

# Entropy of Globalizing World Macroeconomy Time Series Analysis

J. MIŚKIEWICZ<sup>a,b,\*</sup>

<sup>a</sup>*Institute of Theoretical Physics, University of Wrocław, pl. M. Borna 9, PL 50204 Wrocław*

<sup>b</sup>*Department of Physics and Biophysics,*

*Wrocław University of Environmental and Life Sciences, ul. Norwida 25, PL 50375 Wrocław*

Doi: [10.12693/APhysPolA.138.25](https://doi.org/10.12693/APhysPolA.138.25)

\*e-mail: [janusz.miskiewicz@uwr.edu.pl](mailto:janusz.miskiewicz@uwr.edu.pl)

In this paper, the globalisation measure based on the permutation entropy is proposed. The method is applied to four sets of time series: GDP, CPI, employment rates and CO<sub>2</sub> emission rates. It has been shown that the results agree with the general observations of the globalization process. A particular achievement of the study is the proposition of the general method, which can be applied to any aspect of human activity indifference to the existing indexes which are focused on particular parameter or group of parameters.

topics: time series analysis, permutation entropy, econophysics

## 1. Introduction

The globalization is the result of the interaction and integration between people, companies, and governments worldwide. This process was initiated by the inventions facilitating transport, such as railways, cars, aeroplanes and communication — telegraph, telephone, radio, television, and the Internet. One of the most spectacular globalization process in the 20th century was the creation of the European Community, which at the level of political and legal regulation facilitates and improves the functioning of its members, and influence various aspects of EC citizens life.

Since the globalization was recognized as a real process which affects various aspects of human activity there were attempts to measure the globalization quantitatively. In facts, many indexes have been constructed [1–7]. Although it is difficult to summarise those indexes in a few sentences the most common feature is that they are based on some particular aspect of human activity (or some set) [2, 4, 6–14], which is considered representative for the process and measure its state. In fact, the official broadly acceptable globalization index has not been established. Therefore the approach based on the information theory has been proposed [15]. Because the globalization is the result of imposed rules (e.g. by countries agreements or by new facilities) the globalization measure should be capable to observe effects induced by those rules. In statistical physics, the standard measure of the disorder of a system is the entropy. This idea was transferred into information entropy where the classical Shannon entropy

$$S = - \sum_i p_i \ln(p_i). \quad (1)$$

measures the information content of the data ensemble. For the discrete probability distribution the index  $i$  takes the values of the distribution domain. In the case of continuous distribution the differential entropy is defined:

$$S = - \int p(x) \ln(p(x)) dx. \quad (2)$$

The entropy estimation in the case of continuous distribution is a difficult problem. Particularly due to limitations in available data. There are different strategies to estimate sample entropy, e.g., [16–18], but still relatively big amount of data is required.

The entropy is useful because it is sensitive to “randomness” of the system. If the system is “fully random” the entropy is maximal while there are some rules on the system or patterns the entropy is capable to capture it. However, in the case of the macroeconomy, the accessible data sets are of limited size, particularly due to the problem with gathering and processing those data. Most of the macroeconomy parameters are published annually, sometimes quarterly or monthly. In result the analysed time series are of limited size, e.g. five year period is described by five data points for annual frequency or 60 in the case of monthly data. In the case of the standard information entropy measure (1), which is based on probabilities, it is extremely hard to estimate it. A possible solution is to replace the entropy by another measure of disorder, e.g. Thail distance [19–21], or try to estimate the probability density function (e.g. [22]). On the other hand in various financial analysis instead of exact values the symbolic parameters are

used. For example, discussing the recession problem the GDP data are converted into a dichotomous signal of growth — recession states. Similarly, the financial analysis of a market player is finished with the discrete set of decisions: buy, sell and sometimes wait. Therefore the symbolic entropy, i.e., permutation entropy is applied in this paper to measure similarities in development patterns among countries. This approach has one more advantage it is not focused on a particular aspect of human activity but can be applied to any data. In fact, there are areas in which countries strongly cooperate and others which are protected.

## 2. Mathematical background

The proposed method of the globalization measurement is based on the permutation entropy introduced by Brant and Pompe [23]. For the convenience of the reader, the algorithm is shortly recalled here.

Let denote the analysed time series as  $X = x_t$ ,  $t = 1, \dots, N$ . From the time series  $X$  a vector of  $D$ -th subsequent values is taken:

$$(x_s, x_{s+1}, \dots, x_{s+D-1}). \quad (3)$$

Obviously  $D < N$ . This vector is sorted in ascending order, and a permutation pattern  $\pi$  is created.

$$\pi = (r_0, r_1, \dots, r_{D-1}), \quad (4)$$

which fulfils

$$x_{s+r_0} \leq x_{s+r_1} \leq \dots \leq x_{s+r_{D-2}} \leq x_{s+r_{D-1}}. \quad (5)$$

The procedure of permutation pattern creation is repeated to estimate the probability of the pattern occurrence. Finally, the permutation entropy (PE) is

$$PE = - \sum_{i=1}^{D!} \pi_i \ln(\pi_i). \quad (6)$$

$D$  is defined as the permutation entropy order.

For example, let consider a series of eight elements:  $\{3, 2, 4, 6, 3, 1, 5, 2\}$ . In the case of PE of the order  $D = 3$  there are six possible permutations:  $(x_1, x_2, x_3)$ ,  $(x_1, x_3, x_2)$ ,  $(x_2, x_1, x_3)$ ,  $(x_2, x_3, x_1)$ ,  $(x_3, x_1, x_2)$ ,  $(x_3, x_2, x_1)$ . In the considered series there are six tuples of three consecutive elements:  $(3, 2, 4)$ ,  $(2, 4, 6)$ ,  $(4, 6, 3)$ ,  $(6, 3, 1)$ ,  $(3, 1, 5)$ ,  $(1, 5, 2)$ , where  $(2, 4, 6)$  corresponds to the permutation 012 since  $x_{t+0} < x_{t+1} < x_{t+2}$ , while  $(1, 5, 2)$  is the permutation 021 because  $x_{t+0} < x_{t+2} < x_{t+1}$ . Analogously,  $(3, 2, 4)$  and  $(3, 1, 5)$  are of type 102 because  $x_{t+1} < x_{t+0} < x_{t+2}$ ,  $(4, 6, 3)$  is 201 since  $x_{t+2} < x_{t+0} < x_{t+1}$ , and  $(6, 3, 1)$  is 210 because  $x_{t+2} < x_{t+1} < x_{t+0}$ . Finally, the frequencies of the permutations occurrence are:  $\pi_1 = 1/6$ ,  $\pi_2 = 1/6$ ,  $\pi_3 = 2/6$ ,  $\pi_4 = 0$ ,  $\pi_5 = 1/6$ ,  $\pi_6 = 1/6$  and  $PE \approx 1.56$ .

The permutation entropy can be calculated for the given time series or any part of it. The later allows us to analyse the evolution of PE. This can be done in overlapping or non-overlapping time window moving along the time series. Within this work, the overlapping time window will be used.

The estimation of PE is the core of the analysis. However, the globalization measurement requires a comparison of the appropriate time series. The cross-correlation analysis of the complex set of entities usually is performed in two steps firstly the cross-correlation is performed and in the second stage, the correlation matrix is analysed [24]. In this paper the following algorithm is used:

1. Define the time window length  $T$  and the order of PE  $D$ .
2. For each of the time series take the subseries  $i = 1, \dots, T$  and estimate PE.
3. Move the time window by one data point and estimate PE for the consecutive time window. Continue the procedure until the end of the time series is reached. Finally, the set of PE transformed time series is obtained.
4. Define the time window of cross-correlation analysis  $T_M$ .
5. Starting from the beginning of the  $PE$  transformed time series takes the subseries of the length  $T_M$  calculate the cross-correlation matrix based on Manhattan distance.

Denoting the time series as  $A$  and  $B$  and theirs elements as  $a_i$  and  $b_i$  respectively the Manhattan distance between  $A$  and  $B$  in the interval  $i \in (t, t + T_M)$  is defined:

$$MD(A, B)_{t, T_M} = \frac{1}{T_M} \sum_{i=t}^{t+T_M} |a_i - b_i|. \quad (7)$$

The time window is moved at each step by one data point. Finally, the mean value of the cross-correlation matrix is calculated and presented as a final result of the analysis.

The analysis is performed on equal size and simultaneous time series. In the following, the proposed algorithm will be recalled as Permutation Entropy of Globalization Measure (PEGM).

## 3. Macroeconomy data

The globalization is the result of a process which is not general but observed in particular aspect of human activity and of course, it is expected that there are differences in the level of globalization among different groups. To illustrate this fact the analysis is performed on the time series in which the globalization should be visible — gross domestic product, consumer price index, employment rate and CO<sub>2</sub> emission. The time series, except CO<sub>2</sub> emission case, were converted into increment time series to remove trend:

$$R(t) = \frac{p(t) - p(t-1)}{p(t-1)}. \quad (8)$$

Due to the requirement of the same size of the time series length and the simultaneous of the data, some of the time series have not been included in the analysis. Finally, the following databases were used:

- *Gross Domestic Product (GDP)* Annual percentage growth rate of GDP per capita based on constant local currency. Aggregates are based on constant 2010 U.S. dollars. GDP per capita is gross domestic product divided by midyear population. World Bank database [25]. 86 time series. The time series were grouped according to geographical regions: Africa, North America, Central America, South America, Australia and Oceania, Europe.
- *Consumer Price Index (CPI)* Quarterly data from Q4.1997 till Q4.2018, percentage change on the same period of the previous year. OECD database [26]. 18 time series. The time series were grouped according to geographical regions: America, Asia, Europe.
- *Employment rate* Monthly data from 01.1971 till 05.2019, percentage change on the same period of the previous year. OECD database [26]. 24 time series. The time series were grouped according to geographical regions: Europe, non-Europe.
- *CO<sub>2</sub> emission* Yearly data 1970–2017 [27], 153 series. The time series were grouped according to geographical regions: Africa, North America, Central America, South America, Asia, Australia and Oceania, West Europe, East Europe, Europe.

#### 4. Analysis

The proposed method of globalization measure requires three parameters: PE order, PE window size and cross-correlation window size. The determination of these parameters is difficult in principle because one has to reconcile two opposing needs: quality of statistical estimation and data accessibility. Considering the size of the time series the time window for PE analysis could not be too long. For the yearly data, the size of 12 years have been chosen as the longest considering the problem of data stationarity. Therefore, the order of the PE was taken  $D = 4$  yrs.

The main idea of PEGM is to measure similarities, so if in the region or among a group of countries the globalization process went into higher level then the similarities are higher and PEGM is smaller.

The general results of PEGM in the case of GDP time series are presented in Fig. 1. The time series were grouped geographically assuming that the physical localisation has the strongest impact on economical evolution. The following regions have been distinguished: Africa, Asia, North America, Central America, South America, Australia and Oceania and Europe. The first observation (Fig. 1) is that there are different levels of globalization. The lines presenting the evolution of globalization in, e.g. Europe and Australia and Oceania differ significantly. This observation is supported by the fact

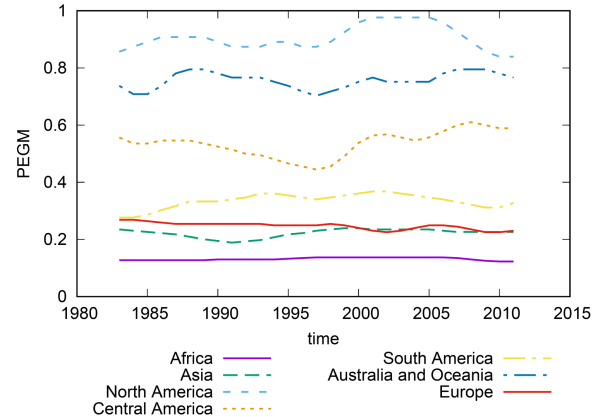


Fig. 1. Permutation Entropy Globalization Measure in the case of GDP increment time series. Parameters of the analysis:  $D = 4$  yrs,  $T = 12$  yrs,  $T_M = 4$  yrs.

that Australia and Oceania is the region where most of the countries are located on islands with natural well-defined borders. Additionally, Australia is much larger than the accompanied countries so the level of cooperation is rather low. In result, it is hard to expect significant similarities. The high value of PEGM agrees with low globalisation in the region. The opposite is in Europe. The countries in this region have a long tradition of trade, cooperation and interaction. European history is full of wars, alliances, trade, cooperation etc. Particularly, after the Second World War during the restoration process, the base of the European Community was settled and finally, the EU was established in 1967. Since then the process of integration, law unification, increase of cooperation is continued. Therefore, the high level of similarities presented as the low values of the mean PGEM (Fig. 1) is not surprising. It is noticeable that a similar level of PGEM is achieved by the Asia region. In some sense, there are several similarities to Europe — the cultural heritage, present high similarities of the level of development, so a low value of PGEM is understandable. However, the detailed analysis of the Asia case is very complicated and goes beyond the scope of this paper.

An interesting result of the analysis is that the top line describes globalization of the North America countries, i.e., Canada, Mexico and the USA. This might be the result that the USA and Canada are so developed that can be really independent. On the other hand, Mexico is on a much different stage of development, so the high differences in economical development results in the high values of PGEM. Special attention is drawing Africa because the PGEM shows a high level of similarities in GDP time series. This is a special case because this is rather due to development of economies rather than the result of the globalization process.

The level of globalization among considered regions differ significantly such that the details of

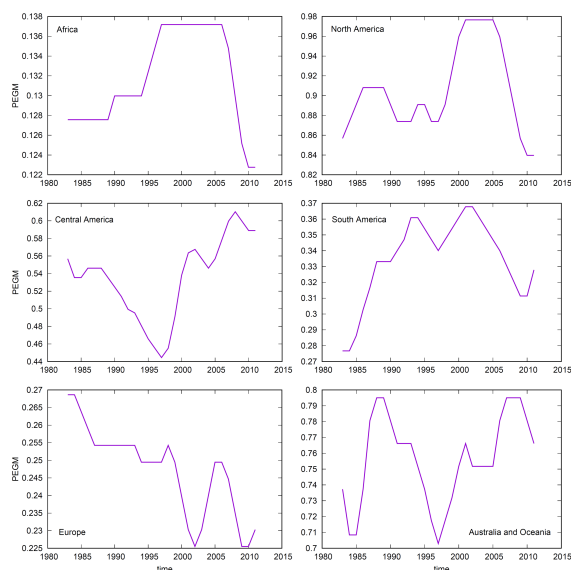


Fig. 2. Permutation Entropy Globalization Measure in the case of GDP increment time series. Parameters of the analysis:  $D = 4$  yrs,  $T = 12$  yrs,  $T_M = 4$  yrs.

the evolution are hidden by the scale, therefore, results of PEGM are also presented on separated plots (Fig. 2). One of the most interesting cases is Europe because the globalization processes are the strongest. It is seen (Fig. 2) that in years 1983–2002 the decrease is almost monotonical. Particularly interesting is the decrease in 2002 which might be related to the Euro introduction process. After that moment some deglobalization might be noticed probably induced by the global crises, which force the countries to undertake special measures to protect their economies, also the responses to crises were individual. Interesting cases are the regions of Africa, South America and North America where recently, i.e., in years 2005–2010 decrease of the mean distance of permutation entropy is observed. The special case is Central America where the deglobalization is observed since 1997. This might be related to the political and economic instabilities of the region.

The CPI analysis was performed on monthly data (Fig. 3). Taking the opportunity of relatively frequent data the number of data points in the time windows was increased up to 24 (for both  $T$  and  $T_M$ ) to improve quality of appropriate parameters estimations. The length of 24 months allows also to smooth seasonal effect on the data. In this case, a limited amount of time series was available therefore only three regions were distinguished: Europe, Asia and America. As in the case of GDP data the European region is characterised by the lowest PEGM values. There is a significant gap between the highest values of Europe and the lowest Asia PEGM suggesting significant difference among these markets regions. In fact, in Europe 19 countries have a common currency which has a strong

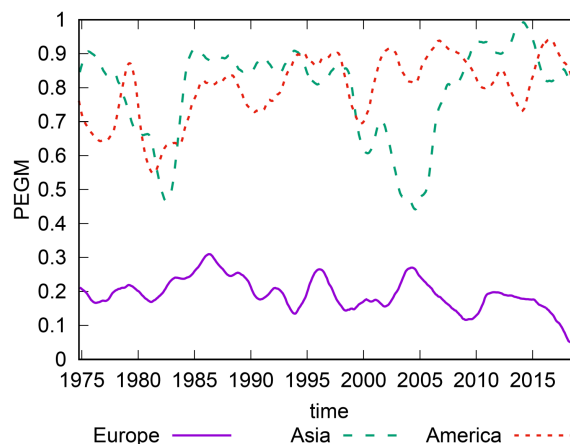


Fig. 3. Permutation Entropy Globalization Measure in the case of CPI increment. Parameters of the analysis:  $D = 4$  mths,  $T = 24$  mths,  $T_M = 24$  mths.

impact on the price level. Additionally, those countries have a common central bank, so the high globalization level is not surprising. The evolution of PEGM can be related to historical events, for example, the period of relatively stable PEGM in the case of CPI around 2000 can be related to Euro introduction. In this process, the inflation was one of the key parameters therefore much attention was paid to the price level which resulted in a stable value of PEGM during that period. Recently, since 2011 the PEGM in Europe region is decreasing which suggest increasing globalization. The opposite situation is observed in Asia where not only high values of PEGM are observed but also significant fluctuations. For example in the interval 2005–2015 PEGM was increased from 0.4 up to 1, so the change was 150%.

The employment rate was analysed to verify the PEGM method if it is not giving false positive answers. In fact, the labour market is usually protected because the unemployment rate is a serious political problem. Even in Europe, there are tendencies to protect local the labour market. For example, when Poland was accessing the European Community the member countries made statements of the labour market protection periods. The European Community is a special case around the world where, in principle, citizens have rights to settle in any country without special permissions. Therefore, after 2007 the constant decrease of PEGM is observed (Fig. 4). The opposite was observed in other countries. In fact, it is difficult to distinguish any particular tendency or trend. The labour market, in general, is not globalizing.

Among the globalization processes, global warming takes recently special attention not only by scientists but also by politicians. Therefore, the analysis of  $\text{CO}_2$  emission rate was included in the analysis. In this case, the most complete database was available so it was possible to distinguish nine regions some of them overlapping —

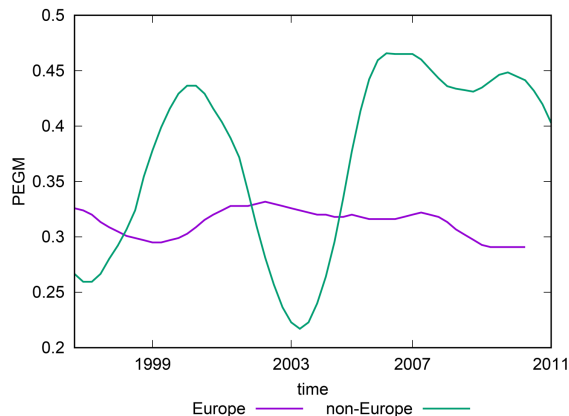


Fig. 4. Permutation Entropy Globalization Measure in the case of employment rate. Analysis parameters:  $D = 4$  qrs,  $T = 12$  qrs,  $T_M = 12$  qrs.

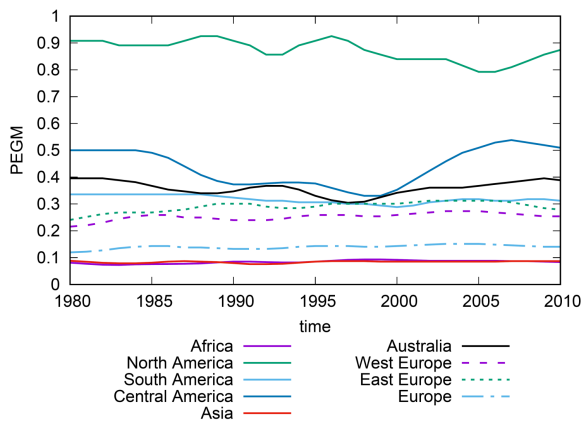


Fig. 5. Permutation Entropy Globalization Measure in the case of CO<sub>2</sub> emission rate. The analysis parameters:  $D = 4$  yrs,  $T = 12$  yrs,  $T_M = 4$  yrs.

Europe and West and East Europe. The overlapping regions were included in the analysis because on one hand some agreements and actions on an European level but there are significant differences in industrial development which influence CO<sub>2</sub> emission therefore not only Europe but also West and East Europe was included in the analysis. The results are presented in Fig. 5. The most noticeable feature in this plot is the separated line at the top of the graph. This is a curve describing the North America region. The result agrees with the fact that the USA never accepted the Kyoto agreement of the CO<sub>2</sub> emission reduction. Analogously Canada which abandon this agreement. In result, there are low similarities in CO<sub>2</sub> emission rate time series. The second region in which there is a little effort on global warming reduction is Central America. The interesting case is Europe where a lot is done for CO<sub>2</sub> at least on the European Community level. It should be stress that the results of PEGM present low values denoting a high level of similarities of CO<sub>2</sub> emission levels. In fact, it might be also the result of similarities in industrial development level.

This conclusion is also supported by the fact the difference is observed between West and East Europe (green and violet dashed lines in Fig. 5). The highest similarities are observed in the case of Asia and Africa, but this is rather due to the low level of industrial development than the intentional effort of these countries.

## 5. Conclusions

In the presented analysis the permutation entropy globalization measure was introduced. The primary aim of this work was to verify if the proposed method properly describes the globalization process. Therefore, two types of time series were analysed those where the globalisation is present and those protected by countries. Particularly spectacular result was obtained in the case of CO<sub>2</sub> emission rates in which countries officially neglecting needs of CO<sub>2</sub> emission reduction are described by a distinct curve. On the other hand, Europe (probability due to the effort of the European Community) presents a high level of entropy similarities in all aspects: GDP, CPI, employment rate and CO<sub>2</sub> emission rates.

Slightly surprising is the result of Africa and Asia where high similarities are observed. This result deserves more detailed studies highly probable is the hypothesis that this the results of a low level of development.

Summarising the presented results it can be stated that the proposed, permutation entropy-based method is capable to describe the globalisation process. The main advantage of the method, indifference to the existing indexes, is that it is not focused on a particular aspect of human activity but the changes introduced by rules imposed on a system. Therefore, the proposed method might be considered more fundamental then existing approaches.

## References

- [1] S. Maslov, *Physica A* **301**, 397 (2001).
- [2] R. Kluver, W. Fu, “The Cultural Globalization Index” in: *Foreign Policy*, 2004.
- [3] A.T. Kearney, Foreign Policy, “The Globalization Index” in: *Foreign Policy* **157**, 74 (2006).
- [4] P. Martens, D. Zywieta, *J. Int. Dev.* **18**, 331 (2006).
- [5] S. Hook Law, W.N.W. Azman-Saini, Hui Boon Tan, *Emerg. Mark. Finance Trade* **50**, 210 (2014).
- [6] H.S. Huh, C.Y. Park, *ADB Economics Working Paper Series No. 587* (2019).
- [7] S. Gygli, F. Haelg, N. Potrafke, J.E. Sturm, *Rev. Int. Organ.* **14**, 543, (2019).
- [8] T.M. Andersen, T.T. Herbertsson, *IZA Discussion Papers, No. 817*, (2003).

- [9] A. Heshmati, *Glob. Econ. J.* **6**, 1850087 (2006).
- [10] M. Caselli, *Globalizations* **5**, 383 (2008).
- [11] D. Axel, G. Noel, M. Pim, *J. Glob. Stud.* **1**, 166 (2010).
- [12] P. Samimi, G.C. Lim, A.A. Buang, *J. Knowl. Manag. Econ. Inf. Technol.* **1**, 197 (2011).
- [13] B.B. Rao, A. Tamazian, K.C. Vadlamanati, *Appl. Econ.* **43(5)**, 551 (2011).
- [14] P. Martens, M. Caselli, P. De Lombaerde, L. Figge, J.A. Scholte, *Globalizations* **12(2)**, 217 (2015).
- [15] J. Miśkiewicz, in: *Theory and Applications of Time Series Analysis*, Eds. O. Valenzuela, F. Rojas, H. Pomares, I. Rojas, Springer International Publishing, 2019 p. 217.
- [16] J. Miśkiewicz, *Physica A* **450**, 473 (2016).
- [17] N. Ebrahimi, K. Pflughoeft, E.S. Soofi, *Stat. Probab. Lett.* **20**, 225 (1994).
- [18] R.B. Govindan, J.D. Wilson, H. Eswaran, C.L. Lowery, H. Preissl, *Physica A* **376**, 158 (2007).
- [19] J. Miśkiewicz, *Physica A* **387**, 6595 (2008).
- [20] J. Miśkiewicz, *Physica A* **389**, 1677 (2010).
- [21] J. Miśkiewicz, M. Ausloos, *Physica A* **389**, 797 (2010).
- [22] J. Miśkiewicz, *Physica A* **450**, 473 (2016).
- [23] C. Bandt, B. Pompe, *Phys. Rev. Lett.* **88**, 174102 (2002).
- [24] J. Miśkiewicz, *Acta Phys. Pol. A* **121**, B-89 (2012).
- [25] The World Bank, *WDI — economy*, 2019.
- [26] *OECD statistics*, 2019.
- [27] M. Muntean, D. Guizzardi, E. Schaaf, M. Crippa, E. Solazzo, J.G.J. Olivier, E. Vignati, “Fossil CO<sub>2</sub> emissions of all world countries: 2018 report”.