INTERCATHEDRA

1(38), 2019, 95–101

http://www.up.poznan.pl/intercathedra/
pISSN 1640-3622
eISSN 2545-2045

http://dx.doi.org/10.17306/J.INTERCATHEDRA.2019.00058
Accepted for print: 24.01.2019

Grzegorz Przekota✉

Koszalin University of Technology, Poland

ESTIMATION OF THE FRACTAL DIMENSION OF FUTURES CONTRACTS FOR WHEAT ON GLOBAL STOCK EXCHANGES

Abstract. Forecasting agricultural prices can be challenging. The price is affected by so many factors that their combined effect may be a random walk process. The purpose of this paper is to verify the random walk hypothesis on the futures contracts for wheat on the CBOT and MATIF exchanges. The research is supposed to identify the type (random, persistent or anti-persistent) of time series represented by futures contracts. The fractal dimension estimated by the surface area division method (SDM) was the tool used to determine the statistical properties of the futures contract time series. This paper presents a simplified calculation scheme for this dimension and a section of significance tables. Usually, research findings do not enable rejecting the hypothesis of a random walk of futures contracts for wheat, which suggests they follow an unpredictable pattern. On the other hand, the CBOT (unlike MATIF) witnessed a slightly different behavior of December and March contracts (which moved closer to a process with a strengthened trend) and of May, June and September contracts (which moved closer to a process of adjustment to the average). This can be used for speculative purposes. However, all the results fell within the limits of random walk. The length of the time series is an interesting question: it turns out that as the time horizon becomes longer, the results are more stable, i.e. the number of inconsistencies with the random walk process decreases. The random walk hypothesis is difficult to reject. It is confirmed that the determinants of agricultural prices drive their random motion.

Keywords: cereal market, futures prices, random walk, persistent series, anti-persistent series

INTRODUCTION

Price volatility in commodity markets is caused by macroeconomic and microeconomic market factors any by unstable natural conditions (Jerzak, 2014). Important aspects affecting price volatility are political factors, including barriers to trade such as customs duties or embargoes, as well as international agreements or production subsidies. In addition, price transmission between products is observed (Hayesc et al., 2011). Also, speculators become increasingly interested in commodity markets. All this means that the commodity market experiences considerable price fluctuations.

Significant price volatility of agricultural products causes various types of difficulties, in particular when it comes to forecasting. This is not conducive to estimating

✉PhD hab. Koszalin University of Technology Professor Grzegorz Przekota, Koszalin University of Technology, Faculty of Economic Science, Kwiatkowskiego 65E St., 65-343 Katowice, Poland, e-mail: grzegorzprzekota@wp.pl

future income and is a barrier to decision making for agricultural producers and processors when considering different investment opportunities. Therefore, the question arises about the statistical properties of price time series. If these are random walk series, then forecasting becomes impossible.

This paper deals with the problem of verifying the random walk hypothesis for wheat futures at the CBOT and MATIF. The daily quotes of all contracts with a maturity date ranging from late 2013 to 2018 were analyzed. Several issues were assessed:

1. What type of time series (random walk, persistent or anti-persistent) is represented by futures contracts for wheat?
2. What effect does the horizon of observation have on the properties of the time series (the length of the time series)?
3. Is it possible to find a connection between the properties of the time series and the contract maturity date?

The fractal dimension estimated by the field division method was chosen as a tool to verify the hypothesis that the futures contracts for wheat follow a random walk.

PRICE FORMATION IN AGRICULTURAL MARKETS

The main determinants of price formation in agriculture include: demand and supply of goods; indirect connection with the consumer market (the demand side); biological and technical nature of agricultural production (supply side); connections with global prices; impact of macroeconomic factors; agricultural and commercial policy of the state; and cross-product and inter-market links (Hamulczuk et al., 2012). Price formation is partially explained by the operation of the free market and by the law of supply and demand which may be disrupted by an institution.

Because products such as cereals and potatoes can be stored in an unprocessed form, the demand also depends on predictions about future prices. The greatest role in this respect is played by intermediaries and traders who undertake warehouse management activities. Another factor of impact on domestic demand is the foreign demand for the relevant product. Under normal circumstances, that demand will be determined by the needs of foreign markets and price relations, including changes in exchange rates. Price relations are a quite serious problem because there are many trading companies on the market who import or export agricultural products, depending on the situation, to take advantage of price differences.

Demand factors are one side the pricing mechanism in agriculture. The other side are supply factors, especially including the biological and technical nature of agricultural production. Natural factors make agricultural producers behave in a manner that translates into the amount and distribution of agricultural prices over time. However, many aspects related to the final effect of an agricultural product are beyond the producer's control.

In recent years, agricultural prices have been strongly influenced by the energy policy. The requirement to use biocomponents in fuels has caused an increase in demand for some agricultural raw materials and has strengthened the link between agricultural production prices and oil prices (Serra and Zilberman, 2011). Rising oil prices make rapeseed cultivation more profitable and increase rapeseed prices. On the other hand, a larger crop area means greater supply which reduces prices (Tyner, 2010). However, the net effect is usually positive (Tyner and Taheripour, 2008). In addition, larger rapeseed crops mean smaller sowing of other plants.

The problem with prices formation in agriculture results from the net effect of the impact of various factors. Each individual factor can be considered in an effort to estimate its impact on the price. However, when there are many factors, the result is difficult to predict. In general, the effect of various factors may be the form of time series of prices which will classify them into one of three classes:

1. Random walk series (white noise), i.e. series in which the previous change in value has no impact on future changes. Events are accidental and uncorrelated with each other; such series are unpredictable.
2. Persistent series (black noise). In this class, trend strengthening is present. This means that – unlike in the anti-persistent series – if an increase (or decrease) was observed between successive terms of the series, then the next term will be more likely to increase (or decrease). These series feature what is referred to as long-term memory which means that correlation exists between the terms.
3. Anti-persistent series (pink noise). In this class, the values tend to converge to the average level. This means that if the value of the series deviates too much up (or down) from the mean, a subsequent downward (or upward) deviation is more likely.

USING THE SURFACE AREA DIVISION METHOD (SDM) TO ESTIMATE THE FRACTAL DIMENSION

Euclidean geometry gives the dimension of the space in which the time series is placed. This space is the Euclidean dimension – 2. Considering the trajectory of the time series as polyline we obtain the Euclidean dimension – 1. Moving away from the Euclidean dimension, it can be noticed that the time series graph does not fill the entire plane where it lies, so its size will be smaller than 2 but larger than 1. The fractal dimension D describes the way the time series fills the space where it is embedded (Peters, 1997).

An example of how the fractal dimension can be used to describe natural phenomena was given by Mandelbrot (Mandelbrot, 1982). The problem consists in measuring the length of the coastline. The result depends on the length of the yardstick: the shorter the yardstick, the more accurate the result, since more irregularities are taken into account. The fractal dimension of the Norwegian and British coastline is 1.52 and 1.26, respectively (Peters, 1997). The coastline of Norway is more jagged than the British coastline.

A popular method of estimating the fractal dimension is the VM variant method (Dubuc et al., 1989) derived from the segment-variation method of SVM (Zwolankowska, 2001). Another popular solution is based on the Hurst exponent which was originally developed to assess the amount of water carried by the Nile River (Hurst, 1951) and was subsequently adapted to financial data by Mandelbrot (1972). It allows to identify long-term relationships. The existence of such a relationship contradicts the hypothesis of random walk. The original method requires a very long time series, and therefore some approximate methods (Hurst, 1951) with debatable reliability are used to assess the shorter series (Katsev and L'Heureux, 2003). A review of the methods and some suggestions for improvement can be found in works by Gloter and Hoffmann (2007).

The author's method of estimating the fractal dimension presented in the study combines elements of the segment-variation method with traditional geometric methods. Let N be the length of a time series divided into $k = 1, 2, ..., N/2$ parts.

The surface area occupied by the series can be defined as:

$$P = N \cdot (y_{\max} - y_{\min}),$$

where: $y_{\max}$ and $y_{\min}$ are the highest and the lowest values in the series, respectively.

After dividing the series into halves, the surface area will be expressed as:

$$p = \frac{N}{2} \cdot \left(y_{\max_1} - y_{\min_1} \right) + \frac{N}{2} \cdot \left(y_{\max_2} - y_{\min_2} \right)$$

There is an inequality between p and P:

$$p \le P.$$

This can be seen in Figure 1.

With any primary division into k parts, the surface area occupied by the series is defined as:

$$P_k = \frac{N}{k} \sum_{i=1}^{k} \left(y_{\max_i} - y_{\min_i} \right)$$

and with a division into $2k$ parts:

$$p_{2k} = \frac{N}{2k} \sum_{i=1}^{2k} \left(y_{\max_i} - y_{\min_i} \right)$$

There is an inequality between P_k and p_{2k}:

$$p_{2k} \le P_k.$$

Therefore:

$$p_{2k} \le 2 \cdot \frac{P_k}{2}$$

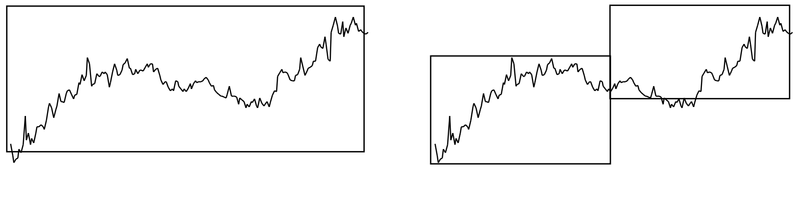

Fig. 1. Time series plotted on the plane; surface areas *P* and *p*
Source: own study.

Then, the following is true for any series:

$$p_{2k} \leq SDM \cdot \frac{P_k}{2}$$

where *SDM* falls within the range <1; 2>. The closer the value of *SDM* is to 1, the more the plot of the series converges to a straight line (i.e. becomes a series with a fixed trend). However, the closer the value of *SDM* is to 2, the more irregular is the shape of the series, i.e. the more often there will be a change in the opposite trend in the series.

The *SDM* value can be estimated as the regression coefficient of the linear function that passes through points ($P_k/2$; p_{2k}) and is expressed as:

$$SDM = \frac{\sum p_{2k} \frac{P_k}{2}}{\sum \left(\frac{P_k}{2}\right)^2}$$

The *SDM* value defined as above can be treated as a measure of time series fraying, i.e. as a fractal dimension of the time series.

Table 1. Critical values of the *SDM* dimension

Length of the series *N*	α = 0.1		α = 0.05	
	Lower limit	Upper limit	Lower limit	Upper limit
200	1.1813	1.5626	1.1447	1.5991
500	1.1935	1.5662	1.1578	1.6019
1 000	1.1942	1.5691	1.1584	1.6050
1 600	1.1985	1.5484	1.1650	1.5819

Source: own calculations.

The tables of critical values for the *SDM* dimension were constructed based on the simulations (Table 1). Depending on the value of *SDM*, three classes of time series can be identified:

1. *SDM* values below the lower limit will imply the existence of a persistent series, i.e. processes with a strengthened trend.
2. *SDM* values between the lower and upper limits will imply the existence of a series which follows a random walk.
3. *SDM* values above the upper limit are characteristic of an anti-persistent series, i.e. processes that tend to return to the average value.

FRACTAL DIMENSION ESTIMATED WITH THE SURFACE AREA DIVISION METHOD (SDM) FOR WHEAT FUTURES CONTRACTS

The random walk hypothesis was verified for wheat contracts traded on the CBOT (Table 2) and MATIF (Table 3). A time series of $N = 200$ data points was analyzed,

Table 2. Values of the fractal dimension of futures contracts for wheat on the CBOT

CBOT	*SDM*	
	N = 200	*N* = 500
2013 Dec	1.3681	1.3579
2014 Mar	1.3454	1.2665
2014 May	1.3707	1.2518
2014 Jul	1.4665	1.2432
2014 Sep	1.4100	1.3000
2014 Dec	1.3258	1.2481
2015 Mar	1.4863	1.4575
2015 May	1.5024	1.4216
2015 Jul	1.5002	1.3281
2015 Sep	1.4915	1.4237
2015 Dec	1.3374	1.4183
2016 Mar	1.2945	1.3260
2016 May	1.4302	1.3880
2016 Jul	1.4254	1.2570
2016 Sep	1.3151	1.2875
2016 Dec	1.1631[b]	1.2286
2017 Mar	1.2494	1.3086
2017 May	1.5936[a]	1.3048
2017 Jul	1.4469	1.4003
2017 Sep	1.3845	1.4648
2017 Dec	1.2516	1.4229
2018 Mar	1.3067	1.4710
2018 May	1.5496	1.5133
2018 Jul	1.4810	1.5281
2018 Sep	1.5425	1.4661

Explanations: a – anti-persistent series (at $p < 0.1$); b – persistent series (at $p < 0.1$).
Source: own calculations.

Table 3. Values of the fractal dimension of futures contracts for wheat on the MATIF

MATIF	*SDM*	
	N = 200	*N* = 500
2013 Nov	1.3182	–
2014 Jan	1.4476	–
2014 Mar	1.5582	1.4411
2014 May	1.3795	1.3799
2014 Nov	1.1775[b]	1.3164
2015 Jan	1.3600	1.3746
2015 Mar	1.4332	1.4791
2015 May	1.5181	1.5176
2015 Sep	1.3360	1.4778
2015 Dec	1.4514	1.6018[a]
2016 Mar	1.2388	1.3964
2016 May	1.2084	1.4189
2016 Sep	1.3411	1.3355
2016 Dec	1.5691[a]	1.2639
2017 Mar	1.5529	1.3506
2017 May	1.6001*	1.3162
2017 Sep	1.3167	1.2821
2017 Dec	1.2842	1.4031
2018 Mar	1.2078	1.3599
2018 May	1.3588	1.3577
2018 Sep	1.1877	1.3615

Explanations: * – anti-persistent series (at $p < 0.05$), a – anti-persistent series (at $p < 0.1$), b – persistent series (at $p < 0.1$).
Source: own calculations.

corresponding approximately to a 10-month quotation period before contract expiry, and a time series of $N = 500$ data points, corresponding approximately to a 2-year observation period. The last listing always coincided with the contract's expiry date. Some contracts (on MATIF) have been traded for less than two years. For such contracts, the *SDM* dimension for $N = 500$ data points was not determined.

When analyzing the results obtained for the time series with $N = 200$ data points recorded on the CBOT, it can be noted that the random walk hypothesis cannot be rejected in any case (at $p = 0.05$). The highest value (1.5936) of the *SDM* dimension corresponds to May 2017. At $p = 0.1$, the series can be considered anti-persistent, i.e. one which tends to return to average levels. In turn, the lowest *SDM* dimension (1.1631) corresponds to December 2016. At $p = 0.1$, the series can be considered persistent, i.e. a process with a strengthened trend.

Interestingly enough, the highest values of the *SDM* dimension for the $N = 200$ series were recorded in May, July and September contracts, and the lowest in December and March contracts.

The results for the 2-year period ($N = 500$) in no case allow to reject the random walk hypothesis (at $p = 0.05$ and also at $p = 0.1$). Based on these findings, it cannot be determined whether the contracts analyzed in a 2-year period (selected by expiry date) are more similar to the persistent or anti-persistent series.

The second exchange covered by this analysis is MATIF. In this case, the observation period was shorter ($N = 200$), and the random walk hypothesis can be rejected at $p = 0.05$ for the May 2017 contract (*SDM* = 1.6001). This means the presence of an anti-persistent series. Also, at $p = 0.1$, the random walk hypothesis may be rejected for December 2016 (*SDM* = 1.5691) and November 2014 (*SDM* = 1.1775) contracts, being an anti-persistent and a persistent series, respectively. No relationship can be identified between the *SDM* dimension and the contract expiry month.

When extending the time horizon ($N = 500$), it turns out that only one contract (December 2015) can be considered an anti-persistent contract (*SDM* = 1.6018) at $p = 0.1$. In other cases, there are no grounds to reject the random walk hypothesis. None of the months can be identified as one with the characteristic values of the fractal dimension.

CONCLUSIONS

Price formation of agricultural products is among the major economic problems. This study analyzed a specific category, the prices of futures contracts for wheat. According to the theory of investing, the futures contract price reflects the complete knowledge of market participants about the future cash price. Various information is coming to the market all the time, which potentially enriches the participants' knowledge, and thus allows for price adjustments.

The analysis presented in this paper allows to conclude that the vast majority of futures contracts for wheat follow a random walk process. The detected deviations from this hypothesis can be treated as incidental, all the more so since they are recorded at a quite high level of significance ($p = 0.1$).

When comparing different markets, it was noticed that only the CBOT exhibited the presence of a relationship between the contract maturity month and the level of the fractal dimension. It turns out that the formation of prices of May, July and September contracts is closer to an anti-persistent process, whereas the formation of prices of December and March contracts is closer to a persistent process (although within the limits of random walk). MATIF was not similar in that respect.

The horizon of observation turns out to be quite important in assessing the statistical properties of the time series of futures contracts for wheat. The longer the horizon of observation ($N = 500$ compared to $N = 200$), the more stable the results, i.e. the smaller the number of cases where the random walk hypothesis can be rejected.

When put in the context of investment practice, these findings mean that even if some deviations from the random walk hypothesis were detected, they did not enable reliable forecasting. These cases are too rare and irregular. Probably speculative possibilities give only the detected regularity of the dependence between the contract's maturity month and the level of the fractal dimension for the CBOT.

The tool used in this study is also worth noting. The fractal dimension estimated by the surface area division method is relatively simple to interpret; it enables classifying time series by relation to the random walk hypothesis. There are no special assumptions about applicability. In this paper, the fractal dimension estimated by the surface area division method was used to evaluate the time series of variable levels, but it can also be used to evaluate a time series of increments.

REFERENCES

Dubuc, B., Roques-Carmes, J. F., Tricot, C., Zucker, S. W. (1989). Evolving the Fractal Dimension of Profiles. Phys. Rev. A, 39,. 1506. https://doi.org/10.1103/PhysRevA.39.1500

Gloter, A., Hoffmann, M. (2007). Estimation of the hurst parameter from discrete noisy data. Ann. Stat., 35 (5), , 1947–1974. http://dx.doi.org/10.1214/009053607000000316

Hamulczuk, M., Gędek, S., Klimowski, C., Stańsko, S. (2012) Prognozowanie cen surowców rolnych na podstawie zależności przyczynowych. Warszawa: IERiGŻ-PIB, Warszawa, p. 11.

Hayesc, D. J., Xiaodong, D., Yub, C. L. (2011) Speculation and volatility spillover in the crude oil and agricultural commodity markets: A Bayesian analysis, Energy Econ., 33, 3, 497–503. https://doi.org/10.1016/j.eneco.2010.12.015

Hurst, H. E. (1951). Long-term storage of reservoirs. Trans. Am. Soc. Civil Eng.,116, 770–799.

Jerzak, M. (2014). Towarowe instrumenty pochodne w zarządzaniu ryzykiem cenowym w rolnictwie. Rocz. Nauk. Ekon. Roln. Rozwoju Obsz. Wiej., 101, 4, 78–84.

Katsev, S., L'Heureux, I. (2003). Are Hurst exponents estimated from short or irregular time series meaningful? Comp. & Geosci, 29, 1085–1089. http://dx.doi.org/10.1016/S0098-3004(03)00105-5

Mandelbrot, B. (1972). Statistcal Methodology for Non-Periodic Cycles: From the Covariance to R/S Analysis. Ann. Econ. Soc. Meas., 1, 259–290.

Mandelbrot, B. (1982). The Fractal Geometry of Nature. New York: W. H. Freeman & Co.

Peters, E. E. (1997). Teoria chaosu a rynki kapitałowe. Warszawa:WIG-Press, p. 58.

Serra, T., Zilberman, D. (2011).Price volatility in ethanol markets. Eur.Rev. Agr. Econ., 38(2), 259–290. http://dx.doi.org/10.1093/erae/jbq046

Tyner, W. E., Taheripour, F. (2008). Policy Options for Integrated Energy and Agricultural Markets. Rev. Agr. Econ., 30, 387–396. https://doi.org/10.1111/j.1467-9353.2008.00412.x

Tyner, W. E. (2010). The Integration of Energy and Agricultural Markets. Agr. Econ., 41, 1, 193–201. https://doi.org/10.1111/j.1574-0862.2010.00500.x

Zwolankowska, M. (2001). Fraktalna geometria polskiego rynku akcji. Szczecin: Rozprawy i Studia Uniw. Szczec., 382, p. 63.

SZACOWANIE WYMIARU FRAKTALNEGO SZEREGÓW KONTRAKTÓW FUTURES NA PSZENICĘ NA GIEŁDACH ŚWIATOWYCH

Abstrakt. Modelowanie i prognozowanie cen w rolnictwie jest procesem niezwykle trudnym. Liczba czynników wpływająca na cenę jest na tyle obszerna, że łącznym efektem ich oddziaływania może być proces błądzenia losowego. Celem artykułu jest weryfikacja hipotezy błądzenia losowego na rynku kontraktów terminowych na pszenicę na giełdach CBOT i MATIF. Badania mają wskazać na typ szeregów czasowych, jaki reprezentują kontrakty terminowe błądzenia losowego: persystentny czy antypersystentny. Podejmuje się także kwestie znaczenia długości ocenianego szeregu czasowego oraz terminem wykupu kontraktu. Narzędziem służącym do określenia właściwości statystycznych ocenianych szeregów czasowych był wymiar fraktalny szacowany metodą podziału pola. W artykule podano skrócony schemat liczenia tego wymiaru oraz fragment tablic istotności. Wyniki badań z reguły nie pozwalają na odrzucenie hipotezy błądzenia losowego notowań kontraktów na pszenicę, co sugeruje ich nieprzewidywalne kształtowanie. Natomiast na giełdzie CBOT zauważono nieco inne zachowanie kontraktów grudniowych i marcowych, którym w krótszym okresie czasu bliżej jest do procesu ze wzmacnianym trendem oraz kontraktów majowych, czerwcowych i wrześniowych, którym bliżej jest do procesu o korekcie w kierunku średniej. Jednak wszystkie wyniki były w granicach błądzenia losowego. Takiej właściwości nie zauważono dla kontraktów notowanych na giełdzie MATIF. Zauważoną właściwość dla giełdy CBOT można próbować wykorzystywać w celach spekulacyjnych. Intersująca okazuje się kwestia długości szeregu czasowego. W opracowaniu badano tylko szeregi długości 200 obserwacji (ok. 10 miesięcy) oraz 500 obserwacji (ok. 2 lata). Okazuje się, że wraz z wydłużaniem horyzontu czasowego wyniki stają się bardziej stabilne, tj. spada liczba przypadków przeczących hipotezie błądzenia losowego. Hipoteza błądzenia losowego jest hipotezą, którą trudno odrzucić. Potwierdza się przypuszczenie, że czynniki kształtujące ceny na rynkach rolnych determinują ich losowy przebieg.

Słowa kluczowe: rynek zbóż, ceny futures, błądzenie losowe, szeregi persystentne, szeregi antypersystentny