Journal of Ecology of Health & Environment
An International Journal

http://dx.doi.org/10.18576/jehe/140102

Surface Solar Radiation and Occupational Radiation Exposure Analysis in Nigeria Using Random Forest: Implications for Renewable Energy Planning and Radiological Safety

Emmanuel Yohanna[1,], Amaechi Onele Azi[2] and Jasini Waida[1]*

[1] Department of Physics, Kashim Ibrahim University, Maiduguri, Borno State, Nigeria
[2] Department of Physics, Federal Polytechnic of Oil and Gas, Bonny, Rivers State, Nigeria

Received: 1 Sep. 2025, Revised: 20 Oct. 2025, Accepted: 25 Nov. 2025.
Published online: 1 Jan. 2026

Abstract: In this study, Surface solar radiation downward (SSRD) estimation and occupational radiation monitoring, both important for environmental planning and radiation protection, are jointly evaluated despite often being assessed separately. Monthly mean SSRD and meteorological variables for 37 stations (2010-2020) were obtained from the NASA Prediction of Worldwide Energy Resources (POWER) database, while occupational annual effective dose records (2012-2016) were obtained from Nigerian Nuclear Regulatory Authority (NNRA) monitoring reports. Random Forest (RF) regression was applied and evaluated against linear regression (LR) using stratified train-test splitting and five-fold cross-validation to analyze the spatial and temporal variability of SSRD across Nigeria's major climatic zones and assess the annual effective dose for regulated worker categories using a machine-learning modeling approach. SSRD showed spatial variability of higher mean in the Sahel (6.5 ± 0.74 kWh m^{-2} day^{-1}), moderate mean in the Savannah (5.8 ± 0.68 kWh m^{-2} day^{-1}), and lower mean values along the Guinea Coast (5.2 ± 0.62 kWh m^{-2} day^{-1}). RF achieved an improved performance for SSRD (R^2 = 0.761, RMSE = 0.420 kWh m^{-2} day^{-1}, MAE = 0.334 kWh m^{-2} day^{-1}) relative to LR. Although annual effective doses were low overall, higher mean doses were observed in Nuclear Medicine and Radiological Research. RF predictive modelling outperformed LR, with a stronger dose prediction of R^2 = 0.872. The findings suggest that predictive modeling may support environmental radiation assessment and occupational dose monitoring by improving the understanding of spatial and temporal variability in radiation-related conditions, especially in settings with limited observational data.

Keywords: Machine learning; Nigeria; Radiation protection; Random Forest; Solar radiation.

1. Introduction

Surface solar radiation downward (SSRD) is a measure of incoming shortwave solar energy that reaches the Earth's surface and is crucial for environmental and energy-related processes[1] [2], [3] The variations in SSRD affect land-atmosphere energy transfer, hydrological conditions, agricultural yield, and the efficacy of solar energy systems, rendering accurate estimation crucial for environmental evaluation and energy strategizing [4], [5]. Although in tropical regions, SSRD is influenced by the interplay of atmospheric factors, notably cloud cover, humidity, and seasonal circulation, which determine the quantity of solar radiation reaching the surface [6], [7], [8]. Nigeria exhibits significant climate variations that are expected to affect SSRD variability spatially and temporally [9], [10]. The southern Guinea Coast is marked by constant humidity and cloudiness, the Savannah serves as a transitional climatic zone, whilst the northern Sahel endures arid conditions and diminished atmospheric moisture [11], [12], [13]. These disparities provide significant gradients in solar radiation availability, affecting agriculture, water balance, and solar energy planning. The significant climatic heterogeneity among places poses obstacles for consistent SSRD prediction, especially where meteorological factors interact in intricate manners.

Estimating SSRD remains challenging as solar radiation is influenced by interacting meteorological variables that vary across space and time. Conventional estimation approaches have applied empirical equations, statistical models, and satellite or reanalysis products, each offering useful approximations under specific climatic conditions [14], [15], [16]. However, model performance may vary where atmospheric controls such as humidity, temperature, and cloudiness interact in nonlinear ways, particularly across regions with contrasting climatic regimes [17]. In a geographical setting such as Nigeria, where climatic gradients are pronounced, linear relationships may not fully represent the complexity of SSRD variability [18], [19]. Machine learning approaches have increasingly been explored to address these challenges by allowing flexible representation of interacting predictors. Random Forest (RF), an ensemble learning method, is often applied in environmental prediction to accommodate nonlinear relationships and variable interactions without strong assumptions about data structure. In this context [20], Linear regression (LR) was included as a benchmark model to provide a transparent comparison and assess the performance of the nonlinear modeling SSRD prediction

*Corresponding author e-mail: yohanna.emmanuel@kiu.edu.ng

under varying climatic conditions.

Occupational radiation monitoring, a component of radiation protection practice, is particularly important for workers involved in medical, industrial, agricultural, and research applications of ionizing radiation [21], [22], [23]. Meanwhile, the annual effective dose is widely used as a monitoring metric because it provides a standardized measure of cumulative occupational exposure and supports routine monitoring under regulatory frameworks [24], [25]. Although occupational doses are generally maintained within recommended exposure limits, periodic assessment remains important for identifying exposure patterns across worker categories and time. While environmental solar radiation and occupational ionizing radiation originate from different physical sources, both involve variability across space and time and require reliable approaches for monitoring and prediction under uncertain conditions. Predictive methods may therefore support environmental characterization as well as occupational dose assessment to identify patterns that are not easily captured through conventional approaches alone [26]. Despite growing use of predictive methods, national studies in Nigeria remain scarce in climatic coverage and integration of environmental and occupational radiation-related assessment perspectives.

This study, therefore, aimed to characterize SSRD variability across Nigeria's major climatic zones and assess occupational annual effective dose across regulated worker categories using RF regression. LR was included as a benchmark model to evaluate differences in predictive performance, while cross-validation was applied to assess uncertainty in model stability. By combining SSRD prediction with occupational radiation exposure analysis within a common predictive framework, the study provides a broader perspective on environmental and radiation-related monitoring that may support assessment and planning in settings with limited observational data.

2. Materials and methods

2.1 Study Area and Climatic Zoning

Nigeria spans several climatic conditions, ranging from humid coastal environments in the south to semi-arid conditions in the north. For this study, the country was grouped into three major climatic zones: the Guinea Coast, the Savannah, and the Sahel, as shown in Fig. 1. The Guinea Coast is characterized by high humidity and rainfall, the Savannah by alternating wet and dry seasons, and the Sahel by relatively dry conditions [27]. Meteorological stations were categorized based on these climatic zones to enhance the regional comparison of solar radiation and environmental variability throughout Nigeria. This zoning method aligns to widely utilized climatic classifications in environmental and meteorological research in West Africa. The climatic zone variable was incorporated as a categorical predictor in the modelling framework to address regional climatic variations influencing surface solar radiation.

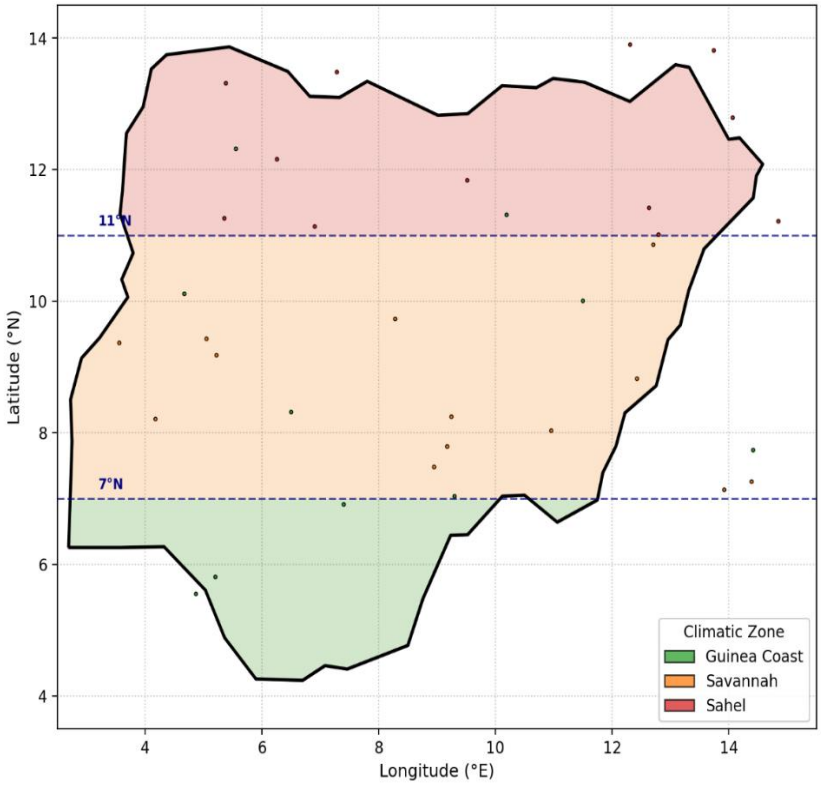

Fig. 1: Map of Nigeria showing the three climatic zones and the stations from NASA Power data

2.2 Data sources

2.2.1 Solar radiation and meteorological variables

SSRD and meteorological variables were obtained from the NASA Prediction of Worldwide Energy Resources (POWER) database for 37 stations across Nigeria between 2010 and 2020. Data were accessed through the NASA POWER Access Viewer (https://power.larc.nasa.gov/data-access-viewer/). To ensure consistency in temporal comparison across stations and climatic zones, we extracted monthly mean values for all variables. The variables extracted included SSRD (kWh m^{-2} day^{-1}), 2m air temperature (°C), relative humidity (%), cloudiness index (dimensionless), and sunshine duration (h day^{-1}). SSRD was utilized as the prediction target in the models, while the meteorological variables were used as predictors. Monthly mean values and climatic zone statistics were derived from the station records following quality screening procedures. A summary of the variables, units, temporal coverage, and derived fields is presented in Table 1.

Table 1: Data sources, variables, units, temporal coverage, and derived fields used in modeling.

Source	Variable	Unit	Temporal Coverage	Spatial Coverage	Role in Study	Derived Fields
NASA POWER	Surface Solar Radiation Downward (SSRD)	kWh m^{-2} day^{-1}	2010-2020	37 stations-Nigeria	Prediction target-SSRD model	Monthly mean; zone stats; IQR-screened series
	2 m Air Temperature	°C	2010-2020	37 stations-Nigeria	Predictor-SSRD model	Monthly mean; IQR-screened
	Relative Humidity	%	2010-2020	37 stations-Nigeria	Predictor-SSRD model	Monthly mean; IQR-screened
	Cloudiness Index		2010-2020	37 stations-Nigeria	Predictor-SSRD model	Monthly mean; IQR-screened
	Sunshine Duration	h day^{-1}	2010-2020	37 stations-Nigeria	Predictor-SSRD model	Monthly mean; IQR-screened
NNRA	Annual Occupational Effective Dose	mSv y^{-1}	2012-2016	National (8 occupational groups)	Prediction target-dose model	Group statistics; time trend (β)
	Occupational Group	categorical	2012-2016	National (8 occupational groups)	Predictor-dose model	Label-encoded integer for RF model

2.2.2 Occupational annual effective dose

Annual occupational effective dose data were obtained from the Nigerian Nuclear Regulatory Authority (NNRA) reports for the period 2012-2016. The dataset included eight occupational groups: industrial radiography, diagnostic radiology, the general monitored workforce category, radiotherapy, nuclear medicine, agricultural clinics, radiological research, and general users. The Occupational effective dose was used as the prediction target for the dose model, while the occupational group was used as the predictor. Occupational categories were encoded as integer labels for the model implementation. Dose measurements were derived from accredited personal dosimetry monitoring systems routinely used for occupational radiation protection, and annual effective dose records were compiled for statistical analysis [28].

2.3 Preprocessing and quality assurance

2.3.1 Outlier Screening of Solar Variables

Monthly mean solar radiation and meteorological variables were cleaned to identify outliers that could influence model performance. Data quality assessment was performed using the interquartile range (IQR) method. For a variable x, The IQR was calculated as presented in Eq. 1:

$$IQR = Q_3 - Q_1 \quad (1)$$

where Q_1 and Q_3 represent the first and third quartiles, respectively.

Potential outliers were identified using Eq. 2 and Eq. 3:

$$x < Q_1 - 1.5(IQR) \quad (2)$$

$$x > Q_3 + 1.5(IQR) \quad (3)$$

Outliers and missing observation rows were also removed during data preparation, and only complete monthly records were retained for analysis during data preprocessing to minimize the influence of anomalous observations on model fitting. Data normalization or scaling was not applied, as RF models tend to be insensitive to differences in variable scale.

2.3.2 Occupational Dose Screening

Similarly, data cleaning was performed for annual occupational effective dose records within the context of internationally recommended occupational exposure limits, where the annual effective dose limit for radiation workers is commonly expressed as an average of 20 mSv y^{-1} over

defined periods to ensure consistency with occupational radiation protection standards [29]. Dose records were screened for inconsistencies, extreme values, and incomplete entries before statistical analysis and model development. Only verified annual records were included in the final dataset.

2.4 Random Forest regression model

RF regression was applied to model SSRD and occupational annual effective dose. RF is an ensemble learning method that combines multiple decision trees trained on bootstrap samples with random feature selection at each split to improve prediction accuracy and reduce overfitting [30]. For a predictor vector x, and the RF prediction is expressed in Eq. 4 as:

$$\hat{y}(x) = \frac{1}{B}\sum_{b=1}^{B} T_b\,(x) \tag{4}$$

where B is the number of decision trees and $T_b(x)$ is the prediction from the b^{th} regression tree.

For the SSRD model, predictor variables included 2m air temperature, relative humidity, cloudiness index, sunshine duration, and climatic zone. Meanwhile, the climatic zone was presented as a variable to account for regional climatic differences across Nigeria, but for occupational dose prediction, annual time and occupational group were used as predictors. Linear regression (LR) was applied as a benchmark for model performance, it was evaluated using the same predictor variables. This comparison was included to assess the performance of the nonlinear structure of RF in improving prediction performance relative to a conventional linear approach. The LR model is expressed in Eq. 5 as:

$$y = \beta_0 + \sum_{i=1}^{n} \beta_i\, x_i + \varepsilon \tag{5}$$

where y represents the predicted response variable, for SSRD or occupational effective dose, β_0 is the intercept, β_i represents the regression coefficients for predictor variables x_i, and ε is the error term.

The RF models were implemented in Python using the scikit-learn library. Hyperparameters were selected and optimized for the predictive performance and overfitting control. The final model configuration included 100 trees ($n_{estimators} = 100$), a maximum tree depth of 10 ($max_{depth} = 10$), a minimum of five samples per leaf ($min_{samples\_leaf} = 5$), and bootstrap sampling is enabled. A fixed random seed ($random\_state = 42$) was used to improve the reproducibility of model training and evaluation.

2.5 Model Validation and Performance Metrics

Model performance was evaluated using a stratified train-test split, where 70% of the data were used for training and 30% for testing. Stratification was applied to preserve climatic zone representation for SSRD modeling and occupational group representation for dose modeling [31]. Model robustness was assessed using five-fold cross-validation. The prediction performance was evaluated using the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). The coefficient of determination was calculated using Eq. 6, given as:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{6}$$

where y_i represents observed values, $\hat{y}_i$ predicted values, and $\bar{y}$ the mean observed value.

The RMSE was calculated using Eq. 7, and the MAE in Eq. 8 was presented as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{7}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{8}$$

were these metrics used to evaluate prediction accuracy and model error in physical units. Additionally, the uncertainty in model performance was evaluated through variability among cross-validation folds and presented as mean ± standard deviation. Feature relevance was assessed using the mean decrease in impurity approach employed in RF to prioritize predictors based on their effectiveness in minimizing prediction error. The mean values were interpolated using spline interpolation to create continuous spatial surfaces across Nigeria for the assessment of spatial patterns in SSRD. Interpolation uncertainty was evaluated qualitatively by analyzing spatial consistency with observed station distributions and climatic zone configurations.

2.6 Statistical Group Comparisons

One-way analysis of variance (ANOVA) was applied to examine the differences in SSRD among climatic zones and occupational effective dose among occupational groups. ANOVA was applied to determine whether group means differed significantly and is expressed in Eq. 9 as:

$$F = \frac{\text{between-group mean square}}{\text{within-group mean square}} \tag{9}$$

where larger F-values indicate stronger separation among groups.

Furthermore, after the significant differences were detected, the Tukey's honestly significant difference (HSD) test was applied for pairwise group comparisons and statistical significance was evaluated at $\alpha = 0.05$, and ANOVA results were reported using the corresponding F-statistics and p-values as presented in the Results and Discussion sections.

3. Results and Discussion

3.1 Temporal Variation of Surface Solar Radiation Downward (SSRD) in Nigeria

The national monthly SSRD demonstrated a distinct and consistent seasonal pattern from 2010 to 2020, characterized by minimal interannual fluctuations, as shown in Fig. 2. The monthly mean SSRD showed a general increase from the beginning of the year and increased between April and June,

and subsequently decreased in the following months. However, monthly climatology revealed mean SSRD values between approximately 5.0 and 6.5 kWh m^{-2} day^{-1}, signifying persistent seasonal variability throughout the study period. The temporal pattern illustrates seasonal variations in atmospheric conditions throughout Nigeria. The observed high SSRD values in the dry and early wet seasons are likely associated with less cloud cover and enhanced air transmissivity, resulting in increased solar energy reaching the surface. Conversely, decreased SSRD during moist months tends to correlate with heightened humidity and cloud formation, especially in the southern areas. Notwithstanding slight annual variations, the overarching temporal trend remained consistent, suggesting a continual seasonal impact on solar radiation availability nationwide.

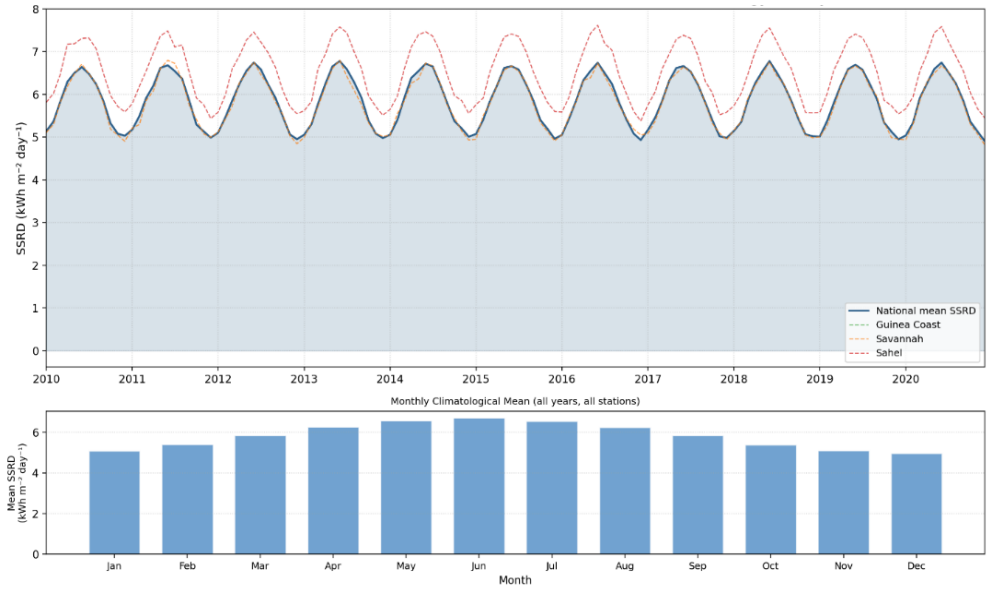

Fig. 2: National SSRD time series (2010-2020) with seasonal cycle shown by monthly climatology overlay.

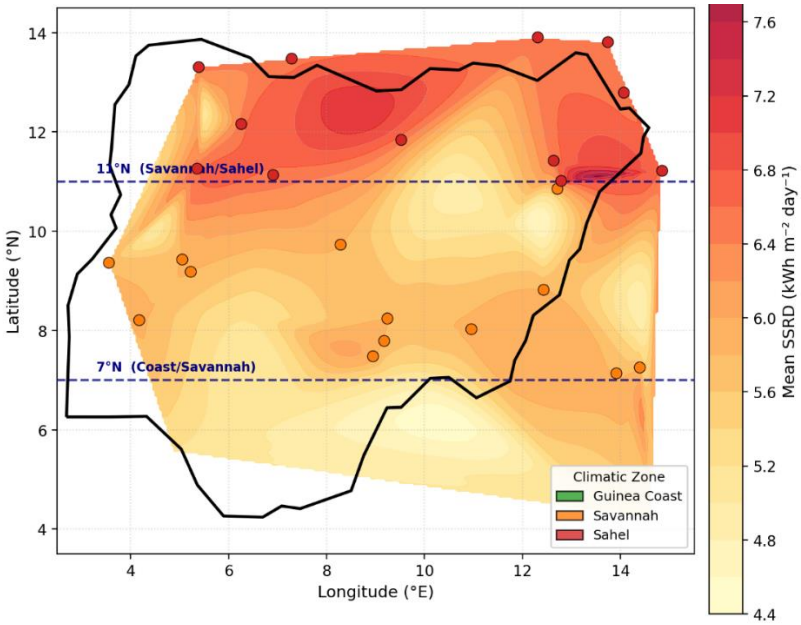

Fig. 3: Spatial interpolation/map of mean SSRD (kWh m^{-2} day^{-1}) with zone boundaries.

3.2 Spatial Distribution of Mean SSRD Across Climatic Zones

The mean SSRD showed a clear spatial gradient across Nigeria, increasing from the humid southern regions toward the drier northern areas, as illustrated in Fig. 3. Lower values were concentrated along the Guinea Coast, with mean SSRD around 5.2 kWh m^{-2} day^{-1}, while higher values occurred in the Sahel, where mean SSRD approached 6.5 kWh m^{-2} day^{-1}. The Savannah displayed intermediate values of about 5.8 kWh m^{-2} day^{-1}, reflecting its transitional climatic setting between the humid south and semi-arid north. The observed

pattern likely corresponds closely with climatic zoning and supported by significant inter-zonal differences in SSRD with ($F(2,4881) = 2.8 \times 10^3$, $p < 0.001$)). The observed higher solar radiation in northern regions may reflect lower atmospheric moisture and reduced cloud cover, while the comparatively lower values in the south are consistent with persistent humidity and cloud formation. Although spline interpolation provides a smooth, continuous surface, it represents an approximation to the underlying field and may reflect sampling density and station geometry.

3.3 Performance of Random Forest and Linear Regression Models for SSRD Prediction

The RF model showed improved predictive performance compared with LR for SSRD estimation, as presented in Table 2. RF achieved an R^2 of 0.761, with a RMSE of 0.420 kWh m^{-2} day^{-1} and a MAE of 0.334 kWh m^{-2} day^{-1}. In comparison, LR achieved a lower predictive accuracy, with an R^2 of 0.469, RMSE of 0.627 kWh m^{-2} day^{-1}, and MAE of 0.502 kWh m^{-2} day^{-1}. The predicted versus observed relationships further highlighted differences between both models as illustrated in Fig. 4. RF predictions clustered more closely around the 1:1 reference line in Fig. 4a, indicating stronger agreement with observed SSRD values. In contrast, the LR showed high dispersion and a wider spread of prediction error in Fig. 4b. The improved RF performance suggests that nonlinear interactions among meteorological variables, including humidity, temperature, and cloudiness, likely contributed to SSRD variability across Nigeria's climatic zones.

Table 2: Performance comparison of RF and LR models for SSRD prediction.

Model	R^2	RMSE	MAE
LR	0.469	0.627	0.502
RF	0.761	0.420	0.334

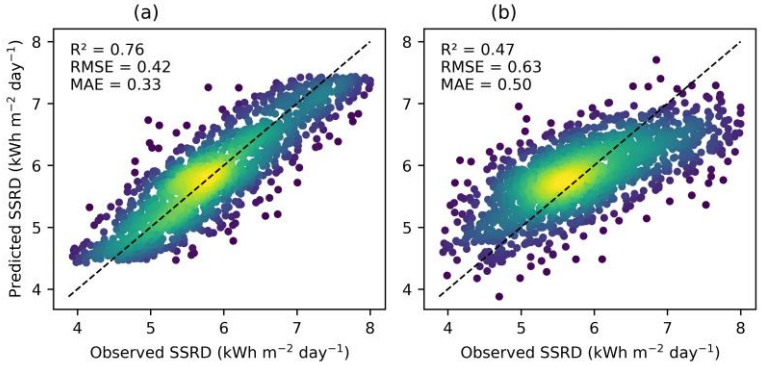

Fig. 4: Predicted versus observed SSRD for (a) RF and (b) LR models.

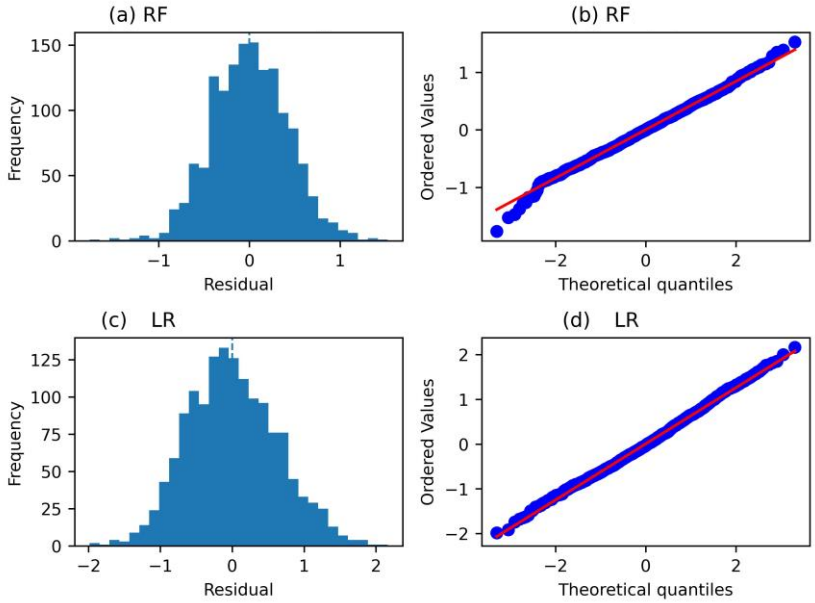

Fig. 5: Residual diagnostics for SSRD prediction models: (a) RF histogram, (b) RF Q-Q plot, (c) LR histogram, and (d) LR Q-Q plot.

3.4 Residual Behavior and Model Reliability

Fig. 5 presents the residual diagnostics, which provided additional insight into the reliability of the RF and LR models for SSRD prediction. RF residuals in Fig. 5a were observed to be more scarcely distributed around zero, indicating lower prediction error and more stable performance across observed SSRD values. In contrast to LR in Fig. 5c, the residuals showed a wider spread, suggesting greater variability in the prediction error and reduced agreement with observations. The Q-Q plots further illustrated that residuals from both models likely followed the theoretical normal distribution reasonably well, although minor departures were observed at the tails. These deviations were more noticeable in the LR model in Fig. 5d, reflecting larger residual extremes relative to RF in Fig. 5b. The narrower RF residual distribution and closer alignment with the reference line support the improved predictive performance observed in Section 3.3. However, some residual variability may remain in both models, indicating that SSRD variability may also be influenced by atmospheric or local factors not fully represented by the selected predictors.

3.5 Influence of Meteorological Predictors on SSRD

RF feature importance showed climatic zone predictor of SSRD, contributing 43.2% of the overall model importance, followed by humidity with 27.2% and temperature (23.6%), as presented in Fig. 6. Cloudiness index (3.8%) and sunshine duration (2.2%) contributed comparatively less to model prediction. This dominant influence of climatic zone suggests that large-scale regional climate conditions shaped SSRD variability across Nigeria. Humidity and temperature also played important roles, likely reflecting their influence on cloud development, atmospheric moisture, and seasonal circulation patterns [32]. The LR model showed broadly consistent relationships as presented in Table 3. Temperature and climatic zone were positively associated with SSRD, while humidity and cloudiness showed negative associations, with statistically significant coefficients of $p < 0.001$. Sunshine duration was observed to be statistically significant with $p = 0.403$, suggesting a weaker linear contribution. The differences between RF importance and LR coefficients suggest that nonlinear interactions among climatic variables likely influenced SSRD variability.

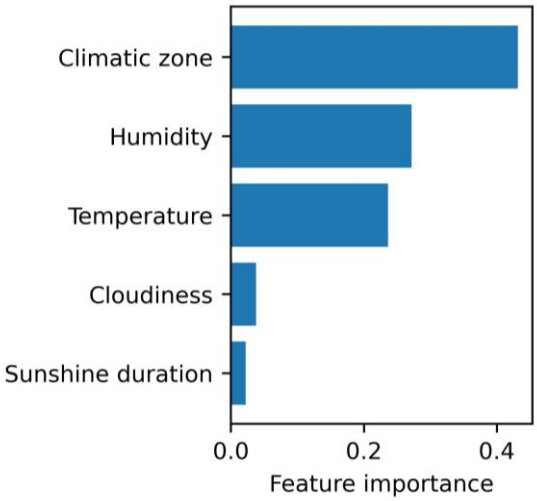

Fig. 6: Random Forest feature importance for SSRD predictors.

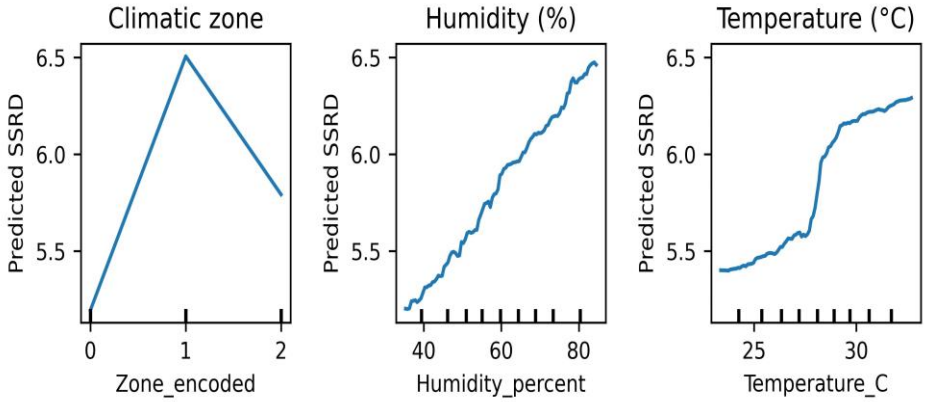

Fig. 7: Partial dependence plots showing the effects of climatic zone, humidity, and temperature on SSRD prediction

Table 3: Linear regression coefficients and statistical significance for SSRD predictors

Predictor	Coefficient	p_value
Temperature	0.1766	0
Humidity	-0.0130	0
Cloudiness	-1.6946	0
Sunshine duration	-0.0160	0.403
Climatic zone	0.1060	0

3.6 Nonlinear Relationships Between Climatic Variables and SSRD

Fig. 7 illustrates nonlinear relationships between specific climatic variables and SSRD, demonstrating the efficacy of RF in comparison to LR. Estimated SSRD showed significant variation across climatic zones, which suggests that regional climatic conditions influenced disparities in solar radiation patterns. The response of climatic zones further supports the importance ranking identified in Section 3.5, wherein zone classification appears to be the most significant predictor. Humidity showed a progressive nonlinear increment in predicted SSRD throughout the observed spectrum, indicating that its effect was not entirely captured by a straightforward linear correlation. Temperature displayed a threshold pattern, characterized by a pronounced increase in predicted SSRD around 28 °C, achieved by a more gradual ascent at elevated temperatures. These patterns suggest that climatic effects on SSRD were inconsistent across varying environmental conditions [32]. The observed nonlinear responses further substantiate the enhanced predictive efficacy of RF, especially in contexts where meteorological variables interact across Nigeria's climatic regions.

3.7 Trends in Occupational Annual Effective Dose (2012-2016)

The occupational annual effective dose varied across worker categories between 2012 and 2016, with noticeable differences in both magnitude and temporal pattern as shown in Fig. 8. Nuclear Medicine and Radiological Research consistently recorded the highest annual doses in Fig. 8a, ranging between 1.8 and 2.1 mSv y^{-1}, whereas lower values were observed in groups such as agricultural clinics and diagnostic radiology. Although dose levels fluctuated slightly over time, most occupational categories showed relatively stable exposure patterns. Additionally, the national mean annual dose showed a gradual increase from approximately 0.92 mSv y^{-1} in 2012 to a peak near 1.12 mSv y^{-1} in 2015, followed by a small decline in 2016. Significant differences among occupational groups were observed in Fig. 8b with ($F(7,32) = 1.3 \times 10^2$) and $p < 0.001$. Despite inter-group variation, annual effective doses remained substantially below the occupational exposure reference level of 20 mSv y^{-1} [33], suggesting dose levels consistent with routine occupational monitoring expectations.

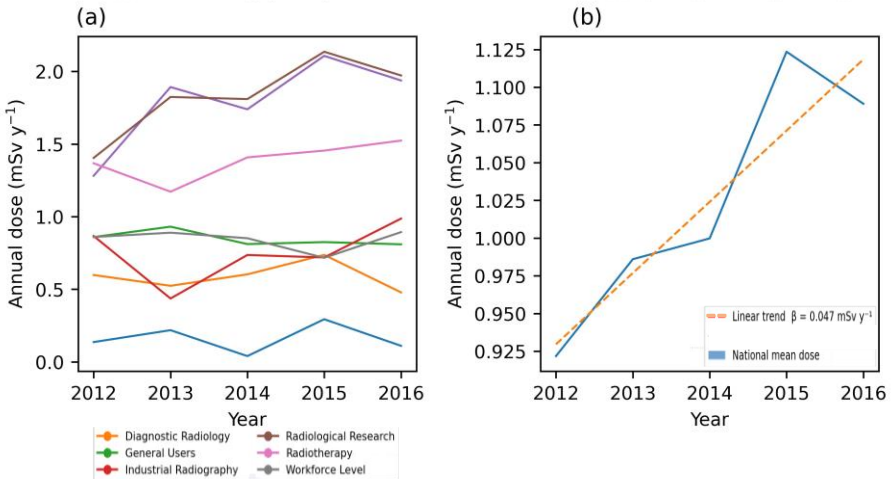

Fig. 8: Occupational annual effective dose trends by group and national mean (2012-2016) (a) mean annual dose by group, (b) National overall trend.

Table 4: Descriptive statistics of occupational annual effective dose by occupational group

Occupational Group	N	Mean (mSv y^{-1})	SD	Min	Median	Max
Agricultural Clinics	5	0.16	0.10	0.04	0.14	0.29
Diagnostic Radiology	5	0.59	0.10	0.48	0.60	0.74
General Users	5	0.85	0.05	0.81	0.82	0.93
Industrial Radiography	5	0.75	0.21	0.44	0.74	0.99
Nuclear Medicine	5	1.79	0.31	1.28	1.89	2.11
Radiological Research	5	1.83	0.27	1.40	1.82	2.14

Occupational Group	N	Mean (mSv y^{-1})	SD	Min	Median	Max
Radiotherapy	5	1.39	0.13	1.17	1.41	1.52
Workforce Level	5	0.84	0.07	0.72	0.86	0.89
Overall	**40**	**1.02**	**0.58**	**0.04**	**0.86**	**2.14**

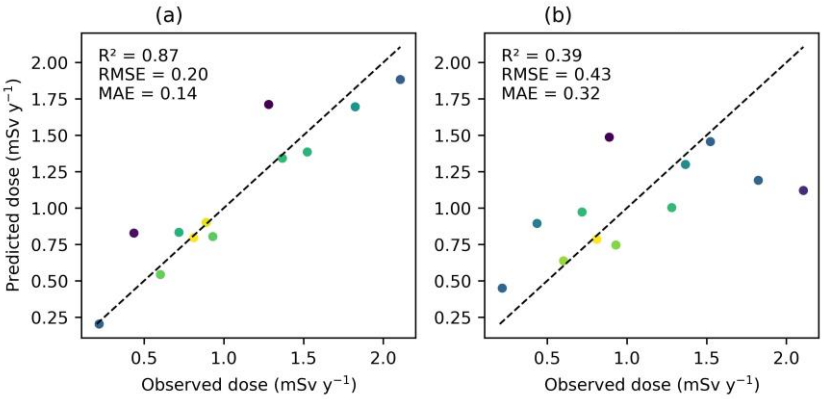

Fig. 9: Predicted versus observed occupational annual effective dose for (a) RF and (b) LR models.

The descriptive statistics of occupational annual effective dose presented in Table 4 further show differences across occupational groups. Higher mean annual doses were observed in Radiological Research (1.83 mSv y^{-1}) and Nuclear Medicine (1.79 mSv y^{-1}), while lower values occurred in Agricultural Clinics (0.16 mSv y^{-1}) and Diagnostic Radiology (0.59 mSv y^{-1}). Intermediate dose levels were recorded for Radiotherapy (1.39 mSv y^{-1}), General Users (0.85 mSv y^{-1}), Workforce Level (0.84 mSv y^{-1}), and Industrial Radiography (0.75 mSv y^{-1}). Despite these differences, the overall mean annual effective dose remained low (1.02 mSv y^{-1}) relative to international occupational exposure reference levels. The observed variation across worker categories may reflect differences in work procedures, source proximity, shielding conditions, and routine exposure patterns. These findings suggest that occupational category remains an important consideration in dose monitoring and optimization of radiation protection practices.

3.8 Occupational Dose Prediction Performance

Fig. 9 shows RF improved predictive performance compared with LR for occupational annual effective dose estimation. RF achieved an R^2 of 0.87, with a root RMSE of 0.20 mSv and a MAE of 0.14 mSv. In comparison, LR achieved a lower predictive accuracy, with an R^2 of 0.39, RMSE of 0.43 mSv, and MAE of 0.32 mSv. The predicted and observed relationships further highlighted differences between the two models as illustrated in Fig. 9. RF predictions aligned more closely with the 1:1 reference line, indicating an agreement with measured dose values, whereas LR showed greater dispersion, especially at higher dose levels. The improved RF performance suggests that occupational dose variability may involve nonlinear interactions between occupational category and time. However, the analysis was based on a five-year dataset and a limited set of occupational groups, which may limit the generalizability of the findings to other time periods or occupations.

3.9 Model Robustness and Uncertainty Assessment

Five-fold cross-validation was applied to assess model stability and uncertainty in predictive performance as presented in Table 5. In SSRD prediction, RF yielded a superior mean R^2 of (0.767 ± 0.010), compared to LR with (0.437 ± 0.025), signifying enhanced and more reliable predictive efficacy across folds. The comparatively low standard deviation noted for RF suggests restricted variability in model performance across various training partitions. Additionally, A comparable trend was noted in the prediction of occupational doses. RF attained a mean R^2 of 0.898 ± 0.034, while LR showed lower performance and increased variability (0.323 ± 0.110). The increased uncertainty noted in occupational dose prediction likely indicates the limited dataset and reduced temporal span. The cross-validation results suggest that RF achieved more consistent performance than LR, although predictive uncertainty persisted in both applications.

Table 5: Cross-validation uncertainty of RF and LR models

Dataset	Model	Mean (R^2)	SD (R^2)
SSRD	RF	0.767	0.010
	LR	0.437	0.025
Occupational Dose	RF	0.898	0.034
	LR	0.323	0.110

3.10 Implications for Environmental Monitoring and Radiation Protection

The findings illustrate the significance of data-driven modelling approaches for environmental and radiological evaluation in Nigeria. The SSRD prediction can provide support in solar resource estimation, environmental planning, and regional energy assessments by providing an understanding of climatic influences on surface radiation variability. Simultaneously, occupational dose estimation offers an attainable framework for tracking exposure trends among worker classifications. Despite originating from different sources, both solar radiation and occupational ionizing radiation analyses illustrate how predictive methodologies can enhance environmental monitoring and refine exposure assessment.

4. Conclusions

This study applied a machine learning approach to estimate surface solar radiation downward (SSRD) across Nigeria's climatic zones and assess annual effective dose across regulated occupational worker categories. The results indicated different spatial and temporal variability in SSRD, with mean values in the Sahel (6.5 ± 0.74 kWh m^{-2} day^{-1}), moderate conditions in the Savannah (5.8 ± 0.68 kWh m^{-2} day^{-1}), and decreasing values along the Guinea Coast (5.2 ± 0.62 kWh m^{-2} day^{-1}). RF consistently outperformed LR when estimating both SSRD and occupational doses, suggesting that nonlinear climatic and occupational relationships contributed to variability.

Climatic zone, humidity, and temperature were observed as significant factors influencing SSRD variability, suggesting the influence of regional climatic conditions and meteorological variability across Nigeria's climatic zones. Occupational annual effective doses consistently remained at low levels, although relatively higher mean values were observed in Nuclear Medicine and Radiological Research. The combined analysis of environmental solar radiation and occupational doses offers complementary insights that may support radiation monitoring and environmental assessment across different operational settings, where predictive models can assist environmental radiation assessment and occupational dose monitoring in settings with limited observational data. These findings were, however, interpreted within the scope of the available datasets and temporal coverage. The absence of explicit aerosol indicators, such as Harmattan dust, and the relatively short duration of occupational monitoring represent important limitations of this study. Future research could improve predictive reliability by expanding spatial coverage, including atmospheric aerosol variables, and extending the duration of occupational dose monitoring.

References

[1] A. Modenese, L. Korpinen, and F. Gobba, "Solar Radiation Exposure and Outdoor Work: An Underestimated Occupational Risk," *International Journal of Environmental Research and Public Health 2018, Vol. 15, Page 2063*, vol. 15, no. 10, p. 2063, Sep. 2018, doi: 10.3390/IJERPH15102063.

[2] M. Yuan, T. Leirvik, and M. Wild, "Global Trends in Downward Surface Solar Radiation from Spatial Interpolated Ground Observations during 1961 2019," *J. Clim.*, vol. 34, no. 23, pp. 9501–9521, Dec. 2021, doi: 10.1175/JCLI-D-21-0165.1.

[3] K. Zhang, L. Zhao, W. Tang, K. Yang, and J. Wang, "Global and Regional Evaluation of the CERES Edition-4A Surface Solar Radiation and Its Uncertainty Quantification," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 15, pp. 2971–2985, 2022, doi: 10.1109/JSTARS.2022.3164471.

[4] N. Galli, M. Curioni, F. Capone, G. Manzolini, and M. C. Rulli, "Co-location of agriculture and solar energy from a global WEFE-Nexus perspective," *EGU25*, Mar. 2025, doi: 10.5194/EGUSPHERE-EGU25-18291.

[5] H. Supe, A. Abhishek, and R. Avtar, "Assessment of the solar energy–agriculture–water nexus in the expanding solar energy industry of India: An initiative for sustainable resource management," *Heliyon*, vol. 10, no. 1, p. e23125, Jan. 2024, doi: 10.1016/J.HELIYON.2023.E23125.

[6] G. Chesnoiu, I. Chiapello, N. Ferlay, P. Nabat, M. Mallet, and V. Riffault, "Regional modeling of surface solar radiation, aerosol, and cloud cover spatial variability and projections over northern France and Benelux," *Atmos. Chem. Phys.*, vol. 25, no. 2, pp. 1307–1331, Jan. 2025, doi: 10.5194/ACP-25-1307-2025.

[7] Q. Wang *et al.*, "Potential Driving Factors on Surface Solar Radiation Trends over China in Recent Years," *Remote Sensing 2021, Vol. 13, Page 704*, vol. 13, no. 4, p. 704, Feb. 2021, doi: 10.3390/RS13040704.

[8] O. J. Matthew, "Estimation of diurnal patterns of global solar radiation, temperature, relative humidity and wind speed from daily datasets at a humid tropical location," *Agric. For. Meteorol.*, vol. 322, p. 109003, Jul. 2022, doi: 10.1016/J.AGRFORMET.2022.109003.

[9] P. Collier, G. Conway, and T. Venables, "SPATIAL AND TEMPORAL VARIABILITY OF 40 YEARS TEMPERATURE AND PRECIPITATION IN THE SAVANNA REGION, NIGERIA," *FUDMA JOURNAL OF SCIENCES*, vol. 3, no. 3, pp. 1–11, 2019, doi: 10.1093/OXREP/GRN019.

[10] A. I. Agbonaye and E. S. Okonofua, "TRENDS AND SPATIAL VARIABILITY OF CLIMATE CHANGE IN NIGERIA'S COASTAL REGION," *Malaysian*

Journal of Civil Engineering, vol. 36, no. 2, pp. 19–32, Jul. 2024, doi: 10.11113/MJCE.V36.21861.

[11] K. A. Julius and R. A. Balogun, "Characteristics and Distribution of Some Radiation Parameters over Nigeria," *European Journal of Environment and Earth Sciences*, vol. 3, no. 4, pp. 32–40, Jul. 2022, doi: 10.24018/EJGEO.2022.3.4.255.

[12] E. Yohanna, W. Jasini, E. Y.-F. J. OF, and undefined 2025, "MAPPING PM2. 5 POLLUTION ZONES IN NIGERIA USING HIERARCHICAL CLUSTERING," *fjs.fudutsinma.edu.ngE Yohanna, W Jasini, EY YusufFUDMA JOURNAL OF SCIENCES, 2025•fjs.fudutsinma.edu.ng*, vol. 9, pp. 429–439, 2025, doi: 10.33003/fjs-2025-0912-4150.

[13] J. Waida *et al.*, "Transfer of K40, Ra226 and Th232 from soil to plants and water resulting from mining activities in Bassa, Plateau State, Nigeria (health implications on the," *researchgate.netJ Waida, U Rilwan, OO Galadima, E Omita, JM Sawuta, PE Ojike, R RebeccaJ. Eco. Heal. Env, 2022•researchgate.net*, vol. 10, no. 2, pp. 5–11, 2022, doi: 10.18576/jehe/100201.

[14] M. S. Okundamiya, J. O. Emagbetere, and E. A. Ogujor, "Evaluation of various global solar radiation models for Nigeria," *Int. J. Green Energy*, vol. 13, no. 5, pp. 505–512, Apr. 2016, doi: 10.1080/15435075.2014.968921.

[15] M. S. Adaramola, "Estimating global solar radiation using common meteorological data in Akure, Nigeria," *Renew. Energy*, vol. 47, pp. 38–44, Nov. 2012, doi: 10.1016/J.RENENE.2012.04.005.

[16] R. I. Maidment, D. I. F. Grimes, R. P. Allan, H. Greatrex, O. Rojas, and O. Leo, "Evaluation of satellite-based and model re-analysis rainfall estimates for Uganda," *Meteorological Applications*, vol. 20, no. 3, pp. 308–317, Sep. 2013, doi: 10.1002/MET.1283;PAGEGROUP:STRING:PUBLICATION.

[17] S. Usta, C. Gençoğlan, and S. Gençoğlan, "Estimation of Daily Average Global Solar Radiation with Nonlinear Regression Models Developed Using Some Meteorological and Geographical Parameters," *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, vol. 13, no. 3, pp. 589–597, Sep. 2022, doi: 10.24012/DUMF.1130793.

[18] M. S. Adaramola, "Estimating global solar radiation using common meteorological data in Akure, Nigeria," *Renew Energ*, vol. 47, pp. 38–44, Nov. 2012, doi: 10.1016/j.renene.2012.04.005.

[19] T. E. Patchali, O. M. Oyewola, O. O. Ajide, O. J. Matthew, T. A. O. Salau, and M. S. Adaramola, "Assessment of global solar radiation estimates across different regions of Togo, West Africa," *Meteorology and Atmospheric Physics 2022 134:2*, vol. 134, no. 2, pp. 26-, Feb. 2022, doi: 10.1007/S00703-021-00856-4.

[20] S. Sambyal, S. Shastri, A. K. Taloor, A. Kumar, and V. Mansotra, "Statistical and machine learning-based bias-correction of ERA5 forcing data: Implications for Himalayan glacier melt modelling," *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 143, p. 104387, Jun. 2026, doi: 10.1016/J.PCE.2026.104387.

[21] N. O. Eddy *et al.*, "Assessment of radiation protection of occupational workers in the radiotherapy facility at National Hospital Abuja, Nigeria," *ajol.info*, vol. 7, no. 6, pp. 1427–1432, Jun. 2025, doi: 10.1007/S42452-025-07069-Z.

[22] M. Dauda, G. Luntsi, N. Chigozie Ivor, P. Ogenyi, and L. Geofery, "Occupational Radiation Monitoring in Tertiary Health Institutions of Northwestern Nigeria," *researchgate.net*, vol. 3, no. 2, pp. 138–144, 2016, Accessed: Jun. 08, 2026. [Online]. Available: https://www.researchgate.net/profile/Prince-Ogenyi/publication/324605949_Occupational_Radiation_Monitoring_in_Tertiary_Health_Institutions_of_Northwestern_Nigeria/links/5ad8e57b458515c60f5a62f0/Occupational-Radiation-Monitoring-in-Tertiary-Health-Institutions-of-Northwestern-Nigeria.pdf

[23] P. Caiphus Rathebe, *Exposure Assessment and Health Effects Due to Ionizing and Non-Ionizing Radiation: An African Perspective*. 2025. Accessed: Jun. 08, 2026. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=b5J3EQAAQBAJ&oi=fnd&pg=PA1965&dq=Occupational+radiation+monitoring+is+an+important+component+of+radiation+protection+practice,+particularly+for+workers+involved+in+medical,+industrial,+agricultural,+and+research+applications+of+ionizing+radiation+in+Nigeria.+&ots=z_inwK-3ES&sig=qasm74iuuiD3KT_lE7havpF81ik

[24] P. Askounis *et al.*, "Exposure dose metrics, instrumentation, and operating procedures for occupational exposure assessment for nanomaterials and nanoparticles: a scoping review," *journals.co.za*, vol. 31, no. 1, pp. 12–24, Mar. 2025, doi: 10.62380/OHSA.2025.31.1.2.

[25] J. W. Hirshfeld *et al.*, "A holistic approach to assessment of population exposure to radiation: challenges and initiatives of a regulatory authority," *journals.lww.com*, vol. 71, no. 24, pp. e283–e351, Jun. 2018, doi: 10.1016/J.JACC.2018.02.016.

[26] T. Ruan, P. Li, H. Wang, T. Li, and G. Jiang, "Identification and Prioritization of Environmental Organic Pollutants: From an Analytical and Toxicological Perspective," *Chem. Rev.*, vol. 123, no. 17, pp. 10584–10640, Sep. 2023, doi: 10.1021/ACS.CHEMREV.3C00056.

[27] I. Diba, M. Camara, and A. B. Sarr, "Impacts of the

Sahel-Sahara Interface Reforestation on West African Climate: Intraseasonal Variability and Extreme Precipitation Events," *Advances in Meteorology*, vol. 2016, no. 1, p. 3262451, Jan. 2016, doi: 10.1155/2016/3262451.

[28] T. Guangul, A. Id, G. Gebre, B. Id, G. Walle, and B. Id, "Occupational radiation exposure dose and associated factors among radiology personnel in Eastern Amhara, Ethiopia," *PLoS One*, vol. 18, no. 5, p. e0286400, May 2023, doi: 10.1371/JOURNAL.PONE.0286400.

[29] D. L. Miller, B. A. Schueler, and S. Balter, "New Recommendations for Occupational Radiation Protection," *Journal of the American College of Radiology*, vol. 9, no. 5, pp. 366–368, May 2012, doi: 10.1016/J.JACR.2012.02.006.

[30] M. Belgiu and L. Drăgu, "Random forest in remote sensing: A review of applications and future directions," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 114, pp. 24–31, Apr. 2016, doi: 10.1016/J.ISPRSJPRS.2016.01.011.

[31] Q. Cao, J. Yan, J. Zhang, M. Wang, L. Long, and H. Lin, "Research on Optimization Method for Emergency Planning Zone Based on a Weather Stratified Sampling Method," pp. 91–101, 2026, doi: 10.1007/978-981-95-3233-9_8.

[32] L. Liu, X. Wang, Y. Li, and W. Wei, "The Effect of Sea Surface Temperature on Relative Humidity and Atmospheric Visibility of a Winter Sea Fog Event over the Yellow-Bohai Sea," *Atmosphere 2022, Vol. 13, Page 1718*, vol. 13, no. 10, p. 1718, Oct. 2022, doi: 10.3390/ATMOS13101718.

[33] C. O. Obasi, B. O. Ijabor, and A. S. Ahmad, "Estimating the Activity Concentration level, Absorbed Gamma Dose Rate, and Radiological Hazards around a Polyvinyl Chloride (PVC) Conduit Pipe factory in Asaba, Nigeria," *Journal of Radiological Protection*, May 2026, doi: 10.1088/1361-6498/AE7384