E-ISSN 2586-9027
Homepage : https://tci-thaijo.org/index.php/SciTechAsia Science & Technology Asia
Vol.30 No.1 January - March 2025 Page: [95-116]
Original research article

Effect of Distance and Radionuclide Activity on Radiation Dose Exposure around the Chernobyl Nuclear Power Plant (ChNPP): Analysis with Poisson Regression Models and Correlation Tests

Nadrah Nadrah[1,*], Ruben Cornelius Siagian[2]

Primary School Teacher Education Study Program, Universitas Muhammadiyah, Makasar 90221, Indonesia
Department of Physics, Faculty of Mathematics and Natural Science, Universitas Negeri Medan, Medan 20221, Indonesia

Received 10 January 2024; Received in revised form 30 January 2025
Accepted 31 January 2025; Available online 24 March 2025

ABSTRACT

Researchers studied radiation exposure around the Chernobyl Nuclear Power Plant (ChNPP) after the accident, focusing on factors like distance from radiation sources and radionuclide activity. The analysis aimed to improve radiation exposure prediction models for mitigation and public health. Pearson correlation examined the relationship between exposure dose and distance, while Poisson regression identified significant factors, such as Zr-95 and Cs-134 activity. The model's validity was confirmed through overdispersion and variance inflation factor (VIF) tests, and residuals analysis assessed bias. Results showed radiation exposure decreased with distance, with a moderate negative correlation. Radionuclide activity significantly affected radiation levels, and a perfect correlation was found between exposure dose and absorbed dose. Poisson regression highlighted distance and radionuclide activity as key factors. This study fills a gap in understanding the complex relationship between distance, radionuclide activity, and radiation exposure around ChNPP. The novelty lies in applying Poisson regression with specific isotopes, providing a more accurate model to predict radiation exposure around nuclear facilities.

Keywords: Chernobyl nuclear power plant; Dose distribution; Radioactive isotopes; Radiation exposure; Spatial analysis

***Corresponding author:** nadrah@unismuh.ac.id **doi:** 10.14456/scitechasia.2025.7

1. Introduction

The accident at the Chernobyl Nuclear Power Plant (ChNPP) on April 26, 1986 became one of the most devastating nuclear disasters in human history [1, 2]. The explosion at Reactor No. 4 released large amounts of radioactive substances into the atmosphere, contaminating a wide area around the site and reaching various countries in Europe [3]. The radiation released resulted in long-term health impacts for the public, ranging from an increase in cancer cases to various genetic disorders. The environment around Chernobyl experienced very serious contamination, mainly due to the deposition of radionuclides such as cesium-137 (Cs-137), strontium-90 (Sr-90), and plutonium-239 (Pu-239), which have long half-lives and have the potential to provide radiation impacts over a very long period of time [4]. More than three decades after the incident, various studies have been conducted to understand the radiation dispersion patterns around Chernobyl, including the distribution of exposure dose levels and radionuclide activities in relation to the distance and direction from the explosion source. However, despite the many studies on the radiation impact of this event, there are still some aspects that require further understanding, especially regarding how the exposure dose level varies with the angle and distance from the reactor center and the specific contribution of certain radionuclides to the radiation dose received in a region. This study was conducted to analyze the distribution pattern of exposure dose levels with respect to angle and distance from ChNPP and to identify the relationship between exposure dose and the activity of major radionuclides such as Zr-95, Nb-95, Ru-106, Cs-134, and Cs-137.

This study aims to analyze the distribution pattern of radiation exposure dose levels against variations in angle and distance from the Chernobyl Nuclear Power Plant (ChNPP), and determine the relationship between exposure dose and the activities of major radionuclides such as Zr-95, Nb-95, Ru-106, Cs-134, and Cs-137. The study identifies the trend of changes in exposure dose levels based on the laws of physics, particularly the inverse square law, in radiation dispersion in the environment around Chernobyl. By investigating the specific contribution of each radionuclide to the radiation dose at various measurement points, the research is expected to provide a deeper understanding of the variation of radiation exposure in the area, which can serve as a basis for mitigation and management of the contaminated environment. The benefits of this research include providing more detailed scientific information on radiation distribution around the Chernobyl accident site and providing a scientific basis for radiation risk mitigation planning for the affected population and environment. This research supports the development of radiation mapping models based on distance, angle, and radionuclide activity variables, which can be used for risk analysis and nuclear safety policy. This research can be applied in studies of the long-term impacts of radioactive contamination. This research is expected to be a reference for further studies in the fields of environmental radiology and nuclear physics, as well as radioactive waste management policies and restoration of contaminated areas.

This research has several limitations that need to be considered. Geographically, the study focused on the area around the Chernobyl Nuclear Power Plant (ChNPP) with coverage based on distance and angle from the explosion center. The analysis was limited to the relationship between

radiation exposure dose levels and distance and angle factors, and only involved major radionuclide isotopes such as Zr-95, Nb-95, Ru-106, Cs-134, and Cs-137. The data used came from secondary sources documented in previous studies as well as field measurements over a period of time, with the analysis method focusing on the correlation between radiation exposure levels and the variables under study. This study does not consider long-term changes in radiation levels, but only within the specified data collection period. External factors such as weather, changes in topography, and decontamination efforts were not analyzed in depth. Therefore, the results of this study apply in a site-specific and time-specific context, so their applicability to other regions or different environmental conditions requires further consideration.

In this study, there are several main hypotheses that will be tested to understand the radiation exposure dose rate distribution pattern around the Chernobyl Nuclear Power Plant (ChNPP). First, the hypothesis regarding the relationship between exposure dose rate and distance from ChNPP tests whether there is a significant relationship between the two variables. The null hypothesis states that there is no significant relationship, while the alternative hypothesis states that the exposure dose decreases with distance. The angular distribution of radiation exposure was also analyzed to determine if there is a significant angular distribution pattern, where the alternative hypothesis states that exposure is uneven in all directions. This study evaluated the correlation between exposure dose rate and the activity of major radionuclides, such as Zr-95, Nb-95, Ru-106, Cs-134, and Cs-137. The null hypothesis states that there is no significant relationship between the two, while the alternative hypothesis argues that some radionuclides have a greater contribution to the exposure rate. The exposure dose rate distribution was also analyzed based on the inverse square law with respect to distance from ChNPP. The alternative hypothesis states that this distribution follows the pattern of the inverse square law, which interprets the characteristics of radiation dispersion in an open environment. This study tests whether there are significant differences in the contribution of each isotope to the exposure dose rate. If the alternative hypothesis is accepted, certain isotopes can be identified that have a dominant role in determining the level of radiation exposure around the study site.

In this study, there are several important theories that discuss the impact of radiation on the environment and health. The first theory is the radiation dispersion theory, which adheres to the principle of the inverse square law of physics, where radiation intensity decreases as the distance from the radiation source increases [5]. The spread of radiation is also influenced by atmospheric factors such as wind direction and speed as well as weather conditions at the time of the event [6]. The second theory concerns radionuclides and exposure doses. Radionuclides such as Cs-137 and Zr-95 have different half-lives and can affect the environment for a long time [7]. Exposure dose measures the amount of radiation received by the human body, which is influenced by distance factors, radionuclide activity in the air, and length of exposure [8]. The third theory is about the health effects of radiation exposure, which can increase the risk of cancer, genetic disorders and other radiation sicknesses [8]. Isotopes such as Cs-137 can cause long-term impacts due to their long half-life [9]. Radiation distribution theory based on distance and direction explains that radiation inten-

sity decreases with distance, and the direction of radiation spread is influenced by wind and other atmospheric factors, resulting in a non-symmetrical distribution pattern. Radiation dispersion models are used to map radiation distribution and estimate exposure doses, taking into account environmental variables such as topography and weather [9]. The model is useful in identifying areas with high exposure doses that require further attention. Contaminated environment rehabilitation theory addresses efforts to reduce the long-term impacts of radiation and restore ecosystems after nuclear disasters such as Chernobyl [10]. An understanding of radiation distribution is necessary to determine safe areas and those to avoid. The theory of radiation and its effect on ecosystems reveals that radiation can affect the flora and fauna around contaminated areas [11]. The impact of radiation on biodiversity, such as genetic mutations and ecosystem changes, is an important focus in understanding the long-term effects of radiation on the environment [12].

Previous studies have extensively mapped the radiation distribution around Chernobyl, revealing that radiation levels decrease as distance from the reactor increases [13–15]. However, there has been no study analyzing the radiation dose distribution based on the angle and direction of spread from the reactor center. Although research on the influence of major radioactive isotopes such as Cs-137 and Sr-90 has been conducted, the contribution of other radionuclides such as Zr-95, Nb-95, and Ru-106 has not been fully explored. Studies have also identified the impact of radiation on health, but have paid less attention to the relationship between radiation dose distribution and specific health variables [16, 17]. On the other hand, although several radiation distribution models exist, many do not take into account geographical and environmental variations in the radiation dose received. Research on the impact of radiation on the environment also lacks emphasis on mapping the radiation dose received by soil and vegetation in detail. Most radiation dose modeling studies also do not consider long-term contamination of certain isotopes as well as the influence of angle and distance from the radiation source. This study fills these gaps with a more detailed and comprehensive approach, including the influence of dispersal angle, isotope distribution, and geographical and environmental variations in more accurately mapping radiation dose around Chernobyl. This research offers several innovative contributions in analyzing the radiation dose distribution around Chernobyl. One of them is the analysis of radiation dose based on the angle and direction of spread, which has not been widely explored in previous studies. This research also focuses on the contribution of radioactive isotopes other than Cs-137 and Sr-90, such as Zr-95, Nb-95, and Ru-106, which have been less discussed. This research will provide a more detailed mapping of radiation dose at various locations, including soil and vegetation, which have not been widely studied. The development of a more comprehensive radiation dose distribution model considering geographical and environmental factors is also a key focus, providing a deeper understanding of radiation dose variations. This research also examines the relationship between radiation dose distribution and more specific health impacts, filling a gap from previous studies that focused more on long-term impacts. Finally, this research proposes a radiation dose model that incorporates the factors of angle, distance, and isotope type, to produce more accurate dose mapping.

2. Methods

2.1 Primary data collection and analysis

The study used a statistical approach to analyze the radiation exposure dose distribution around the Chernobyl Nuclear Power Plant (ChNPP) by considering variables such as distance, observation angle, exposure dose, and radionuclide activity. The data analysis process was conducted using R Studio software, which was used to describe the radiation pattern and the relationship between variables [18].

Data was obtained from the results of a study conducted by the Ukrainian Institute of Agricultural Radiology (UIAR), which is part of the National University of Life and Environmental Sciences of Ukraine [19]. The research was conducted in the period May to June 1987, which was the time after the nuclear accident, to measure and analyze the level of radionuclide contamination around the ChNPP, the research area can be seen in Fig. 1. The data collection process was carried out by creating a network of radiation sampling points at different distances from the NPP, ranging from 5 km to 60 km with an interval of 10° for each direction. Of the 540 sample points determined, 489 points were sampled (including soil) with the core sampling method using a 14 cm diameter drill and a depth of 5 cm. Soil samples were analyzed at the UIAR laboratory using a high-quality germanium detector (GEM-30185, ORTEC, USA) and an ADCAM-300 multichannel meter (ORTEC, USA) to detect concentrations of gamma-emitting radionuclides, including Zirconium-95 (95Zr), Niobium-95 (95Nb), Ruthenium-106 (106Ru), Caesium-134 (134Cs), Caesium-137 (137Cs), and Cerium-144 (144Ce).

At each measurement point, the radiation dose level was measured at a height of 1 meter above ground level. Soil samples were analyzed to determine radionuclide activity concentrations using a gamma detector [20]. The extraction procedure was performed using a 1M NH4Ac solution at pH 7, which was used to dissolve the exchangeable cesium in the soil (134Cs and 137Cs) [21]. The solution is analyzed to determine the amount of exchangeable cesium, which provides information on the potential radiation to be absorbed or transferred by the environment. The data obtained is used to calculate the density of soil contamination by converting radionuclide activity concentrations to Becquerels per square meter (Bq/m^2), which is more representative for assessing radiation exposure in soil.

Fig. 1. Study Area.

2.2 Mathematical methods for analysis

2.2.1 Exposure Dose Rate (mR/hr)

Electromagnetic radiation or particles from a radioactive source spread isotropically (evenly in all directions) in three-dimensional space [22]. When particles or radiation travel from the source, the intensity of radiation received by a point decreases with increasing distance from the

source [23].

The inverse square law states that the radiation intensity (or exposure dose) at a given distance from the source is inversely proportional to the square of that distance [24]. This means that if you double the distance from the source, the exposure dose will be reduced by four times. Mathematically, this law can be written as:

$$I \propto \frac{1}{R^2}, \tag{2.1}$$

where I is the radiation intensity (or exposure dose), R is the distance from the radiation source, and radiation source activity, A is how much radiation is emitted by the source per unit time. Activity is measured in units of Bq (Becquerel), which indicates the number of decays per second. The radiation intensity, I at a given point will depend on the activity of the radiation source and its distance. Greater radiation source activity will result in higher intensity at the same measurement point. The exposure dose rate, D at a given point will be directly proportional to the radiation source activity, A and inversely proportional to the square of the distance, R from the radiation source. Therefore, the relationship between exposure dose and distance can be written in the form of the formula:

$$D = \frac{A}{R^2}, \tag{2.2}$$

where D is the exposure dose rate at the measurement point (in mR/hour or μGy/hour, depending on the units used), A is the activity of the radiation source (in Bq), and R is the distance from the radiation source (in km or meters, depending on the measurement used). The exposure dose is measured in units such as milliroentgen per hour (mR/hour) or microsievert per hour (μSv/hour), depending on the type of radiation and the units used. In this case, we assume that the exposure dose rate is measured in units of mR/hr.

2.2.2 Correlation between exposure dose rate and distance

To calculate the correlation between exposure dose rate and distance, it is necessary to determine several important values. These include the amount of data (n), the sum of the distance variable, the sum of the exposure dose variable, the sum of the product of the distance and dose value pairs, and the sum of the squares of each variable. Data consist of pairs of observations for distance ($x_1, x_2, ..., x_n$), and radiated dose ($y_1, y_2, , y_n$). Each pair of data will be calculated for its contribution to the sum of the required values. Components in the Pearson formula include the sum of the distance values, the sum of the dose values, the sum of the multiplication of the distance and dose values, and the sum of the squares of each variable. Once all the quantities have been calculated, the values are substituted into the Pearson formula:

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{([n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}. \tag{2.3}$$

The correlation coefficient r is between -1 and 1, where ($r = 1$) is a perfect positive relationship, ($r = -1$) indicates a perfect negative relationship, and ($r = 0$) represents no linear relationship. A positive value r is when x increases, y also increases, while a negative value (r) indicates that when x increases, y decreases.

2.2.3 Absorbed dose (μGy/hour)

Radiation from a source spreads isotropically in space, which means that the radiation intensity decreases according to the inverse square law, as shown in Eq. (2.1). So, we can write the relationship be-

tween intensity and activity:

$$I = \frac{A}{4\pi R^2}. \tag{2.4}$$

Since in practical conditions, the intensity measured on a particular surface without taking into account the spread throughout the space, the factor of (4π) soften ignored in empirical calculations. Absorbed dose (D_s) is the energy absorbed by a medium per unit mass. The general relationship between absorbed dose and radiation intensity is given by:

$$D_s = I \times T. \tag{2.5}$$

Substitute Eq. (2.4), then:

$$D_s = \frac{A}{R^2} \times T. \tag{2.6}$$

The absorbed dose (D_s) in units of microGray per hour (μGy/h) is calculated based on the radionuclide activity (A) in Becquerel (Bq), which is the number of nuclear decays per second. This dose value depends on the distance from the radiation source (R) in kilometers (km), where the radiation intensity decreases according to the inverse square law. A time conversion factor (T) is used to adjust the dose based on the duration of exposure in a given unit of time.

2.2.4 Activity radionuclide (Bq/m^2)

Measurement in an area of (A_t). If the total number of radiation particles detected in the area is (N), then the activity per unit area can be calculated as

$$A = \frac{N}{A_t}, \tag{2.7}$$

where N is the number of radiation particles detected in the measurement area and A_t is the measurement area in square meters (m^2). Eq. (2.7) is useful for calculating the activity of radionuclides in units of Bq/m^2. This formula is derived from the basic definition of radionuclide activity which is calculated based on the number of decays that occur within a certain area.

2.2.5 Interquartile range (IQR)

Quartiles divide the sorted data into four equal parts, where the first quartile (Q1) is the point that divides the first 25% of the data, the second quartile (Q2 or median) divides the dataset into two equal parts (50%), and the third quartile (Q3) divides 75% of the dataset. The determination of quartiles depends on the number of data n, either odd or even. If n is odd, the position of the first quartile is calculated by:

$$Q_1 = x_{(\frac{n+1}{4})}, \tag{2.8}$$

and the third quartile:

$$Q_3 = x_{(\frac{3(n+1)}{4})}. \tag{2.9}$$

If n is even, the position of Q1 is obtained from the average of the two center values with the formula:

$$Q_1 = \frac{x_{(\frac{n}{4})} + x_{(\frac{n}{4}+1)}}{2}. \tag{2.10}$$

Q3 is calculated as:

$$Q_3 = \frac{x_{(\frac{3n}{4})} + x_{(\frac{3n}{4}+1)}}{2}. \tag{2.11}$$

The interquartile range (IQR) is then obtained from the difference between the third quartile and the first quartile.

2.2.6 Standard deviation

Standard Deviation (σ) is defined as the square root of the variance [25]. Variance itself measures the average squared distance of each data point from the mean:

$$\sigma = \sqrt{Var(X)}, \tag{2.12}$$

where $Var(X)$ is the variance of the data (X) [25]. Variance is calculated with the following formula:

$$Var(X) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2. \quad (2.13)$$

However, in the calculation of population standard deviation, where (n) is the divisor, while for samples (if the data used is a sample of the population), ($n-1$) is used as the divisor to get an unbiased estimate. The first step in calculating standard deviation is to calculate the average of the data. The average (μ) is the sum of all data values divided by the total number of data (n):

$$\mu = \frac{1}{n}\sum_{i=1}^{n}x_i. \quad (2.14)$$

2.2.7 Poison regression

A Poisson regression model aims to model the relationship between the dependent variable (radiation dose) and one or more independent variables (distance from ChNPP, Zr-95 activity, and Cs-134 activity) [26]. It is assumed that the value of (Y) (radiation dose) follows a Poisson distribution, where the mean (λ) depends on the independent variables ($X_1, X_2, , X_n$). The Poisson regression model is in the form:

$$\log(\mu) = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n, \quad (2.15)$$

where (μ) a is the predicted value for radiation dose (average or incident intensity), ($\beta_0, \beta_1, , \beta_n$) is the estimated regression coefficient, and ($X_1, X_2, , X_n$) are the independent variables (Distance from ChNPP, Zr-95 activity, and Cs-134 activity). In the Poisson regression model, it is assumed that the expected (μ) (average radiation incidence) can be influenced by the independent variables [27]. Since (Y) follows a Poisson distribution:

$$\mu = E[Y] = \exp(\beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n). \quad (2.16)$$

Use of the natural logarithm function ($\log(\mu)$) converts the model into a linear form, which is easier to estimate with linear regression techniques [28]. In this case, the logarithm function transforms the regression parameter ($\beta_0, \beta_1, , \beta_n$) into influences that are easier to interpret. The likelihood function for the Poisson distribution, with observed data, is given by:

$$L(\beta_0, \beta_1, , \beta_n) = \Pi_{i=1}^{n}\frac{(\mu_i)^{y_i}e^{-\mu_i}}{y_i!}, \quad (2.17)$$

where $\mu_i = \exp(\beta_0 + \beta_1 X_{1i} + \cdots + \beta_n X_{ni})$ is the expectation or average for each observation (i). To obtain parameter estimates ($\beta_0, \beta_1, , \beta_n$), maximizes the likelihood function by calculating its log-likelihood derivative and setting it equal to zero. After obtaining the estimated values for the regression coefficients ($\hat{\beta}_0, \hat{\beta}_1, , \hat{\beta}_n$), the effect of each independent variable on radiation dose was interpreted. Specifically, the regression coefficient (β_i) measures the log change of expected radiation dose (μ) to a one-unit change in the variable (X_i) For relative changes in (μ), exponentiate the coefficient (β_i):

$$\exp(\beta_i). \quad (2.18)$$

Interpretation as a multiplying factor that shows how changes in X_i affect radiation dose.

2.2.8 Overdispersion test

Deviance (D) in Poisson regression measures how far the fitted model differs from the perfect model (saturation) [29]. Deviance is defined as:

$$D = 2\sum_{i=1}^{n}[y_i \log(\frac{y_i}{\mu_i}) - (y_i - \hat{\mu}_i)], \quad (2.19)$$

where (Y_i) is the actual value of observation (i), $\hat{\mu}_i$ is the value predicted by the model for observation (i), n is the total number of observations. The degrees of freedom in the model are calculated as:

$$df = n - p, \tag{2.20}$$

where n is the total number of observations and p is the number of parameters in the model (including the intercept). Dispersion statistics (ϕ) are calculated by comparing deviance with degrees of freedom [30]:

$$\phi = \frac{D}{df}, \tag{2.21}$$

where ϕ is the dispersion statistic, D is the deviance, and df is the degrees of freedom. If $\phi \approx 1$, then the Poisson model is appropriate. If $\phi > 1$, then overdispersion occurs.

2.2.9 Variance Inflation Factor (VIF)

Multiple linear regression with the model:

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + \epsilon. \tag{2.22}$$

β_i has the variance given by:

$$Var(\hat{\beta}_i) = \sigma^2 (X^T X)^{-1}_{ii}. \tag{2.23}$$

where σ^2 is the variance of the error (ϵ), and ($X^T X$) is the information matrix of the independent variables. If multicollinearity is present, the diagonal elements of matrix $(X^T T)^{-1}$ will increase, which leads to an increase in the variance of the regression coefficient [31]. To understand how (R_i^2) influence the variance, regression of independent variables against other independent variables has been considered:

$$\begin{aligned} X_i = \quad & \alpha_0 + \alpha_1 X_1 + \cdots + \alpha_{i-1} X_{i-1} \\ & + \alpha_{i+1} X_{i+1} + \cdots + \alpha_k X_k + \epsilon. \end{aligned} \tag{2.24}$$

This regression measures the extent to which (X_i) can be explained by a linear combination of the other variables. If R_i^2 is close to 1, then X_i are almost completely predicted by the other variables, which indicates high multicollinearity [32]. It is known that the variance of the regression coefficient (β_i) can be expressed as:

$$Var(\hat{\beta}_i) = \frac{\sigma^2}{1 - R_i^2} (X^T X)^{-1}_{ii}. \tag{2.25}$$

$(X^T X)^{-1}_{ii}$ is the basic factor of variance in regression, we define the variance inflation factor as:

$$VIF_i = \frac{1}{1 - R_i^2}. \tag{2.26}$$

VIF_i measures how much variability the coefficient (β_i) increased due to multicollinearity [33]. If R_i^2 is close to 1, then $1 - R_i^2$ is close to zero, so VIF_i becomes very large, indicating high multicollinearity. If R_i^2 is low, then VIF_i close to 1, which means there is no significant multicollinearity. As a rule of thumb, $VIF > 10$ is very high multicollinearity, VIF between 5 and 10 is moderate multicollinearity, while $VIF < 5$ indicates no serious problems with multicollinearity [32].

2.2.10 Deviance analysis with Chi-Square Test (ANOVA)

In regression analysis, there are two types of deviance used to evaluate the model, namely the Deviance Model and the Deviance Null Model [30]. Model Deviance is the deviance value of the model that includes predictor variables or the full model, which is the error rate after considering the variables in the model. The Null Model deviance is the deviance of a model that uses only the intercept without any predictor variables, thus representing the error

rate of the model without additional information [34]. If the Model Deviance value is much smaller than the Null Model Deviance, then the model with predictor variables is considered better in explaining the data [35].

In deviance analysis, the Chi-Square test is used to determine whether the difference between the model with and without the predictor variable is significant or just occurs by chance [36]. The null hypothesis (H_0) states that a model without predictor variables is sufficient to explain the data, while the alternative hypothesis (H_1) states that a model with predictor variables better explains the data. The Chi-Square test statistic is calculated using the formula:

$$\chi^2 = D_{null} - D_{model}, \tag{2.27}$$

where D_{null} is the deviance of the model without predictor variables, and D_{model} is the deviance of the model with predictor variables. Values (χ^2) this follows the Chi-Square distribution with degrees of freedom (df) calculated as the difference between the number of parameters in the full model (p_{model}) and the number of parameters in the null model (p_{null}). If the value of χ^2 obtained is significant, then the model with predictor variables is considered better in explaining the data. Deviance (D) is a measure of how well a regression model can explain variation in the data compared to a saturated model [37]. Mathematically, deviance is defined as:

$$D = 2\sum_{i=1}^{n}[y_i \log(\frac{y_i}{\hat{y}_i}) + (1-y_i)\log(\frac{1-y_i}{1-\hat{y}_i})], \tag{2.28}$$

where y_i is the actual value of observation (i), $\hat{y}_i$ is the value predicted by the model, and (n) is the amount of data.

2.2.11 Principal Component Analysis (PCA)

We have a dataset with p variables and n observations represented in the form of a matrix X of size $n \times p$. To get PCA, it is necessary to standardize the data by calculating the mean and standard deviation of each variable [38] as follows:

$$Z = \frac{X - \bar{X}}{\sigma}. \tag{2.29}$$

As for $\bar{X}$ is the average of each variable and σ is the standard deviation of each variable. If the data has been standardized, the covariance matrix C can be calculated as follows:

$$C = \frac{1}{n-1}Z^T Z. \tag{2.30}$$

Or a correlation matrix if the variables have different scales. For the main component, the eigenvalue can be found λ_i dan vector own (v_i) from matrix covarians (C):

$$Cv_i = \lambda_i v_i, \tag{2.31}$$

where λ_i is the eigenvalue indicating the variance of the i-th principal component and (v_i) is the eigenvector that shows the direction of the principal component. Since each eigenvalue λ_i expresses the variance explained by the i-th principal component, then the proportion of variance explained by the i-th principal component is given by:

$$\text{Proportion of Variance} = \frac{\lambda_i}{\sum \lambda_i}, \tag{2.32}$$

where λ_i is the eigenvalue of the (i)th principal component and $\sum \lambda_i$ is the total eigenvalue of all the principal components. Eigenvalue λ_i is the variance explained by the i-th principal component, where the component with the largest eigenvalue captures the largest variance in the data [39]. The sum of all eigenvalues $\sum \lambda_i$

represents the total variance of the dataset, so PCA only changes the representation of the data without reducing the information. The calculated variance proportion of $\lambda_i / \sum \lambda_i$ helps determine the extent to which each principal component explains the variability in the data.

2.2.12 Variogram

The empirical estimate is an unbiased estimate of the definition of variogram statistics. Starting from the definition:

$$\gamma(h) = \frac{1}{2} E[z(x) - z(x+h)^2]. \quad (2.33)$$

Assume $n(h)$ sample pairs $(x_i, x_i + h)$, so the expectation can be approximated by the sample mean:

$$E[(z(x) - z(x+h)^2)] \approx \frac{1}{n(h)} \sum_{i=1}^{n(h)} (z(x) - z(x+h))^2. \quad (2.34)$$

So

$$\gamma(h) = \frac{1}{2} \times \frac{1}{n(h)} \sum_{i=1}^{n(h)} [z(x) - z(x+h)]^2, \quad (2.35)$$

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} [z(x) - z(x+h)]^2. \quad (2.36)$$

This empirical formula for the variogram is a valid estimation based on its statistical definition. If the value of $\gamma(h)$ is small, the variability between $z(x)$ and $z(x+h)$ is low, while the spatial correlation is high. Conversely, if $\gamma(h)$ larger, the variability increases, indicating that the values at locations x and $x + h$ are increasingly unrelated. At very large distances, $\gamma(h)$ tends to reach a sill, which indicates that spatial correlation no longer exists.

2.2.13 K-Mean Clustering

Each data point is placed into a cluster (C_k) which has a nearby centroid (C_k). This means:

$$C_k = \{x_i | \|x_i - c_k\|^2 \leq \|x_i - c_j\|^2, \forall j \neq k\}. \quad (2.37)$$

Which means that the data points (x_i) will be included in the cluster (C_k) After the cluster division is done, update the centroid to be the average of all points in the cluster:

$$c_k = \frac{1}{|C_k|} \sum_{x_i \in C_k} x_i. \quad (2.38)$$

As for $|C_k|$, the number of members in the cluster (C_k). $\sum_{x_i \in C_k} x_i$ is the number of all points in the cluster. The minimized SSD objective function is:

$$SSD = \sum_{k=1}^{K} \sum_{x_i \in C_k} \|x_i - c_k\|^2. \quad (2.39)$$

At the optimal value of the centroid derived SSD against (c_k):

$$\frac{\partial}{\partial c_k} \sum_{x_i \in C_k} \|x_i - c_k\|^2 = \sum_{x_i \in C_k} 2(x_i - c_k) = 0, \quad (2.40)$$

$$\sum_{x_i \in C_k} x_i - |C_k| c_k = 0,$$

$$c_k = \frac{1}{|C_k|} \sum_{x_i \in C_k} x_i. \quad (2.41)$$

This shows that K-Means always updates the centroid with the average of the points in the cluster, which is mathematically the best way to minimize SSD.

The K-means algorithm works iteratively by randomly assigning an initial centroid, then clustering the data points to the nearest centroid (E-step), and updating the centroid based on the average of the points in the cluster (M-step) [40]. This process repeats until the centroid stabilizes or changes

very little. Since the SSD is always decreasing or fixed, the K-means algorithm is guaranteed to converge, ensuring optimal clustering by minimizing the distance within the cluster [41].

3. Results and Discussion

In the analysis, the distribution of exposure dose levels with respect to the angle and distance from the Chernobyl Nuclear Power Plant (ChNPP) was observed, as shown in Figs. 2-3. From the graphical results, it can be seen that the points representing measurements at various angles show variations in exposure dose levels with respect to the distance from the ChNPP. As the distance from the ChNPP increases, the exposure dose level tends to decrease, as seen in the color change of the dot from blue to red. This implies a higher exposure dose in areas closer to the radiation source. The difference in dot size also illustrates the difference in exposure dose size, which becomes clearer as the different colors indicate different distances from the ChNPP.

For the angle variable, the mean is 185.5 degrees with the radiation distribution concentrated around 360 degrees, slightly skewed towards the semicircle. The median angle of 190 degrees and standard deviation of 105.26 degrees show high variation. The angles range between 10 and 360 degrees, with the majority of the data distributed between 90 and 280 degrees. For the distance from the ChNPP, the mean is 23.62 km, with the majority of observations at a relatively close distance from the ChNPP. The median distance of 17 km with a standard deviation of 17.51 km shows considerable variation. The distance ranges between 5 and 60 km, with most of the data falling between 8.3 km and 37.5 km. The mean exposure dose rate was 1.64 mR/hr, with a median of 0.4 mR/hr and a standard deviation of 3.18 mR/hr, implying considerable variation in doses. The dose range was between 0.01 to 26 mR/h, with most observations showing low doses. The correlation between Exposure Dose Rate and Distance from ChNPP was -0.3734, which shows a moderate negative correlation, means that the greater the distance from ChNPP, the lower the exposure dose. The correlation between Exposure Dose Rate and Angle was 0.1221, a very weak positive correlation, meaning that the angle of observation hardly affected the exposure dose, with distance being more dominant. The average angle is about 185.78 degrees with a standard deviation of 105.14, showing a fairly wide distribution around the center of the circle. For Distance from ChNPP, the average distance was 23.66 km with a standard deviation of 17.53 km, showing a fairly wide variation in the distance measurement data. In the Absorbed Dose Rate variable, the average absorbed dose was 14.22 microGray/hour with a standard deviation of 27.71, showing a very large variation in dose at different locations. The lowest absorbed dose value was 0, and the highest value reached 227 microGray/hour, a measurement point with very high radiation.

For the isotope Nb-95, the average activity was 403,684.8 Bq/m^2 with a large variation (standard deviation 799,006.4 Bq/m^2). The highest activity value was recorded at 7,070,000 Bq/m^2 and the lowest at 740 Bq/m^2. The median was 126,000 Bq/m^2 and the interquartile range (IQR) was 359,200 Bq/m^2, implying a wide distribution of activity. For Ru-106, the mean activity was 1,376,814 Bq/m^2 with a very high standard deviation (2,509,827 Bq/m^2), meaning a very large variation. The maximum value was recorded as 20,000,000 Bq/m^2 and the minimum as 7,400 Bq/m^2.

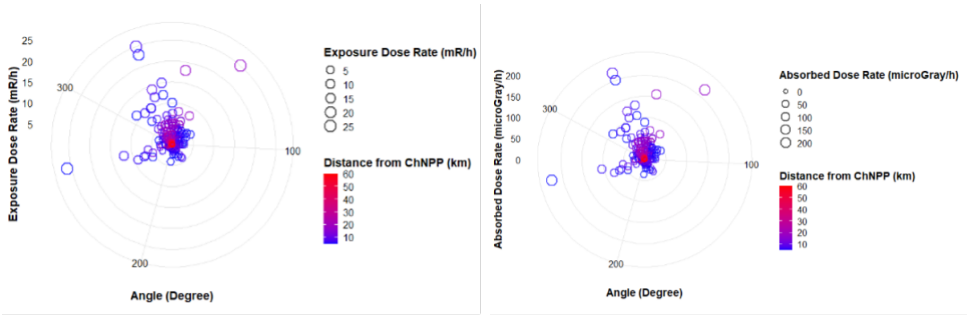

Fig. 2. Distribution of absorbed dose and exposure in polar coordinates.

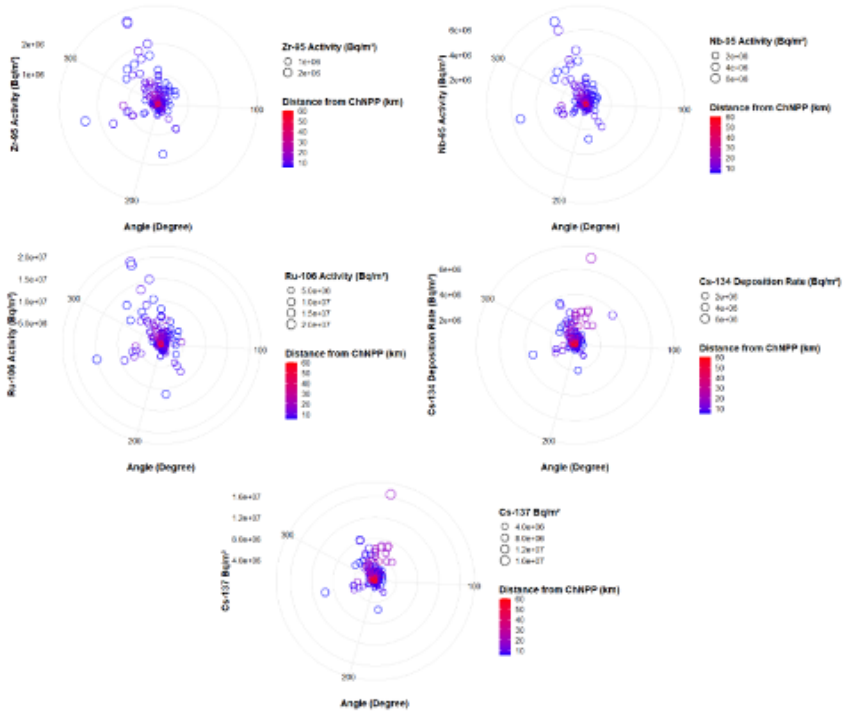

Fig. 3. Distribution of radionuclide activity (Zr-95, Nb-95, Ru-106, Cs-134, and Cs-137) around the Chernobyl Nuclear Power Plant (ChNPP) within a radius of 60 km.

The median of 459,000 Bq/m^2 and IQR of 1,124,000 Bq/m^2 means a fairly wide distribution. For Cs-134, the mean activity was 379,388.3 Bq/m^2 with a lower standard deviation (688,317.9 Bq/m^2). The lowest activity value was recorded at 3,700 Bq/m^2 and the highest at 6,960,000 Bq/m^2. The median of 133,000 Bq/m^2 and IQR of 324,000 Bq/m^2 means a more centralized distribution than the other isotopes. Cs-137 has a mean activity of 976,701.6 Bq/m^2 with a high standard deviation (1,659,719 Bq/m^2). The minimum value was recorded as 7,400 Bq/m^2 and the maximum as 16,600,000 Bq/m^2. The median of 366,000 Bq/m^2 and the IQR of 879,000 Bq/m^2 represented considerable variation around the mean of value. The difference be-

tween the mean and median for Nb-95 is 277,684.8 Bq/m^2, showing a distribution skewed towards higher values, with some measurement points having very high activity. The difference between the maximum and minimum values is 7,069,260 Bq/m^2, meaning a very large range and uneven distribution of radiation. In Ru-106, the difference between the mean and median (917,814.3 Bq/m^2) is larger, the distribution is more skewed with a significant difference between the mean and median. The difference between the maximum and minimum values of 19,992,600 Bq/m^2 shows a very large variation, with high radiation concentrations at some points and low values in most areas. In Cs-134, the difference between the mean and median is 246,388.3 Bq/m^2, smaller than the other isotopes, showing a more concentrated distribution. Although there is still variation, the difference between the maximum and minimum (6,956,300 Bq/m^2) is not as large as in Ru-106, indicating a more consistent distribution. In Cs-137, the mean-median difference is 610,701.6 Bq/m^2, larger than in Cs-134 and similar to Nb-95 and Ru-106, showing a skewed distribution. The maximum-minimum difference of 16,592,600 Bq/m^2 implies a very large range, almost equivalent to Ru-106, with high radiation concentrations in some spots and many spots with low activity.

The results of the Pearson correlation analysis between the two variables, Exposure Dose rate and Absorbed dose rate, show a correlation coefficient of 1.00, as shown in Fig. 4. This is a very strong and perfect relationship between the two variables. Changes in the Exposure Dose rate in mR/h are followed proportionally by changes in the Absorbed Dose rate in microGray/h, without any significant deviation or fluctuation between the two. This correlation of 1 indicates that the two variables move together in the same direction, with a very strong linear relationship. Fig. 5 shows the activity distribution of radionuclides (Zr-95, Nb-95, Ru-106, Cs-134, and Cs-137) around the Chernobyl Nuclear Power Plant (ChNPP) within a radius of 60 km, using polar coordinates. The x-axis represents the angle in degrees, while the y-axis represents the radionuclide activity in Bq/m^2. Each data point is colored based on its distance from the ChNPP, from blue (near) to red (far), with the size of the point representing the radionuclide activity level. The coefficient for dis-

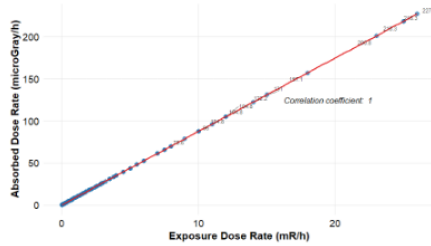

Fig. 4. Correlation of adsorbed dose rate to exposure dose rate.

tance from ChNPP shows a highly significant value (-0.8637 with p-value < 2e-16), which means that the greater the distance from ChNPP, the lower the level of radiation dose exposure received. This is consistent with the effect of distance on the decrease in radiation intensity, and is similar to the inverse square law which describes the decrease in radiation intensity as the distance from the radiation source increases. The activity of Zr-95 and Cs-134 also had a positive effect on the radiation dose exposed. The coefficient for Zr-95 is 0.2085 (p-value < 2e-16) and for Cs-134 is 0.3117 (p-value < 2e-16), which means that the higher the activity of both isotopes, the greater the radiation dose exposed to the area. This means that Zr-95 and Cs-134

Table 1. Correlation of adsorbed dose rate to exposure dose rate.

Factor	Koefisien (Estimate)	Std. Error	z-value	*p*-value
Intercept	-0.00263	0.065003	-0.041	0.968
Distance from ChNPP in KM	-0.86372	0.09201	-9.387	<2e-16 ***
Zr-95 activity in Bq/m^2	0.208493	0.022284	9.356	<2e-16 ***
Activity of Cs-134 in	0.311708	0.020854	14.947	<2e-16 ***

contribute significantly to the radiation levels around ChNPP, with an increase in the activity of these isotopes being directly proportional to the increase in radiation dose received. There is a strong relationship between the factors tested and the radiation dose exposed, with deviance meaning a significant improvement from the null model (1266.29 to 319.43). Table 1 summarizes the results of the Poisson regression analysis conducted to identify factors affecting radiation exposure dose around ChNPP.

Based on the overdispersion test on the Poisson regression model, the statistical value is 0.8849. The overdispersion test is used to check whether the data used in the model has greater variability than expected based on the Poisson distribution. In this case, a test statistic higher than 1 usually means overdispersion, which is a condition where the variation of the data exceeds what is predicted by the Poisson model, which generally assumes that the variance and mean are equal. However, the statistical value obtained (0.8849) is lower than 1, indicating that this data does not suffer from overdispersion. This means that the variation in the measured radiation data can be explained reasonably well by the Poisson model, without any additional unexplained variability. The Poisson distribution is therefore an appropriate approximation for these data, as there is no indication that the model fails to capture the variation in the observed radiation dose.

Fig. 5 on the left shows the relationship between the fitted values and the residuals (the difference between the observed and predicted values). The blue dots on this graph are scattered around the red horizontal line that defines zero residuals, indicating that no systematic pattern remains in the residuals after the model is applied. This indicates that the Poisson regression model does a good job of predicting the data, without any suspicious patterns or bias in the prediction error. The right image is a histogram of the residuals, this is the distribution of the model error. Most of the residuals are symmetrically distributed around zero, with the highest frequency of residuals close to zero. It can be seen that most of the model predictions are fairly accurate, although there are some more extreme residual values.

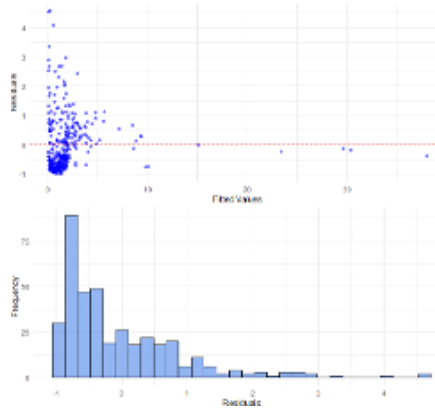

Fig. 5. Relationship between predicted value and residuals (left) and histogram of residuals distribution (right) in poisson regression model.

The Variance Inflation Factors (VIF) analysis results obtained from the

Table 2. Variance Inflation Factors (VIF) dan analisis deviance menggunakan uji Chi-Square (anova).

Variable	VIF Value	Deviance Reduction	*p*-value
Jarak dari ChNPP dalam km	1.296	413.09	<2.2e-16
Zr-95 in Bq/m^2	1.975	407.05	<2.2e-16
Cs-134 Bq/m^2	1.611	126.72	<2.2e-16

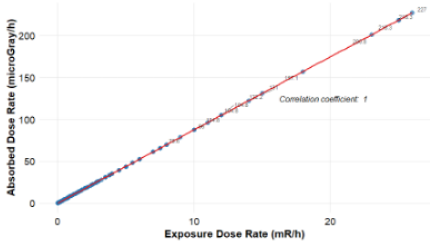

Fig. 6. Correlation of adsorbed dose rate to exposure dose rate.

Poisson regression model show the VIF values for the variables used, namely, Distance_from_ChNPP_km, Zr95_Bqm2, and Cs134_Bqm2. The VIF values are 1.29597, 1.975426, and 1.610596, respectively, which means that there is no significant multicollinearity among the independent variables. VIF values below 5 indicate that this model does not experience multicollinearity problems that can significantly affect the interpretation of regression coefficients. The results of the deviance analysis through the Chi-Square (anova) test mean that each independent variable included in the model makes a significant contribution to reducing the deviance model. The first variable, Distance_from_ChNPP_km, reduced the deviance by 413.09 with a very small *p*-value (< 2.2e-16), shows that the distance from ChNPP greatly affects the level of radiation dose exposure. Zr95_Bqm2 reduces deviance by 407.05, with a very small *p*-value (< 2.2e-16), where Zr-95 activity is significantly related to the exposed radiation dose. The addition of the variable Cs134_Bqm2 reduces deviance by 126.72, and the *p*-value is very small (< 2.2e-16) which means that Cs-134 activity also has a significant relationship with radiation dose. Table 2 summarizes the results of the Variance Inflation Factors (VIF) analysis and deviance analysis using the Chi-Square (anova) test.

Table 2 shows that the VIF values for each independent variable are low (all below 5), so there is no significant multicollinearity problem. The deviance analysis (using the Chi-Square test) showed that each variable-distance from ChNPP, Zr-95 activity, and Cs-134 activity-made a significant contribution to the model's deviance reduction.

In the principal component analysis (PCA) applied to the radionuclide activity dataset around ChNPP, the results show that the first principal component (PC1) explains most of the variability in the data. With a standard deviation of 2.0638, PC1 contains about 85.19% of the total variance, indicating that the variables in the dataset are strongly correlated and most of the information can be explained by this one principal component. The second principal component (PC2) is also significant, although its contribution is much smaller than PC1, accounting for 13.61% of the variance. This means that while most of the information has been explained by PC1, there is still considerable variation that can be captured by the second component. The third (PC3) and subsequent (PC4 and PC5) principal components have very small contributions to the total variance, explain-

ing only 0.69%, 0.36%, and 0.16% respectively, indicating further information is becoming increasingly limited.

Variogram analysis was conducted to observe the spatial pattern of radiation dose exposure levels around ChNPP. The variogram analysis showed that the average gamma value (variance) for radiation dose fluctuated slightly as the measurement distance increased, with gamma values ranging between 6.938 and 9.620, indicating variations in radiation dose levels in the area. In this analysis, the measurement distances show a range from about 1.1 km to 110 km from the radiation source, with most measurements being within the range of 30 to 80 km from the ChNPP site.

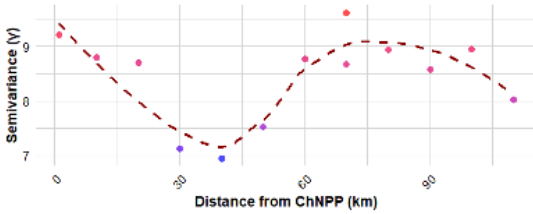

Fig. 7. Fluctuation of radiation dose semivariance based on measurement distance around chernobyl with peak point.

There is a trend seen in Fig. 6 that the farther the measurement distance, the greater the semivariance value recorded, although there are small fluctuations at some distances. Data points farther away from the radiation source mean greater variation in radiation dose, which could be due to uneven radiation dispersion or other environmental factors. The highest semivariance value was recorded at a distance of about 70 km, meaning that at this distance there is a significant change in the level of radiation dose exposure.

The results of applying the K-means algorithm to data including two variables, radiation exposure dose rate and absorbed dose rate, resulted in three distinct clusters. This clustering is based on the proximity of data points in the two-dimensional space, with the aim of identifying patterns that may not be directly visible. From the classification results, there are three clusters with different numbers of observations: Cluster 1 has 316 observations, Cluster 2 has 60 observations, and Cluster 3 has only 9 observations. Cluster 1 defines the group with the largest number of observations, which indicates that most of the data have similar characteristics in terms of exposure dose level and absorbed dose. In contrast, Cluster 3, with only 9 observations, defines a smaller group and implies dose values that are much higher or lower than the average of the other groups. While most of the data points are at lower exposure and dose levels (clustered in Cluster 1), the smaller clusters correspond to areas with very high (Cluster 3) or very low (Cluster 2) radiation exposure.

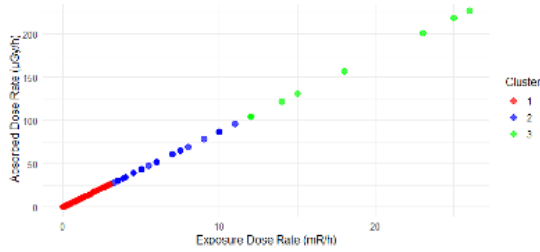

Fig. 8. Variation by dose group with colors representing exposure levels.

Fig.7 shows that the red color represents the group with the higher dose level, while the blue and green colors represent the group with the lower dose. There is significant variation between data points in terms of radiation exposure and absorbed dose received by objects or individuals around the area.

Fig. 8 is the estimated radiation intensity based on the distance from the radiation source (ChNPP), the function used

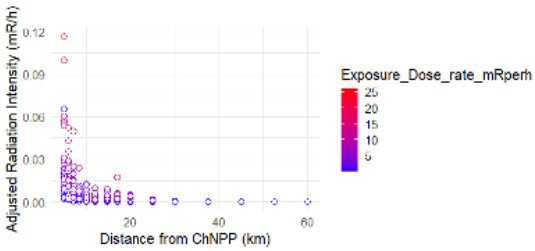

Fig. 9. Radiation Intensity Based on Distance from Radiation Source (ChNPP) with Moderate Negative Correlation (-0.349).

Fig. 10. Fluctuation of total absorbed radiation dose over time in the vicinity of the Chernobyl Nuclear Power Plant (ChNPP).

calculates the radiation intensity based on distance, where the initial intensity at the 5 km point is considered as a reference. The calculated correlation between the distance from ChNPP and the adjusted intensity shows a correlation value of -0.349. This value indicates a negative relationship between the two variables, albeit with a moderate strength. This negative relationship is consistent with the physical law of radiation scattering, which states that radiation intensity decreases as the distance from the radiation source increases. Although the correlation is not very strong, the value of -0.349 means that distance plays a role in reducing the radiation dose received, which is important for protection against radiation exposure around ChNPP. For a point located at a distance of 5 km from ChNPP, the intensity after attenuation is calculated to be about 606.53, which means that the radiation intensity decreases exponentially as the distance from the radiation source increases.

Fig. 9 represents the total absorbed radiation dose over time, measured at various time intervals around the ChNPP. It can be seen that the total radiation dose fluctuates over time, which interprets the variation in radiation exposure received at various time points. An increase or decrease in the dose may indicate changes in the intensity of radiation exposed due to factors such as changes in radionuclide activity or environmental conditions around the ChNPP.

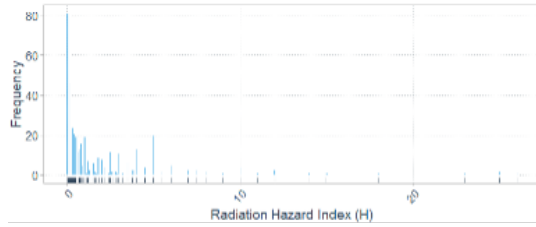

Fig. 11. Distribution of radiation hazard index around the Chernobyl Nuclear Power Plant (ChNPP) and its trend toward safe radiation levels.

Fig. 10 shows that most of the Radiation Hazard Index values are centered around low values, with most of the data below the value of 1, meaning relatively safe radiation levels at a greater distance from ChNPP. However, there are some higher values, with the maximum reaching 26, indicating measurements with significant radiation doses at some points closer to the radiation source. The vertical line is the average Radiation Hazard Index giving an idea of the general trend of the data, which mostly shows values below the prescribed safe limit (1 mR/h).

4. Conclusion

The theoretical conclusion of the study shows that there is a strong relation-

ship between the distance from ChNPP and the level of radiation dose exposure, in accordance with the inverse square law, where the greater the distance, the lower the radiation dose received. The activity of isotopes such as Zr-95, Cs-134, and Ru-106 is significantly related to the radiation dose received, with an increase in isotope activity leading to higher radiation exposure. The exposure dose and absorbed dose have a very strong linear correlation with no deviations. The use of Poisson regression models was effective in analyzing the influence of factors such as distance and isotope activity on radiation dose, with appropriate results and well-explained data variance.

Practical conclusions from the study show the importance of considering the distance from the ChNPP in radiation risk management. Reduction of radiation exposure in areas farther from the plant can be used to establish safe zones. Activity monitoring of isotopes such as Zr-95 and Cs-134 is key to understanding radiation distribution and identifying areas of high risk. The absorbed radiation dose needs to be monitored continuously to protect public health, and health policies should be adjusted in light of these findings. Poisson regression models can help in radiation dose analysis, while public education on the importance of maintaining a safe distance and complying with radiation regulations is needed.

Future research can focus on temporal analysis of radiation dose distribution around ChNPP to understand long-term changes in radiation exposure due to different isotope activities. In addition, it is important to develop a more accurate radiation dose prediction model by considering other environmental factors such as weather conditions and topography. Further research is also recommended to explore the impact of long-term radiation exposure on public health, with a focus on epidemiology and the development of more effective mitigation strategies.

Acknowledgements

We would like to express our sincere gratitude to Dr. V. Kashparov, Dr. S. Levchuk, Dr. M. Zhurba, Dr. V. Protsak, Dr. N.A. Beresford, and Dr. J.S. Chaplow for providing the spatial radionuclide deposition data from the 60 km radial area around the Chernobyl Nuclear Power Plant, collected in 1987. This dataset, available under the Open Government Licence (OGL), has been instrumental in completing this extended research. We also acknowledge the Ukrainian Institute of Agricultural Radiology and the NERC Environmental Information Data Centre for their support in making this data available for analysis.

References

[1] Sarkisov AA, Vysotskii VL. The largest nuclear accident in the history of the nuclear fleet. Reconstruction of events and analysis of the accident consequences to assess the risks and hazards of small nuclear power facilities. Nuclear Engineering and Design. 2021;384:111440.

[2] Michelot L. Chernobyl: The Path to Healing Human and Ecological Scars. In: Sustainability Stories: The Power of Narratives to Understand Global Challenges. Springer; 2024. p. 297–305.

[3] Iqbal J, Howari FM, Mohamed AMO, Paleologos EK. Assessment of radiation pollution from nuclear power plants. In: Pollution Assessment for Sustainable Practices in Applied Sciences and Engineering. Elsevier; 2021. p. 1027–53.

[4] Masson O, Romanenko O, Saunier O, Kirieiev S, Protsak V, Laptev G, et al. Europe-wide atmospheric radionuclide

dispersion by unprecedented wildfires in the Chernobyl exclusion zone, April 2020. Environmental Science & Technology. 2021;55(20):13834–48.

[5] Hyvärinen T. Lower extremity dose dependency on chest dose, distance from patient, and imaging settings in cardiology and radiology staff. 2023.

[6] Akinyoola JA, Oluleye A, Gbode IE. A Review of Atmospheric Aerosol Impacts on Regional Extreme Weather and Climate Events. Aerosol Science and Engineering. 2024;8(3):249–74.

[7] Corbacho JÁ, Guillén J. Long-term radiological assessment of a Mediterranean freshwater ecosystem surrounding a nuclear power plant. Environmental Science and Pollution Research. 2024;1–15.

[8] Tapio S, Little MP, Kaiser JC, Impens N, Hamada N, Georgakilas AG, et al. Ionizing radiation-induced circulatory and metabolic diseases. Environment international. 2021;146:106235.

[9] Okhrimchuk D, Hurtevent P, Gonze MA, Simon-Cornu M, Roulier M, Carasco L, et al. Long-term behaviour of Cs-137, Cs-133 and K in beech trees of French forests. Journal of Environmental Radioactivity. 2024;277:107450.

[10] Yang H, Feng Q, Xu W, Tang Y, Bai G, Liu Y, et al. Unraveling the nuclear isotope tapestry: Applications, challenges, and future horizons in a dynamic landscape. Eco-Environment & Health. 2024.

[11] Ludovici GM, Chierici A, de Souza SO, d'Errico F, Iannotti A, Malizia A. Effects of ionizing radiation on flora ten years after the Fukushima Dai-ichi disaster. Plants. 2022;11(2):222.

[12] Mothersill C, Rusin A, Elliott A, Seymour C. Radiation and chemical induced genomic instability as a driver for environmental evolution. In: Genome Stability. Elsevier; 2021. p. 639–58.

[13] West A, Tsitsimpelis I, Licata M, Jazbec A, Snoj L, Joyce MJ, et al. Use of Gaussian process regression for radiation mapping of a nuclear reactor with a mobile robot. Scientific reports. 2021;11(1):13975.

[14] Li CY, Xia XB, Cai J, Zhang ZH, Zhang GQ, Wang JH, et al. Radiation dose distribution of liquid fueled thorium molten salt reactor. Nuclear Science and Techniques. 2021;32(2):22.

[15] Cao P, Hao C, Ma C, Yang H, Sun R. Physical field simulation of the ultrasonic radiation method: An investigation of the vessel, probe position and power. Ultrasonics Sonochemistry. 2021;76:105626.

[16] Boice JD, Cohen SS, Mumma MT, Hagemeyer DA, Chen H, Golden AP, et al. Mortality from leukemia, cancer and heart disease among US nuclear power plant workers, 1957–2011. In: The Million Person Study of Low-Dose Radiation Health Effects. CRC Press; 2024. p. 141–62.

[17] Ohba T, Tanigawa K, Liutsko L. Evacuation after a nuclear accident: Critical reviews of past nuclear accidents and proposal for future planning. Environment international. 2021;148:106379.

[18] Kronthaler F, Zöllner S. Data analysis with RStudio. Data Analysis with RStudio. 2021.

[19] Labunska I, Levchuk S, Kashparov V, Holiaka D, Yoschenko L, Santillo D, et al. Current radiological situation in areas of Ukraine contaminated by the Chornobyl accident: Part 2. Strontium-90 transfer to culinary grains and forest woods from soils of Ivankiv district. Environment International. 2021;146:106282.

[20] Ahmed RS. A review on soil radionuclide distribution in Iraq analysed using gamma ray spectroscopy. Environmental Forensics. 2021;22(1–2):91–8.

[21] Zhu L. Radiochemical analysis of hard-to-measure radionuclides (135Cs and 126Sn) for decommissioning and environmental tracer studies. 2021.

[22] Qin H, Gong R, Liu X, Shen M, Wei Z, Yu F, et al. Forward and backward information retention for accurate binary neural networks. In 2020. p. 2250–9.

[23] Gonoskov A, Blackburn T, Marklund M, Bulanov S. Charged particle motion and radiation in strong electromagnetic fields. Reviews of Modern Physics. 2022;94(4):045001.

[24] Karmaker N, Maraz KM, Islam F, Haque MM, Razzak M, Mollah M, et al. Fundamental characteristics and application of radiation. GSC Advanced Research and Reviews. 2021;7(1):064–72.

[25] Kao S. A robust standard deviation control chart based on square A estimator. Quality and Reliability Engineering International. 2022;38(5):2715–30.

[26] Lee SW. Regression analysis for continuous independent variables in medical research: statistical standard and guideline of Life Cycle Committee. Life cycle. 2022;2.

[27] Takahashi H, Yasumura S, Takahashi K, Ohira T, Shimura H, Ohto H, et al. Detection of thyroid cancer among children and adolescents in Fukushima, Japan: a population-based cohort study of the Fukushima Health Management Survey. EClinicalMedicine. 2024;75.

[28] Braga A da S, Cordeiro GM, Ortega EM, Silva GO, Vasconcelos JC. A random-effects regression model based on the odd log-logistic skew normal distribution. Journal of Statistical Theory and Practice. 2022;16(2):33.

[29] Alhassan S, Usman A. Fit A Poisson Regression, Negative Binomial Regression And Poisson Log-Normal Regression Models: An Application To Road Accident Data In North Central Nigeria. International Journal of Modeling and Applied Science Research. 2024;

[30] Di Mari R, Ingrassia S, Punzo A. Local and overall deviance r-squared measures for mixtures of generalized linear models. Journal of Classification. 2023;40(2):233–66.

[31] Lecca P, Carpentieri B. Regression. In: Introduction to Mathematics for Computational Biology. Springer; 2023. p. 197–232.

[32] Kalnins A, Praitis Hill K. Additional caution regarding rules of thumb for variance inflation factors: extending O'Brien to the context of specification error. Quality & Quantity. 2024;1–24.

[33] Gokmen S, Dagalp R, Kilickaplan S. Multicollinearity in measurement error models. Communications in Statistics-Theory and Methods. 2022;51(2):474–85.

[34] Ebrahim EA, Cengiz MA, Terzi E. The Best Fit Bayesian Hierarchical Generalized Linear Model Selection Using Information Complexity Criteria in the MCMC Approach. Journal of Mathematics. 2024;2024(1):1459524.

[35] Ramos B, Menéndez T, Nolivos I. Statistical modeling of the climatic influence on a 5 GHz microwave link: A tropical weather case empirical study. IEEE Access. 2024

[36] Ailobhio D, Ikughur J. A Review of Some Goodness-of-Fit Tests for Logistic Regression Model.

[37] KANG YT, Him NC. The Poisson Regression and Quasi-Poisson Regression Analysis on FIFA World Cup Games. Enhanced Knowledge in Sciences and Technology. 2024;4(2):349–58.

[38] Gewers FL, Ferreira GR, Arruda HFD, Silva FN, Comin CH, Amancio DR, et al. Principal component analysis: A natural

approach to data exploration. ACM Computing Surveys (CSUR). 2021;54(4):1–34.

[39] Picci G. Principal Component Analysis. In: An Introduction to Statistical Data Science: Theory and Models. Springer; 2024. p. 273–305.

[40] Saha S. Prevention of Empty Clusters and Incomplete Data Problems using Modified K-Means and Gaussian Mixture Model. 2023.

[41] Nigro L, Cicirelli F. Improving Clustering Accuracy of K-Means and Random Swap by an Evolutionary Technique Based on Careful Seeding. Algorithms. 2023;16(12):572.