International Science Journal of Engineering & Agriculture
2024; 3(4): 30-37
https://isg-journal.com/isjea
doi: 10.46299/j.isjea.20240304.03
ISSN: 2720-6319

Diagnosing the stability of large-scale processes using fractal structure analysis of time series

Nodar Abelashvili
Faculty of Informatics and control systems, Georgian Technical University, Tbilisi, Georgia

Nona Otkhozoria
Faculty of Informatics and control systems, Georgian Technical University, Tbilisi, Georgia
ORCID 0000-0002-5837-5345

Vano Otkhozoria
Faculty of Informatics and control systems, Georgian Technical University, Tbilisi, Georgia
ORCID 0009-0008-6028-3758

Eka Akhlouri
Faculty of Informacis and control systems, Georgian Technical University, Tbilisi, Georgia

To cite this article:
Abelashvili Nodar, Otkhozoria Nona, Otkhozoria Vano, Akhlouri Eka. Diagnosing the stability of large-scale processes using fractal structure analysis of time series. International Science Journal of Engineering & Agriculture. Vol. 3, No. 4, 2024, pp.30-37. doi: 10.46299/j.isjea.20240304.03.
Received: 06 17, 2024; **Accepted**: 07 23, 2024; **Published**: 08 01, 2024

Abstract: This paper aims to study large-scale processes that persist over time to describe their stability, and to monitor and diagnose negative changes. By utilizing fractal structure analysis of time series, the research investigates the applicability of the Hurst exponent in diagnosing the stability of various natural and man-made systems. The findings highlight the limitations of standard Gaussian statistics and the effectiveness of fractal analysis in revealing hidden patterns and long-term dependencies in complex systems.

Keywords: Fractal analysis, Hurst exponent, time series, large-scale processes, stability diagnosis, non-periodic cycles, long-term memory.

1. Introduction

Ensuring the stability of measurement results is crucial in sectors requiring accurate and reliable measurement of large-scale process parameters, such as production, transportation, energy, and environmental monitoring. These processes often exhibit complex dynamic changes in both time and space. Minor fluctuations or unforeseen effects under such conditions can significantly impact measurement results, leading to loss of stability and accuracy (Butakov & Grakovsky, 2005).

Researchers and engineers face various challenges, primarily stemming from disturbances (errors) and external influences affecting the measurement process. These include environmental changes, electromagnetic interference, mechanical vibrations, and other uncontrollable factors. Moreover, large-scale processes evolve over time due to equipment wear, material composition changes, and environmental variations, further diminishing measurement stability and data quality.

To address these issues, effective methods for diagnosing measurement stability must be developed. These methods enable quick identification and mitigation of unwanted effects, as well as real-time monitoring and control of measurement processes. Fractal analysis of time series emerges as a promising tool in this regard, uncovering hidden patterns and structural features that classical methods may overlook (Falconer, 2013) (Otkhozoria, Azmaiparashvili, Petriashvili, Otkhozoria, & Akhlouri, 2023).

Diagnostics play a pivotal role in ensuring measurement reliability and accuracy across various industries, scientific disciplines, and technological applications. They facilitate early detection of anomalies and deviations from norms during measurements, preemptively averting potential errors. Diagnostics also assess the quality and stability of measurement results, essential for reliability and accuracy assurance. Analyzing statistical data characteristics and noise levels further aids in determining measurement reliability.

Results from diagnostics inform the development and implementation of management and correction methods for measurement systems. These may involve automated equipment calibration, compensation for external influences, or optimizing measuring device parameters to enhance accuracy and reliability. In industrial and manufacturing contexts, measurement accuracy directly impacts final product/service quality. Compliance with quality standards and regulatory requirements, driven by diagnostics, ensures customer satisfaction and regulatory compliance.

Thus, developing diagnostic mechanisms for observing, monitoring, controlling impacts, and preventing adverse consequences of global environmental changes and large-scale technological processes is a pressing task of our time. Establishing theoretical foundations and practical implementations in diagnostics is critical for meeting modern challenges effectively (Chkheidze, Otkhozoria, & Narchemashvili, 2021).

2. Object and subject of research

The object of this research is large-scale processes that exhibit complex dynamics over extended periods. The subject of this research is the stability and behavior of these processes as analyzed through the fractal structure of their time series data.

3. Target of research

The aim of this paper is to study large-scale processes that persist over time, to describe the stability of their condition, and to monitor and diagnose negative changes. By employing fractal analysis techniques, this research seeks to uncover underlying patterns and provide insights into the stability and potential anomalies within these processes.

4. Literature analysis

Standard Gaussian statistics are valid based on the following assumptions. The central limit theorem states that as the number of trials increases, the marginal distribution of a random system will approach a normal distribution. The events must be independent and identically distributed (i.e., they must not influence each other and must have the same probability of occurrence).

When studying complex systems, it is common to assume that the system follows a normal distribution so that standard statistical analysis can be applied. However, systems studied in practice (such as sunspots, annual precipitation averages, financial markets, and time series of economic indicators) often do not follow a normal distribution. Hurst proposed the normalized range method (R/S analysis) for analyzing such systems. This method distinguishes between random and fractal time series and draws conclusions about the existence of non-periodic cycles, long-term memory, and other characteristics.

5. Research methods for Fractal analysis of time series of multi-year monitoring results

Fractal analysis of time series from multi-year monitoring data at Vardzia's complex monitoring system considers temperature observations recorded every three hours in December from 2017 to

2023 by an internal space sensor of a boiler. The monthly observations total 240 (30 days × 8 measurements per day = 240, denoted as n=240). Over the entire seven-year period, this results in a total of N=1680 observations.

These observations are categorized by year, resulting in fragments of size n=240 (1680 observations ÷ 7 years = 240), indexed by i. Each fragment is further divided into daily elements, with m=8 measurements per day. Each individual measurement is denoted by the index u.

Table 1 presents the initial results of seven years of observation from December 1-15, 2017, segmented into fragments and daily elements.

Table 1. Data for the month of December 2017

01-15 December 2017															
	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
$\bar{X}_m$	5.65	6.56	6.80	5.25	5.67	5.43	4.98	4.87	4.65	5.38	5.83	5.78	5.58	5.23	5.04
$X(m,$	9.85	8.73	11.23	10.59	10.54	11.74	11.89	13.23	9.05	8.59	10.13	9.09	9.05	8.59	10.13
S_m	0.65	0.59	0.76	0.69	0.74	0.68	0.55	0.94	0.86	0.93	0.84	0.64	0.56	0.71	0.58
16-31 December 2017															
	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
$\bar{X}_m$	6.24	5.54	5.24	5.65	6.24	5.55	4.88	4.23	4.38	4.23	4.79	4.32	4.58	5.12	5.24
$X(m,$	9.54	10.74	10.99	12.23	14.55	14.66	14.04	13.27	10.85	9.73	10.23	11.59	10.87	10.85	12.52
S_m	0.428	0.565	0.74	0.72	0.75	0.68	0.74	0.58	0.52	0.63	0.63	0.78	0.65	0.76	0.76
01-30 December 2017 $\boldsymbol{\bar{X}_{2017} = 5.297}$ $\boldsymbol{S_n}$ =0.659 $\boldsymbol{X(m,n)}$ =4.475 $\boldsymbol{R_n}$ =11.03															

The daily values of the data are presented as arithmetic averages derived from the entire day's observations (8 measurements each day):

$$\bar{X}_m = \frac{1}{n}\sum_{i=0}^{8} x_i(t_i°\mathrm{C}), \tag{1}$$

And the arithmetic average for each year (fragments) of the December data from 2017 to 2023 is as follows:

$$\langle x(n)\rangle = \sum_{i=1}^{n} x(i) = \frac{1}{7}(\langle x_{2017}\rangle + \langle x_{2018}\rangle + \langle x_{2019}\rangle + \langle x_{2020}\rangle + \langle x_{2021}\rangle + \langle x_{2022}\rangle + \langle x_{2023}\rangle) = \frac{1}{7}(5{,}265 + 5{,}839 + 6{,}142 + 6{,}19 + 6{,}202 + 6{,}23 + 6.95) = 6{,}1;$$

The total deviations from the average value are calculated using the formula:

$$X(m,n) = \sum_{u=1}^{m}[X(u) - \langle x(n)\rangle] \tag{2}$$

where $\langle x(n)\rangle = \bar{x}(n)$ represents the average value of the complete December 2017 dataset; $X(u)$ - denotes the arithmetic average of daily data for December of the respective year; Accordingly, $X(m,n)$ is the cumulative sum of deviations from the average value for the same year and month.

The data for December 16-30, 2017, and each subsequent year are calculated similarly, except that the data from the first and second halves of December of each year are aggregated. For instance, the combined value for the first half of December 2017 is $X(m,n) = \sum_{u=1}^{m}[X(u) - \langle x(n)\rangle]$, and for the second half, it would be:

$$X(m,n)_{dec} = (1 - 15\ dec.)X(m,n)_{dec} + (16 - 30\ dec.)X(m,n)_{dec} = 0{,}92$$

We calculate the mean square deviation of the observation data:

$$S_m = \sqrt{\frac{(\bar{X}_m - X_u)^2}{m-1}} \tag{3}$$

Also, for each fragment, we calculate the range of deviations, also known as the swing or amplitude change:

$$R(n) = max_m X(m,n) - min_m X(\mathrm{m,n}) \tag{4}$$

According to the observation results, the schedule depicting changes in average temperature values for the month of December is provided in Table 2 and Figure 1.

Table 2. Average temperatures for the month of December

years of observation	2017	2018	2019	2020	2021	2022	2023
Arithmetic average	5.265	5.84	6.142	6.19	6.202	6.23	6.95

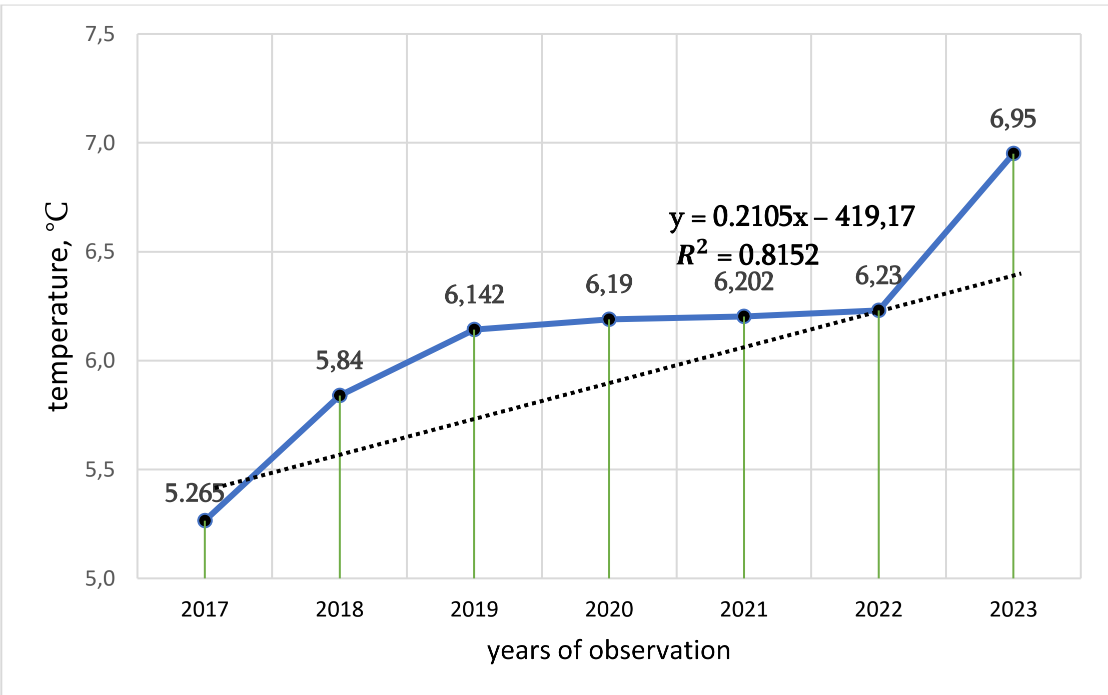

Figure 1. The change in the average value of the temperature in the month of December.

The change in the average value of the temperature in the month of December. Based on these data, it is possible to identify a range of temperature values where a change in temperature is expected with high probability, and any deviation from this range can be considered an anomalous event. In our case, this range is ±0.5°C.

The algorithm for calculating the obtained data will be formed as follows:

1. We divide N (observations over 7 years) into fragments (the month of December every year) and elements (daily observations of the month with 3-hour intervals), totaling 8 measurements per day x 30 days x 7 years = 1680 measurements.

2. We introduce notation for data:

$\{x(i)\}, i = 1, 2, \dots, n,$ which represents the arithmetic mean:

$$\langle x(n)\rangle = {}^{1}\!/_{n} \sum_{i=1}^{n} x(i); \tag{5}$$

over the years of observation (the mean of the means), approximately $\langle(6{,}117)\rangle \approx 6{,}1$.

3. To calculate the standard deviation:

$$S(n) = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}[x(i) - \langle x(n)\rangle]^2} \tag{6}$$

Calculated for each year, the standard deviation of the average temperatures in December (for December 2018) is:

$$S_{\bar{\bar{X}}} = \sqrt{\frac{(\bar{X}_n - X_i)^2}{n-1}} = 0.595 \tag{7}$$

The sum of deviations from the mean value after m measurements is

$$X(m,n) = \sum_{u=1}^{m}[\bar{X}_m - \langle x(n)\rangle] \tag{8}$$

The difference between the maximum and minimum values of the received data, or deviation, is $R(n) = max_m X(m,n) - min_m X(m,n)$

Hurst's empirical law applies to a large number of natural phenomena and processes:

$$\left(\frac{n}{2}\right)^H = \frac{R(n)}{S_n} \tag{9}$$

The steps are:

1. For each element (one day) in a fragment (one month), calculate the sum of deviations from the average.

2. Calculate the displacement for all elements.

3. Divide the deviation by the mean square deviation of this fragment.

The results of the obtained data are presented in Table 3.

Table 3. Fragments according to the years 2017-2023

№	Years	$R(n)$	S_n	$R(n)/S_n$	$log\, R(n)/S_n$	$\left(\frac{n}{2}\right)^H$	$H = log\, n/2$
1	2017	12.63	1.33	9.49	0.97	$({}^{30}\!/_{2})^H$	1.176
2	2018	11.45	0.62	18.46	1.267	$({}^{60}\!/_{2})^H$	1.47

3	2019	11.17	0.64	17.45	1.24	$(^{90}/_{2})^{H}$	1.95

Continued table 3

4	2020	12.12	0.602	20.13	1.29	$(^{120}/_{2})^{H}$	1.78
5	2021	13.07	1.05	12.44	1.11	$(^{150}/_{2})^{H}$	1.87
6	2022	12.67	0.746	17.12	1.23	$(^{180}/_{2})^{H}$	1.95
7	2023	13.21	0.686	19.25	1.28	$(^{210}/_{2})^{H}$	2.02

For all seven fragments, plotting $\left[\log\frac{R(n)}{S(n)}, \log\frac{n}{2}\right]$ gives a set of points.

Using the method of least squares, we find the mathematical model of the regression line equation, whose angular coefficient is the estimate of the Hurst index.

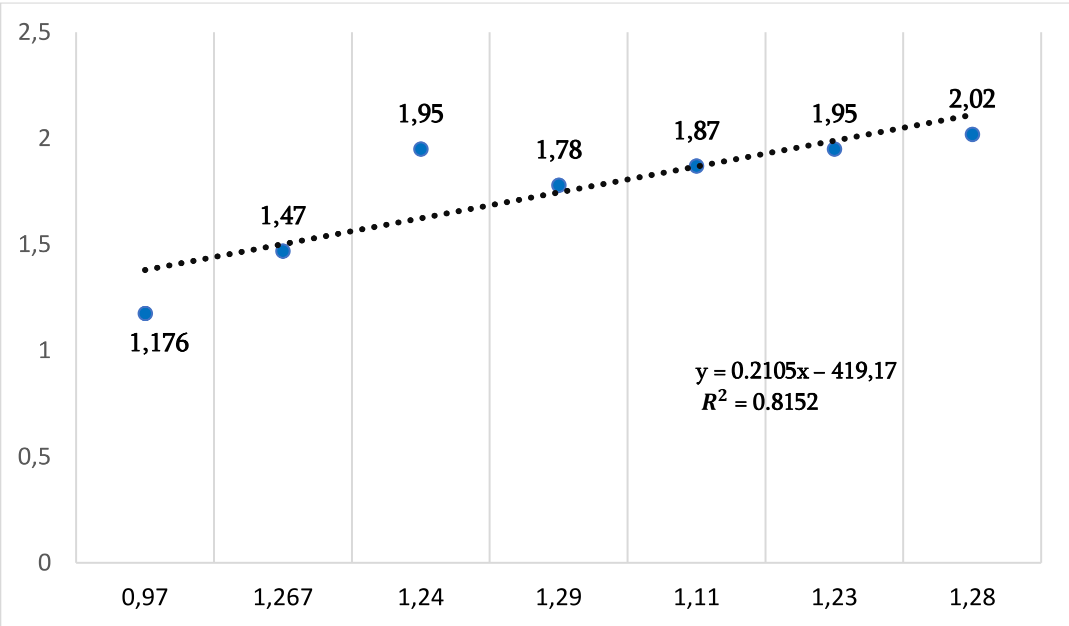

Figure 2. Logarithmic values of coordinates for fragments by years.

The regression equation is:

$$Y = 0.2105x + 419.17; \tag{10}$$

The correlation coefficient is $R^2 = 0.8152$

The coefficients of the obtained equation were tested for adequacy based on the t-Student criterion. The coefficients were found to be adequate. A high correlation coefficient and a P-value of less than 0.05 provide a basis to conclude that the obtained equation adequately describes the process and can be used for prediction.

The value of the Hurst coefficient $H = 0.2105 < 0.5$ indicates that the process is stable, and extreme deviations are not expected. This approach can be used for the early diagnosis of long-term monitoring results.

The Hurst coefficient is related to the level of self-similarity or the long-term trend (persistent) behavior of stochastic processes. The range $0 < \mathrm{H} < 0.5$ corresponds to an anti-persistent series, where growth in the previous period is likely to be followed by a decrease in the next period. The range $0.5 < \mathrm{H} < 1$ corresponds to a persistent series, where growth in the previous period is likely to continue as a trend in the future. A Hurst coefficient close to 0.5 indicates a greater number of random detectors in the series, with a less pronounced growth trend.

In addition to serving as a classification criterion, the Hurst index can directly record the degradation of an object before destructive processes occur. For example, in the temperature monitoring system of Vardzia, the level of temperature deviation due to environmental changes shows a clear growth trend, potentially influencing the humidity regime inside the boiler. Despite the lack of observational data, it is clear that these changes could have cumulative effects, such as rock deformation or structural changes. The transition from the localization zone of plastic deformation to the rupture zone is characterized by a sharp increase in the Hurst index.

8. Conclusions

The fault deformation and structural change diagnosis algorithm is efficient even with limited computing resources, enabling the creation of a unified core for measurement data collection and processing in the monitoring system. The adequacy of the procedure increases with the number of analytical quantities.

References:

1) Butakov, V., & Grakovsky, A. (2005). “Estimation of the level of stochasticity of time series of an arbitrary origin using the Hurst exponent. Computer Modeling and New Technologies, 27-32.

2) Esling, P., & Agon, C. Time series data mining. *ACM Comput. Surv. 2012, Vol.45,* (no.1), p.1232-1247.[

3) Falconer, K. (2013). Fractals: A Very Short Introduction. Published by Oxford University Press, 138-152 page.

4) Korus, Ł., & Piórek, M. (2015). "Compound method of time series classification",. *Nonlinear Analysis: Modelling and Control, vol. 20,*(no.4), pp. 545-560, 2015.

5) Azmaiparashvili, Z., & Otkhozoria, N. M. (2016). Identification of Two Sorts of Processes and Determining of Their Differences Criteria. Journal of Technical Science and Technologies,. https://doi.org/10.31578/jtst.v5i2.106

6) Dmowska, R., & Saltzman, B. (1999). *Long-Range Persistence in Geophysical Time Series,.* Academic Press, p.175. doi:https://doi.org/10.1016/s0065-2687(08)x6022-4

7) Esling, P., & Agon, C. Time series data mining. *ACM Comput. Surv. 2012, Vol.45,* (no.1), p.1232-1247.[Google Scholar] [CrossRef].

8) Feder, J. (1988). *“Fractals".* Plenum Press,: New York,. doi:https://doi.org/10.1007/978-1-4899-2124-6

9) Ivanisenko, I., Kirichenko, L., & Radivilova, T. "Investigation of multifractal properties of additive data stream,". *2016 IEEE First International Conference on Data Stream Mining & Processing (DSMP), Lviv, 2016*, pp. 305-308. doi:10.1109/DSMP.2016.7583564

10) Coelho, A., & Lima, C. "Assessing fractal dimension methods as feature extractors for EMG signal classification,". *Engineering Applications of Artificial Intelligence, Vol.36,* pp.81-98, 2014.

11) Chkheidze, I., Otkhozoria, N., & Narchemashvili, M. (2021). EVALUATION OF MEASUREMENT QUALITY USING THE MONTE-CARLO METHOD. Universum, 65-70. DOI: 10.32743/UniTech.2021.84.3-4.65-70

12) Otkhozoria, N., Azmaiparashvili, Z., Petriashvili, L., Otkhozoria, V., & Akhlouri, E. (2023). Labview In The Research Of Fractal Properties Of The Topology Of Networks And Stochastic Processes. World Science. doi:https://doi.org/10.31435/rsglobal_ws/30092023/8020

13) Otkhozoria, N., Otkhozoria, V., Narchemashvili, M. (2021) Fractality Of Measurements Of Quantities And Real Processes. International Trends In Science And Technology, Engineering Sciences June 2021 Doi: https://doi.org/10.31435/rsglobal_conf/30062021/7620

14) Azmaiparashvili, Z., Otkhozoria, N. (2021)Mathematical Model for Studying the Accuracy Characteristics of Devices for Measuring the Resonant Frequency of Oscillatory Systems New Approaches in Engineering Research” 121-131 https://doi.org/10.9734/bpi/naer/v5/10242D

15) Chkheidze, I., Kharatishvili, L., & Karoyan, I. (2008). ,,The modern concept of uncertainty of the result of measurement as a form of development and generalization of the concept of error". *Works 11, Appendix to the Journal of the Academy of Educational Sciences of Georgia "Moambe," Tbilisi, 2008*, p. 287-291.