# Development of a program to assess the correlation between environmental radiation according to diverse geographical regions and elevations

Chae Hyun Lee [a], Han Young Joo [a], Jeong Yeon Lee [a], Sang Yun Lee [a], Joo Hyun Moon [a*], Si Hyun Lee [b]
[a]*The Dankook Univ., Engineering Hall, 119, Dandae-ro, Dongnam-gu, Cheonan-si, Republic of Korea*
[b]*Gwangju Institute of Science and Technology, 123, Cheomdangwagi-ro, Buk-gu, Gwangju, Republic of Korea*
[*]*Corresponding author: jhmoon86@dankook.ac.kr*

***Keywords :** environmental radiation, linear regression analysis, correlation evaluation, radiological emergency*

## 1. Introduction

Government agencies and affiliated organizations collect and archive environmental radiation data nationwide. Leveraging this information to conduct correlation analysis between radiation levels in various locations—such as the atmosphere, air, soil, and ocean—can prove advantageous in promptly identifying causes during radiation emergencies, like nuclear accidents or leaks of radioactive materials. Such analysis assists in efficiently managing accidents, containing damage, and implementing preventative measures for future incidents.

The aim of this study is to swiftly examine the connection between environmental radiation data across different elevations and localities, aiming to pinpoint significant correlation patterns. These patterns will serve as a basis for establishing algorithms derived from the gathered results or for exploring practical application scenarios.

## 2. Methods and Results

This section summarizes the overview of the configuration and developmental directions of the program aimed at deriving correlations among environmental radiation data.

For the testing phase of the program, we utilized the 2021 environmental radiation data sourced from the open API provided by the Korea Atomic Energy Research Institute [1].

### *2.1 Program Configuration Overview*

The program utilizes the Python programming language to conduct correlation analysis via regression.

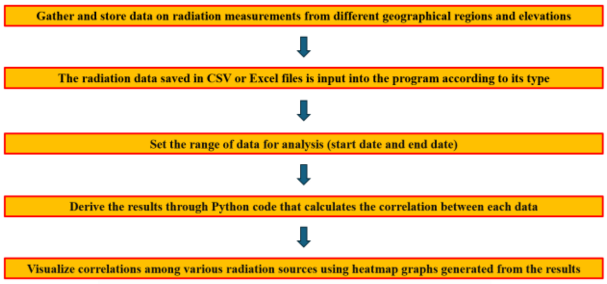

Figure 1. Analysis Program Operation Flow Chart

Initially, the environmental radiation data slated for analysis undergoes conversion into CSV or Excel files. Subsequently, users utilize the user interface depicted in Figure 2 to select the desired type of environmental radiation data (A, B, C) and specify the analysis range (start date, end date). The program then reads the CSV or Excel files in Data Frame format, ensuring the presence of all necessary items while eliminating unnecessary elements. This process involves the removal of duplicate data and handling of missing values to enhance the suitability and integrity of the data analysis outcomes. Additionally, it performs checks on function performance and potential failures to ensure the smooth execution of the analysis process.

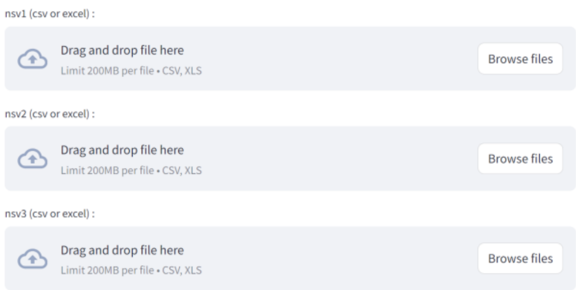

Figure 2. Radiation dose data input UI

The program conducts regression analysis ((A, B), (A, C), (B, C)) between the data, displaying the top five results upon completion, as shown in Figure 3.

| | date | nSv_1 | nSv_2 | nSv_3 |
|---|---|---|---|---|
| 0 | 2021-01-01 00:00:00 | 0.0117 | 0.0106 | 0.0145 |
| 1 | 2021-01-01 01:00:00 | 0.0117 | 0.0106 | 0.0144 |
| 2 | 2021-01-01 02:00:00 | 0.0118 | 0.0107 | 0.0145 |
| 3 | 2021-01-01 03:00:00 | 0.0118 | 0.0107 | 0.0145 |
| 4 | 2021-01-01 04:00:00 | 0.0118 | 0.0108 | 0.0145 |

Figure 3. Top 5 data UI with high analytical values

Furthermore, for correlation visualization, the numerical results are transformed into a heatmap graph and depicted as demonstrated in Figure 4.

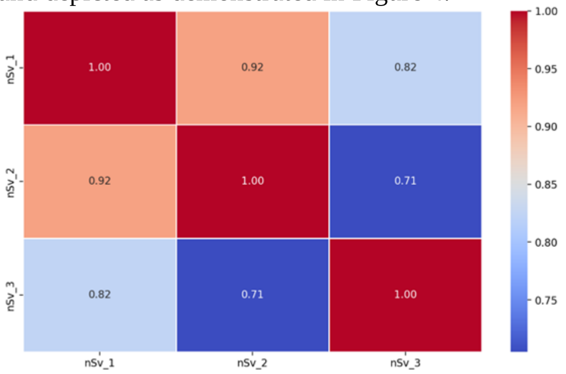


Figure 4. Data Analysis Results and Correlation Heatmap UI

After saving the results, users have the option to either begin a new analysis or exit the program. By leveraging "Streamlit," we converted the program into a webpage format, offering advantages in sharing, distributing, and modifying applications. Through its intuitive user interface, depicted in Figure 1, this method offers benefits in both data analysis and visualization.

```
if columns_exist:
    # 결과 확인 버튼
    if st.button("데이터 분석하기"):
        progress_text = "Operation in progress. Please wait."
        my_bar = st.progress(0, text=progress_text)

        for percent_complete in range(100):
            time.sleep(0.01)
            my_bar.progress(percent_complete + 1, text=progress_text)

        # 각 데이터프레임을 합치는 과정을 진행합니다.
        merged_df = pd.merge(df1, df2, on='date', how='inner')
        merged_df = pd.merge(merged_df, df3, on='date', how='inner')

        # 합친 데이터프레임 생성
        time.sleep(1)  # 가상의 지연

        # 결과를 출력합니다.
        st.subheader('상위 5개 데이터 확인')
        st.dataframe(merged_df.head(5))  # 상위 5개의 행을 출력

        st.subheader('데이터 범위:')
        st.write(f"시작 날짜: {merged_df['date'].min()}, 끝 날짜: {merged_df['date'].max()}")


        st.subheader('상관관계 시각화')
        last_index = merged_df.columns[0]
        corr_values = merged_df.corr()['nSv_1'].drop(['date'])
        st.write(corr_values)

        # 히트맵 그래프
        st.subheader('상관관계 히트맵')

        # Matplotlib figure 생성
        fig, ax = plt.subplots(figsize=(10, 6))
```

Figure 5. Python code: From data analysis to heatmap graph

generation

After the data is inserted, the Python code from the execution of the analysis to the processing of the derived value into a heatmap graph is as follows.

### *2.2 Direction of development*

The current program is in its early development stages and faces challenges in collecting and analyzing real-time data. Following the current protocol, the first goal is to incrementally increase the number of manual analyses to construct a correlation algorithm for environmental radiation doses using machine learning techniques. The ultimate objective is to establish an artificial intelligence (AI)-based system capable of receiving and storing large volumes of real-time radiation data, conducting autonomous analysis, and learning. Future plans involve leveraging existing public data, such as through open APIs, to validate analysis results and further develop the system.

## 3. Conclusions

Examining the correlation between environmental radiation doses and formulating algorithms holds significant potential in the realm of radiation emergency preparedness and disaster prevention. For instance, in the event of a radiation incident within a specific region, understanding the correlation between radiation levels in the affected area and those in neighboring regions can facilitate the optimization of evacuation routes. Areas with lower correlation may be prioritized for evacuation, while measures for disaster prevention in areas with high correlation can be intensified. Moreover, by establishing a correlation algorithm between environmental radiation and altitude in proximity to the site, comprehensive management of environmental radiation within the reference area can be achieved. With the acquisition of various data sources capable of conducting real-time analysis and learning from large datasets, along with the development of the program, it is anticipated that optimal radiation protection methods can be derived through predicting radiation levels based on correlation data. Additionally, a system capable of accurately and swiftly discerning and alerting of radiation emergencies can also be established.

## REFERENCES

[1] Current Status of Environmental Radiation Monitoring at the Korea Atomic Energy Research Institute (Jan.2021-Dec.2021).

[2] Lee. SiHyun, Lee. HongYeon, and Yeom. JungMin, Machine Learning Based Model Development and

Optimization for Predicting Radiation, RADIATION INDUSTRY, 2023.