right tail)



a) b)

Source: [own study with usage of R programme]

Figure 5.6. Diagram of dispersion of residue [fig. a)] quantile diagram [fig. b)] for wheat (left tail)



a) b)

Source: [own study with usage of R programme]

On the basis of quantile diagrams presented on figures 5.1. – 5.6. we may state that a test based on residuals does not give a basis for rejecting the null hypothesis about matching of generalized extreme value distribution with tails of empirical distribution for each of the examined agricultural products.

Conclusions

In the above article the author attempted to present statistical properties of distribution of prices for futures contracts for agricultural products such as corn, soy and wheat, listed on the Commodity Exchange in Chicago. Particular emphasis was placed on modelling extreme values of the examined futures contract for agricultural products on the basis of block maxima.

On the basis of empirical research one may draw conclusions about the existence of heavy tails of price distribution for agricultural products. The author believes, however, that there is no basis for stating existence of asymmetry of left vs right tail, as it is often the case with other financial instruments, particularly actions where probability of significant loss overweighs probability of significant profit. This observation is nevertheless entirely consequent with expectations. We must remember, however, that in case of futures contracts, or in derivative law in general, identifying the left side of distribution with losses and the right side of distribution with profit, will not find sufficient ground. Eventually it is the type of investor's position (short of long) that decides whether price growth is a profit or a loss.

It is worth emphasising that testing based on residuals did not give sufficient grounds for rejecting the null hypothesis about matching of generalized extreme value distribution with tails of empirical distribution for each of the examined agricultural products.

Bibliography

1. Aparicio F. M., Estrada J. (2001), *Empirical distribution of stock returns*: *European securities markets 1990-1995*, The European Journal of Finance, vol. 7: p. 1-21
2. Blattberg R., Gonedes N. (1974), *A comparison of the stable and student distributions as statistical models for stock prices*, Journal of Business, vol. 47: p. 244-280
3. Broca D. S. (2002), *The distribution of Indian Stock Returns*: *a tale of two tails*, Decision, vol. 29: p. 131-137
4. Clark P. (1973), *A subordinated stochastic process model with finite variance for speculative prices,* Econometrica, vol. 41: p. 135-155
5. Coles S. (2001), *An introduction to statistical modeling of extreme values*, Springer, London
6. Danielsson J., de Vries C. G. (1997), *Tail index estimation with very high frequency data*, Journal of Empirical Finance, vol. 4: p. 241-257
7. Davison A. C. (1984), *Modelling excesses over high thresholds, with an application*, [in:] *Statistical extremes and applications*, J. de Oliveira (eds.), D. Reidel, p. 461-482
8. Davison A. C., Smith R. L. (1990), *Models for exceedances over high threshold*, Journal of Royal Statistic Society, vol. 52(3): p. 393-442
9. Deng S., Jiang W., Xia Z. (2002), *Alternative statistical specifications of commodity price distribution with fat tail*, Advanced Modelling and Optimization, vol. 4: p. 1-8
10. Embrechts P., Resnick S., Samorodnitsky G. (1999), *Extreme value theory as a risk management tool*, North American Actuary Journal, vol. 3: p. 30-41

11. Fama E. (1965), *The behavior of stock market prices*, Journal of Business, vol. 38: p. 34-105
12. Gencay R., Selcuk F. (2004), *Extreme value theory and value at risk: relative performance in emerging markets*, International Journal of Forecasting, vol. 20: p. 287-303
13. Gray J. B., French D. W. (1990), *Empirical comparisons of distributional model for stock index prices,* Journal of Business Finance and Accounting, vol. 17: p. 451-459
14. Jin H. J. (2007), *Heavy-tailed behavior of commodity price distribution and optimal hedging demand*, The Journal of Risk and Insurance, vol. 74: p. 863-881
15. Katz R. W., Parlange M. B., Naveau P. (2002), *Statistics of extremes in hydrology*, Advances in Water Resources, vol. 25: p.1287-1304
16. Malik G., *Statystyczne własności rozkładu cen produktów rolnych na rynku terminowym na przykładzie giełdy towarowej w Chicago* [w:] *Nauka i gospodarka w dobie destabilizacji*, Praca zbiorowa pod redakcją naukową: J. Teczke, J. Czekaj, B. Mikuła & R. Oczkowska, Biuro Projektu Nauka i Gospodarka, Kraków 2011, s. 81-92
17. Mandelbrot B. (1963), *The variation of certain speculative prices*, Journal of Business, vol. 36: p. 394-419
18. McNeil A. J. (1997), *Estimating the tails of loss severity distributions using extreme value theory*, Astin Bulletin, vol. 27(2): p. 117-137
19. McNeil A. J. (1999), *Extreme value theory for risk managers*, [in:] *Internal Modeling and CAD II*, Risk Publication, p. 93-113
20. Peiro A. (1994), *International evidence on the distribution of stock returns*, Applied Financial Economics, vol. 4: p. 431-439
21. Praetz P. D. (1972), *The distribution of share price changes*, Journal of Business, vol. 45: p. 49-55
22. Weron A., Weron R., *Inżynieria finansowa*, Wydawnictwo Naukowe PWN, Warszawa 2021
23. Zeliaś A., *Metody Statystyczne*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2000