UDC 338.48:519.2

Akhmedova Olena
akhmedovayelena@gmail.com, ORCID ID: 0000-0003-1573-7710
Doctor of Sciences in Public Administration, Associate Professor at the Department of Tourism and Hotel and Restaurant Business,
Kyiv National University of Technologies and Design, Kyiv

Svoievolin Ihor
svoievolin.ihor@hneu.net, ORCID ID: 0009-0004-4865-385X
Postgraduate Student, Simon Kuznets Kharkiv National University of Economics, Kharkiv

EVALUATING INBOUND TOURISM TRENDS THROUGH LINEAR REGRESSION ANALYSIS

Abstract. *The dynamics of tourist flows and forecasts changes in demand using linear regression analysis have been considered in the paper. The period from 2011 to 2022 was taken as a basis. The analysis covers 18 countries that were selected both in terms of traditional popularity among tourists and in terms of growth in interest in lesser-known destinations. The R software environment and the tidyverse, dplyr, ggplot2 and gt libraries were used to process the data, which ensured efficient information processing, graphing and classification of countries by trends. The data was downloaded in CSV format, pre-processed and checked for missing values. Only records with non-zero values for the number of tourists were taken into account for the analysis. Additional variables were also created to classify countries by trends. The linear regression method allows determining the direction and intensity of changes in the number of tourists in each country. The slope coefficient shows the growth, stability or decline in tourist flows. Based on the quartile distribution, categories A, B, C and D were determined, reflecting the level of tourism development and the rate of its growth or decline. The results of the analysis show that due to the significant growth of various types of tourism in rapidly developing countries, most European countries are experiencing a decline in tourist flows, with the greatest losses observed in countries with high dependence on international tourism or geopolitical problems. At the same time, some countries, such as Ireland and Greece, are demonstrating positive trends, indicating successful practices in attracting tourists and opportunities for further research. The results obtained make it possible to identify promising regions for the development of tourism infrastructure and investment, optimise the use of natural resources and reduce the negative impact of tourism on the environment. The article demonstrates that the use of linear regression methods is an effective tool for forecasting demand, planning tourism and supporting the sustainable development of tourist regions, which is important for government agencies, businesses and tourism service operators.*

Keywords: inbound tourism, tourism industry, tourist region, linear regression analysis, sustainable tourism, tourist flows, data analysis.

Ахмедова О. О.
akhmedovayelena@gmail.com, ORCID ID: 0000-0003-1573-7710
д.н.держ.упр., доцент кафедри туризму та готельно-ресторанного бізнесу,
Київський національний університет технологій та дизайну, Київ

Своєволін І. О.
svoievolin.ihor@hneu.net, ORCID ID: 0009-0004-4865-385X
аспірант, Харківський національний економічний університет імені Семена Кузнеця, Харків

ОЦІНКА ТЕНДЕНЦІЙ В'ЇЗНОГО ТУРИЗМУ ЗА ДОПОМОГОЮ ЛІНІЙНОГО РЕГРЕСІЙНОГО АНАЛІЗУ

Анотація. *У статті розглянуто динаміку туристичних потоків та прогнозування змін попиту з використанням лінійного регресійного аналізу. За основу було взято період з 2011 по 2022 рік. Аналіз охоплює 18 країн, які були обрані як з точки зору традиційної популярності серед туристів, так і за показниками зростання інтересу до менш відомих напрямків. Для обробки даних застосовано програмне середовище R та бібліотеки tidyverse, dplyr, ggplot2 і gt, що забезпечило ефективне опрацювання інформації, побудову графіків і класифікацію країн за трендами. Дані були завантажені у форматі CSV, попередньо оброблені і перевірені на наявність пропущених значень. Для аналізу враховувались лише записи з ненульовими значеннями кількості туристів. Також було створено додаткові змінні для класифікації країн за трендами. Метод лінійної регресії дозволяє визначити напрямок і інтенсивність змін у кількості туристів у кожній країні. Величина коефіцієнта нахилу (Slope) показує приріст, стабільність або зменшення туристичних потоків. На основі квартильного розподілу були визначені категорії A, B, C і D, що відображають рівень розвитку туризму та темпи його зростання або зниження. Результати аналізу свідчать, що завдяки значному зростанню різних видів туризму у країнах, що швидко розвиваються, більшість європейських країн демонструє зниження туристичних потоків, причому найбільші втрати спостерігаються в країнах із високою залежністю від міжнародного туризму або геополітичними проблемами. Водночас деякі країни, наприклад Ірландія та Греція, демонструють позитивні тенденції, що вказує на успішні практики залучення туристів та можливості для подальших досліджень. Отримані результати дозволяють ідентифікувати перспективні регіони для розвитку туристичної інфраструктури та інвестицій, оптимізувати використання природних ресурсів та зменшити негативний вплив туризму на навколишнє середовище. Стаття демонструє, що застосування методів лінійної регресії є ефективним інструментом для прогнозування попиту, планування туризму та підтримки сталого розвитку туристичних регіонів, що має значення для державних органів, бізнесу та операторів туристичних послуг.*

Ключові слова: в'їзний туризм, туристична галузь, туристичний регіон, лінійний регресійний аналіз, стійкий туризм, туристичні потоки, аналіз даних.

JEL Classification: L83, L14

DOI: https://doi.org/10.32782/tourismhospcee-13-1

Statement of the problem. Over the past decade, the tourism industry in Europe has undergone significant transformations, with a steady increase in international flights and a shift towards more sustainable and nature-based tourism.

Europe has long been a dominant player in the global tourism industry, receiving slightly less than half of the world's international arrivals. As the most visited tourist region, the European Union has played a decisive role in driving and shaping these trends.

However, as a result of the tourism growth in many rapidly growing economies, Europe has gradually lost its share of the global inbound tourism market. Although the tourism industry in Europe has grown steadily by 3–4.5% over the last ten years, this trend is expected to continue in the coming decades. The Mediterranean region, in particular, expects tourism to double over the next 30 years [1].

To better understand changes in inbound tourism to Europe, especially the one connected to sustainable green tourism, the linear regression analysis has been used as a statistical tool. Linear regression is a powerful method that allows analysing the relationship between independent variables, such as economic indicators, and a dependent variable, which in this case is the number of inbound tourists. By applying linear regression, researchers can identify key factors that significantly influence the growth or decline of tourism in the European region and gain valuable insights into the main drivers of these trends. The analysis presented covers the period from 2011 to 2022, providing a complete understanding of inbound tourism trends during this period.

Analysis of recent research and publications. World Tourism Organization (UNWTO) presents data on the tourism inbound flows on the regular basis [2; 3; 4]. It provides governments, scientists and professionals in tourism sphere with deep statistical information on non-resident arrivals, overnight states, accommodation but lacks comparative

ISSN: 2786-4812 (Print), 2786-4820 (Online)

analysis among countries and regions especially in the context of the specific tourism time.

With the help of linear regression analysis scientists make an effort to forecast tourists' arrivals to a certain territory (Johny A., 2022) [5], to understand tourism market trends ensuring stability of the research results (Wei Min, 2014) [6], to predict competition trends in various destinations and to learn the influence of tourism on sustainability and prosperity of the territory (Trišić I. et al, 2023 Gelbman A., 2021). [7; 8].

Identification of previously unresolved aspects of the problem. The sustainability in tourism is of great importance nowadays. As problems in tourism management arise, especially in the major urban destinations, the governments and businesses start searching for innovative and creative solutions that are impossible without thorough mathematical calculations and understanding of trends and careful assessment of flows and consumer behaviour.

Thus, the paper presents the results of a study conducted using the R programming environment. The potential of this method is demonstrated not only in traditional tourism, but also for forecasting demand for eco-tourism, that in today's world, where sustainability and environmental conservation are becoming priorities, the ability to predict changes in tourist interest in eco-routes, environmentally friendly transport and sustainable hotels is extremely important.

Setting the task. The objective of the study is to analyse the tendencies of inbound tourism using linear regression analysis in order to forecast its dynamics in the future. To achieve the stated aim, the study proposes to set and address the following tasks:

– to structure the data on 18 countries;

– to apply the linear regression method to estimate the trends for each zone using slope coefficients (growth or decline);

– to conduct clustering using the quantile distribution method;

– to classify zones according to definitions A, B, C, D depending on the trend.

Summary of the main research material. The results of linear regression analysis give the understanding in which countries there is a downward or an upward trend, which will allow to analyse the reasons for growth or decline [5].

To gain a complete understanding of this phenomenon, data on inbound tourism for 18 countries was selected based on their popularity as tourist destinations, as well as their relative obscurity in the global tourism landscape [2; 3; 4].

While some of the selected countries, such as France, Germany, Greece and Italy, are known for their thriving tourism industry, others, such as Ireland, Finland, Cyprus and Ukraine (prior to the onset of the full-scale war initiated by the Russian Federation), have gradually gained popularity as attractive tourist destinations in recent years [1].

The analysis also highlights the challenges and opportunities facing the European tourism sector and points the way forward for further research.

The R programming language was used to analyse inbound tourism trends, as it provides a wide range of tools for data processing and statistical analysis [9]. The analysis included the following main stages:

– the data was downloaded in CSV format, preprocessed and checked for missing values;

– only records with non-zero values for the number of tourists were included in the analysis;

– additional variables were created to classify countries by trends.

Modelling using linear regression:

$$Y = aX + b,$$

where Y is the number of tourists and X is the year. The slope coefficient (a) determines the direction of the trend (growth or decline).

For each country separately, an assessment of trends in the number of tourists was performed using the linear regression method.

To visualise the trends, graphs and tables were created to reflect changes in tourist flows. *R* libraries such as *tidyverse, dplyr, ggplot2,* and *gt* were used to ensure efficient data processing and the creation of clear graphs.

Slope calculation:

```
trend_analysis <- data %>%
group_by (Country, Zone) %>%
summarise(
Slope = if (n() > 1) coef(lm(Tourists ~ Year))[2]
else NA_real_,
.groups = "drop"
)
```

Classification of countries by trends:

```
trend_analysis <- trend_analysis %>%
mutate(Category = case_when(
Slope >= quantile (Slope, 0.75, na.rm = TRUE) ~ 'A',
Slope >= quantile (Slope, 0.50, na.rm = TRUE) ~ 'B',
Slope >= quantile (Slope, 0.25, na.rm = TRUE) ~ 'C',
TRUE ~ 'D'
))
```

Slope is the main indicator that measures the dynamics of changes in the number of tourists over time. This indicator is the result of linear regression and determines the rate (gradient) of growth or decline in tourist flows.

$$Slope = \Delta Tourists / \Delta Year$$

Slope values enable the assessment of both the direction and the magnitude of the trend:

High positive value: significant growth.

High negative value: significant decline.

Value close to zero: stability.

The category determines the relative position of a country in the overall distribution of trends. The use of quartiles makes it easy to compare countries with different dynamics:

• A (Top 25%) – Leaders with high tourist growth.

• B (50–75%) – Countries with moderate growth.

• C (25–50%) – Countries with stability or a slight decline.

• D (Bottom 25%) – Countries with the largest losses in tourist flow.

This approach is particularly important for visualisation, as it allows making quick identification of "hot" and "problem" points.

After completing the tasks and successfully running the code, the results were analysed.

Fig. 1. presents the following: X-axis – Mean Slope is an economic indicator of the dynamics of changes in inbound tourism for each tourist area. Negative values indicate a decrease in the number of tourists, values closer to 0 indicate stability, and positive values, which are absent in this case, would indicate growth. Y-axis – tourist areas (Europe, Americas, Africa, etc.).

All trends presented above are negative, indicating a decline in the popularity of regions as destinations for inbound tourism, with the European tourist area ranking first in the overall standings among our 18 European countries. Based on this analysis, it is impossible to say for sure why tourist flows in Europe are falling, but we can accurately record it and set ourselves new directions for assumptions and research.

The next stage of the analysis is to study the trend for each country separately. Fig. 2. demonstrates a depiction of the mean slope for inbound tourism by country. The values reflect the dynamics of tourist numbers over time, where the Y-axis represents countries (grouped by tourism dynamics) and the X-axis represents the mean slope. Negative values (red) indicate a decrease in the number of tourists (negative trend), while positive values (blue) indicate an increase in the number of tourists (positive trend).

The graph clearly shows that the largest decline is observed in countries with high dependence on international tourism or geopolitical problems, and conversely, countries with low dependence on international tourists demonstrate stability. While Fig. 1. presents combined data for all countries within each

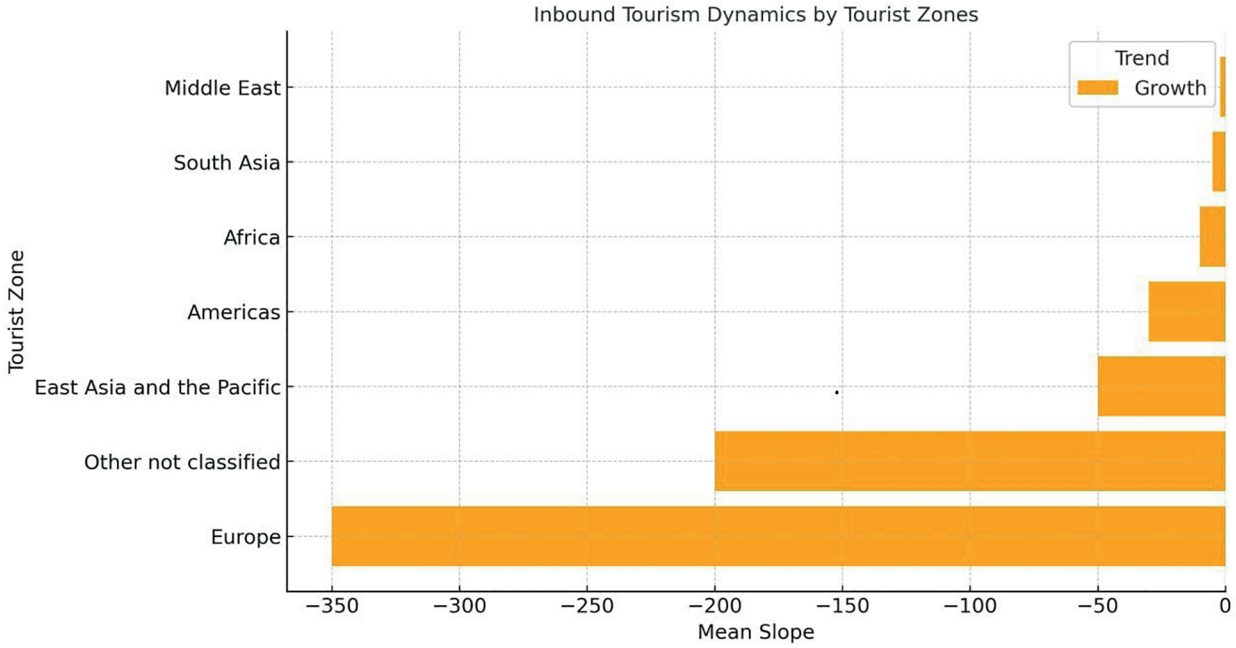

Figure 1. Trend in tourist areas

Source: compiled by the authors based on statistical data [2–4]

ISSN: 2786-4812 (Print), 2786-4820 (Online)

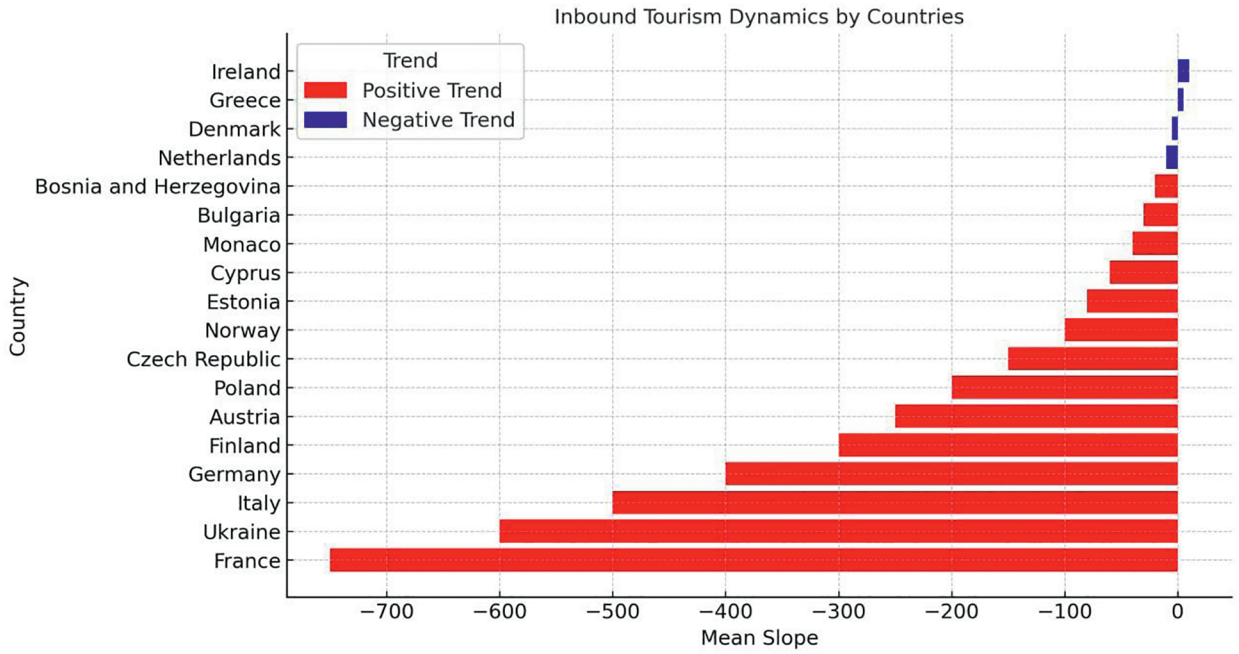

Figure 2. 18 European countries' trends

Source: compiled by the authors based on statistical data [2–4]

tourist zone, and the negative trend in the European zone has offset any growth observed in individual countries, Fig. 3, gives understanding of each country separately and clearly provides their individual trends. For example, Ireland and Greece have positive trends that were "hidden" in the negative trend of the entire European zone. Such research provides a direction for further assumptions and research, i.e., how countries have maintained positive growth.

Next, the classification of zones according to definitions A, B, C, D, depending on trends were considered [6]. Fig. 3 demonstrates the classification of tourist areas (categories A, B, C, D) for each country. The colours reflect the dynamics of tourist zone popularity based on the mean slope for inbound tourism:

A (green) – areas with the highest growth in tourist numbers;

B (blue) – areas with stable or moderate growth;

C (orange) – areas with a slight decline or stability;

D (red) – areas with the largest decline in tourist numbers.

With this classification, it is clear that the African zone has a decline in tourist flows for only three countries but overall maintains an insignificant decline. The American zone is generally the same, with a significant decline visible in three countries – Italy, Monaco and Ukraine. East Asia and the Pacific zone will receive a predominantly D (red) rating, indicating a decrease in tourist flows from this region. The European zone also has a predominantly D (red) rating, indicating a significant decline in intra-European tourism. The Central Asia and South Asia zones exhibit a slight decline.

Conclusions and prospects for further research. The use of linear regression to analyse tourism trends demonstrates the potential of this method not only in traditional tourism, but also for forecasting demand for other types of tourism. In today's world, where sustainability and environmental conservation are becoming priorities, the ability to predict changes in tourist interest in eco-routes, environmentally friendly transport and sustainable hotels is extremely important.

The use of linear regression allows for the analysis of key factors influencing the growth or decline in demand for certain type of tourism, including changes in climatic conditions, socio-economic factors, and government policy in the field of sustainable development. Analysing data using this method allows to identify regions with the greatest potential for tourism development and to identify patterns that will contribute to the effective planning of tourism initiatives.

Classifying areas by their level of tourism development based on trends, similar to the analysis of inbound tourism, will help attract investors and direct resources to the most promising areas. Thanks to the forecasts obtained, it is possible to optimise the use of environmental resources, reduce the

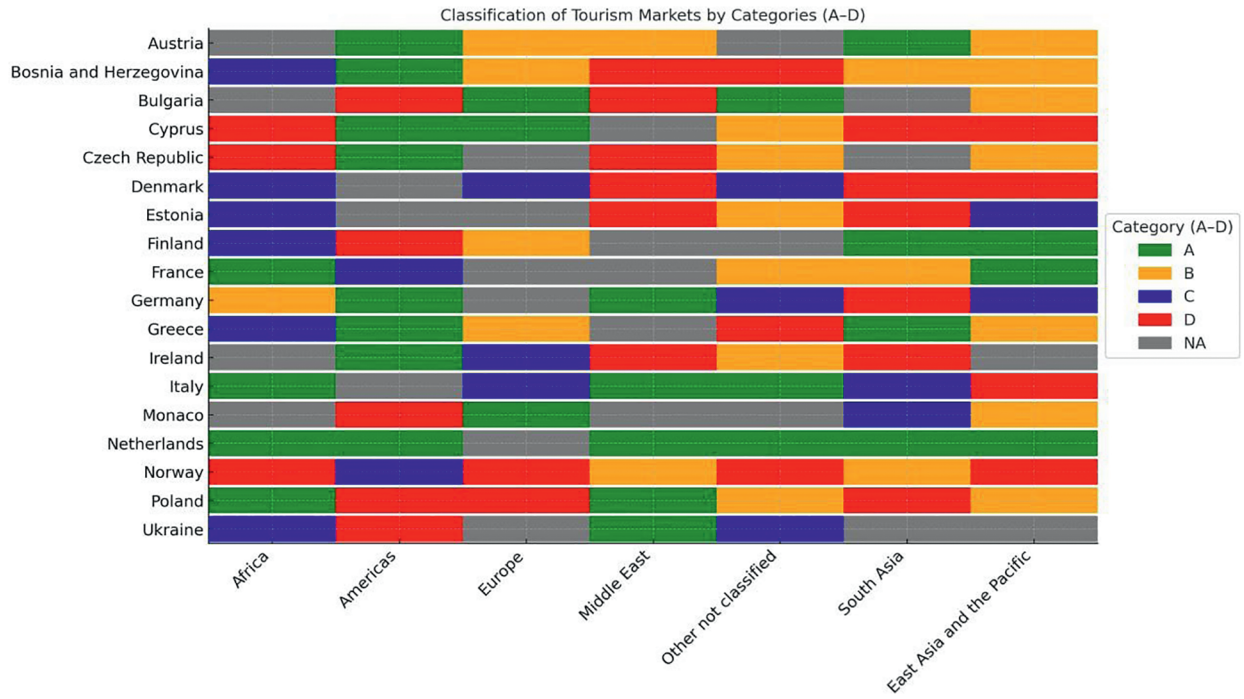

Figure 3. Categories of tourist zones for each country

Source: compiled by the authors based on statistical data [2–4]

negative impact of tourism on nature, and develop infrastructure that meets modern standards of sustainable development.

Thus, the methods used in this study can become an effective tool for tourism planning and management, helping government agencies, businesses, and tour operators adapt to new challenges and promote the harmonious development of tourist regions in accordance with the principles of sustainability.

REFERENCES

1. UN Tourism: official website, availble at: https://www.unwto.org

2. World Tourism Organization (2017), Yearbook of Tourism Statistics, Data 2011–2015, 2017 Edition, UNWTO, Madrid, DOI: https://doi.org/10.18111/9789284418428

3. World Tourism Organization (2022), Yearbook of Tourism Statistics, Data 2016–2020, 2022 Edition, UNWTO, Madrid, DOI: 10.18111/9789284423576

4. World Tourism Organization (2023) Yearbook of Tourism Statistics, Data 2017–2021, 2023 Edition, UNWTO, Madrid, DOI: https://doi.org/10.18111/9789284424139

5. Johny, A. (2022), A Forecasting model of tourists in tourist destinations using regression technology, *Journal of Research in Science and Engineering (JRSE),* Volume-4, Issue-1, DOI: https://doi.org/10.53469/jrse.2022.04(01).17

6. Wei Min (2014), A research on tourism market forecast and evaluation based on the application of linear regression model, *International Journal of u- and e- Service, Science and Technology,* Vol. 7, No. 6, pp. 213–226, DOI: http://dx.doi.org/10.14257/ijunesst.2014.7.6.19

7. Trišić I., Nechita F., Milojković D. & Štetić S. (2023), Sustainable tourism in protected areas – Application of the prism of sustainability model, *Sustainability,* no. 15(5148), DOI: https://doi.org/10.3390/su15065148

8. Gelbman A. (2021), Tourist Experience and Innovative Hospitality Management in Different Cities, *Sustainability,* no, 13(12), DOI: https://doi.org/10.3390/su13126578

9. Paradis E. (2006), Analysis of Phylogenetics and Evolution with R., Springer New York, NY, 211 p., DOI: https://doi.org/10.1007/978-0-387-35100-1

Стаття надійшла: 07.07.2025
Стаття прийнята: 22.07.2025
Стаття опублікована: 06.10.2025

ISSN: 2786-4812 (Print), 2786-4820 (Online)