International Refereed Journal of Engineering and Science (IRJES)
ISSN (Online) 2319-183X, (Print) 2319-1821
Volume 14, Issue 6 (Nov. – Dec. 2025), PP. 180-183

The integrated approach to identifying anomalies and trends in climate data

Fedosov Vladimir Andreevich
** (Student, Department of Industrial Safety and Industrial Ecology, FSBEO HE Ufa University of Science and Technology, Russia, Ufa)*

Elizariev Alexey Nikolaevich
*** (Associate Professor, Candidate of Geographical Sciences, Department of Industrial Safety and Industrial Ecology, FSBEO HE Ufa University of Science and Technology, Russia, Ufa)*

Nutfullina Elizaveta Ildarovna
**** (Graduate student, Department of Industrial Safety and Industrial Ecology, FSBEO HE Ufa University of Science and Technology, Russia, Ufa)*

Abstract: *The article presents an analysis of existing methods for assessing climate risks, which include quantitative and qualitative approaches to forecasting the impact of climate factors. The effectiveness of methods such as linear trend, Student's and Fisher's criteria, harmonic analysis, amplitude smoothing, random forest method, time series clustering, and network models is assessed according to general criteria. Particular attention is paid to the possibility of using these methods as an algorithmic basis for solving decision-making theory problems in conditions of uncertainty caused by climate change. Approaches to the integration of climate data and the results of their analysis are considered. The study demonstrates that climate risk assessment methods can serve as a source of structured input data and analytical models for decision support systems, enabling more informed choices of adaptation strategies. The results of the work can be used to develop and expand the functionality of software packages designed to study and apply decision-making methods in the context of climate instability.*

Keywords: *climate risks, method evaluation, cluster analysis, anomalies, time series, network models.*

Date of Submission: 11-12-2025 Date of acceptance: 22-12-2025

I. INTRODUCTION

In recent decades, interest in global climate issues has grown significantly due to increasing uncertainty and intensifying climate risks. Contemporary research is increasingly focused not only on identifying the facts of climate change, but also on developing methods for assessing its consequences and tools for formulating sound adaptation strategies. According to the analysis of the World Economic Forum, 'The BRICS Climate Agenda in the Current Context' (Moscow, 29–30 August 2024) [1], climate risks remain one of the key global threats for the next ten years, which is confirmed by the results of the forum. The forum emphasised the need to integrate climate data, improve methodological approaches to its analysis, and expand the algorithmic foundations for decision support systems. Particular attention was paid to the development of tools that allow the results of climate research — including trend assessment, statistical criteria, machine learning and clustering methods — to be used to create more sustainable development models, improve energy efficiency and transition to a low-carbon economy.

Climate change poses not only an environmental threat, but also a serious economic one. Yield losses in key agricultural regions could reach 10–25% by 2030, increasing financial risks and threatening food security in the agricultural sector [2].

Making informed choices about adaptation measures requires a deep understanding of the dynamics of climate change and methods for analysing it. To develop effective solutions, it is necessary to study scientific approaches that allow for the rapid assessment of climate risks and the prediction of their consequences.

II. MATERIALS AND METHODS

There are various approaches to calculating climate risks that help to study how climate characteristics change over time. One of the most common methods is the linear trend [3], which allows us to track how steadily the temperature or amount of precipitation changes over many years. If, over decades, the average temperature in a particular region has been steadily increasing [4], this method will show the presence of such a trend. At the same time, it is possible to statistically assess how significant this increase is, which helps to understand whether it is a random change or a certain pattern.

Statistical methods using criteria such as Student's and Fisher's [5] are used to analyse sudden changes. They allow us to check whether there have been any sharp jumps in climate data for a given period caused by a sudden increase in precipitation or temperature anomalies. In this case, the time series of data is divided into several periods and checked to see if they behave the same in each of these segments. The method helps to assess whether there have been sudden climate changes, such as a sudden increase in average winter temperatures over a decade compared to the previous year [6].

Harmonic analysis [7] is used to understand cyclical climate change: if a region experiences periods of drought or heavy rainfall every 5–7 years, this method allows us to identify these cycles and understand their structure (for example, how recurring climatic phenomena affect agriculture or water resource management).

In cases where it is necessary to analyse not only long-term trends but also short-term changes, the ‘cycle amplitude trimming’ method [8] is used. This approach allows various fluctuations (short-term and long-term) to be sequentially excluded from the time series. In this case, it is worth first excluding interannual changes in order to highlight ten-year or century-long trends (similar to filtering noise in data, where only the most significant fluctuations remain).

Cycle amplitude smoothing [9] is also used, which works on a similar principle but focuses on smoothing the amplitudes of fluctuations rather than subtracting them, which can be useful for analysing smoother climate changes, such as gradual warming or cooling over decades. Thus, by considering a series of temperatures, it is possible to exclude short-term fluctuations (sudden cooling) and identify the overall warming trend over the past 50 years.

The random forest method [10,11] is a machine learning tool based on the construction of multiple decision trees. The method is used for forecasting based on large amounts of historical data, which makes it effective in situations where it is necessary to analyse and predict climate change. Random forest effectively resolves both linear and nonlinear dependencies, allowing complex relationships between climate parameters such as temperature and precipitation to be identified. Scikit-learn libraries are mainly used to work with it, providing ready-made tools for training models.

In addition to machine learning methods, an important approach for grouping data with similar characteristics is time series clustering [12,13]. This approach, used in climatology, allows regions or time periods with similar climatic conditions to be linked in order to identify patterns and anomalies. Clustering methods, such as k-means or DBSCAN, are widely used in time series analysis. With their help, it is possible to group different regions with similar weather conditions or separate periods of sudden weather changes. For these purposes, the Scikit-learn library, which supports various clustering methods, is also mainly used.

In addition, graph-based models [14] are used to model complex relationships between different climate indicators, allowing the creation of graphs in which different climate parameters (e.g., temperature, precipitation, wind speed) are treated as network nodes, and their interactions are treated as connections between nodes. Such models help to better understand the links between global climate change and local weather phenomena, and the NetworkX library, specifically designed for working with graph structures in Python, is used to create and analyse graphs.

Most methods work effectively with long-term data (Table), suggesting that methods such as linear trend, harmonic analysis, and random forest make it possible to identify stable climate changes and trends using large amounts of data. However, Student's and Fisher's criteria analyse stationarity and are more limited in their use, as they analyse data in specific subsamples, which may limit their use in longitudinal studies.

Table An assessment of the effectiveness of methods based on general criteria

Method	Cycles and seasonality	Anomaly detection	Nonlinear dependencies	Long-term trends	Flexibility of application
Linear trend	1	1	1	5	4
Study/Fish.	1	5	1	4	2
Harmonic analysis	5	1	2	5	4
Smoothing	5	2	2	4	4
Random forest	1	5	4	5	5
Clustering	5	5	2	5	5
Network models	4	5	5	5	5

For a preliminary comparison of the methods used to climate data analysis, a binary and categorical evaluation system was chosen, as it provides a structured and unambiguous division of methods according to key functional characteristics (table). This approach is widely used in applied climatology, machine learning, and decision support systems when it is necessary to quickly determine the scope of applicability of each algorithm in the presence of a large number of criteria [15]. The scoring was carried out by experts in the field, taking into account both the theoretical rigour of the mathematical methods and the requirements for their practical application to solve environmental monitoring and natural and man-made risk assessment problems.

The evaluation of methods based on a set of key criteria (long-term trends, anomalies, seasonality, non-linearity and flexibility) shows that each approach has its place in the structure of climate analysis, and the distribution of scores reflects the strengths and weaknesses of each.

The highest scores (5 points) were given to the linear trend method, random forest, time series clustering, and network models, which highlights their ability to consistently reproduce long-term patterns. The Student's and Fisher's criteria were rated slightly lower (4 points) because they confirm the statistical significance of trends but do not provide a complete model of change over time. Harmonic analysis and smoothing methods (4 points) also showed high results, but they are sensitive to the structure of the time series and require more stringent conditions for applicability.

The highest scores (5 points) were given to Student's and Fisher's criteria, random forest, clustering, and network models. This is because these methods either directly test the statistical significance of deviations (Student/Fisher) or use algorithms that can detect rare events by analysing the data structure.

The smoothing method (2 points) only indirectly indicates anomalies through sharp deviations from the smoothed curve. Harmonic analysis and linear trend, which received 1 point, work worst with extreme events, as they smooth out variations and focus on average system behaviour patterns.

Harmonic analysis and amplitude smoothing received the highest scores (5 points), which is logical: these approaches are specifically designed to search for cyclicality and regular fluctuations. Time series clustering (5 points) also shows high efficiency, as different seasonal patterns can form separate clusters.

Network models (4 points) are capable of identifying seasonal patterns through changes in the connectivity between variables. At the same time, linear trend, Student's/Fisher's criteria, and random forest show low results: seasonality is a weak or ‘blurred’ characteristic for them.

Network models have the maximum score of 5 points, reflecting their ability to identify complex and changing relationships between climate variables.

Random forest (4 points) also works well with nonlinearities, allowing for a large number of factors and interactions to be taken into account.

Harmonic analysis and amplitude smoothing (2 points) partially reflect nonlinear processes but remain within the confines of limited models.

Linear trends and statistical criteria are characterised by the lowest score (1 point) as their capabilities are limited by the assumption of linearity or stationarity.

Random forest, clustering, and network models (5 points each) offer the greatest flexibility. These methods work well with both small samples and large climate archives, scale across space and time, and adapt to different types of data.

Linear trend, harmonic analysis, and smoothing (4 points each) have medium flexibility—their application is limited by model assumptions and the requirement for a specific series structure.

Student's and Fisher's criteria (2 points) demonstrate low flexibility, as they are oriented towards strict statistical conditions, require preliminary data processing, and are applicable mainly in stationary conditions.

A comparative analysis of climate risk assessment methods, including classical statistical approaches and modern machine learning algorithms, revealed that a stable and informative description of long-term climate processes can be achieved through the combined use of different types of methods. It became clear that clusters, anomalous modes, and hidden patterns are most clearly manifested when the results of statistical analysis are supplemented with clustering and outlier detection algorithms. To integrate these approaches into a single system and ensure comprehensive analysis of long-term series of climate parameters. A software package called ‘Software package for cluster analysis and anomaly detection in long-term environmental data’ [16] has been developed. Its appearance was the logical conclusion of the research work: comparison of methods allowed us to determine the optimal algorithms, the identified limitations of existing approaches demonstrated the feasibility of their integration into a single application, and analysis of the requirements for practical interpretation of climate risks demonstrated the need to develop a tool capable of supporting decision-making processes in conditions of climate uncertainty.

III. CONCLUSIONS

The analysis of climate risk assessment methods showed that no single method provides comprehensive coverage of all aspects of climate variability, but their combination allows for the effective identification of long-term trends, anomalies, seasonal cycles, and nonlinear dependencies. Classic statistical approaches, such as linear trends and Student's and Fisher's criteria, are well suited for confirming the significance of changes and detecting extreme events, while harmonic analysis and amplitude smoothing methods are effective for identifying cycles and seasonal patterns, and modern machine learning algorithms – random forests, time series clustering, and network models – demonstrate high efficiency when working with complex and nonlinear relationships between climate variables.

Based on the results of a comparative analysis of methods, 'Software package for cluster analysis and detection of anomalies in long-term environmental data' was developed, which combines the best algorithms of statistical and machine analysis, provides automated processing of long-term climate series, visualisation of results and support for decision-making in conditions of climate uncertainty. Thus, the integration of climate risk assessment methods into the software environment allows for the creation of structured analytical models, the identification of anomalous patterns and long-term trends, and the establishment of a reliable basis for the development and expansion of the functionality of decision support systems in conditions of climate instability.

REFERENCES

[1]. The BRICS Climate Agenda in Modern Conditions Forum [Electronic resource]. Access mode: https://brics-russia2024.ru/events/forumy-konferentsii/forum-klimaticheskaya-povestka-briks-v-sovremennykh-usloviyakh/

[2]. Agriculture, forestry and fisheries, value added (% of GDP) [electronic resource]. Access mode: https://data.worldbank.org/indicator/NV.AGR.TOTL.ZS?name_desc=false

[3]. Kiktev D. B., Kryzhov V. N. On the comparison of different methods of linear trends // Meteorology and hydrology. 2004. №11. P. 27-38.

[4]. Vasiliev D. Yu., Semenov V. A., Chibilev A. A. Climatic changes in the temperature regime in Russia in the early XX – early XXI century // Geography and natural resources. 2023. №2. P. 15-23.

[5]. Lobanov V. A., Akbarov Kh., Nzambimana E., Spasenkov Ya. M. Application of statistical methods to assess the uniformity and increase the amount of long-term hydrometeorological information // Proceedings of the International Scientific and Practical Conference. 2024. P. 331-348.

[6]. Vasiliev D. Y. Features of climate change in the Southern Urals: causes and consequences. 2023. P. 285.

[7]. Ushkov A. L., Meltsaykin E. A. Analysis of biharmonic and harmonic models by the methods of iterative extensions // Mathematical modeling. 2022. №85. P. 51-66.

[8]. Teplukhin A.V. Short-range potential functions in computer modeling of water and aqueous solutions // Journal of Applied Chemistry. 2016. №8. P. 1723-1752.

[9]. Lobanov V. A., Mami M. T., Grigorieva A. A. Statistical methods of decomposition of series of hydrometeorological characteristics // Innovative methods of mathematics and physics in ecological and hydrometeorological research. 2023. P. 274-288.

[10]. Kartiev S. B., Kureichik V. M. Classification algorithm based on the principles of a random forest for solving the forecasting problem // Software products and systems. 2016. P. 1-5.

[11]. Zarova E. V., Dubravskaya E. I. The "Random Forest" method in the study of the influence of macroeconomic indicators of regional development on the level of informal employment // Questions of statistics. 2020. №6. P. 37-55.

[12]. Melnikova Yu. V. Segmentation of multidimensional non-stationary time series using clusterization methods // Agricultural science in the XXI century: problems and prospects. 2015. P. 471-474.

[13]. Dulya I. S. Application of deep learning methods to the task of clustering time series // Alley of Science. 2021. №5(56). P. 974-978.

[14]. Efanova N. V., Slesarenko I. V. Competence model as a tool for network economic system HR sustainability assessment // Modern economic: problems and solutions. 2021. №11(143). P. 28-37.

[15]. Rybak V. A., Shokr Ahmad. An analytical review and comparison of existing decision support technologies // System analysis and applied Informatics. 2016. №3

[16]. Fedosov V.A., Elizariev A.N., Tarakanov D.A., Vaganova D.E. Software package for cluster analysis and detection of anomalies in long-term environmental data: certificate of state registration of the computer program №RU 2025685589, Russia, 2025.