

COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS, AND INFORMATICS

**Radon Behaviour in the Atmosphere of Slovakia and Its Potential as
a Tracer for Assessing Environmental Processes**

Dissertation Thesis

Bratislava, 2025

Mohammad Alem Sultani

COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS, AND INFORMATICS



**Radon Behaviour in the Atmosphere of Slovakia and Its Potential as
a Tracer for Assessing Environmental Processes**

Dissertation Thesis

Study program: Nuclear and Subnuclear Physics

Fields of study: Nuclear and Subnuclear Physics

Department: Department of nuclear physics and biophysics

Supervisor: prof. RNDr. Jozef Masarik, DrSc.

Bratislava, 2025

Mohammad Alem Sultani



THESIS ASSIGNMENT

Name and Surname: Mohammad Alem Sultani
Study programme: Nuclear and Subnuclear Physics (Single degree study, Ph.D. III. deg., full time form)
Field of Study: Physics
Type of Thesis: Dissertation thesis
Language of Thesis: English
Secondary language: Slovak

Title: Radon Behavior in the Atmosphere of Slovakia and Its Potential as a Tracer for Assessing Environmental Processes

Annotation: Radon (^{222}Rn) is a naturally occurring radioactive noble gas produced in the uranium the decay series of uranium (^{238}U). Uranium is widely present in rocks and oil with half-life of billions of years, ultimately decaying to the stable isotope lead (^{206}Pb). As a result, radon is continuously released into the environment. With a half-life of 3.8 days, radon can travel significant distances from its source and accumulate in enclosed spaces. It decays through the emission of alpha particles, forming a series of short-lived radioactive progeny (^{218}Po , ^{214}Pb , ^{214}Bi , and ^{214}Po). These decay products are electrically charged and readily attach to fine aerosols in the atmosphere, which can be inhaled and deposited in lung tissue. Radon, with its progeny account for more than half of the global average effective dose from natural radiation, primarily due to indoor exposure. Although outdoor radon contributes less, its levels can be locally elevated in areas with high ^{238}U or ^{226}Ra content in the subsoil and low atmospheric dispersion. Since outdoor radon can enter indoor environments and significantly influence indoor concentrations, it is important for both radiation protection and environmental research. Moreover, radon is increasingly used as a tracer in atmospheric studies, including vertical mixing, greenhouse gas emissions and air mass transport. In Slovakia, continuous outdoor radon monitoring has been conducted since the 1990s, though mainly at a single site in Bratislava. Despite the presence of multiple uranium deposits and radon-prone areas, comprehensive nationwide research is still lacking. This dissertation aims to address this gap by investigating the behaviour of outdoor radon and its progeny in the atmosphere of Slovakia, analysing their interactions with environmental factors, and exploring their potential as tracers.

Aim: The aim of this dissertation is to enhance understanding of radon behavior in the atmosphere of Slovakia and to evaluate its potential as a tracer in environmental processes. To achieve this, continuous measurements of outdoor radon and its short-lived decay products will be conducted using different measurement techniques. The resulting time series data will be analyzed to identify the key environmental and meteorological factors influencing radon variability, employing both statistical and machine learning methods. Furthermore, the data will be used to explore the application of radon as a tracer in environmental studies, including the determination of the mixing layer height and the estimation of aerosol residence time.



Comenius University Bratislava

Faculty of Mathematics, Physics and Informatics

Literature: List of literature to be studied by the doctoral student in preparation for the dissertation exam:

Introductory nuclear physics (K. S. Krane) Environmental radioactivity (M. Eisenbud et al.)
Radon and its decay products in indoor air (W. Nazaroff et al.)
Radon in the environment (M. Wilkening)
Environmental radon (R. Cothorn et al.)
Radon: A tracer for geological, geophysical and geochemical studies (M. Baskaran)
Environmental radionuclides (K. Froehlich et al.)
Environmental isotopes in the hydrological cycle (W.G. Mook et al.)
Introduction to climate modeling (T. Stocker)
Reports - UNSCEAR 82, UNSCEAR 93, UNSCEAR 2000
Environmental physics (C. Smith)
Radionuclides in the environment (P. Povinec et al.)
Analysis of environmental radionuclides (P. Povinec et al.)
Modelling radioactivity in the environment (M. Scott et al.)
Scientific publications

Comment: Dissertation exam subjects (according to the study program):

1. Dissertation project
2. Radiation environmental physics and dosimetry

Tutor: prof. RNDr. Jozef Masarik, DrSc.

Consultant: RNDr. Martin Bulko, PhD.

Consultant: doc. RNDr. Monika Müllerová, PhD.

Department: FMFI.KJFB - Department of Nuclear Physics and Biophysics

Head of prof. RNDr. Jozef Masarik, DrSc.

department:

Assigned: 30.01.2021

Approved: 08.02.2021

prof. RNDr. Jozef Masarik, DrSc.

Guarantor of Study Programme

.....
Student

.....
Tutor

Contents

Abstract.....	iv
Abstract (Slovak version).....	vi
Acknowledgment.....	viii
Preface.....	x
List of Abbreviations.....	xii
List of Figures	xiv
List of Tables.....	xvi
Introduction and objectives	1
1. Theoretical background	4
1.1. Basic properties of radon.....	4
1.2. Radon sources.....	7
1.3. Radon emanation	8
1.4. Radon migration in soil	9
1.4.1. Radon transport model.....	11
1.5. Radon exhalation.....	13
1.6. Outdoor radon.....	14
1.6.1. Temporal variation of outdoor radon	16
1.6.2. Spatial variation of outdoor radon	18
2. Material and methods	19
2.1. Radon measurement techniques	19
2.2. Radon and its progeny sampling	21
2.3. Continuous measurement of radon	21
2.4. Radon measurement using AlphaGUARD	25
2.5. Radon progeny measurement	26
2.6. ^{210}Pb measurement	26
2.7. The potential use of radon as a tracer	27
2.7.1. Boundary layer height	27
2.7.2. MLH based on radon	28
2.7.3. Aerosols residence time.....	32
2.8. Conventional data analysis techniques.....	35
2.8.1. Fast Fourier analysis.....	35

2.8.2. Autocorrelation function.....	35
2.8.3. Canonical correlation analysis	36
2.8.4. Principal component analysis	36
2.8.5. Multiple linear regression	37
2.8.6. Generalised additive model.....	37
2.9. Machine learning in radon studies.....	38
2.9.1. Gradient boosting machine.....	39
2.9.2. Extreme gradient boosting.....	40
2.9.3. Random forest.....	40
2.10. Regression models evaluation.....	41
2.11. Data preprocessing	42
3. Time series analysis of RAC	45
3.1. Descriptive statistics of RAC	45
3.2. Data interpolation.....	45
3.3. Data smoothing - results	46
3.5. Periodicities in RAC time series	50
3.6. Diurnal cycle of RAC	52
3.7. Seasonal variation of RAC.....	54
4. Factors governing radon variability	55
4.1. Factors governing RAC at synoptic scales	58
4.2. Factors governing diurnal variation of RAC	60
4.3. Factors governing seasonal variations of RAC	63
4.4. Principal component analysis	65
4.5. Multiple linear regression analysis	67
4.6. Generalised additive model analysis	69
4.7. Machine learning-based regression model	74
4.8. Evaluation of regression models.....	76
4.9. SHAP analysis of the XGBoost model	77
4.10. Partial dependence plots of the XGBoost model.....	79
5. Radon decay products	81
5.1. Descriptive statistics of radon progeny.....	82
5.2. Radon equilibrium factor	84
5.3. Radon and its progeny ratios.....	86
5.4. Radon progeny interaction with particulate matter.....	87
5.5. Regression analysis of radon progeny	92

5.6. Multiple linear regression applied to radon progeny.....	93
5.7. Generalised additive model applied to radon progeny	94
5.8. Machine learning methods in radon progeny analysis	96
5.9. SHAP analysis of the random forest model	99
5.10. SHAP dependence plots for the random forest	101
5.11. Partial dependence plot for the random forest.....	101
6. Radon as a tracer in atmospheric research	102
6.1. MLH based on radon - results	102
6.2. Comparison of MLH with BLH-ERA5	105
6.3. Aerosols residence time - results	108
Conclusion	111
Future research directions	114
List of Publications.....	115
References	117

Abstract

Radon (^{222}Rn) is a naturally occurring radioactive noble gas produced in the decay series of uranium (^{238}U), which is ubiquitously present in rocks, soils and some building materials. Radon from soil is being continuously released into the atmosphere. It can also accumulate in enclosed spaces, where it can pose a significant health risk upon inhalation. With a half-life of 3.8 days, ^{222}Rn undergoes further radioactive decay through a series of short- and long-lived progeny. Beyond its radiological significance, outdoor radon and its progeny serve as effective environmental tracers.

This dissertation presents a comprehensive, multi-year study of outdoor radon dynamics in Bratislava, Slovakia, with a focus on its temporal variability, environmental influences, and radon potential as a tracer of environmental process. The study spans six years (2018-2023) of continuous radon measurements using a scintillation detector, and three years (2020-2022) of short-lived progeny measurements collected via alpha spectrometry. Low average radon concentrations ($5.6 \pm 3.9 \text{ Bq}\cdot\text{m}^{-3}$) were detected in the atmosphere of Bratislava during the monitored period, which is significantly lower than the global average of $10 \text{ Bq}\cdot\text{m}^{-3}$.

Diurnal cycles of outdoor radon showed maximum concentrations in the early morning and minimum concentrations in the afternoon, while seasonal trends showed the lowest concentrations in April and the highest in November. These patterns are governed by changes in boundary layer height (BLH) and radon flux dynamics. Conventional regression and machine learning models were both applied to hourly data to identify the dominant environmental drivers of radon variability. The machine learning models significantly outperformed the traditional approaches, consistently identifying BLH as the dominant predictor of outdoor radon concentration and precipitation as the least influential factor.

Further analysis of the interactions between radon progeny, particulate matter (PM), meteorological variables, and BLH demonstrated a strong, non-linear relationship between radon progeny and PM, suggesting enhanced attachment under polluted conditions.

A box model incorporating outdoor radon concentration and radon flux was used to estimate the mixing layer height. The estimated mixing layer height effectively captures diurnal and seasonal boundary-layer dynamics, despite some limitations due to low radon concentration measurements and uncertainties in the radon flux data.

Furthermore, the mean aerosol residence time was derived from activity ratios of $^{210}\text{Pb}/^{222}\text{Rn}$, $^{210}\text{Pb}/^{214}\text{Pb}$, and $^{210}\text{Pb}/^{214}\text{Bi}$, averaging 3.15 days (ranging from 0.35–6.73 days), which is consistent with findings in the existing literature.

Overall, this work underscores the importance of long-term outdoor radon monitoring for both atmospheric research and public health. It demonstrates that radon and its progeny are effective, low-cost tracers of atmospheric mixing and aerosol dynamics.

Abstract (Slovak version)

Radón (^{222}Rn) je prirodzene sa vyskytujúci rádioaktívny vzácny plyn vznikajúci v premenovom rade uránu (^{238}U), ktorý je všadeprítomný v horninách, pôdach a niektorých stavebných materiáloch. Radón je kontinuálne uvoľňovaný z pôdy do atmosféry, hromadí sa tiež v uzavretých priestoroch a jeho inhalácia môže predstavovať významné zdravotné riziko. ^{222}Rn sa s dobou polpremeny 3,8 dňa ďalej rozpadá prostredníctvom série krátko a dlhožijúcich produktov premeny. Okrem svojho rádiologického významu slúžia radón a jeho dcérske produkty ako účinné stopovacie látky v životnom prostredí.

Táto dizertačná práca predstavuje komplexnú, viacročnú štúdiu dynamiky radónu v atmosfére Bratislavy (Slovensko), so zameraním na jeho časovú variabilitu, vplyvy životného prostredia a na potenciál radónu ako stopovača environmentálnych procesov. Štúdia zahŕňa šesť rokov (2018–2023) kontinuálnych meraní radónu pomocou scintilačného detektora a tri roky (2020–2022) meraní krátkožijúcich dcérskych produktov získaných pomocou alfa spektrometrie. Za sledované obdobie boli v atmosfére Bratislavy zistené nízke priemerné koncentrácie radónu ($5,6 \pm 3,9 \text{ Bq}\cdot\text{m}^{-3}$), čo je výrazne menej ako je celosvetový priemer $10 \text{ Bq}\cdot\text{m}^{-3}$.

Denné cykly koncentrácie radónu dosahovali maximá v skorých ranných hodinách a minimá popoludní, zatiaľ čo sezónne trendy sa vyznačovali najnižšími koncentraciami v apríli a najvyššími v novembri, čo je spôsobené zmenami výšky hraničnej vrstvy atmosféry (BLH) a dynamikou exhalácie radónu z pôdy. Na identifikáciu hlavných environmentálnych faktorov ovplyvňujúcich variabilitu radónu boli na hodinové dáta aplikované konvenčné regresné modely a modely strojového učenia. Modely strojového učenia výrazne prekonal tradičné modely a konzistentne identifikovali BLH ako dominantný prediktor koncentrácií radónu, zatiaľ čo zrážky sa ukázali ako najmenej významný faktor.

Ďalšia analýza interakcií medzi dcérskymi produktmi radónu, polietavými prachovými časticami (PM), meteorologickými premennými a BLH preukázala silný, nelineárny vzťah medzi produktmi premeny radónu a PM, čo vypovedá o ich zvýšenom zachytení na prachových časticách v prípade atmosférického znečistenia.

Na odhad výšky zmiešavacej vrstvy atmosféry bol využitý tzv. box model, ktorého vstupné parametre boli aktivita radónu vo vonkajšej atmosfére a exhalačná rýchlosť radónu z pôdy. Odhadnutá výška zmiešavacej vrstvy účinne zachytáva denné a sezónne variácie hraničnej

vrstvy, a to aj napriek určitým obmedzeniam spôsobeným meraním nízkych koncentrácií radónu a neistotám v určení jeho exhalačnej rýchlosti.

Priemerný čas zotrvania aerosólov bol odhadnutý z pomerov aktivít $^{210}\text{Pb}/^{222}\text{Rn}$, $^{210}\text{Pb}/^{214}\text{Pb}$ a $^{210}\text{Pb}/^{214}\text{Bi}$, pričom jeho priemerná hodnota bola na úrovni 3,15 dňa (s rozsahom 0,35-6,73 dňa), čo je v dobrej zhode s hodnotami uvádzanými v literatúre.

Táto práca ako celok zdôrazňuje význam dlhodobého monitorovania radónu pre atmosférický výskum a verejné zdravie a dokazuje, že radón a jeho produkty premeny sú účinnými a nízkonákladovými stopovačmi atmosférického premiešavania a dynamiky prachových častíc v atmosfére.

Acknowledgment

I would like to express my heartfelt gratitude to my official supervisor, **Prof. RNDr. Jozef Masarik, DrSc.**, for his continuous support, guidance, and encouragement throughout my doctoral studies. His consistent support and willingness to assist at every stage of my PhD studies have been a great source of motivation.

I am deeply thankful to my late supervisor, **doc. RNDr. Karol Holý, CSc.**, whose kindness, support, and encouragement had a profound impact on both my academic and personal life. He was not only a dedicated mentor but also a genuinely compassionate person who believed in me from the very beginning. His decision to accept me as a PhD student was truly life-changing, opening doors to opportunities I had only dreamed of. After his passing, I felt the loss of not only a great academic guide but also a deeply valued supporter. I will always carry gratitude for the trust he placed in me and the foundation he helped build for my future. *May he rest in peace.*

I would like to express very special thanks to **RNDr. Martin Bulko, PhD.**, who has been not only an invaluable consultant but also the closest mentor and friend throughout my PhD journey. Since the passing of my late supervisor, he has supported me academically and personally, always being there for advice, discussion, and encouragement. Our friendship has grown through countless chats, conversations, and even games of badminton. I am especially grateful for the warmth and generosity he has shown, including inviting me into his home. His presence has meant a great deal to me, both professionally and personally.

I also sincerely appreciate the support of my consultant, **doc. RNDr. Monika Müllerová, PhD.** Although I did not have many opportunities to discuss my work with her in person or learn from her directly, I value her presence as the current head of our research group and appreciate her role in supporting the academic environment in which this work was carried out.

Special thanks go to my fellow PhD candidates in our department - **Mgr. Adam Sitarčík, PhD.**, and **Mgr. Jozef Mišt** - for their support and encouragement throughout my doctoral journey.

I gratefully acknowledge the faculty and administrative staff of the Faculty of Mathematics, Physics and Informatics for their support, as well as Comenius University in Bratislava for the financial support through scholarships and grants. I would especially like to thank the

Doctoral Studies Office for their continuous assistance, and particularly **Prof. RNDr. Robert Jajcay, DrSc.**, for his trust, support, and encouragement throughout my doctoral journey.

I also appreciate the kind support of the **Mlyny-UK dormitory staff**, who made my stay comfortable.

Above all, I am deeply indebted to my family for their unwavering support, love, and inspiration. This work is a reflection of their selfless encouragement.

With sincere thanks,

Mohammad Alem Sultani

Preface

My interest in radon research began when I first learned about its unique properties. Radon is a fascinating radionuclide for researchers because of its distinctive characteristics i.e., on one hand, posing health risks due to its radioactive nature, and on the other hand, offering potential as a tracer in environmental processes. Working with radon provides an excellent opportunity for those passionate about interdisciplinary research, as it involves aspects of nuclear physics, atmospheric science, and environmental monitoring. I have been fortunate to engage in this interdisciplinary field.

Initially, my plan was to focus on outdoor radon measurement using passive detectors, such as etched track detectors, in different regions of Slovakia. However, due to limited laboratory equipment, my research direction shifted toward continuous radon measurements and their application as environmental tracer. I found this path more compelling and promising. Fortunately, our faculty has a long-standing history of collecting outdoor radon data, spanning more than three decades.

Ultimately, I set the objective of my dissertation to study the behaviour of outdoor radon and explore its potential as a tracer for environmental processes. To achieve these aims, I analysed long-term radon activity concentration (RAC) and used it to estimate the mixing layer height (MLH). The MLH represents the lowest part of the troposphere, ranging from hundreds of meters to several kilometres, and is a critical parameter in atmospheric research as it influences pollutant dispersion in the lower atmosphere. Since radon data for atmospheric studies require high accuracy and traceability, special care was taken during data preprocessing due to some limitations.

My research objectives further extended to analysing the variability of radon and its progeny concerning key meteorological factors, boundary layer height (BLH), and particulate matter (PM). I utilized various statistical and machine learning techniques to gain deeper insights into the behaviour of radon and its governing factors. Understanding these interactions is essential for effective application of radon as a tracer and for mitigating associated health hazards. In particular, I developed an interest in investigating the interaction between radon progeny and PM, as both are carcinogenic and have been relatively underexplored.

The structure of this dissertation is organized as follows:

Chapter 1: This chapter presents the theoretical background of the dissertation, including the basic properties of radon, its production in soil, transportation mechanisms, exhalation processes, and behaviour in the atmosphere.

Chapter 2: This chapter describes the experimental methods used for measuring outdoor radon and its progeny, as well as the data analysis techniques employed.

Chapter 3: This chapter presents the initial findings of the research. It includes a descriptive statistical analysis of the radon measurement data, complemented by an in-depth time series analysis to reveal temporal patterns and trends.

Chapter 4: This chapter explores the environmental factors affecting radon temporal variability. It discusses various statistical and machine learning approaches used to identify the factors governing radon levels.

Chapter 5: This chapter presents the study of radon progeny and its relationship with particulate matter and other meteorological parameters.

Chapter 6: This chapter presents the application of radon and its progeny as tracers in atmospheric research. Specifically, it details the results for the retrieval of mixing layer height using radon measurements. Furthermore, this chapter presents the estimation of aerosol residence time, derived from the activity ratios of radon and its progeny.

List of Abbreviations

ACF	Autocorrelation Function
adj. R^2	adjusted coefficient of determination
AI	Artificial Intelligence
ANN	Artificial Neural Network
BLH	Boundary Layer Height
BLH-ERA5	Boundary Layer Height based on ERA5 reanalysis dataset
CBL	Convective Boundary Layer
CCA	Canonical Correlation Analysis
CDF	Cumulative Distribution Function
CosDoY	Cosine of Days of the Year
CosH	Cosine of Hour of the Day
CV	Cross Validation
ECMWF	European Centre for Medium-Range Weather Forecasts
EEC	Equilibrium Equivalent Concentration
Feq	Radon Equilibrium Factor
FFT	Fast Fourier Transform
FMPI CU	Faculty of Mathematics, Physics, and Informatics, Comenius University
GAM	Generalised Additive Model
GBM	Gradient Boosting Machine
HPGe	High Purity Germanium
ICRP	International Commission on Radiation Protection
IQR	Interquartile range
LSCH	Large Scintillation Chamber
MA	Moving Average
MAE	Mean Absolute Error
ML	Machine Learning
MLH	Mixing Layer Height
MLH-Rn	Mixing Layer Height based on Radon
MLR	Multiple Linear Regression
MSE	Mean Square Error

NORM	Naturally Occurring Radioactive Materials
P	Pressure
PCA	Principal Component Analysis
PDP	Partial Dependence Plot
PM	Particulate Matter
P-P	Probability-Probability plot
Prec	Precipitation
Q1	First quartile
Q3	Third quartile
RAC	Radon Activity Concentration
Radon	^{222}Rn
RF	Random Forest
RH	Relative Humidity
RL	Residual Layer
RMSE	Root Mean Square Error
SBL	Stable Boundary Layer
SHAP	SHapley Additive exPlanations
SinDoY	Sine of the Days of the Year
SinH	Sine of Hour of the Day
SSCH	Small Scintillation Chamber
T	Temperature
Thoron	A radon isotope ^{220}Rn
T_R	Aerosol Residence Time
UNSCEAR	United Nations Scientific Committee on the Effects of Atomic Radiation
VIF	Variance Inflation Factor
WD	Wind Direction
WS	Wind Speed
XGBoost	Extreme Gradient Boosting

List of Figures

Fig. 1.1 The decay series of radon.....	6
Fig. 1. 2 The schematic view of the radon emanation and exhalation from the soil.	7
Fig. 1. 3 Accumulation chamber setting used for the measurement of radon and CO ₂ fluxes.	14
Fig. 1. 4 An example of RAC composite diurnal cycle averaged over May-2022.....	17
Fig. 2.1 Radon measurement techniques based on electricity consumption.	20
Fig. 2.2 Sampling location.	21
Fig. 2.3 Schematic diagram of sampling instrumentation for continuous measurement.	22
Fig. 2.4 Comparison of RAC measured by scintillation chambers and AlphaGUARD.....	24
Fig. 2.5 Alpha GUARD portable radon monitor and its ionization chamber.....	25
Fig. 2.6 The schematic diagram of the radon progeny measurement system.....	25
Fig. 2.7 Single HPGe detector (PGT, 270 cm ³) and the large low-background shield.	26
Fig. 2.8 Schematic diagram of the vertical structure of boundary layer height.	28
Fig. 2.9 The schematic view of the box model.	29
Fig. 2.10 An example of the temporal evolution of MLH based on radon measurements.....	31
Fig. 2.11 Illustration of random forest trees.	41
Fig. 3.1 An example of cubic spline interpolation.	46
Fig. 3.2 Comparison of data smoothing using MA and FFT.....	47
Fig. 3.3 Correlation between raw data and smoothing based on different methods.....	48
Fig. 3.4 RAC data fitting with normal distribution function.	49
Fig. 3.5 Cullen and Frey graph.....	49
Fig. 3.6 Periodogram of RAC time series based on FFT.	51
Fig. 3.7 Periodogram of the ACF of the RAC time series.	51
Fig. 3.8 Composite diurnal cycle of RAC for each month.	53
Fig. 3.9 Composite diurnal cycle of RAC for each season.	53
Fig. 3.10 Seasonal boxplots of RAC.	54
Fig. 4.1 Hourly RAC, BLH and major meteorological parameters.....	56
Fig. 4.2 Hourly RAC against BLH and meteorological factors with moving average.....	57
Fig. 4.3 Pearman's Correlation coefficient analysis.....	59
Fig. 4.4 Composite diurnal cycles of RAC versus BLH and key meteorological factors	62
Fig. 4.5 Heatmap correlation of RAC, BLH and meteorological diurnal variation.	62
Fig. 4.6 Composite monthly boxplots of RAC against influencing factors	64
Fig. 4.7 Pearson correlation between composite monthly values	64
Fig. 4.8 PCA results for radon.....	67
Fig. 4.9 Three-dimensional PCA plot of the time-related predictors	71

Fig. 4.10 Estimated smooth functions of predictors from GAM analysis.	73
Fig. 4.11 Scatter plots of measured RAC against predicted based on different regression model. .	75
Fig. 4.12 Feature importance plots of regression models.	76
Fig. 4.13 Regression models performance evaluation metrics.	77
Fig. 4.14 Summary plot of SHAP analysis.	78
Fig. 4.15 SHAP dependence plots of each feature used in XGBoost	79
Fig. 4.16 Partial dependency plots of each feature used in XGBoost model.....	80
Fig. 5.1 Diurnal cycles of short-lived radon progeny.	82
Fig. 5.2 Seasonal boxplots of short-lived radon progeny.	83
Fig. 5.3 Composite diurnal cycles of F_{eq} and radon progeny ratio.	85
Fig. 5.4 Composite monthly boxplot of F_{eq} and radon progeny ratio.....	85
Fig. 5.5 Hourly EEC, PM _{2.5} and PM ₁₀	88
Fig. 5.6 Pearson correlation heatmap among radon progeny, F_{eq} and PM.....	88
Fig. 5.7 Monthly variation of radon, its progeny and PM.	89
Fig. 5.8 Pearson's correlation heatmap for radon progeny.	90
Fig. 5.9 Graphical representation of canonical correlation analysis.	92
Fig. 5.10 Estimated smooth functions of the predictors in the GAM for radon EEC.	96
Fig. 5.11 Scatter plots of the regression models	98
Fig. 5.12 Feature importance from four regression models used in radon progeny analysis.....	98
Fig. 5.13 SHAP summary plot for the RF model.....	99
Fig. 5.14 SHAP dependence plot for the RF model used for prediction of EEC.....	100
Fig. 5.15 PDPs plots of RF model used in prediction of radon EEC.....	100
Fig. 6.1 Hourly MLH and RAC evolutions for the month of August 2020.	103
Fig. 6.2 Composite diurnal cycles of MLH for each month of the year.	104
Fig. 6.3 MLH monthly boxplot created based on data covering the entire study period.....	105
Fig. 6.4 Composite diurnal cycles of a) BLH-ERA5 and b) MLH-Rn for each season.	106
Fig. 6.5 Correlation between MLH-Rn and BLH-ERA5 composite diurnal cycles.	107
Fig. 6.6 Monthly variations of MLH-Rn compared to BLH-ERA5.	107
Fig. 6.7 Composite monthly values of MLH vs BLH-ERA5.	108
Fig. 6.8 Pearson correlation heatmap between monthly radon and its progeny.	108
Fig. 6.9 Frequency distribution of aerosol residence times	109

List of Tables

Table 1.1 Physical and chemical properties of Radon.....	6
Table 3.1 Descriptive statistics of 2-hourly RAC.....	45
Table 4.1 PCA result.....	65
Table 4.2 Loadings of each variable on the first three principal components of PCA.	65
Table 4.3 Multicollinearity check based on Variation Inflation Factor.....	67
Table 4.4 The summary of MLR-radon.....	68
Table 4.5 Concurvity results of predictors used in the GAM (part 1).	70
Table 4.6 PCA result of time-related predictors.	70
Table 4.7 Concurvity results of predictors used in the GAM (part 2).	71
Table 4.8 Summary of GAM analysis of RAC.....	72
Table 4.9 Regression model evaluation metrics.	76
Table 5.1 Descriptive statistics of measured and calculated variables.	82
Table 5.2 Results of the canonical correlation analysis (part 1).	91
Table 5.3 Results of the canonical correlation analysis (part 2).	92
Table 5.4 Multicollinearity check of radon progeny predictors.....	92
Table 5.5 Summary of MLR for radon progeny.....	93
Table 5.6 Multicollinearity check among radon EEC predictors.....	94
Table 5.7 GAM result for radon progeny	95
Table 6.1 Descriptive statistics of TR based on three ratios of radon and its progeny.	109
Table 6.2 Comparison of aerosol residence times with those published in the literature.	110

Introduction and objectives

Radon (^{222}Rn) is a naturally occurring radioactive noble gas produced in the uranium (^{238}U) decay series. Uranium is widely present in rocks and soil with a half-life of billions of years, ultimately decaying to the stable isotope of lead (^{206}Pb). As a result, radon is continuously released into the environment. With a half-life of 3.8 days, radon can travel significant distances from its source and accumulate in enclosed spaces. It decays through the emission of alpha particles, forming a series of short-lived radioactive progeny. Most of Radon decay products are electrically charged and readily attach to fine aerosols in the atmosphere, which can be deposited in lung tissue when inhaled.

According to the UNSCEAR (2000), radon and its decay products (especially indoor radon) are the leading source of natural radiation exposure to the public, accounting for nearly half of the global mean effective dose. Exposure to radon is therefore a significant health concern and has been identified as the second leading cause of lung cancer after smoking (ICRP, 1993).

Radon in the atmosphere primarily originates from the soils and rocks near the Earth's surface, groundwater and anthropogenic sources such as mining and the use of naturally occurring radioactive materials (NORM). While outdoor radon concentrations ranges (typically $1\text{--}100\text{ Bq}\cdot\text{m}^{-3}$) are much lower than indoor levels (Čeliković et al., 2022). Although the effective dose from outdoor radon is relatively low, it still contributes to total natural radiation exposure. Furthermore, outdoor radon monitoring provides a valuable reference for distinguishing natural background levels from elevated anthropogenic emissions.

Beyond health implications, radon has proven to be an effective environmental tracer due to its unique physical and chemical properties. These include its exclusive terrestrial origin, removal primarily through radioactive decay, inert nature as a noble gas, low water solubility (which limits washout by precipitation), and a half-life similar to the residence time of atmospheric aerosols. Importantly, its source strength shows minimal spatio-temporal variability (Karstens et al., 2015). These features make radon highly suitable for tracing various environmental processes. Some of the key applications of outdoor radon as a tracer of environmental processes include:

- i) Assessing atmospheric vertical mixing and stability (e.g., Chambers et al., 2015, 2016; Gregorič et al., 2020; Kikaj et al., 2019; Perrino et al., 2001; Sesana et al., 2003; Sultani et al., 2023; Vecchi et al., 2005; Zahorowski et al., 2004).
- ii) Estimating greenhouse gas emissions (Grossi et al., 2018; Levin et al., 2021; Schmidt et al., 2001).
- iii) Tracing air mass transport (Gupta et al., 2004).
- iv) Hydrological studies (Adyasari et al., 2023; Baskaran, 2016; Sukanya et al., 2022).

Tracing dynamic processes such as earthquake prediction and volcanic activity (Friedman, 2001; Hwa Oh & Kim, 2015; Utkin & Yurkov, 2010; Woith, 2015; Zmazek et al., 2003). The primary motivation of this dissertation is to improve understanding of the behaviour of outdoor radon and its short-lived progeny, particularly in relation to the environmental and atmospheric factors that govern their variability. Although numerous studies worldwide have examined atmospheric radon and its progeny, drawing definitive conclusions remains difficult due to the complex interplay of meteorological conditions, atmospheric mixing, source strength variability, and aerosol dynamics. This complexity poses ongoing challenges in predicting radon behaviour with high confidence as well as using it as a tracer of environmental process.

A specific focus of this research is the interaction between short-lived radon progeny and particulate matter (PM), an area where current scientific understanding remains limited. Both radon progeny and PM are recognized carcinogens, and their potential synergistic effects on human health are still poorly understood. By leveraging long-term, continuous measurement data and applying advanced statistical and machine learning (ML) approaches, this study aims to analyse these interactions and provide deeper insights into their behaviour under varying atmospheric conditions.

Additionally, this dissertation explores the application of radon as a tracer in environmental studies. While radon has been widely used for estimating parameters such as mixing layer height (MLH) and aerosol residence time, significant methodological challenges and uncertainties persist. For example, although various studies have used radon to estimate MLH, the reliability and accuracy of these models remain questionable and require further investigation. Similarly, the use of radon progeny ratios to determine aerosol residence time has yielded inconsistent results across different studies. This dissertation seeks to contribute

to the assessment of existing models, address their limitations, and offer improved methodologies for using radon as an atmospheric tracer.

Given these diverse applications and the need for more detailed outdoor radon studies, this dissertation investigates outdoor radon and its progeny within the Slovak environment with the following objectives.

Objectives of the dissertation

The primary objectives of this dissertation are:

1. To conduct continuous measurements of outdoor radon and its short-lived decay products in the atmosphere of Slovakia using multiple measurement techniques.
2. To analyse the time series of outdoor radon and its progeny and investigate the key environmental factors influencing their variability, using both statistical and machine learning approaches.
3. To assess the potential use of radon as a tracer for atmospheric studies, including the determination of MLH and the estimation of aerosol residence time.

Chapter 1

1. Theoretical background

This chapter presents the fundamental properties and behaviour of radon in outdoor environments, covering its physical and radioactive characteristics, sources, transport mechanisms and atmospheric variations.

1.1. Basic properties of radon

Radon was first discovered in 1898 by Friedrich Ernst Dorn, a German scientist (Dorn, 1900). It is colourless, odourless and tasteless noble gas. Radon is the heaviest among the noble gases, which is a group of chemical elements known for their very low chemical reactivity due to their full valence electron shells. Some of the physical and chemical properties of radon are summarized in **Table 1.1**. Radon can readily be absorbed on charcoal, silica gel and similar materials, a property which can be used to separate it from other gases. Two isotopes of radon are well known, found in the environment, and considered radiologically significant:

- **^{222}Rn (radon):** A decay product of ^{226}Ra from the ^{238}U decay series, with a half-life of approximately 3.82 days. It is the most abundant and long-lived radon isotope in the environment.
- **^{220}Rn (thoron):** Originating from the ^{232}Th decay series, it has a half-life of approximately 55.6 seconds. Due to its short half-life, thoron can travel only a few tens of centimetres in air and is usually negligible at a distance.

In this study, only ^{222}Rn (hereafter referred to as radon) is studied. Radon decays by emission of alpha particles to ^{218}Po and initiating a decay chain that includes both short-and-long-lived daughter products (^{218}Po , ^{214}Pb , ^{214}Bi , ^{214}Po , ^{210}Pb , ^{210}Bi , ^{210}Po), until it reaches the stable element ^{206}Pb (**Fig. 1.1**). Four short-lived isotopes of radon are heavy metals, each with a half-life of under 0.5 hours. These isotopes remain closely associated with radon and contribute significantly to its radiological health effects. After these short-lived radionuclides, the next element in the decay chain is ^{210}Pb , which has a much longer half-life of 21 years. ^{210}Pb eventually decays into stable ^{206}Pb through intermediate decay steps involving ^{210}Bi and ^{210}Po . The radioactive equilibrium between radon and its daughters is

Theoretical background

considered to be reached after seven half-lives of the daughter product (Prichard & Gesell, 1984). The activity of the daughter product can be described by following relation:

$$A_2(t) = A_{10} \frac{\lambda_2}{\lambda_2 - \lambda_1} (e^{-\lambda_1 t} - e^{-\lambda_2 t}) + A_{20} e^{-\lambda_2 t} \quad (1.1)$$

where A_{10} is the initial activity of the parent nuclide at time zero, A_{20} is the initial activity of the daughter nuclide, and λ_1 , λ_2 are the decay constants of the parent and daughter nuclides, respectively.

The general solution describing the activity $A_n(t)$ of the n -th member in a radioactive decay chain is given by the Bateman equations (Bateman, 1910):

$$A_n(t) = A_{10} \sum_{i=1}^n C_i e^{-\lambda_i t}, \quad (1.2)$$

Where the coefficient C_i is defined as:

$$C_i = \prod_{k=2}^n \lambda_k / \prod_{j=2, j \neq i}^n (\lambda_j - \lambda_i) \quad (j \neq i) \quad (1.3)$$

Once the equilibrium is achieved, the activity of all members of the decay chain declines with the half-life of the parent radionuclide (radon). The radon daughters are metals that behave differently than radon. Most of the daughter products acquire a positive net charge which gets associated with molecules of water vapour or other atmospheric constituents, forming ions. These ions tend to get attached to aerosols or adjacent surfaces. Radon decay products exist in two forms: an unattached fraction, with diffusion diameters ranging from 0.5 to 5 nm, and a fraction that attaches to aerosol particles, with diameters between 5 and 3000 nm (Abdelfatah Mostafa et al., 2020; Porstendörfer, 1994). Once attached, radon progeny is removed from air by deposition. Heavier particles settle by gravity, while diffusion carries tiny unattached clusters to surfaces. In the respiratory tract, both attached and unattached progeny can deposit notably, the unattached fraction (very small clusters) has a much higher probability of reaching deep lung (Butterweck et al., 2005). According to the (UNSCEAR, 2000), radon and its progeny are the main sources of public exposure from natural radioactivity, contributing to nearly half of the global mean effective dose to the public. Considering the low level of outdoor radon, it will yield to a few tenths of a mSv radiation exposure to the population per year, whereas indoor radon (with significantly higher levels) yield much more. Nonetheless, outdoor radon and its progeny contribute to the collective lung dose of a population, especially in rural or mining regions. Therefore, the

Theoretical background

exposure to radon increases the radiation risk to the population and is considered the second leading cause of lung cancer after smoking (ICRP, 1993). WHO notes that radon causes roughly 3–14% of lung cancers in populations depending on average levels (WHO, 2009).

Table 1.1 Physical, and chemical properties of ^{222}Rn taken from (Baskaran, 2016).

Property	Values
Atomic number	86
Standard atomic weight	222
Outer shell electron configuration	$6s^2 6p^6$
Density	9.73 kg m^{-3} (at 0°C , $1.013 \times 10^5 \text{ Pa}$)
Melting point ($^\circ\text{K}$)	202
Normal boiling point ($^\circ\text{K}$)	208.2
Heat of fusion (kJ mol^{-1})	3.247
Heat of vaporization (kJ mol^{-1})	18.0
First ionization enthalpy (kJ mol^{-1})	1037
Oxidation states	0, 2, 6
Electronegativity	2.2 (Pauling scale)
Covalent radius (nm)	0.150
van der Waals radius (nm)	0.220
Half-life ($T_{1/2}$)	3.823 d
Decay constant (λ)	$2.098 \times 10^{-6} \text{ s}^{-1}$
Diffusion coefficient in air (D_a)	$1 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$
Diffusion coefficient in water (D_w)	$1 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$

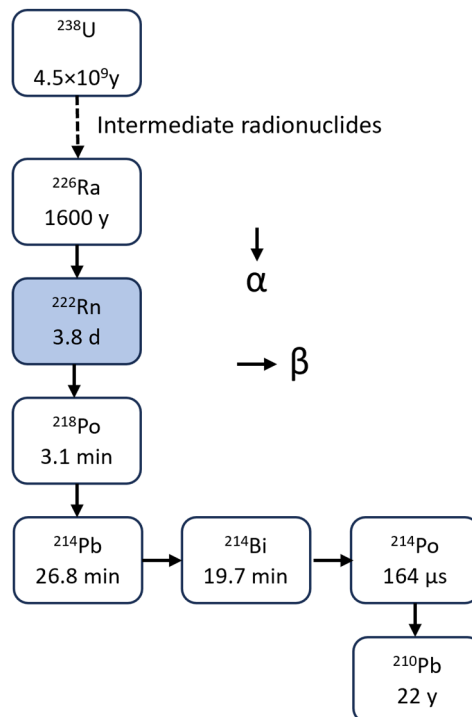


Fig. 1.1 The radioactive decay series of ^{222}Rn .

1.2. Radon sources

Radon in the environment mainly originates from the soil and rocks through the decay of Radium (^{226}Ra) which is a decay product of ^{238}U . Radium with a half-life of 1620 years is usually in secular equilibrium with ^{238}U . Radium is widely distributed in crustal rocks, soils, and building materials, and its presence in soil is the primary source of radon gas. Radium content in surface soils typically ranges from 10 to 100 Bq.kg^{-1} , in localized areas near uranium mining or tailings reaching values as high as 1700 Bq.kg^{-1} (Nazaroff, 1992). Materials such as phosphate rock, granite, and shale often contain higher levels of uranium and radium and thus serve as significant radon sources (Appleton, 2007). Additional sources of radon include groundwater, oceans and building materials especially from deep wells. The rate at which radon is released from the soil, or exhaled, depends on radium content and various meteorological factors, such as moisture, temperature, pressure (Čeliković et al., 2022). In the atmosphere, radon primarily enters from the soil through diffusive and advective transport (Nazaroff, 1992). The concentration of radon in the lower atmosphere is influenced by the amount of ^{238}U and ^{226}Ra in source materials, radon exhalation rate, and atmospheric mixing conditions driven by meteorological factors (e.g., Porstendörfer, 1994).

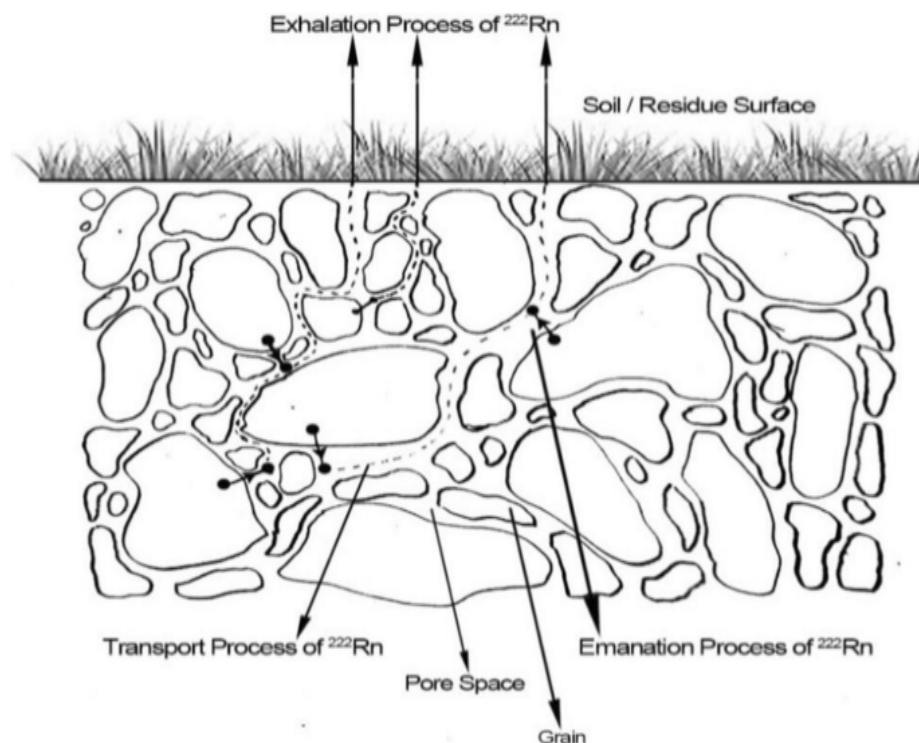


Fig. 1.2 The schematic view of the radon emanation and exhalation from the soil to the atmosphere taken from (Hassan et al., 2014).

1.3. Radon emanation

The soil can be described as a porous medium composed of organic material, mineral particles, and pores that contain water and soil gases. Emanation refers to the process by which radon atoms produced inside mineral grains escape into the pore spaces of soil or rocks. This process is critical because only the radon that escapes into pore spaces is available for migration and enter the atmosphere. Not all radon atoms produced by the decay of radium contained in the rock or soil grains are released into pore space. Some of radon gas remains embedded in the same grain, some travels within a pore space and become embedded in the adjacent grain, and some is released to the pore space (**Fig. 1.2**). According to (Appleton, 2007) usually 20 – 40 % of newly formed radon in soil and up to 70 % in clay emanate to the pore space, where they are mixed in the soil gas or water in the pores.

The radon emanation coefficient (or emanation power) is defined as the ratio of number of radon atoms that escape from a soil grain to the total number of radon atoms formed in radioactive decay (Schumann & Gundersen, 1996). It is the most essential physical parameter for determining the behaviour of radon in materials. Radon emanation occurs through a process known as alpha recoil. This happens when an atom of radium decays into radon, which is accompanied by the immediate emission of an alpha particle. The energy released during this decay is distributed between the radon atom and the alpha particle and is inversely proportional to their masses. Consequently, due to the momentum conservation, the alpha particle and radon atom move in opposite directions. The radon atom continues to travel through the surrounding material until all of its energy has been dissipated. The distance travelled depends on the density and the composition of the materials. The range of recoiling of radon is 0.02 – 0.07 μm in common minerals, 0.1 μm in water, and 63 μm in air (Nazaroff, 1992). Only radium atoms lying within this recoil distance from grain surfaces can contribute via direct recoil; atoms deeper inside require additional pathways to escape.

Several factors influence the radon emanation coefficient i.e., the distribution of radon parent atoms in the solid, the soil grain size, soil porosity, and the moisture content and temperature (Nazaroff, 1992; Hassan et al. 2009; Phong Thu et al., 2020). Soil moisture affects the radon emanation coefficient in two opposing ways: it reduces the mobility of radon atoms by slowing their diffusion through soil pores (i.e., the diffusion coefficient of radon in water is about three orders of magnitude lower than in air and it can enhance emanation by dissolving radon atoms trapped in intergranular pores or embedded within soil particles (Edsfeldt, 2001; Wilkening, 1990). Smaller particles (higher specific surfacer area) show higher emanation

coefficients while the emanation coefficient decreases with increasing the soil grain size (Phong Thu et al., 2020; Sakoda et al., 2011). The emanation coefficient is directly proportional to the temperature, as an increase in temperature aids the release of gases from solid particles (BARRETTO, 1973; Iskandar et al. 2004).

1.4. Radon migration in soil

Following emanation from soil grains, radon atoms migrate through the water and air contained in soil pores toward the soil surface, where they may be released into the atmosphere (**Fig. 1.2**). Radon transport in soil primarily occurs via two mechanisms: diffusion and advection.

Diffusion refers to the movement of gas molecules from regions of high concentration to regions of low concentration, driven by concentration gradients. In the case of radon, this occurs due to the much higher concentration of radon in soil gas compared to atmospheric air typically about three orders of magnitude higher. Consequently, radon tends to diffuse upward to the atmosphere and reduce the gradient. Long-term evidence suggests that radon migration from soil to the atmosphere is predominantly governed by molecular diffusion (Nazaroff, 1992). The random molecular diffusion of radon in the vertical direction (z-axis) can be described quantitatively by Fick's First Law, modified for porous media (Savović & Djordjević, 2008):

$$J_d = -D_e \nabla C_{Rn} \quad (1.4)$$

Where:

- J_d [Bq.m⁻².s⁻¹]: radon diffusive flux density,
- D_e [m².s⁻¹]: effective molecular diffusion coefficient of radon in soil,
- C_{Rn} [Bq.m⁻³]: radon activity concentration in the interstitial space,
- ∇ [m⁻¹]: gradient operator, and
- the negative sign indicates that the radon diffuses from high to low concentration.

The effective diffusion coefficient D_e accounts for the impact of soil structure, notably porosity and tortuosity, and is given by:

$$D_e = \varepsilon \tau D_0 \quad (1.5)$$

Where:

ε [-]: porosity (fraction of soil volume occupied by pores),

Theoretical background

τ [-]: tortuosity factor (pathway complexity),

D_0 [$1.2 \times 10^{-5} \text{m}^2 \text{s}^{-1}$]: diffusion coefficient in air.

The effective diffusion coefficient D_e depends on the types of the soil and moisture content of the soil typically ranges from 10^{-10} to 10^{-5} (Nazaroff, 1992; Nielson et al., 1994).

A related parameter is the diffusion length L_d , representing the characteristic distance radon travels before decaying:

$$L_d = \sqrt{\frac{D_e}{\lambda_{Rn}}} \quad (1.6)$$

Where λ_{Rn} [s^{-1}] is the radon decay constant.

In addition to diffusion, advection can transport radon via bulk soil gas flow, driven by pressure gradients in the subsurface. The advective flux in the vertical direction is:

$$J_a = v \cdot C_{Rn} \quad (1.7)$$

Where:

- J_a [$\text{Bq} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$]: advective flux of radon,
- v [$\text{m} \cdot \text{s}^{-1}$]: superficial velocity of the soil gas (Nazaroff, 1992),

The velocity v is determined by Darcy's Law:

$$v = -\frac{k}{\mu} \nabla P \quad (1.8)$$

Where:

- k [m^2]: permeability of the soil matrix,
- μ [$\text{Pa} \cdot \text{s}$]: dynamic viscosity of the gas phase of the soil pores,
- ∇P [$\text{Pa} \cdot \text{m}^{-1}$]: pressure gradient, and
- the negative sign indicates flow from high to low pressure.

The total radon flux from soil to atmosphere is the sum of diffusive and advective contributions:

$$J_{\text{total}} = -D_e \nabla C_{Rn} + \left(-\frac{k}{\mu} \nabla P \right) C_{Rn} \quad (1.9)$$

Equation (1.9) provide the way to model the transport of radon in the soil. The model relates the transport of radon in the soil to the physical properties of soil which is the result of

Theoretical background

laminar flow and molecular diffusion. The total soil-to-air radon flux (J_{total}) depends critically on the physical properties of the soil (particularly porosity, tortuosity, and permeability) as well as on additional parameters such as moisture content, temperature, grain size, and radium content. High porosity and permeability enhance both diffusive and advective transport, while increased tortuosity and water saturation hinder gas movement (Chen et al., 1995; Nazaroff, 1992; Nunes et al., 2023; Phong Thu et al., 2020).

1.4.1. Radon transport model

Several models have been proposed to describe radon transport in soil, notably by Wilkening (1990) and Nazaroff (1992). The generalized radon transport model by Nazaroff (1992) integrates diffusion, advection, radioactive decay, and radon production, with partitioning among gas, water, and sorbed phases. This model does not yield a general analytical solution and typically requires numerical methods for solution.

$$\begin{aligned} (\epsilon_a + K\epsilon_w + \rho K) \frac{\partial C_{\text{Rn}}}{\partial t} \\ = \epsilon_a D_e \nabla^2 C_{\text{Rn}} + \frac{k}{\mu} \nabla C_a \cdot \nabla P - \lambda_{\text{Rn}} C_{\text{Rn}} (\epsilon_a + K\epsilon_w + \rho K) \\ + f A_{\text{Ra}} \lambda_{\text{Rn}} \rho \end{aligned} \quad (1.10)$$

Where:

- ϵ_a : Air-filled soil porosity
- ϵ_w : Water-filled soil porosity,
- P [kg.m^{-3}]: Bulk density of soil,
- K [$\text{m}^3.\text{kg}^{-1}$]: Sorption partition coefficient,
- D_e [$\text{m}^2.\text{s}^{-1}$]: Effective diffusion coefficient of radon in soil
- C_{Rn} [Bq.m^{-3}]: Radon concentration in soil air,
- k [m^2]: Intrinsic permeability of the soil,
- μ [Pa.s]: Dynamic viscosity of air,
- P [Pa]: Pressure,
- λ_{Rn} [s^{-1}]: Decay constant of radon,
- f : Radon emanation coefficient,
- A_{Ra} [Bq.kg^{-1}]: Radium content.

Theoretical background

The right-hand side of Equation (1.10) represents, in order, the diffusive transport, the advective transport, the radioactive decay of radon, and finally, the radon production term. Nazaroff's radon transport model assumes that radon migrates only through the gas phase of soil, with diffusion following Fick's law and advection governed by Darcy's law. The soil is considered homogeneous and isotropic, with constant properties such as porosity, permeability, and diffusivity. Radon is assumed to partition at equilibrium among gas, water, and sorbed phases, with no transport occurring in the liquid or sorbed states. Air is treated as incompressible due to small pressure gradients, and mechanical dispersion is neglected. Radon generation from radium decay is treated as a uniform, steady source, and no chemical reactions are considered.

A simplified version of equation (1.10) is considered by (Wilkening, 1990). It is a steady-state equation describing the behaviour of radon gas in a homogeneous porous material that extends infinitely in one direction:

$$\frac{D}{\varepsilon} \frac{d^2 C_{Rn}}{dz^2} - \frac{1}{\varepsilon} \frac{d(vC_{Rn})}{dz} - \lambda_{Rn} C_{Rn} + \phi = 0 \quad (1.11)$$

Where:

- $D [m^2.s^{-1}]$: bulk diffusion coefficient,
- ε : soil porosity
- $v [m.s^{-1}]$: transport velocity
- $\lambda_{Rn} [s^{-1}]$: radon decay constant,
- $\phi [Bq.m^{-3}.s^{-1}]$: radon production rate.

The terms from right to left in equation (1.11) represent diffusion transport, advection transport, radioactive decay, and production rate of radon, respectively. Under purely diffusive conditions with boundary conditions ($C_{Rn} = 0$ at $z = 0$ and $C_{Rn} = \frac{\phi}{\lambda_{Rn}}$ at $z \rightarrow -\infty$) the solution of (1.10) becomes:

$$C_{Rn}(z) = \frac{\phi}{\lambda_{Rn}} \left[1 - \exp \left(-z \sqrt{\frac{\varepsilon \lambda_{Rn}}{D}} \right) \right] \quad (1.12)$$

This relationship illustrates that radon concentration increases exponentially with depth, reaching equilibrium after a few meters. The molecular diffusion of radon is typically limited to a depth of a few meters of surface soil, depending on the diffusion coefficient and porosity.

If the velocity profile $v(z)$ is known (e.g., from Darcy's law), the total radon flux is:

$$J = -D \frac{dC_{Rn}}{dz} + v \cdot C_{Rn} \quad (1.13)$$

1.5. Radon exhalation

The radon exhalation rate quantifies the rate of radon gas released per unit area and time from the surface of a porous material (e.g., soil, rock, or building material) into the atmosphere. It is governed by the interplay of three key processes: (1) the decay of radium to radon, (2) the emanation of radon into pore spaces, and (3) the transport of radon through the material via diffusion and advection (UNSCEAR, 2000). The radon exhalation rate is measured in $[Bq \cdot m^{-2} \cdot s^{-1}]$. The radon exhalation rate mainly depends on the concentration of uranium and radium in the soil and soil properties such as grain size, porosity, permeability, and moisture contents (Appleton, 2007; Tchorz et al., 2018). Environmental factors such as moisture content, temperature, and atmospheric pressure modulate the radon exhalation rate by altering the emanation and diffusion of radon. For instance, moisture saturation reduces diffusion of radon by blocking pore pathways, while temperature gradients enhance diffusive transport (Sakoda et al., 2011). The exhalation of radon from the soil is an important phenomenon that significantly affects the radon concentrations in both indoor and outdoor environments (Čeliković et al., 2022).

Different methods have been developed to measure the radon exhalation rate from the soil directly or estimate it indirectly. Most of these methods are based on the accumulation or build-up of radon in a closed chamber (**Fig. 1.3**). The direct measurement of radon flux is performed in the so-called accumulation chamber placed on the surface of the soil, and the flux is calculated based on the following relation (Jonassen, 1983):

$$C_{Rn}(t) = C_0 \exp(-\lambda t) + \frac{\varphi S}{V\lambda} (1 - \exp(-\lambda t)) \quad (1.14)$$

Where:

- $C_{Rn}(t) [Bq \cdot m^{-3}]$: radon concentration at time t ,
- $C_0 [Bq \cdot m^{-3}]$: initial radon concentration at time $t = 0$,
- $\lambda_{Rn} [s^{-1}]$: radon decay constant,
- $\varphi [Bq \cdot m^{-2} \cdot s^{-1}]$: radon exhalation rate,
- $S [m^2]$: horizontal cross-sectional surface area of the accumulation chamber, and
- $V [m^3]$: accumulation chamber volume.

Radon concentrations can be measured using various techniques, including integrated, continuous, and instantaneous sampling methods.



Fig. 1.3 Accumulation chamber setting for the measurement of radon and CO₂ fluxes at the campus of the Faculty of Mathematics, Physics, and Informatics, Comenius University (FMPI CU) in Bratislava, Slovakia.

1.6. Outdoor radon

Once exhaled from the soil, radon behaves as an inert, radioactive gas subject to atmospheric transport, turbulent mixing, and radioactive decay. Although its molecular weight (222 amu) is approximately five times greater than that of air, radon does not gravitationally settle. Instead, it is efficiently mixed by turbulent eddies and wind-driven advection (Cothorn & Smith, 1987; Williams et al., 2011). Because it is chemically inert, its only significant removal process in the atmosphere is radioactive decay.

Outdoor radon concentrations are significantly lower than indoor values, typically averaging around 10 Bq·m⁻³ globally, but can range from 1 to 100 Bq·m⁻³ depending on location and meteorological conditions (Čeliković et al., 2022). Despite these relatively low levels, outdoor radon plays an important role in atmospheric physics, such as tracing boundary layer dynamics, validating dispersion models, and serving as a natural tracer for surface exchange processes. Radon's distribution in the atmosphere is governed by four key physical processes:

1. **Advection** – transport by wind,
2. **Turbulent diffusion** – vertical and horizontal mixing,
3. **Radioactive decay** – with a half-life of 3.82 days for ²²²Rn,
4. **Surface emission** – flux from soils and rocks.

Theoretical background

These processes are described through mathematical transport models that range from simple analytical formulations under steady-state conditions to complex numerical models accounting for spatiotemporal variability.

A one-dimensional transport model which can be applied to the radon is given by (Minato, 1988) as follow:

$$\frac{\partial C_{Rn}}{\partial t} + U \frac{\partial C_{Rn}}{\partial l} = \frac{\partial}{\partial z} \left(K \frac{\partial C_{Rn}}{\partial z} \right) - \lambda_{Rn} C_{Rn} + S \quad (1.15)$$

Where:

- C_{Rn} [Bq·m⁻³]: Radon concentration,
- T [s⁻¹]: Time
- U [m.s⁻¹]: horizontal wind speed (advection),
- l [m]: distance which radon travelled,
- z [m]: height from the ground,
- K [m².s⁻¹]: vertical diffusion coefficient,
- λ_{Rn} [s⁻¹]: Radioactive decay constant of radon ($\sim 2.1 \times 10^{-6}$ s⁻¹),
- S [Bq·m⁻³.s⁻¹]: Radon source flux from soil or surface.

In the original study by (Minato, 1988), this equation was used with steady-state condition to estimate the monthly mean values of outdoor radon.

Wilkening (1990) developed a vertical diffusion–advection–decay model that provides an analytical solution to the vertical distribution of radon under steady-state atmospheric conditions. The model assumes radon is transported vertically by eddy diffusion and vertical wind and is simultaneously removed via radioactive decay. The governing equation is:

$$\frac{\partial C_{Rn}}{\partial t} = \frac{\partial}{\partial z} \left(K \frac{\partial C_{Rn}}{\partial z} \right) - w \frac{\partial C_{Rn}}{\partial z} - \lambda C_{Rn} \quad (1.16)$$

Where:

- C_{Rn} [Bq.m⁻³]: radon concentration in the atmosphere,
- z [m]: height above the ground,
- K [m².s⁻¹]: vertical diffusion coefficient,
- w [m.s⁻¹]: vertical wind speed,
- λ_{Rn} [s⁻¹]: radon decay constant.

Assuming a steady state condition ($\partial C_{Rn}/\partial t = 0$) and constant value of K , the integration of equation (1.15) gives the decrease of radon concentration with increasing height:

$$c = c_0 \exp \left[- \left(\frac{\lambda}{K} \right)^{1/2} z \right] \quad (1.17)$$

This solution indicates that radon concentration decreases exponentially with increasing height above the ground. The model is especially useful for describing nocturnal boundary layer behaviour or stable stratification conditions, where vertical mixing is suppressed, and radon accumulates near the surface. This model does not include horizontal processes, assumes constant eddy diffusivity, and neglects time-dependent dynamics, making it more suitable for idealized or short-range vertical studies (Wilkening, 1990).

1.6.1. Temporal variation of outdoor radon

Outdoor radon concentration varies across different timescales, including diurnal, synoptic, seasonal, and annual-reviewed in (Čeliković et al., 2022). Diurnal variations of radon are mainly governed by the diurnal variation of meteorological parameters (e.g., temperature, wind velocity, precipitation, snow cover) and atmospheric mixing processes. A typical diurnal variation of RAC is shown in **Fig. 1.4**. Usually, the diurnal cycle of RAC has a sinusoidal shape, with the maxima occurring in the early morning and the minima in the late afternoon. This behaviour can be well explained by the changes in atmospheric stability and mixing. Under clear and calm conditions, stable nocturnal stratification leads to the formation of a temperature inversion layer that traps radon near the ground, resulting in peak concentrations just before sunrise (Čeliković et al., 2022). As solar heating begins, convective turbulence develops, enhancing vertical mixing in the boundary layer and diluting radon concentrations. This mixing generally causes radon levels to reach a minimum by late afternoon or early evening. The diurnal cycle is most distinct on cloudless, low-wind days with large temperature gradients (e.g., Chambers et al., 2015). Assuming a constant radon flux, Jacobi et al. (1963) observed minimum RACs during strong atmospheric mixing conditions. In contrast, during weak atmospheric mixing conditions, the authors observed radon concentrations that were 100 times higher. The diurnal variation of outdoor radon is more pronounced on cloudless summer days with large temperature gradients than on cloudy days with smaller temperature gradients (Sesana et al. 2003). The diurnal variation of RAC is also influenced by the geological location of the measurement station. For instance, radon concentration often varies greatly in river valleys due to strong nocturnal inversions, whereas

Theoretical background

in the surrounding hillside areas, radon concentration varies much less due to weaker inversions during the night (Porstendorfer et al. 1994).

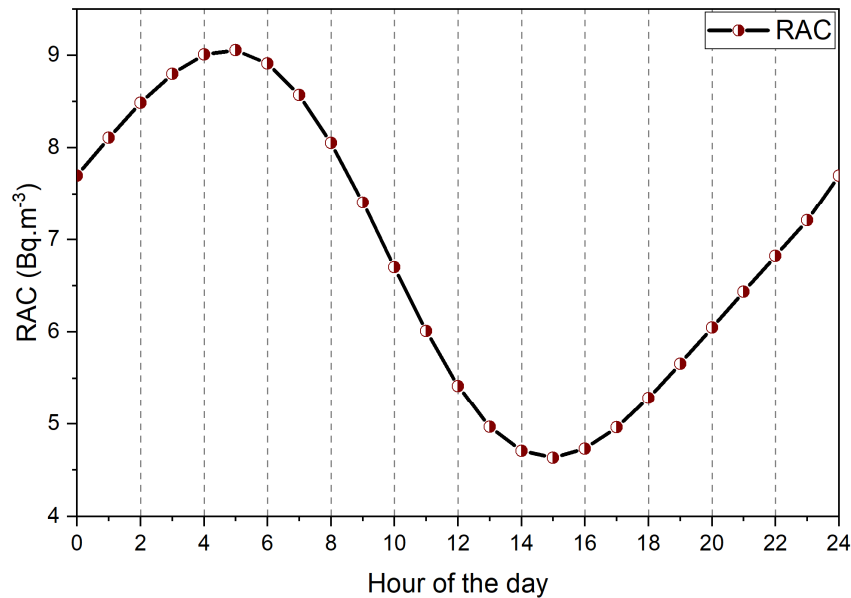


Fig. 1.4 An example of the composite diurnal cycle of RAC measured at FMPI CU Bratislava in May 2022.

Outdoor radon concentrations at a given measurement site can show varying seasonal patterns based on local weather conditions. These seasonal variations are primarily influenced by changes in nighttime duration, soil radon emissions, and atmospheric mixing and meteorological factors. In many mid-latitude regions, higher radon levels are often observed during the colder seasons (e.g., Sesana et al., 2003; Oikawa et al., 2003; Özen et al., 2018), largely due to longer nights and more frequent temperature inversions that suppress vertical mixing and allow radon to accumulate near the surface. One long-term study, for example, found that fall–winter radon averages exceeded spring–summer values by approximately 20–30% in a valley city, attributing this pattern to persistent nocturnal stability and inversion layers (Kubiak & Zimnoch, 2022). However, some studies have reported the lowest radon levels in winter, linking this to mitigating factors such as snow cover, frozen ground, high humidity, and stronger winds that reduce both radon flux and accumulation (Žunic et al., 2007). Snow and frost act as insulating barriers, significantly suppressing radon exhalation from the soil (Čeliković et al., 2022). Conversely, in summer, drier soils and vegetation transpiration may increase radon emission, although these effects are often offset by stronger convective mixing, shorter nights, and higher solar angles, which tend to dilute radon concentrations. Field campaigns have noted that the largest diurnal amplitudes often occur in summer due to strong day–night contrasts, even if average

concentrations are not at their peak (Sultani et al., 2024). Ultimately, the net seasonal pattern results from the interplay between meteorological stability and soil emission dynamics. As summarized by Zimnoch et al. (2014), three key factors influence seasonal radon variation: (a) the seasonal variation in atmospheric stability, with prolonged inversion periods typically occurring in winter and autumn; (b) the seasonal changes in radon exhalation rate from the soil; and (c) the so-called fetch effect, referring to the origin of air masses reaching the site. For instance, maritime air masses generally carry significantly lower radon levels compared to continental ones, as radon flux over the ocean is 2–3 orders of magnitude lower than that over land (Chambers et al., 2015).

1.6.2. Spatial variation of outdoor radon

Outdoor radon concentrations exhibit significant spatial variability due to a complex interplay of geographic, geologic, topographic, and land-use factors (Petermann & Bossew, 2021; Čeliković et al., 2022). Coastal regions typically display lower radon levels compared to inland areas; a phenomenon attributed to the reduced radon flux from oceanic surfaces relative to terrestrial environments (Aquilina & Fenech, 2019). This maritime influence is well-documented in regions such as Ireland and Germany, where coastal zones consistently exhibit lower radon concentrations than adjacent inland areas (Gunning et al., 2014; Kümmel et al., 2014). Conversely, elevated radon levels are frequently observed in inland regions underlain by uranium-rich bedrock, such as granite formations or volcanic soils, where enhanced radon generation and exhalation occur (Križman & Stegnar, 1992).

Topography exerts a strong control on local radon dynamics. Valleys and basins often act as radon traps during stable atmospheric conditions, particularly under nighttime temperature inversions and cold air pooling, whereas elevated terrains such as hilltops benefit from increased atmospheric mixing and lower concentrations (Križman & Stegnar, 1992). For instance, continuous monitoring in Slovakia revealed divergent radon trends between flatland and hilly sites under identical meteorological conditions, underscoring terrain-mediated modulation of local weather patterns and radon behaviour (Holý et al., 2016).

Chapter 2

2. Material and methods

This chapter discusses the methods for measurement of outdoor radon and its progeny, followed by their potential applications as tracers in atmospheric research. Additionally, it outlines the data analysis techniques used in this study.

2.1. Radon measurement techniques

The measurement of radon in outdoor environments can be conducted through various methodologies, which are broadly categorized based on the principles of detection and operational requirements. Radon can be measured either directly, or indirectly by detecting its decay products. Radon and some of its progeny (^{218}Po , ^{214}Po , ^{210}Po) emit alpha particles during radioactive decay, while other progeny such as ^{214}Pb , ^{214}Bi , ^{210}Pb , ^{210}Bi are beta emitters, often accompanied by gamma radiation. The techniques and instruments employed for radon measurement are primarily designed to detect these alpha, beta, or gamma emissions. The selection of an appropriate measurement method depends on several factors, such as the type of radiation being detected, the duration of measurement, the portability of the equipment, its applicability in field conditions, and cost considerations. Radon measurement methods can be classified into two main categories:

Based on electricity consumption for operation: This classification divides radon measurement techniques into active and passive methods.

Based on measurement duration or sampling type: This classification includes grab sampling, continuous monitoring, and integrated sampling methods.

Classification Based on Electricity Consumption

The radon measurement techniques classified based on the electricity consumption i.e., active methods and passive methods is shown in **Fig. 2.1**.

Active Methods: Active radon measurement techniques require electrical power to operate and are typically suitable for short-term measurements. These devices provide real-time measurements by simultaneously sampling and analysing radon or its progeny in air samples. This approach is particularly advantageous in scenarios where radon activity concentrations

Experimental background

exhibit significant and rapid fluctuations. Active methods are also useful for determining average RAC and assessing the impact of ventilation on indoor radon levels. Commonly used active instruments include scintillation chambers and commercial detectors such as AlphaGUARD and RAD7 (Bertin Technologies, n.d.; DurrIDGE, n.d.).

Passive Methods: In contrast, passive radon measurement techniques do not require electrical power and are well-suited for long-term field measurements. These methods are particularly valuable for studying the combined effects of seasonal variations, weather conditions, and environmental factors on radon levels. Long-term integrated measurements are also preferred for estimating annual average radon concentrations and assessing potential health risks to humans. Prominent examples of passive techniques include solid-state nuclear track detectors (SSNTDs), charcoal detectors, and thermoluminescence detectors.

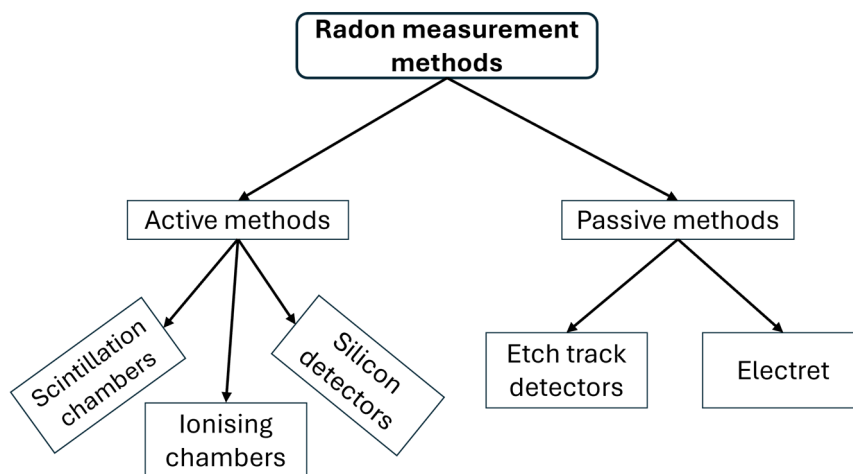


Fig. 2.1 Radon measurement techniques based on electricity consumption for operation.

Classification based on measurement duration or sampling type

Instantaneous/Grab Sampling: This method involves collecting a large number of samples over a very short period, typically in the order of minutes. The samples are subsequently analysed in a laboratory. Grab sampling is commonly used to assess radon levels in groundwater, freshwater, and air samples at specific locations.

Continuous Monitoring: Continuous radon measurement devices sample and count radon concentrations in the atmosphere in real time. These sensitive instruments provide data that can be utilized for the study of environmental processes.

Integrated Sampling: This technique is employed to determine the cumulative radon concentration over an extended period, ranging from months to years. Integrated sampling

Experimental background

is particularly useful for assessing large-scale monthly or annual average radon levels. Passive detectors are the most widely used instruments for this purpose.

This study employs continuous measurement of outdoor radon and its progeny at FMPI CU in Bratislava, Slovakia. The following sections present a detailed overview of the instrumentation and methodologies used for the measurement of radon and its progeny.

2.2. Radon and its progeny sampling

Radon and its decay products (^{218}Po , ^{214}Pb , ^{214}Bi , ^{210}Pb) were monitored at the Faculty of Mathematics, Physics, and Informatics (FMPI) campus of Comenius University in Bratislava, Slovakia (Latitude: $48^\circ 9' 4''\text{N}$; Longitude: $17^\circ 4' 14''\text{E}$) (**Fig. 2.2**). This campus is situated about 3 km northwest of the city centre and is surrounded by urban infrastructure. Bratislava, an industrial hub in southwestern Slovakia, has a population of around 500,000. The city experiences a continental climate, characterized by warm, humid summers and cold winters. Average air temperatures range from approximately -5°C in winter, with January being the coldest month, to around 20°C in summer, with July typically being the warmest.



Fig. 2.2 Sampling location: a) Location of the city of Bratislava, Slovakia within Europe; b) Monitoring stations in Bratislava from which the data were collected (Sultani et al. 2023, 2024).

2.3. Continuous measurement of radon

The schematic view of radon measurement system is shown in **Fig. 2.3**. The air was sampled from a height of 1.5 m through a PVC pipe and pumped to the detection system at a flow rate of $\sim 0.5\text{ L min}^{-1}$. A 10 L delay volume was incorporated in the intake lines to ensure that the air entering the chamber will no longer contain thoron (^{220}Rn) with a half-life of 56 s. Detected count rates were automatically recorded and stored in a computer's memory.

Experimental background

To protect the detector from humidity, two columns were fitted to the intake line: a freezing column that was kept at a temperature of $-20\text{ }^{\circ}\text{C}$ and a column filled with silica gel, trapping residual humidity in the air samples. Right before entering the scintillation chamber, the air passed through a membrane filter that trapped aerosols and radon daughters. This chamber has a volume of 4.5 L. The inner walls of this cylindrical chamber are coated with ZnS (Ag). The inner space of the chamber is divided into 9 sectors to achieve good geometrical conditions for efficient radon detection. The scintillations were collected through two glass windows placed at the far ends of the chamber, and the emitted light signals are collected by two TESLA 65 PK 423 photomultipliers with a 130 mm diameter. The analysed air first entered each external sector, continued to the central sector and finally exited the chamber. The chamber has a flow rate of 0.5 L m^{-1} . Both halves of the chamber function as independent detectors and the photomultiplier signals are summed and processed in a standard manner. Detected counts are automatically recorded and stored in a computer's memory. Radon detection efficiency is 1.7 counts per 1 decay of ^{222}Rn . The lower limits of detection at the 95% confidence level (Currie, 1968; Zahorowski et al., 2004) is 2.8 Bq.m^{-3} (background count rate is 0.0166 s^{-1} and detector sensitivity are $0.0026\text{ s}^{-1}/1\text{ Bq.m}^{-3}$). This measurement system allows obtaining almost 80% of radon activity concentration data in the outdoor atmosphere with an error less than 30% (Putman, 1962).

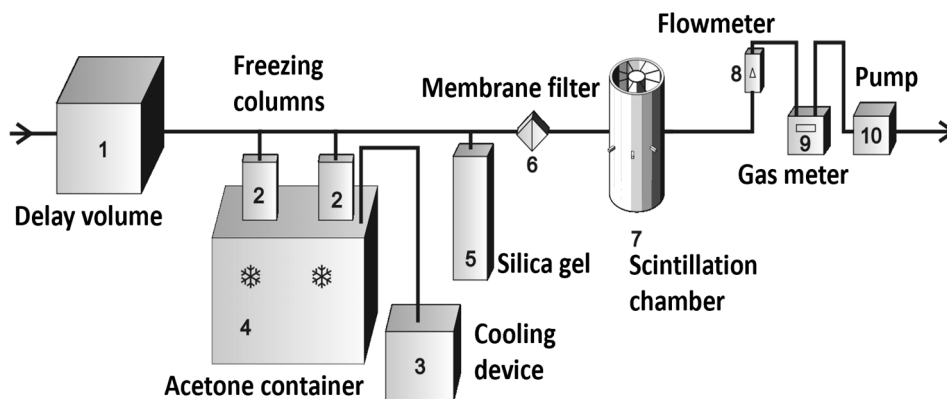


Fig. 2.3 Schematic diagram of sampling instrumentation for continuous measurement: 1) Delay volume, 2) Freezing columns kept at $-20\text{ }^{\circ}\text{C}$ 3) Freezing device, 4) Freezing liquid, 5) Silica gel column, 6) Membrane filter, 7) Scintillation chamber, 8) Flow meter, 9) Gas meter, 10) Air pump (Bulko, 2010).

Subsequently, the RAC, belonging to 2-h intervals, are calculated using the Ward and Borak method (Ward et al., 1991). This method is based on the determination of the so-called normalized detector response function, which characterizes the response of the detector to radon-laden air over time. This response function incorporates the effect of all detector

Experimental background

parameters and operating conditions such as flow rate, humidity, counting efficiency, detector volume and plate-out effect, thus eliminating the need to estimate these effects independently. The detector calibration procedure is as follows: for the first measurement interval (2 hours in our case), air with a known RAC is drawn through the detector. The detector counts rates in this interval increase rapidly due to the increased RAC within the sensitive volume of the detector as well as due to associated buildup of radon progeny within the volume and at inner surface of the detector. At subsequent intervals, radon-free air is passed through the detector. In these subsequent intervals, detected count rates gradually decrease to background levels as ^{222}Rn , its decay products are flushed from the detector volume and radon progeny deposited on the detector inner walls are allowed to decay. The calibration outputs are the τ_i coefficients; their mathematical derivation is described in detail in the original study (Ward et al, 1991).

The resulting formula for calculating the radon activity concentration $C(t)$ is:

$$C(t) = \tau_0(N_0 - N_B) - \sum_{i=1}^m \tau_i(N_i - N_B), \quad (2.1)$$

where N_0 are the detected gross counts in the last measuring interval, N_B is the detector background rate, N_i are the detected gross counts in the previous intervals and τ_i are the inverting coefficient determined as follows:

$$\begin{aligned} \tau_0 &= \frac{1}{\varphi_0}, \\ \tau_1 &= \tau_0(\tau_0\varphi_1), \\ \tau_2 &= \tau_0(\tau_0\varphi_2 - \tau_1\varphi_1), \\ \tau_3 &= \tau_0(\tau_0\varphi_3 - \tau_1\varphi_2 - \tau_2\varphi_1), \\ &\vdots \\ &\vdots \\ \tau_m &= \tau_0(\tau_0\varphi_m - \tau_1\varphi_{m-1} - \dots - \tau_{m-1}\varphi_1). \end{aligned} \quad (2.2)$$

φ_i is the coefficient of the nonlinear detector response function. Since the absolute values of τ_i coefficients for higher values of i approach zero, it is sufficient to consider up to 3 preceding intervals, e.g. $m = 3$. The uncertainty corresponding to each count is calculated based on the following formula:

$$\sigma_c^2 = \sum_{i=0}^m N_i^2 \sigma_{\tau_i}^2 + \sum_{i=0}^m \tau_i^2 (N_i + N_B) + N_B \sum_{i=0}^m \tau_i^2 \quad (2.3)$$

The uncertainties of the first four inverting coefficients appearing in equation (2.3) can be calculated using the following relations:

$$\begin{aligned} \sigma_{\tau_0}^2 &= \left(\frac{1}{\varphi_0^2} \sigma_{\varphi_0} \right)^2, \\ \sigma_{\tau_1}^2 &= (2\tau_0\varphi_1)^2 \sigma_{\tau_0}^2 + \tau_0^4 \sigma_{\varphi_1}^2, \\ \sigma_{\tau_2}^2 &= (2\tau_0\varphi_2 - \tau_1\varphi_1)^2 \sigma_{\tau_0}^2 + (\tau_0\varphi_1)^2 \sigma_{\tau_1}^2 + (\tau_0\tau_1)^2 \sigma_{\varphi_1}^2 + \tau_0^4 \sigma_{\varphi_2}^2, \\ \sigma_{\tau_3}^2 &= (2\tau_0\varphi_3 - \tau_1\varphi_2 - \tau_2\varphi_1)^2 \sigma_{\tau_0}^2 + (\tau_0\varphi_2)^2 \sigma_{\tau_1}^2 + (\tau_0\varphi_1)^2 \sigma_{\tau_2}^2 + (\tau_0\tau_2)^2 \sigma_{\varphi_1}^2 \\ &\quad + (\tau_0\tau_1)^2 \sigma_{\varphi_2}^2 + \tau_0^4 \sigma_{\varphi_3}^2. \end{aligned} \quad (2.4)$$

The reliability of the detectors was tested by simultaneous measurements of RAC by LSCH with another independently calibrated scintillation chamber (SSCH) as well as by a commercially available radon detector AlphaGUARD, which served as an independent reference standard. The comparison took place during a period of two weeks in the summer of 2018 in a well-ventilated room, with RAC ranging from 10 to 70 Bq.m⁻³ and exhibiting distinctive diurnal cycles. The time series of RAC obtained by all three detectors were almost identical; the corresponding linear regression coefficients for all combinations of detector pairs were at the level of $R^2 > 0.90$ (Fig. 2.4).

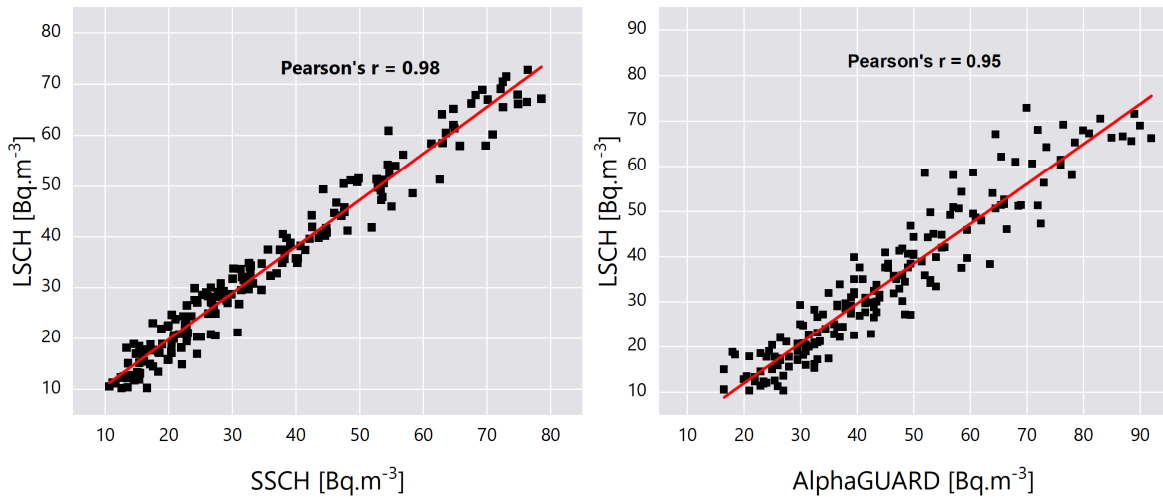


Fig. 2.4 Comparison of RAC measured by scintillation chambers and AlphaGUARD.

2.4. Radon measurement using AlphaGUARD

In this study, the commercial radon detector called AlphaGUARD is also used for monitoring atmospheric radon. AlphaGUARD contains a cylindrical pulse ionization chamber with an active volume of 0.56 L (Bertin Technologies, n.d.). The AlphaGUARD has a linear response ranging from 2 Bq. m⁻³ to 2 MBq. m⁻³ of radon concentration. This device can also monitor air temperature (−10 to + 50), air pressure (700 mbar to 1100 mbar), and air humidity (varying from 0% to 99%). The AlphaGUARD can run for 10 days on its internal battery in either flow or diffusion mode, and collect data in 1, 10 and 60-min intervals. The schematic view of AlphaGUARD is shown in **Fig. 2.5**.

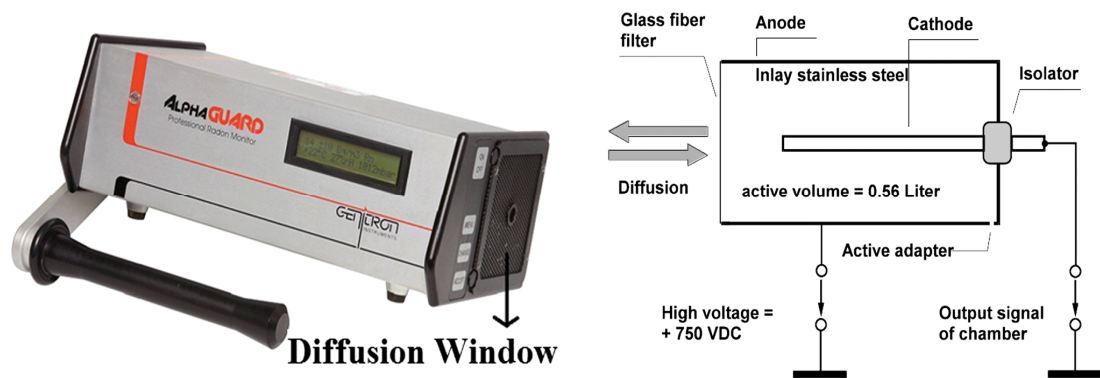


Fig. 2.5 AlphaGUARD portable radon monitor (left) and its schematic view (right) (Pant et al., 2016).

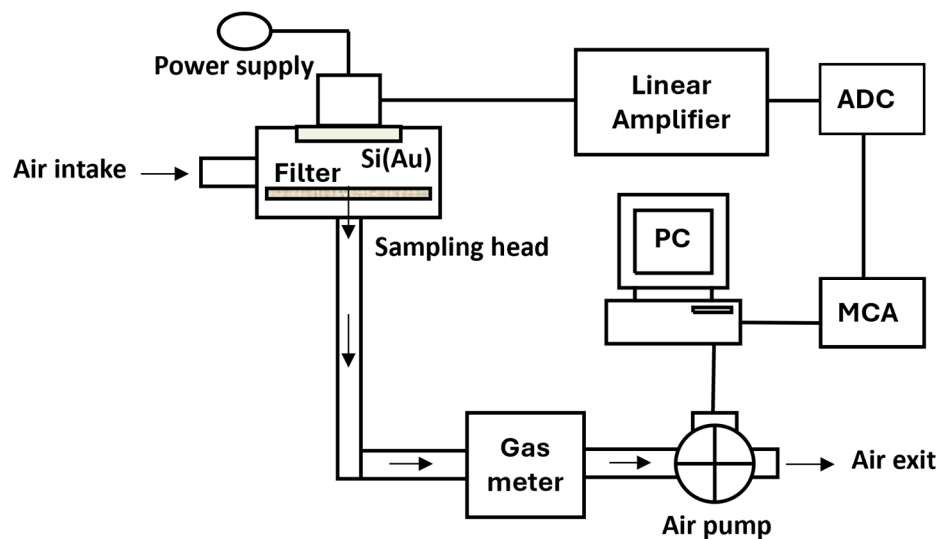


Fig. 2.6 The schematic diagram of the radon progeny measurement system.

2.5. Radon progeny measurement

For the measurement of short-lived radon decay products (^{218}Po , ^{214}Pb , ^{214}Bi), an alpha spectrometric method was used (Stanys et al., 1997). The sampled air was passed through an RW 19 filter with a diameter of 47 mm and a pore size of $1.2\ \mu\text{m}$ at a flow rate of about $35\ \text{L}\cdot\text{min}^{-1}$. The alpha particles emitted by the radon decay products trapped on the filter are registered by an Si (Au) planar detector with a sensitive area of $450\ \text{mm}^2$. The detector and the filter are located in the sampling head (**Fig. 2.6**). The measurement time of one sample is 2 hours; this period is divided into three intervals. During the first interval, which lasts 20 minutes, the air is passed through the filter while the activity of the filter is measured. The subsequent second and third intervals (of 50 and 70 minutes, respectively) consist of spectrometric measurements of the filter without any air circulation. For ^{218}Po , ^{214}Pb and ^{214}Bi , the lower limit of detection in the outdoor atmosphere is about $0.4\ \text{Bq}\cdot\text{m}^{-3}$ for a measurement uncertainty of 30 %. The air entering each of these two detection systems was sampled at a height of 1.5 m above the ground. In total, twelve values representing the concentration of ^{218}Po , ^{214}Pb and ^{214}Bi were obtained each day (i.e. one value every 2 hours).

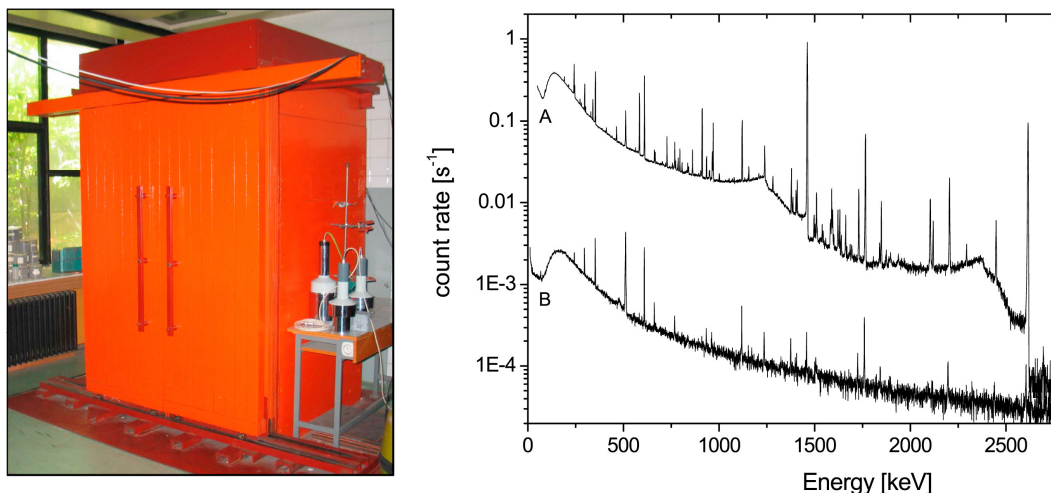


Fig. 2.7 Picture on the left: Low-background shield containing the PGT HPGe detector. Chart on the right: Comparison of gamma-ray background spectra measured by the PGT HPGe detector when placed outside (A) and inside (B) the low-background shield (Sýkora et al., 2017).

2.6. ^{210}Pb measurement

The aerosol samples for the measurement of ^{210}Pb were collected using a high-volume sampler operating at a flow rate of approximately $80\ \text{m}^3\cdot\text{h}^{-1}$, positioned 2.85 meters above the ground. Further details on the sampling method can be found in (Povinec et al., 1988),

Sýkora et al. (2017), and Sýkora & Povinec, (2020). Nitrocellulose membrane filters (PRAGOPOR 4) with 0.85 μm pores ensuring nearly 100% particle collection efficiency were used for sampling. Each sampling period lasted about one week, during which between 11,000 and 15,000 m^3 of air was filtered. Adjustments for air temperature and atmospheric pressure were applied to accurately determine the volume of air processed. The exposed filters were analysed in the low-background gamma spectrometry laboratory located in the faculty's basement using two HPGe detectors: a Canberra GX 4020 and PGT IGC65-DI 845 (Fig. 2.7).

2.7. The potential use of radon as a tracer

2.7.1. Boundary layer height

Boundary layer height (BLH) is the lowest part of troposphere which is directly influenced by the interaction with the Earth's surface and responds to surface-driven forces within short time scales, typically an hour or less (Stull, 2012). The troposphere is the Earth's lowest atmospheric layer, containing approximately 80% of the atmosphere's total mass. It extends from the Earth's surface to an altitude of about 11 kilometres and is characterized by a gradual decrease in temperature with increasing altitude. The depth of the boundary layer, also known as mixing layer height (MLH) varies temporally and spatially and ranges from hundreds of meters to kilometres. It is one of the key factors that governs the dispersion and distribution of atmospheric compounds. BLH serves as a valuable indicator of dispersion conditions and vertical mixing within the lower troposphere, offering insights into the degree of pollution dilution near the Earth's surface (Omori & Nagahama, 2016).

The BLH evolves continuously due to the heating and cooling of the Earth's surface, and it goes through a distinct cycle. Fig. 2.8 presents a schematic view of the BLH evolution in high-pressure zones over land taken from (Stull, 2012). The changes in BLH over time is largely controlled by air temperature and the energy exchange between the Earth's surface and adjacent atmosphere. Daily fluctuations in surface heating and cooling, combined with changes in solar radiation, lead to variations in BLH. As illustrated in Fig. 2.8, the inversion layer at the top of the boundary layer during the daytime acts as a barrier, inhibiting further vertical movement of atmospheric air. Following sunrise, the BLH starts growing, and a so-called convective mixed layer develops as a result of solar heating of the Earth's surface. The convective mixed layer continues to rise throughout the morning, mixing and retaining the air in the upper atmosphere, reaching heights of 1 to 2 km by mid-afternoon (Kaimal &

Experimental background

Finnigan, 1994). From mid-afternoon onwards, the BLH decreases as the intensity of solar radiation diminishes. Just before sunset, a stable (nocturnal) boundary layer (SBL) forms, leaving behind a residual layer (RL); this stable layer persists throughout the night. The RL is formed above the SBL and below the bottom of inversion layer and is not directly affected by the turbulence-generating surface perturbations. Both the RL and SBL are rapidly destroyed after sunrise with the emergence of the mixed layer, particularly on days with intense solar radiation. The nocturnal SBL lacks a clearly defined upper boundary, unlike the daytime mixed layer, which is capped by the inversion layer. The top of the nocturnal layer is typically identified as the height where turbulence intensity diminishes to a small fraction of its value at ground level, since the strongest mechanical turbulence occurs near the ground due to surface friction and wind shear. A common guideline suggests that the height of the nighttime layer can be determined as the level where turbulence drops to approximately 5% of its surface intensity. Alternatively, it can be described as the average elevation of the inversion layers. Therefore, the estimated BLH or MLH, which exhibits a continuous profile, consist of the height of the convective mixed layer during the daytime and as the height of the SBL during the nighttime.

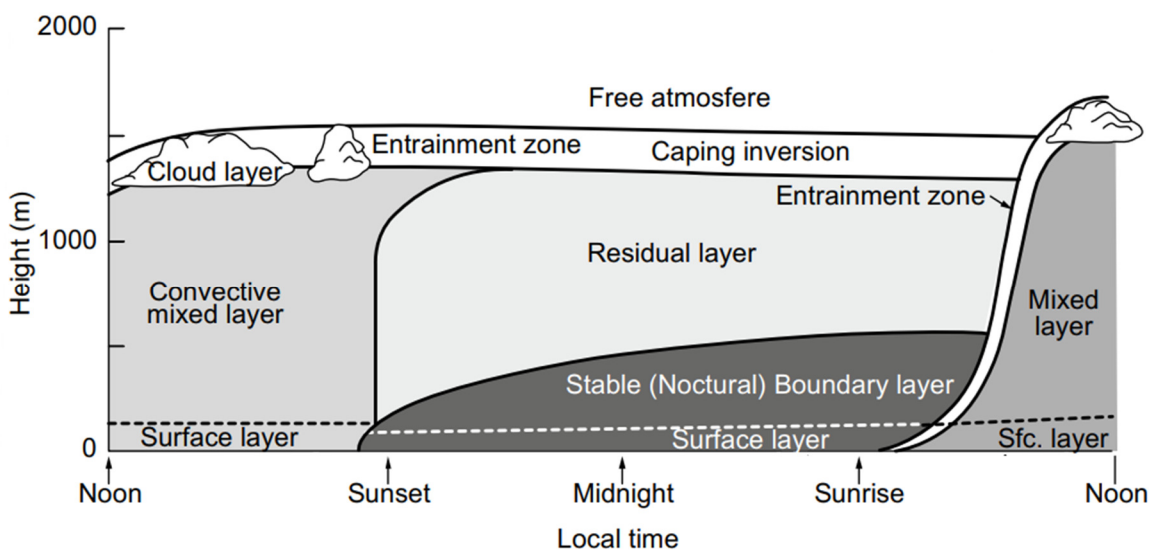


Fig. 2.8 Schematic diagram of the vertical structure of boundary layer height (Stull, 2012).

2.7.2. MLH based on radon

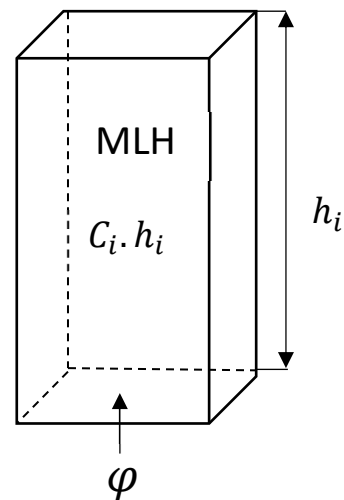
Meteorological stations do not measure the MLH directly; instead, it is estimated using various methods such as meteorological radiosondes, ground-based remote sensing techniques (including sodar, lidar, and Doppler radar), and aircraft surveys, as summarized by Seibert et al., (2000). Despite these advancements, determining the MLH accurately and

Experimental background

automatically with high resolution remains a challenge. One alternative approach involves using outdoor radon to estimate MLH. Radon is considered a valuable tracer in atmospheric studies due to its distinct characteristics: it originates primarily from the soil, has a suitable half-life comparable to the residence time of most of atmospheric compounds, can be measured with high accuracy, and, as a noble gas, it does not undergo chemical reactions. Since its main source is the soil and local radon flux variations are minimal compared to those driven by atmospheric mixing, radon serves as a reliable indicator of atmospheric dynamics. This study employs a so-called box model to determine MLH based on radon concentrations. This model was originally introduced by Fontan et al. (1979); the model was later refined by Sesana et al. (2003) and Vecchi et al. (2018). This method has been effectively applied in several studies (e.g., Allegrini et al., 1994; Salzano et al., 2016; Vecchi et al., 2018; Griffiths et al., 2013).

The schematic representation of the box model is illustrated in **Fig. 2.9**, with the ground surface serving as its base and the mixing layer as its height (h_i). Radon gas, with a concentration C_i and an emission rate φ_{Rn} from the surface, enters the box and is assumed to be uniformly mixed within it.

Fig. 2.9 The schematic view of the box model, having ground surface as its base, and mixing layer as its height (h_i).



This model is based on the mass balance relation, under the following assumptions:

- radon exhalation rate is considered constant during a short time interval (Δt),
- radon concentration varies only as a function of vertical stability (i.e., horizontal variation does not influence the radon concentration),
- the mixing of radon within the box is homogenous,
- radon in the residual layer is conserved.

Experimental background

Bearing in mind these assumptions the budget of layer – integrated trace gas (radon) inside the box can be written as:

$$\frac{\partial(C)}{\partial t} = \varphi_{Rn} - hC\lambda + C^r \frac{\partial h}{\partial t} \quad (2.5)$$

Where:

- $C [Bqm^{-3}]$: measured radon concentration,
- $\varphi_{Rn} [Bqm^{-2}s^{-1}]$: radon exhalation rate,
- $h [m]$: boundary layer height,
- $\lambda [h^{-1}] = 0.0076$ is the radon decay constant,
- $C^r [Bqm^{-3}]$: is the radon concentration in the residual layer.

The residual layer forms as the mixing layer decays in the late afternoon. The radon concentration in residual layer C^r is assumed to be conserved and can be modelled as follows:

- 1) when $\partial h / \partial t > 0$ (MLH increases), the C^r is the concentration of radon in the encroached residual layer which can be modelled as $C = C_{min}e^{-\lambda(t-t_{min})}$, where C_{min} is the measured radon concentration in the late afternoon of the previous day.
- 2) when the $\partial h / \partial t < 0$ (MLH decreases), the $C^r = C$, i.e., the measured radon concentration in the current time t .

Considering the mentioned assumptions, equation (2.5) can be solved analytically for discrete time intervals during which MLH and RAC are assumed to be constant. If we know the MLH at the current time step h_i , and have data on the RAC time series measured at time steps i and $i + 1$, then the MLH in the subsequent time step h_{i+1} can be derived from the following formula:

$$C_{i+1}h_{i+1} = \frac{\varphi_{Rn}}{\lambda}(1 - e^{-\lambda\Delta t}) + C_i h_i e^{-\lambda\Delta t} + C_i^r (h_{i+1} - h_i) e^{-\lambda\Delta t} \quad (2.6)$$

Equation (2.6) for this box model comprises three primary terms: emission, legacy, and encroachment. On the right-hand side of equation (2.6), the first term corresponds to the emission component, where radon is supplied to the box due to its exhalation from the soil. The second term represents the remaining legacy (radon remaining in the box from the previous time step), while the final term accounts for the encroachment effect caused by the expansion or contraction of the MLH during the current time interval. This describes the scenario where, when the MLH is increasing, the mixing layer and the residual layer become

Experimental background

coupled. As a result, the radon concentration remaining in the residual layer from the previous day encroaches into the box. Conversely, when the MLH is decreasing, it decouples from the residual layer, and the radon concentration from the residual layer is no longer added to the box. The box model is initialised in the late afternoon, when the MLH begins to shrink (equation 2.7).

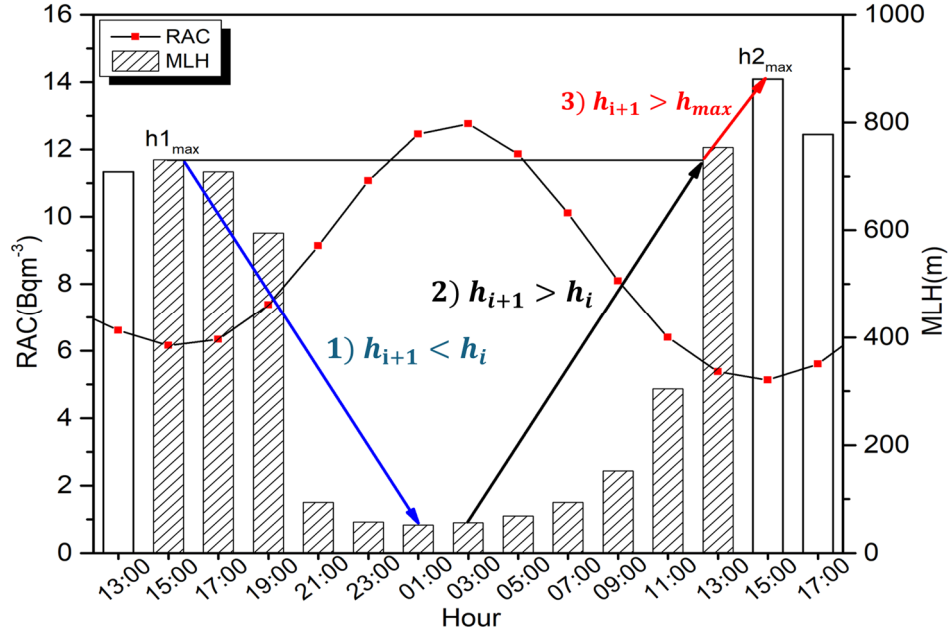


Fig. 2.10 An example of the temporal evolution of MLH based on radon measurements, illustrating the box model procedure. The plot shows MLH values derived for 13:00 on the first day to 17:00 on the second day. $h1_{max}$ and $h2_{max}$ represent the highest MLH on day one and day two, respectively.

To describe the full diurnal cycle of MLH, the following three distinct conditions can be derived from the rearrangement of equation (2.6), also illustrated in **Fig. 2.10**:

1. When the MLH is shrinking ($h_{i+1} < h_i$) (usually in the late afternoon), the residual layer becomes decoupled from the mixing layer ($C_i^r = C_i$):

$$h_{i+1} = \frac{\varphi_{Rn}(1 - e^{-\lambda\Delta t})}{\lambda(C_{i+1} - C_i e^{-\lambda\Delta t})} \quad (2.7)$$

2. when the MLH growing ($h_{i+1} > h_i$) (usually after the sunrise), residual layer containing a volume of air is incorporated into the mixing layer:

$$h_{i+1} = \frac{\frac{\varphi_{Rn}}{\lambda}(1 - e^{-\lambda\Delta t}) + h_i e^{-\lambda\Delta t}(C_i - C_i^r)}{C_{i+1} - C_i^r e^{-\lambda\Delta t}} \quad (2.8)$$

Experimental background

3. when the $h_{i+1} > h_{\max}$ (h_{\max} is the maximum height recorded in the previous day), the condition 2 is modified as:

$$h_{i+1} = \frac{\frac{\varphi_{Rn}}{\lambda} (1 - e^{-\lambda \Delta t}) + h_i C_i e^{-\lambda \Delta t} + (h_{\max} - h_i) C_i^r e^{-\lambda \Delta t}}{C_{i+1}} \quad (2.9)$$

If this modification were not implemented, the model would add radon-free space in the box above the h_{\max} , which would lead to uncertainty in the determination of MLH.

The radon concentration was recorded at 2-hour intervals, specifically for the odd hours of the day (e.g., 1:00, 3:00, 5:00, etc.). To enhance the accuracy of the MLH calculation by increasing the number of data points, the original 2-hour RAC data were interpolated using the cubic spline method to estimate values for the even hours (e.g., 2:00, 4:00, 6:00, etc.). This interpolation ensures that the MLH time series derived from radon data more accurately captures variations in atmospheric conditions, such as shifts in sunrise and sunset times throughout the year. In this research, the daily mean radon flux was obtained from the European radon flux map, which is based on GLDAS-Noah v2.1 soil moisture data (Karstens et al., 2023). Since the model exhibits high sensitivity to RAC fluctuations, the measured RAC values, recorded at 2-hour intervals, were initially smoothed using the Fast Fourier Transform (FFT) to filter out fluctuations with periods shorter than 8 hours. Following this preprocessing step, the MLH was determined using equations (2.7), (2.8), and (2.9).

2.7.3. Aerosols residence time

The residence time of aerosols (T_R) is a crucial factor that indicates the average duration aerosols remain in the atmosphere before settling on the Earth's surface. This metric is essential for assessing the atmospheric concentration of various substances and their potential effects on the lower atmosphere and surface environment (Rangarajan, 1992). Notably, the T_R is inversely related to the atmospheric removal rate constant, a key parameter in atmospheric models that can be used to study the transport and fate of airborne pollutants. Various experimental and modelling approaches have been employed to estimate aerosols residence time. A comprehensive review of the models used for this purpose is provided by (Kristiansen et al., 2016). However, significant uncertainties persist in modelled T_R values due to limited data on atmospheric removal processes. One method relies on atmospheric mixing height and deposition velocity, but its accuracy is constrained by the scarcity of reliable deposition velocity data (Crova et al., 2021). An alternative technique involves using natural radionuclides, such as ^7Be and radon/thoron decay products, as tracers. These

Experimental background

radionuclides attach to fine aerosols, making them effective tracers for studying aerosol dynamics (e.g., Porstendörfer et al., 2000). This method is based on analysing the radioactive disequilibrium between pairs of radionuclides in the decay chain. The degree of disequilibrium provides insights into the atmospheric aerosol's lifetime before deposition. This method has been used extensively to determine the aerosols mean residence time. For example, studies using different radionuclide pairs (e.g., $^{210}\text{Po}/^{210}\text{Pb}$, $^{210}\text{Bi}/^{210}\text{Pb}$, $^{210}\text{Pb}/^{222}\text{Rn}$, $^{210}\text{Pb}/^{214}\text{Pb}$, $^{210}\text{Pb}/^{214}\text{Bi}$, $^{214}\text{Pb}/^{222}\text{Rn}$, $^{214}\text{Bi}/^{222}\text{Rn}$) have reported a wide range of mean T_R , from a few hours to several weeks (Barba-Lobo et al., 2023; Baskaran & E. Shaw, 2001; Mohery et al., 2016; Porstendörfer et al., 2000; Turekian et al., 1977). The choice of radionuclide pair significantly influences the calculated T_R . For instance, ratios involving long-lived radon progeny, such as $^{210}\text{Po}/^{210}\text{Pb}$ and $^{210}\text{Bi}/^{210}\text{Pb}$, tend to yield longer T_R , often in the range of weeks (Długosz et al., 2009; Marley et al., 2000; Papastefanou, 2009). In contrast, ratios combining long-lived and short-lived decay products (e.g., $^{210}\text{Pb}/^{214}\text{Pb}$, $^{210}\text{Pb}/^{214}\text{Bi}$, $^{210}\text{Pb}/^{222}\text{Rn}$) typically result in shorter T_R , ranging from a few days to a week (Crova et al., 2021; Mohery et al., 2016; Sýkora et al., 2017). Notably, ratios of immediate radon progeny, such as $^{214}\text{Pb}/^{222}\text{Rn}$, suggest even shorter residence times, often less than a day (Barba-Lobo et al., 2024). Another independent approach, based on the growth rate of ^7Be aerosols, estimated a residence time of approximately 8 days (Papastefanou, 2009a). The wide variability in reported T_R values highlights the challenge of selecting the most reliable radionuclide pair for accurate estimation. Ratios involving long-lived radon progeny, such as ^{210}Po , ^{210}Pb , and ^{210}Bi , are often less reliable due to the influence of external sources of these radionuclides, including soil dust resuspension, volcanic activity, biomass burning, and fossil fuel combustion (Długosz-Lisiecka & Bem, 2012; Lambert et al., 1983; Papastefanou, 2009a; Poet et al., 1972). Recent findings by Crova et al. (2021) suggest that the $^{210}\text{Pb}/^{214}\text{Bi}$ pair offers a more robust estimate of aerosol residence time compared to methods based on deposition velocity and MLH.

The T_R is calculated based on the ratios of radon decay products e.g., (Turekian et al., 1977; Baskaran & Shaw, 2001; Crova et al., 2021; Mohery et al., 2016; Sýkora et al., 2017; Moore et al., 1973). In this model it is assumed that a set of data obtained from the analysis of air samples is a representative sample of a well-mixed box into which radon is continuously injected. If the radon flux entering the box remains constant over a short period of time, it is possible to obtain the aerosols residence time based on the ratios of radon and its progeny.

Experimental background

The concentration of radon decay products as a function of time can be derived from the following equation:

$$\frac{dN_d}{dt} = \lambda_p N_p - \lambda_d N_d - \lambda_r N_d \quad (2.10)$$

Where N_p and N_d are the parent and daughter products, λ_p and λ_d are their respective radioactive decay constants. The term $\lambda_r N_d$ accounts for the atmospheric removal of the daughter nuclei attached to aerosol particles by all processes, e.g., by scavenging, gravitational settling, or removal by coagulation (all these removal processes together are characterized by the removal constant λ_r). Under steady-state conditions, equation (2.10) yields relation (2.11), which gives the mean aerosols residence time.

$$T_R = \frac{1}{\lambda_r} = \frac{C_d}{\lambda_d(C_p - C_d)} \quad (2.11)$$

The $T_R (= \frac{1}{\lambda_r})$ is the aerosols residence time and C_p and C_d are the activity concentrations of the parent and daughter radionuclides in units of Bq.m^{-3} . This model has certain intrinsic limitations, such as the assumption of steady-state conditions and constant radon flux, which are not always met. However, studies have reported a strong correlation between the ratios of radionuclide pairs used to determine aerosol residence time (e.g., between ^{210}Pb and short-lived radon progeny), suggesting the presence of steady-state conditions (Kim et al., 2000). Additionally, given the small spatiotemporal variability of radon flux, it can be reasonably assumed to remain approximately constant over short periods, as also suggested by Crova et al. (2021). The aerosols residence time calculated in this way can be considered representative for particles with an aerodynamic diameter between 0.1 and 1 μm , since this is the size range of aerosols to which radon progeny preferentially attach (Porstendörfer et al., 2000). The residence time was calculated from relation (2.11) using the weekly ratios of $^{210}\text{Pb}/^{214}\text{Pb}$, $^{210}\text{Pb}/^{214}\text{Bi}$ and $^{210}\text{Pb}/^{222}\text{Rn}$. In the case of the $^{210}\text{Pb}/^{222}\text{Rn}$ pair, relation (2.11) must be adjusted to include the equilibrium factor F_{eq} :

$$T_R = \frac{1}{\lambda_r} = \frac{C_d}{\lambda_d(F_{eq} \cdot C_p - C_d)} \quad (2.12)$$

For this calculation, weekly values of F_{eq} data from the present study were used.

2.8. Conventional data analysis techniques

In this study, a combination of statistical and machine learning (ML) techniques was employed to preprocess the dataset and model radon time series. These methods were selected to effectively capture underlying trends and relationships between RAC and its influencing factors. The analytical approaches include time series analysis, regression and correlation techniques, as well as various machine learning models for regression analysis. Each method was chosen based on its suitability for specific aspects of the data and the objectives of the study. The techniques applied are outlined and discussed in detail below.

2.8.1. Fast Fourier analysis

The Fast Fourier Transform (FFT) was applied to analyse the RAC time series in order to explore its underlying frequency components. FFT is a widely used and efficient algorithm for performing spectral analysis on time series data (Morrison, 1994). This technique transforms the original time-domain signal into the frequency domain, allowing for the identification of periodic patterns and the contribution of various time scales to the overall signal. In this study, the built-in FFT function in OriginLab Pro¹ was used for RAC spectrum analysis. Considering an input series x_i of length N , its discrete Fourier transform is given by:

$$F_n = \sum_{i=0}^{N-1} x_i e^{-\frac{2\pi j}{N} ni}, \quad n = 0, 1, \dots, N-1 \quad (2.13)$$

where F_n represents the complex amplitude of the frequency component at index n with the j being the imaginary unit.

The output of the FFT in OriginLab includes frequency components along with the corresponding complex-transformed values. Additionally, it can yield various analytical results such as magnitude, amplitude, phase, power density, and other related computations.

2.8.2. Autocorrelation function

Autocorrelation Function (ACF) is a mathematical equation used to measure the correlation of a time series with a delayed version of itself over different time lags. It is widely used in time series analysis, signal processing and statistics to understand the repeated trends and pattern. In the other words, the correlation between two points in a time series is called ACF. Mathematically, the ACF for a time series y_t is given by:

¹ Help Online - Origin Help - Fast Fourier Transform

$$\text{Corr}(y_t, y_{t-1}), \quad k = 1, 2, \dots \quad (2.14)$$

The value of k is the time gap being considered between two values of time series and is called lag.

2.8.3. Canonical correlation analysis

Canonical Correlation Analysis (CCA) is a multivariate statistical technique used to examine the relationship between two groups of variables (Uurtio et al., 2017). The core principle of CCA is to generate two composite variables from each set, which effectively summarize the overall correlation between the two groups. These composite variables help quantify the association between the sets of indicators. CCA results are typically represented through canonical correlation, canonical loadings, cross-loadings, and both standardized and unstandardized correlation coefficients. Canonical correlation and loadings provide insight into the overall relationship between the two sets of variables and the individual contribution of each variable to the corresponding canonical variates. In this study, the CCA analysis was conducted using the canonical correlation function in the Statistical Package for Social Sciences (SPSS, Version 26).

2.8.4. Principal component analysis

The Principal Component Analysis (PCA) is a well-established multivariate statistical technique primarily used as an explanatory tool to reveal the complex relationships between multi-high dimensional variables. PCA is particularly effective for exploratory and descriptive analyses of high-dimensional data (Jolliffe, 2002). Its primary strength lies in its ability to perform dimensionality reduction by transforming a set of potentially correlated variables into a smaller set of uncorrelated variables, known as principal components (PC). These components are linear combinations of the original variables, ordered in such a way that the first few components capture most of the variance present in the original dataset, thereby minimizing information loss. The PCA has been used in several radon studies (e.g., Al-Shboul, 2023; Banrion et al., 2023; Sabbarese et al., 2022). In this research, hourly data for all variables, except for wind direction and pressure were subjected to PCA. In this study, the PCA was applied to assess the relationship between RAC and its potential influencing factors.

2.8.5. Multiple linear regression

Multiple Linear Regression (MLR) is a statistical technique used to explore how multiple independent variables collectively influence a single dependent variable (MONTGOMERY & FRIEDMAN, 1993). Building upon simple linear regression, MLR employs the least squares method to determine the relationships among variables. The general form of the MLR model is expressed as:

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 \dots b_nX_n \quad (2.15)$$

where Y is the dependent variables, $b_0, b_1, b_2, b_3 \dots b_n$ are the unknown regression coefficients and $X_1, X_2, X_3 \dots X_n$ are the independent variables (predictors).

Key outputs of MLR typically include:

R value: Measures the direction and strength of the overall linear relationship between predictors and the dependent variable.

Adjusted R²: Reflects the proportion of variance explained by the model, adjusted for the number of predictors used.

Beta (β) coefficients: Indicate the individual effect of each predictor while controlling for others.

Hypothesis tests: These include t-tests for individual predictors and the F-test for evaluating the model; these tests help to assess statistical significance.

By examining these metrics, MLR offers valuable insights into how well the model explains or predicts the dependent variable.

2.8.6. Generalised additive model

The Generalized Additive Model (GAM) is a flexible extension of the Generalized Linear Model (GLM) that allows for non-linear relationships between the response variable and predictor variables. While a GLM models the expected value of the response as a linear combination of the predictors through a specified link function, a GAM replaces this linear combination with a sum of smooth, non-parametric functions. These smooth functions are unspecified and are estimated from the data, with inference focusing on their shapes and effects. Unlike GLM, which assume a linear relationship between the predictors and the response (e.g., βx), GAM allow for complex, non-linear associations by incorporating smooth functions (e.g., $f(x)$). This flexibility makes GAM particularly well-suited for

Experimental background

modelling ecological, environmental, and biomedical data, where relationships between variables are often inherently non-linear (Hastie, 1992). The general form of the GAM is given by:

$$g(E(Y)) = \beta_0 + f_1(x_1) + f_2(x_2) + \dots + f_n(x_n) \quad (2.16)$$

where g is the link function (e.g., identity, log, logit), Y is response variable, β_0 is the intercept, $f_i(x_i)$ are non-parametric smooth functions estimated from the independent variables (x_i).

In this thesis, GAM is employed to infer information about the effects of factors governing RAC. GAM is considered a suitable modelling approach in this context because it can effectively accommodate non-linear relationships between predictor variables and RAC. Unlike traditional linear models, GAMs do not require a fixed functional form for each predictor; instead, they estimate smooth functions from the data, allowing for greater flexibility and a more accurate representation of complex environmental processes. These smooth functions provide valuable insights into the nature and shape of the relationships between predictors and the response variable. GAMs offer several advantages i.e., their flexibility enables them to model non-linear effects without overfitting, and their interpretability allows each predictor's effect to be visualized and analysed individually. Additionally, GAMs maintain the statistical rigor of GLM while expanding their applicability to a wider range of data patterns commonly observed in environmental studies.

2.9. Machine learning in radon studies

ML is a subfield of Artificial Intelligence (AI) that involves algorithms which learn patterns from data to make decisions or predictions, without being explicitly programmed for specific tasks. These models can analyse new data, perform classifications, and generate predictions. A key use of ML is in regression analysis, where it often outperforms traditional statistical techniques. Commonly employed ML models for regression analysis include artificial neural networks (ANN), gradient boosting machines (GBM), extreme gradient boosting (XGBoost), and random forests (RF). These approaches have found extensive use in radon-related studies. Examples include geogenic radon mapping (e.g., Elío et al., 2023; Petermann et al., 2021, 2024; Rezaie et al., 2022, 2023), exploring the variables that influence radon levels (Al-Shboul, 2023; Dicu et al., 2023; Elío et al., 2023; Naskar et al., 2023; Yang et al., 2025), managing incomplete or irregular time series data (Janik et al., 2018), and identifying anomalies potentially linked to seismic activity (e.g., Mir et al., 2021, 2022; Tareen et al.,

2019). In general, ML techniques tend to deliver more accurate and robust results than traditional methods, underscoring their value in radon research and analysis. Some key advantages of ML regression over conventional regression methods include the ability to model complex and nonlinear relationships without requiring explicit specification, allowing for greater flexibility in capturing real-world patterns. Unlike traditional approaches such as MLR, which assume linearity and often rely on assumptions like normality, homoscedasticity, and particularly stationarity in time series data, ML models can operate effectively without these strict prerequisites. This makes ML regression more suitable for dynamic, non-stationary datasets where the statistical properties change over time. Additionally, ML models are capable of automatically detecting interactions and dependencies among variables, and they tend to be more robust to multicollinearity compared to traditional statistical methods and handle high-dimensional or large-scale datasets more efficiently. While conventional methods offer simplicity and interpretability, ML regression provides a powerful and adaptable alternative for more complex and less predictable data like outdoor radon. In this study, we employed the GBM, XGBoost and RF algorithms for regression analysis to identify the factors influencing radon temporal variability. These ML approaches are explained in detail below.

2.9.1. Gradient boosting machine

GBM is a powerful ensemble ML technique, mainly used in regression and classification problems. Rather than constructing a single predictive model, GBM creates an ensemble by incrementally adding multiple weak learners (like decision trees) where each learner is trained to correct the errors made by the previous ensemble (Natekin & Knoll, 2013). This iterative process involves optimizing a differentiable loss function using a form of gradient descent in function space, where new learners are aligned with the negative gradients of the loss function. To enhance predictive performance and mitigate overfitting, several regularization strategies are incorporated, including learning rate reduction (shrinkage), random subsampling of training data, and early stopping based on validation error. Since their foundational development by Friedman (2001), GBM has achieved widespread success across diverse ML applications and are now considered a cornerstone method in predictive modelling.

2.9.2. Extreme gradient boosting

XGBoost is an advanced version of the gradient boosting framework designed to improve performance and scalability in ML tasks (Chen & Guestrin, 2016). It addresses the limitations of traditional gradient boosting by incorporating several key improvements. One of the main improvements in XGBoost is regularisation, which prevents overfitting by adding a penalty term to the objective function and controlling model complexity. XGBoost supports both L1 (Lasso) and L2 (Ridge) regularisation. L1 regularisation promotes sparsity by driving some variable coefficients to zero, effectively performing feature selection and excluding irrelevant features. L2 regularisation, on the other hand, prevents large coefficients, which promotes stability and ensures that the model doesn't become overly sensitive to fluctuations in the data. This combination helps to keep a balance between simplicity and stability, improving the model's ability to generalise to unseen data. Another key feature is the use of second-order optimisation, which takes into account both the first and second derivatives of the loss function. This results in faster convergence and improved accuracy compared to standard gradient boosting, which uses only first-order information. XGBoost also incorporates a post-pruning technique that allows the model to limit the depth and complexity of the tree, thereby reducing overfitting. In addition, the algorithm uses parallel processing and distributed computing, which speeds up training and makes it suitable for large datasets. The flexibility of hyperparameter tuning further enhances XGBoost's adaptability, providing options for adjusting learning rates, boosting rounds and subsampling ratios. This has made XGBoost popular for regression and classification in various fields.

2.9.3. Random forest

The RF algorithm is an ensemble learning method that combines multiple decision trees, where each tree is constructed using a randomly selected subset of data and features (Breiman, 2001). These trees operate independently, following the same probability distribution (**Fig. 2.11**). As the number of trees increases, the overall error stabilizes toward a fixed limit. The generalization error of a random forest depends on both the strength of individual trees and the correlation between them. By aggregating the predictions of multiple trees (i.e., through majority voting for classification or averaging for regression) random forests enhance accuracy while reducing overfitting. They are effective for handling high-dimensional data, identifying important features, and improving predictive performance.

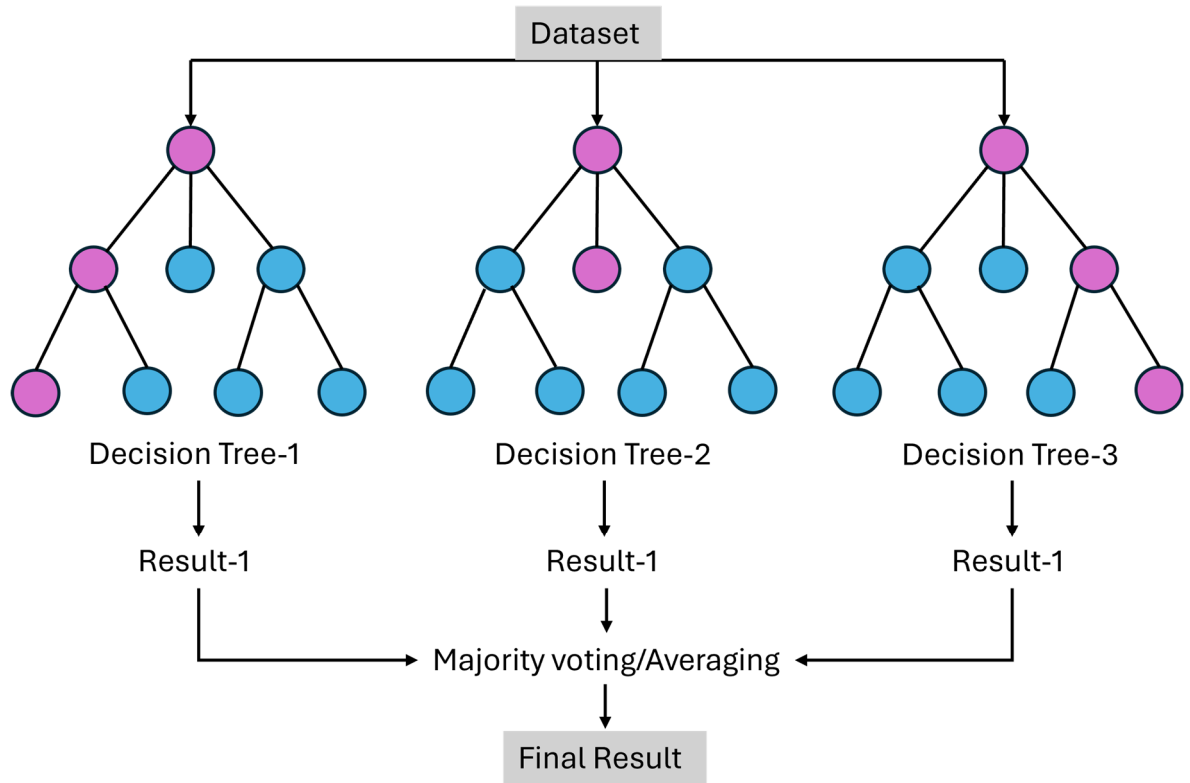


Fig. 2.11 Illustration of random forest trees.

2.10. Regression models evaluation

To compare and evaluate the model's performance, certain metrics were calculated, including mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE) and adjusted coefficient of determination (adj. R^2). These metrics are widely used as standard metrics for evaluation and comparison of regression models in different fields including radon science (Janik et al., 2018). These metrics are described mathematically by the following equations:

MSE measures the average of the squares of the errors:

$$MSE = \frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2 \quad (2.17)$$

where O_i are the observed values, P_i are the predicted values and the n is the number of observations.

RMSE is the square root of MSE, giving error in the same units as the target variable:

$$\text{RMSE} = \sqrt{\text{MSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2} \quad (2.18)$$

The MAE measures the absolute differences:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |O_i - P_i| \quad (2.19)$$

Adjusted R^2 is a modified version of R^2 that adjusts for the number of predictors in the model:

$$\text{adj. } R^2 = 1 - (1 - R^2) \cdot \frac{n - 1}{n - k - 1} \quad (2.20)$$

where R^2 is the coefficient of determination, and k is number of independent variables.

2.11. Data preprocessing

To ensure the accuracy and reliability of the regression and classification models, several preprocessing steps were applied to the dataset. These include outlier detection and treatment, data smoothing, handling missing values, normalization, testing for stationarity, feature selection, feature engineering, and splitting the dataset.

Outlier detection: Outliers were identified manually by visually inspecting time series plots of each variable. Abnormal values, often caused by instrument malfunction, were excluded to maintain data integrity.

Data smoothing: To reduce noise, the RAC data was smoothed using a 4-point FFT. This acts as a low-pass filter, effectively removing high-frequency components (i.e., fluctuations with periods shorter than 8 hours) and preserving longer-term trends.

Normalization: Since some regression models are sensitive to feature scales, all variables were normalized using Min–Max scaling to ensure uniformity and improve model performance.

Stationarity: In time series analysis, stationarity is a key, yet complex, concept. A time series is deemed stationary when its statistical properties (i.e., mean, variance, and autocorrelation) do not change over time. This assumption is essential in many forecasting and regression techniques and its violation can lead to inaccurate or misleading conclusions. For instance, in the presence of non-stationarity, models might incorrectly interpret patterns

Experimental background

like trends or seasonality as meaningful relationships, thereby compromising the reliability of statistical tests. However, environmental datasets including RAC frequently exhibit seasonal patterns and long-term trends indicating possible non-stationarity. To assess the stationarity of a time series, several statistical tests are commonly employed, each based on different assumptions and hypotheses. The Augmented Dickey-Fuller (ADF) test (Dickey & Fuller, 1979) is a widely used method for detecting unit roots in time series data. It tests the null hypothesis that the series is non-stationary due to the presence of a unit root against the alternative that it is stationary. Complementing this, the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test (Kwiatkowski et al., 1992) adopts the opposite hypothesis framework: it tests the null hypothesis that a series is stationary (either around a level or a deterministic trend), with the alternative being that it is non-stationary. Employing both ADF and KPSS tests in parallel allows for a more robust diagnosis of stationarity, as their opposing null hypotheses can reveal different types of non-stationary behaviour. The ADF test was applied to the RAC time series containing six years of hourly data. The null hypothesis of the ADF test is that a unit root is present in a time series sample. The alternative hypothesis is usually stationarity or trend-stationarity. When applied to the entire dataset, the test returned a p-value of 0.0, indicating that the series is stationary. However, when the ADF test was applied to shorter time periods, the p-values exceeded 0.05, suggesting non-stationarity in the short term. This discrepancy arises because short-term data often exhibit diurnal and seasonal patterns, while the complete time series tends to fluctuate around a relatively stable long-term mean. A similar observation was made by Bossew et al. (2024) in their study of indoor radon time series in Berlin: although short-term segments displayed cyclic behaviour, the full time series was found to be approximately stationary. They emphasized that detecting trends within short temporal windows does not necessarily imply non-stationarity, as the broader, long-term pattern may reflect a different structure. In other words, what appears as a trend on a short timescale may be embedded within a stable, stationary process when viewed over a longer horizon. In this study, the full dataset was used for regression analysis under the assumption of long-term stationarity.

Feature selection and engineering: Feature selection and feature engineering are processes in machine learning that involve choosing and transforming data features to improve model performance. Feature selection focuses on identifying the most relevant features, while feature engineering involves creating new features or modifying existing ones to enhance the model's accuracy and effectiveness.

Experimental background

For our dataset, feature selection involved analysing the pairwise correlations between each predictor and RAC. This resulted in most of the independent variables being included. An exception was wind direction, which was excluded due to its cyclic nature. Although wind direction can be transformed into sine and cosine components, its contribution to the regression models was found to be negligible.

Feature engineering was performed to enhance the predictive power of the time-related variables, including day of the year (DOY; 1–365) and hour of the day (H; 0–23). These variables capture seasonal and diurnal effects, which are important for time series modelling. To account for their cyclic behaviour, these parameters were transformed into sine and cosine features using the following formulas:

$$\text{Sin}_{\text{feature}} = \sin\left(\frac{2\pi t}{T}\right), \quad \text{Cos}_{\text{feature}} = \cos\left(\frac{2\pi t}{T}\right) \quad (2.21)$$

where t is the time value (e.g., hour or day of year), and T is the period (e.g., 24 hours or 365 days).

This transformation ensures continuity at cycle boundaries (from 24:00 to 01:00 and from December 31 to January 1), enabling models to capture smooth periodic patterns. Incorporating these engineered time features significantly improved model performance, particularly for regression approaches involving machine learning.

Data Splitting: The final step in the data preprocessing involved splitting the dataset into training and testing subsets. The training subset comprised 80% of the data, whereas 20% of the remaining data were used for model evaluation. Data preprocessing and regression modelling were performed using the SPSS software package (IBM Corp., 2021), the H2O machine learning platform (H2O.ai, 2020), and R programming language (R Core Team, 2023).

Chapter 3

3. Time series analysis of RAC

This chapter presents the time series analysis of RAC, including data preparation (interpretation and smoothing), descriptive statistics, frequency distribution, periodicity, and hourly, diurnal and seasonal variations.

3.1. Descriptive statistics of RAC

The descriptive statistics of 2-hourly RAC measurements collected over a six-year period (January 2018 to December 2023) are summarized in **Table 3.1**. A total of 25,326 valid observations were recorded, with RAC values ranging from 0.00 to 31 Bq.m⁻³ and a mean of 5.6 Bq.m⁻³. The mean RAC value is lower than the global average of RAC 10 Bq.m⁻³ (Tirmarche et al., 2010). The standard deviation of 3.9 indicates moderate variability in the data, while the standard error of the mean (Std.EM = 0.024) suggests that the sample mean is a reliable estimate of the population mean. Approximately 2.8% of the data were missing, which is within an acceptable range for long-term environmental monitoring. Additionally, 465 measurements were identified as statistical outliers, defined as values falling outside the range of (Q1 – 1.5×IQR, Q3 + 1.5×IQR, where Q1 and Q3 represent the 25th and 75th percentiles, respectively, and IQR is the interquartile range), indicating occasional extreme RAC levels likely due to episodic events or anomalous conditions.

Table 3.1 Descriptive statistics of 2-hourly RAC measured during (2018 – 2023).

N	Minimum	Maximum	Mean	Std.EM	Std. Deviation	Missing%	No. of Extremes ²
25326	0.00	31	5.6	0.024	3.9	2.8	465

3.2. Data interpolation

To improve the temporal resolution of the data and synchronize it with other variables, the 2-hourly measured RAC data were interpolated to 1-hourly intervals using the cubic spline option of the interpolate/extrapolate method in OriginLab (OriginLab Corporation, 2025). Unlike simpler interpolation methods, such as linear or polynomial interpolation, cubic

² Number of cases outside the range Q1 – 1.5×IQR, Q3 + 1.5×IQR.

spline interpolation employs a series of piecewise cubic polynomials between each pair of data points, ensuring both smoothness and continuity (McKinley & Levine, 1998). An example of the interpolation result is shown in **Fig. 3.1**. As can be seen, this method generates additional points between each pair of measurements while preserving the overall smoothness and continuity of the data trend.

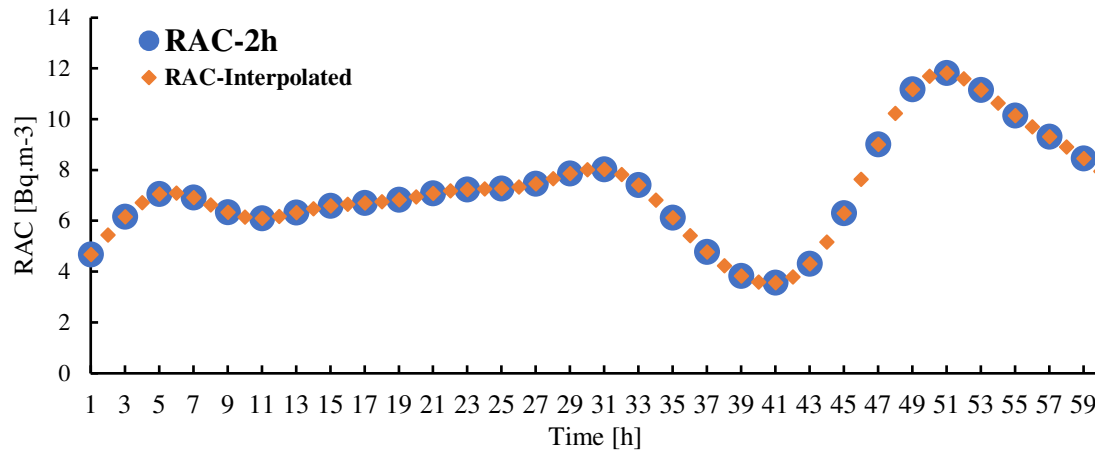


Fig. 3.1 An example of cubic spline interpolation. The blue dots show the 2-hourly RAC data, and the orange squares represent the interpolated points.

3.3. Data smoothing - results

The outdoor radon signals recorded in the detection systems often includes noise and extreme fluctuations due to the harsh environmental factors and instrumental error. Data smoothing techniques are used to reduce the noise and extreme fluctuations in the time series while preserving the trends and patterns. In this study, the radon time series were smoothed using the moving average (MA) and FFT smoothing techniques. The MA is the simplest, widely used technique for smoothing time series data. In this technique, the average of a fixed number of consecutive observations is considered the smoothed value of the central point. The number of observations is called the window size and determines how much smoothing is applied to the data set. The window size is crucial in order to achieve the desired smoothing effect. The other smoothing technique is FFT, which uses the Fourier transform to filter out the high-frequency noise from a time series, leaving behind the low-frequency components that represent the underlying trends and pattern (Kimball, 1974). The FFT algorithm transforms the time series data into the frequency domain, cutting off the higher frequencies according to a specific threshold. OriginLab software was used to smooth the radon time series data using both MA and FFT smoothing techniques. The result is shown in **Fig. 3.2**. The frequency threshold (f_{cutoff}) is calculated using the following equation:

$$f_{\text{cutoff}} = \frac{1}{2n\Delta t} \quad (3.1)$$

Where n represents the number of points in the window and Δt is the time interval between two consecutive data points. By considering the points in the window as described in equation (3.1), certain frequencies, or in other words, certain periods can be excluded. Since we used 2-hourly data ($\Delta t = 2\text{h}$), $n = 1, 2, 3$ corresponds to frequencies of $0.25/\text{h}$, $0.125/\text{h}$, and $0.083/\text{h}$, respectively, leading to the exclusion of any periodic fluctuations with periods shorter than 4h, 8h, and 12h.

It was found that the smoothing using 5-points MA and 1-point FFT do not reduce the noise sufficiently (**Fig. 3.2**). However, the 2-pts (8h) and 3-pts (12h) FFT smoothing reduce the noise reasonably well and reveal the expected trends. For all smoothing methods, strong correlation with Pearson's $r \geq 0.85$ was observed between raw and smoothed data (**Fig. 3.3**). For further analysis we considered the 2-pts FFT smoothing.

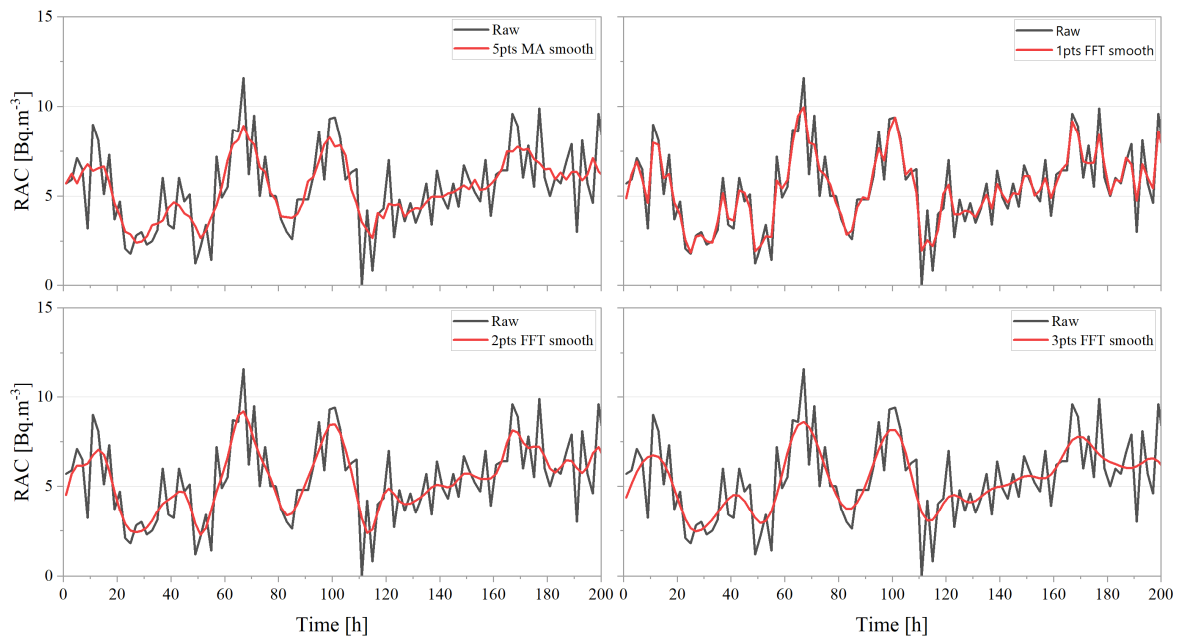


Fig. 3.2 Comparison of data smoothing using MA and FFT with different frequency cutoffs.

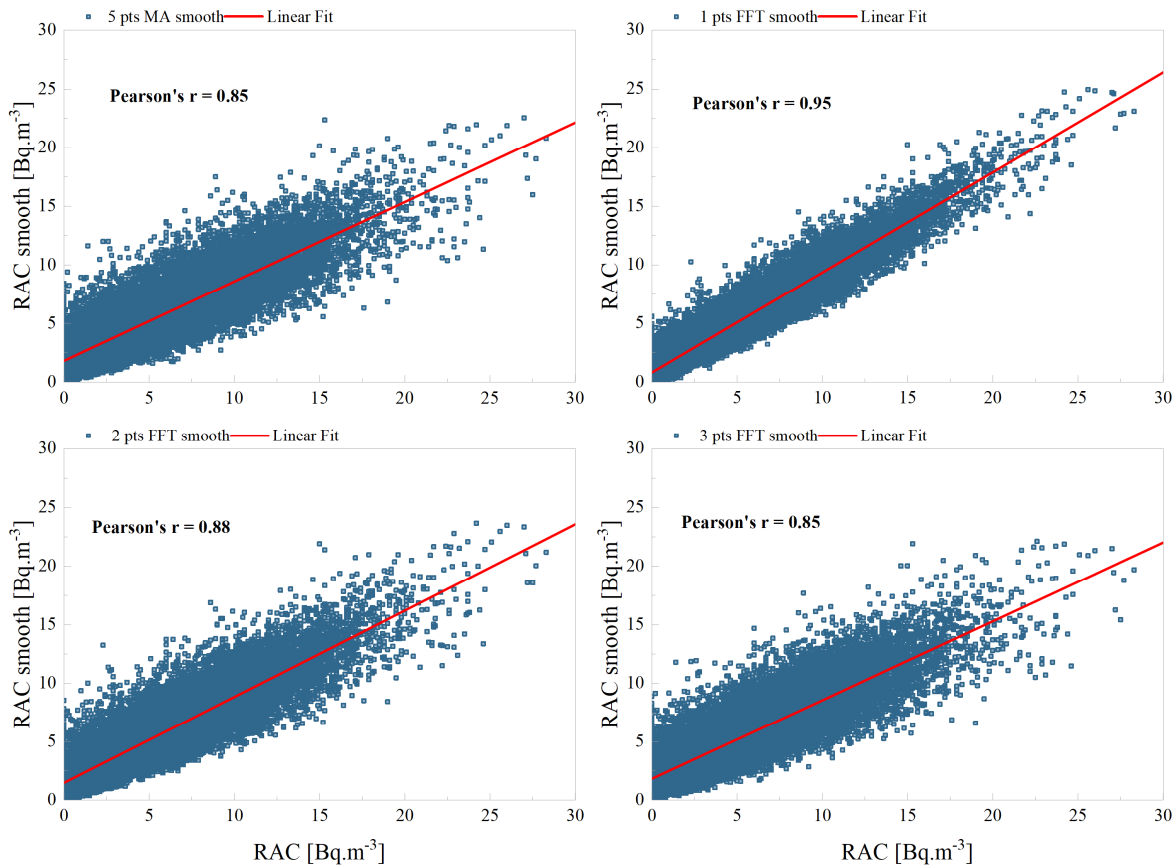


Fig. 3.3 Correlation between the raw data and the data smoothed using FFT smoothing.

3.4. Frequency distribution of RAC

To understand the statistical distribution of the RAC data, we applied various statistical methods. Understanding the data distribution is crucial for capturing the true nature of the data and selecting appropriate analysis techniques. Different statistical tests often assume specific forms of data distribution (e.g., normal, binomial, or gamma distributions). Choosing the correct test based on the actual distribution of the data ensures the accuracy and validity of the analysis. Furthermore, recognizing the distribution type helps in data transformation, outlier detection, and the selection of suitable modelling techniques.

As an initial step, the distribution of the data was assessed by fitting it to several theoretical distributions, including the normal, gamma, and Gaussian mixture models. The fit to the normal distribution is illustrated in **Fig. 3.4**. Notable deviations in the probability–probability (P–P) and quantile–quantile (Q–Q) plots indicate that the data does not conform to a normal distribution. In contrast, based on the P–P plots and empirical cumulative distribution function (CDF), the RAC data was found to align most closely with a gamma distribution.

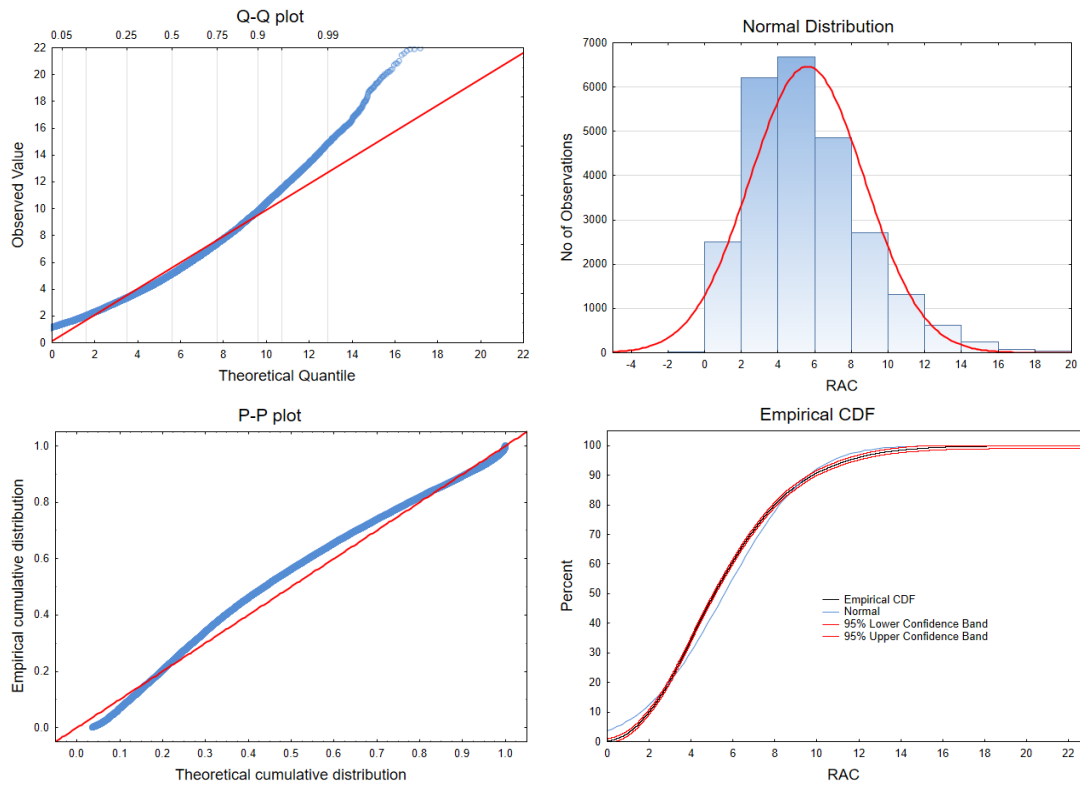


Fig. 3.4 RAC data fitting with normal distribution function.

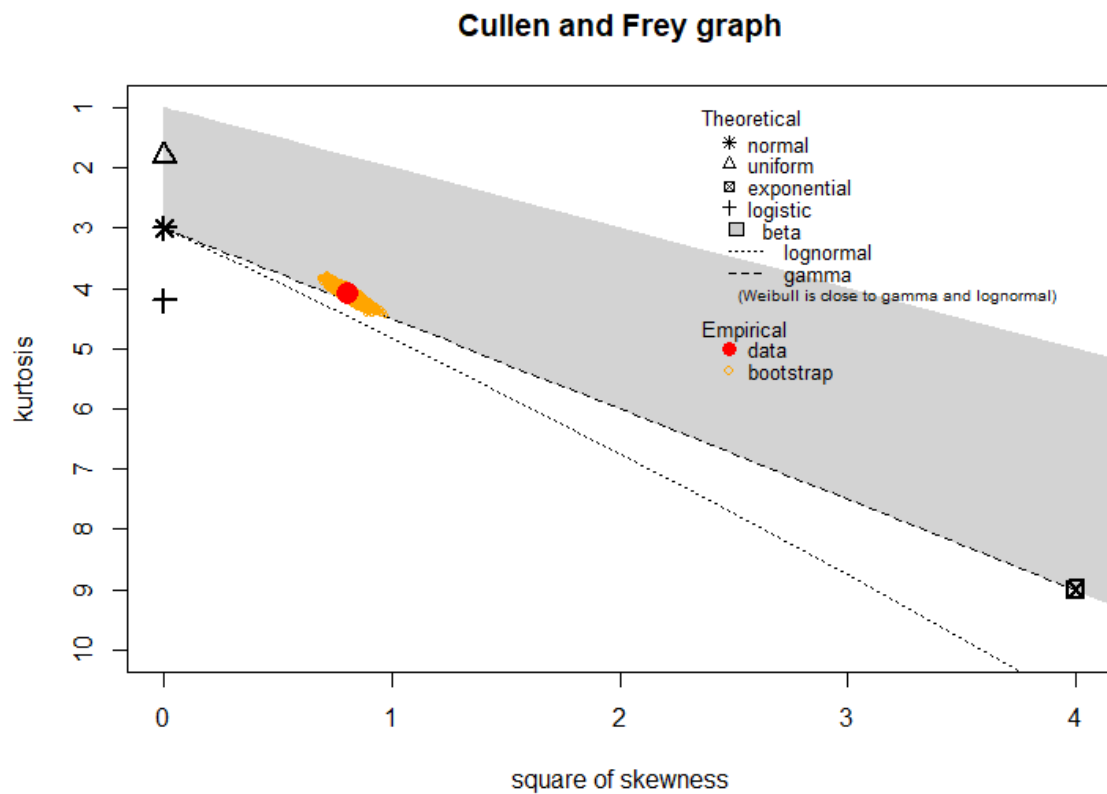


Fig. 3.5 Cullen and Frey graph, plotting the kurtosis against the square of the skewness for the RAC dataset (red circle) and the bootstrapped RAC data (orange smudge).

The second method involved using the *Cullen and Frey graph*, a visual tool used to assess the skewness and kurtosis of a dataset, helping to determine the appropriate probability distribution for the data. It plots these two statistical measures, allowing for a quick evaluation of the data's characteristics (Cullen & Frey, 1999). This graph was generated using the 'fitdistrplus' package in R (Delignette-Muller & Dutang, 2015), with the data bootstrapped 1000 times. Different regions of the graph correspond to different families of distributions, allowing for a visual assessment of the likely distribution of the data. As shown in **Fig. 3.5**, the RAC data distribution falls along the line associated with the gamma distribution.

3.5. Periodicities in RAC time series

A time series consists of three main components: trend, periodic, and random components. Understanding each of these components is crucial for accurately characterizing the true nature of time series data and making reliable predictions. Periodicities refer to patterns or cycles that occur at regular intervals within a time series. The length of time after which the pattern repeats is called the period. Mathematically, periodicity is defined as follows:

$$x(t) = x(t + T), \quad (3.2)$$

where $x(t)$ is a signal with a period T for all time t . In this study two well-known techniques i.e., FFT and autocorrelation function (ACF) were used to assess the periodicities in RAC time series.

The periodogram of the radon time series based on FFT analysis is shown in **Fig. 3.6**. The periodogram reveals dominant periodicities at 8 hours, 12 hours, and the expected 24 hours. To further investigate the periodic behaviour of the radon time series, a second method based on the ACF was applied. The ACF periodogram (**Fig. 3.7**) clearly shows 24-hour and 12-hour periodicities; however, the 8-hour periodicity has not appeared. The 24-hour cycle is well understood and is primarily attributed to daily cycles in meteorological factors, particularly solar radiation. In contrast, the origins of the 12-hour and 8-hour periodicities are less clear, though they may be related to Earth's tidal effects and atmospheric mixing conditions. Similar periodicities have been reported in previous studies (Bossey et al., 2024; Kumar & Nagaraja, 2024; Siino et al., 2020).

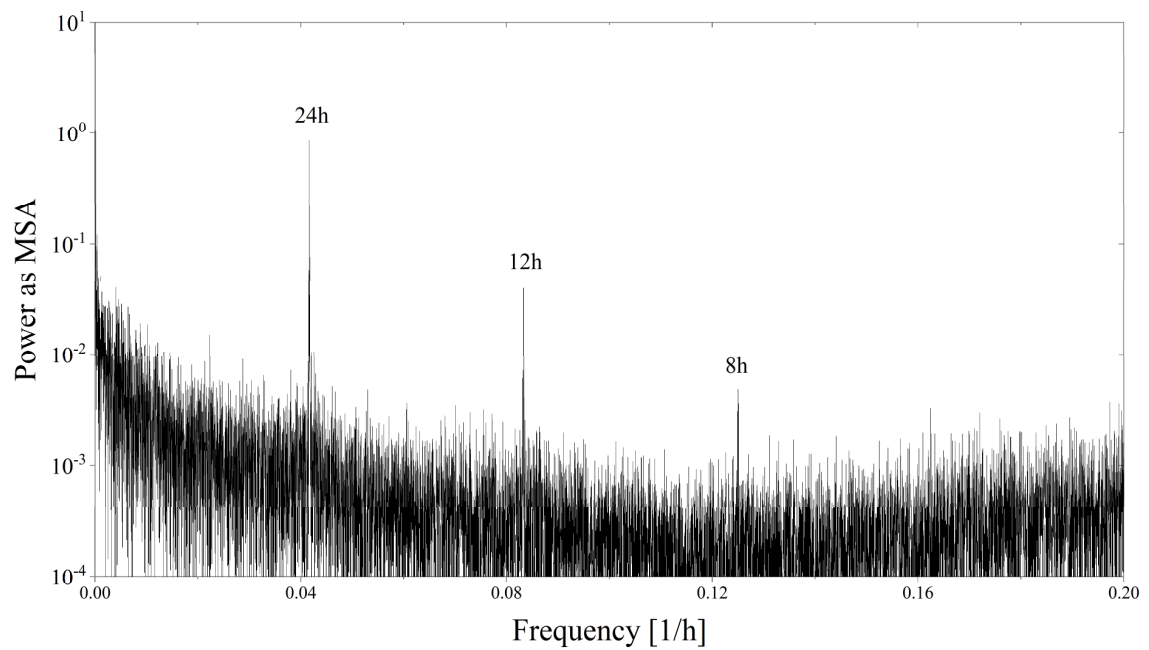


Fig. 3.6 Periodogram of RAC time series based on FFT.

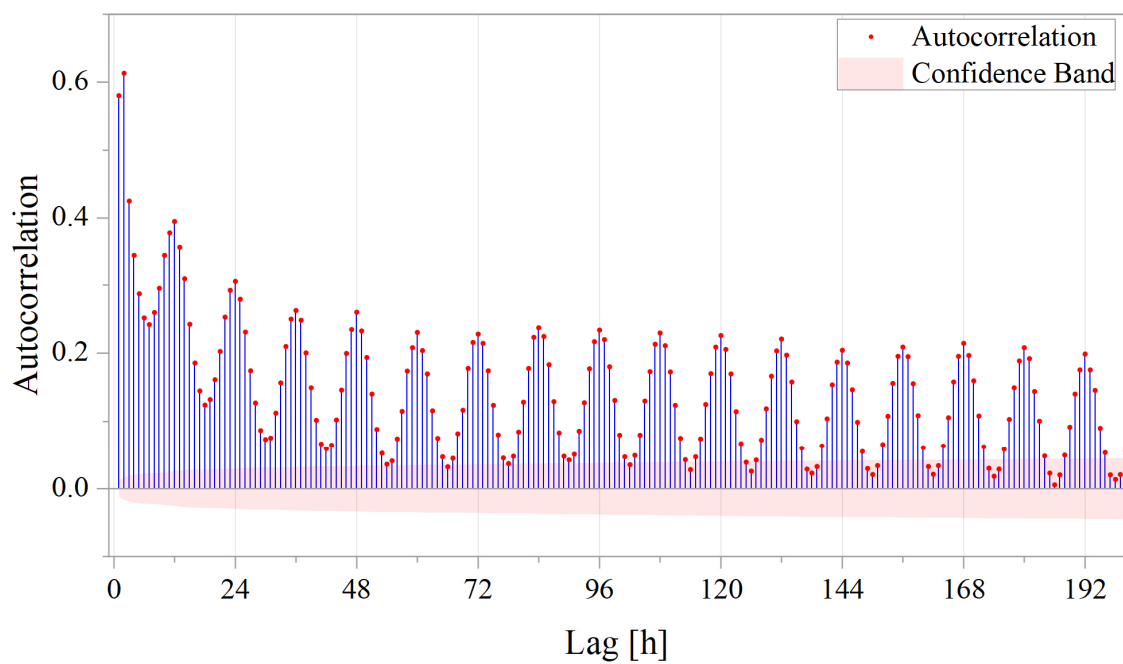


Fig. 3.7 Periodogram of the ACF of the RAC time series.

3.6. Diurnal cycle of RAC

According to the periodicity analysis of outdoor radon data, it was revealed that the 24-hour period is the most dominant in the radon time series. The composite diurnal cycles provide valuable insights into the daily evolution of RAC, as random fluctuations tend to cancel each other out. The RAC exhibited a clear diurnal cycle, with the highest concentration in the early morning (03:00–05:00) and the lowest during the afternoon (14:00–16:00). The same diurnal cycle was observed elsewhere (Griffiths et al., 2013; Sultani et al., 2023; Zahorowski et al., 2004). This pattern is primarily governed by changes in atmospheric mixing conditions/BLH (e.g., Pal et al., 2015; Zimnoch et al., 2014). After sunrise, solar heating warms the Earth's surface, warming near-surface air parcels and generating turbulent mixing. This turbulence vertically expands the boundary layer, forming a convective boundary layer (CBL) whose depth peaks in mid-afternoon. As solar intensity diminishes towards evening, the CBL begins to collapse. After sunset, the surface cools rapidly via thermal radiation, creating a temperature inversion (colder air near the surface, warmer air above). This inversion stabilizes the atmosphere, suppressing turbulence and reducing the BLH to a shallow stable boundary layer, typically 200–300 m deep. During this stable phase, radon accumulates near the surface due to limited vertical mixing, reaching peak concentrations by early morning. After sunrise, solar heating erodes the nocturnal inversion, restoring turbulent mixing and deepening the boundary layer. This dilutes surface concentrations of radon and its progeny by vertically dispersing them, leading to the observed mid-afternoon minimum.

Distinct diurnal cycle of RAC was observed for different months and seasons. The highest amplitudes occur in the summer (Jun–Aug), followed by spring (Mar–May) and autumn (Sep–Nov) while the lowest amplitude was observed in winter (Dec–Feb). These variations are likely due to the seasonal changes in radon exhalation rates and meteorological conditions. In warmer months, intense atmospheric mixing results in more pronounced diurnal fluctuations of radon while in winter, more stable atmospheric conditions contribute to less fluctuations in radon concentrations.

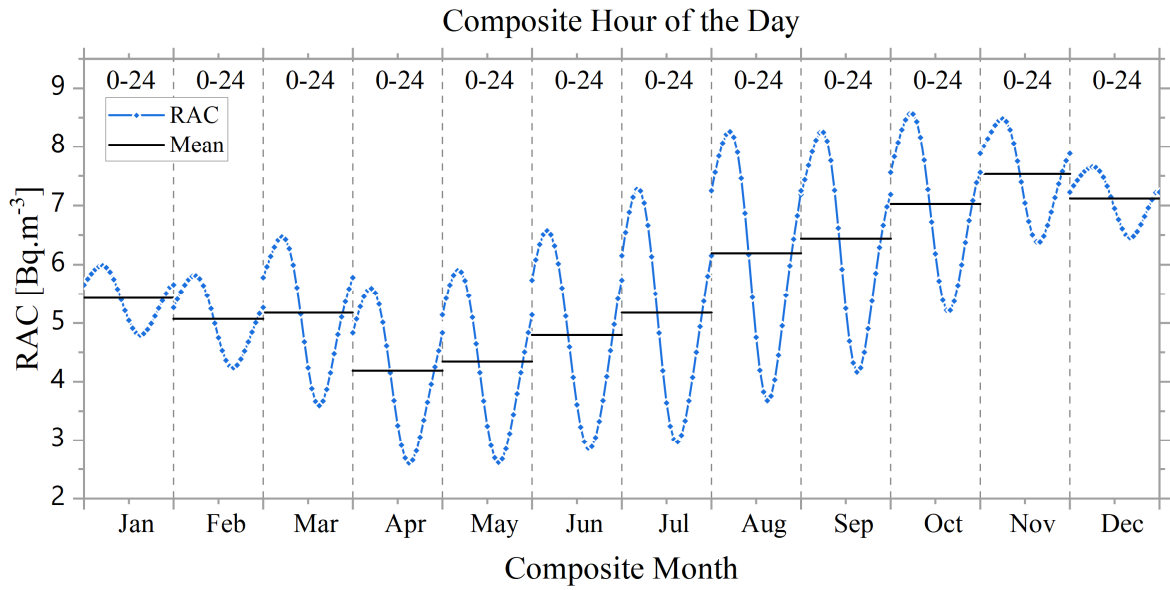


Fig. 3.8 Composite diurnal cycle of RAC for each month. The data are aggregated over the entire study period (2018 – 2023). The blue curves with dots illustrate the 24-h cycle of RAC, whereas the horizontal lines represent the monthly averages.

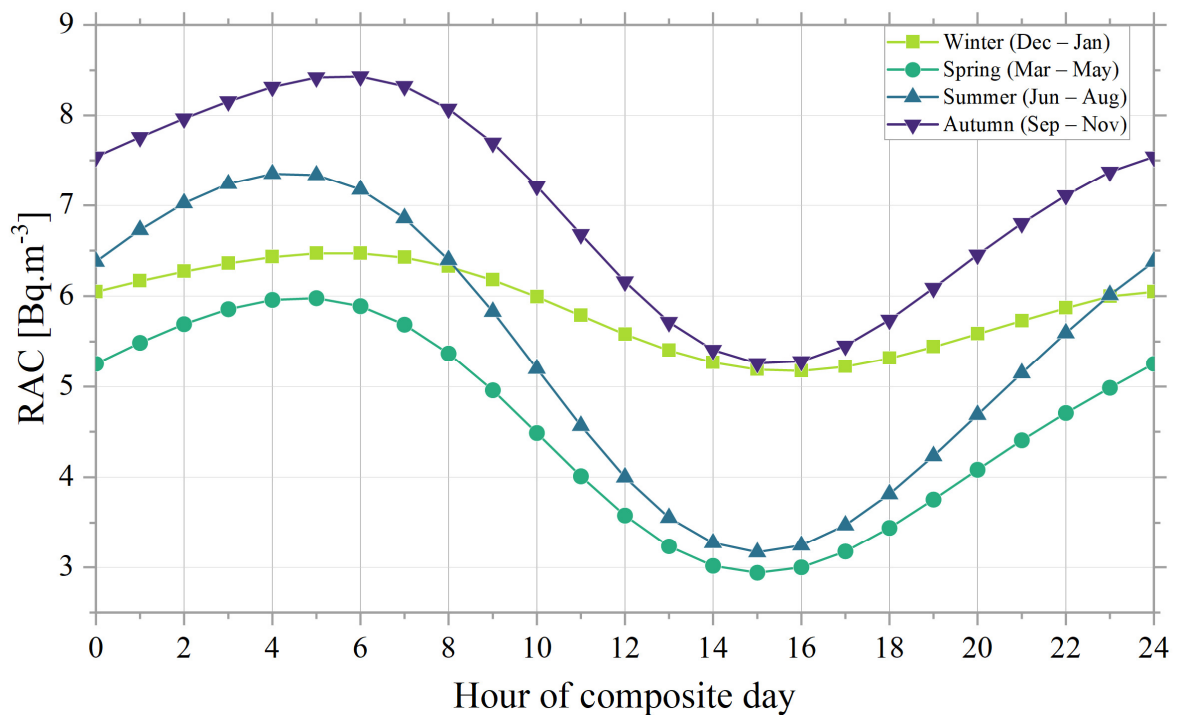


Fig. 3.9 Composite diurnal cycle of RAC for each season. The data are aggregated over the entire study period (2018 – 2023).

3.7. Seasonal variation of RAC

Fig. 3.10 shows the seasonal variation of RAC, presented as composite monthly boxplots. The RAC exhibits a clear seasonal trend, with the lowest levels observed in April and May and the highest levels in October and November. The mean RAC is lowest in April and then steadily increases, reaching its maximum in November, before gradually decreasing from December to April. Throughout the year, the mean remains consistently higher than the median, indicating a right-skewed distribution likely influenced by occasional high values. This seasonal pattern can be attributed to several factors, primarily seasonal changes in atmospheric stability, air mass history, and variations in radon exhalation rates (Zimnoch et al., 2014). During autumn and winter, more frequent and prolonged atmospheric inversions lead to greater atmospheric stability, favouring the accumulation of radon near the ground. In contrast, during spring and summer, stronger atmospheric mixing enhances the dispersion of radon, resulting in lower near-surface concentrations.

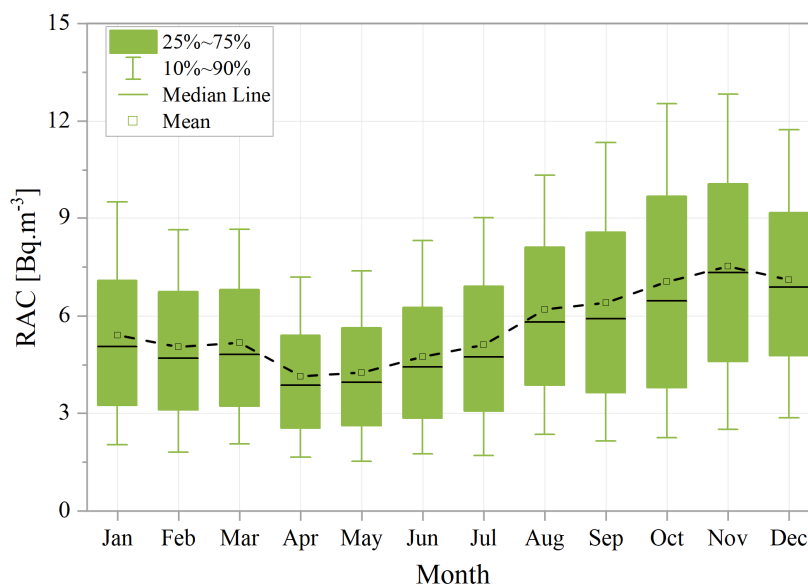


Fig. 3.10 Seasonal boxplots of RAC. The data are aggregated over the entire study period (2018 – 2023).

Chapter 4

4. Factors governing radon variability

Ground-level radon variability is influenced by several environmental factors, including boundary layer height (BLH) and a range of meteorological parameters. Among these factors, BLH is the most significant factor in controlling the dispersion and accumulation of radon near the surface. In this study, BLH derived from the ERA5 reanalysis dataset, along with key meteorological parameters measured simultaneously, were analysed to assess the variability of outdoor RAC. The meteorological parameters included temperature (T), relative humidity (RH), wind speed (WS), wind direction (WD), atmospheric pressure (P), and precipitation (Prec), all of which were measured hourly at a nearby sampling station. The BLH based on the ERA5 reanalysis dataset produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) provides a high-resolution global atmospheric reanalysis dataset dating back from 1950 to the present day (Hersbach et al., 2020). The BLH is calculated using the Bulk Richardson number, has a spatial resolution of $0.25^{\circ} \times 0.25^{\circ}$ and a temporal resolution of 1 hour. The BLH-ERA5 dataset has been widely used in atmospheric research worldwide. Several studies have compared BLH-ERA5 with experimental datasets obtained from ground-based LIDAR, sodar, ceilometer, and radiosonde measurements. Despite some differences, there was generally good agreement between the BLH-ERA5 data and the experimental results (e.g., Allabakash et al., 2020; Madonna et al., 2021; Guo et al., 2021; Sinclair et al., 2022; Li et al., 2023). For our analysis, the hourly BLH-ERA5 data for our locality (latitude: $48^{\circ}9'N$, longitude: $17^{\circ}4'E$) were extracted, covering the years 2018 to 2023, which aligns with the period during which radon and its progeny concentrations were measured.

Fig. 4.1 represents hourly variation of RAC, BLH and key meteorological factors for a period of 10 days for the month of June 2018. The RAC data were smoothed using FFT smoothing with 2-pts frequency cutoff; other parameters were not smoothed. Generally, the RAC shows a clear diurnal cycle influenced by the changes in the diurnal cycles of BLH and meteorological factors. Usually, the highest RAC coincides with the shallowest BLH, lowest temperature and wind speed. On the other hand, the lowest RAC is observed to occur when BLH is very deep, and the temperature and wind speed are high.

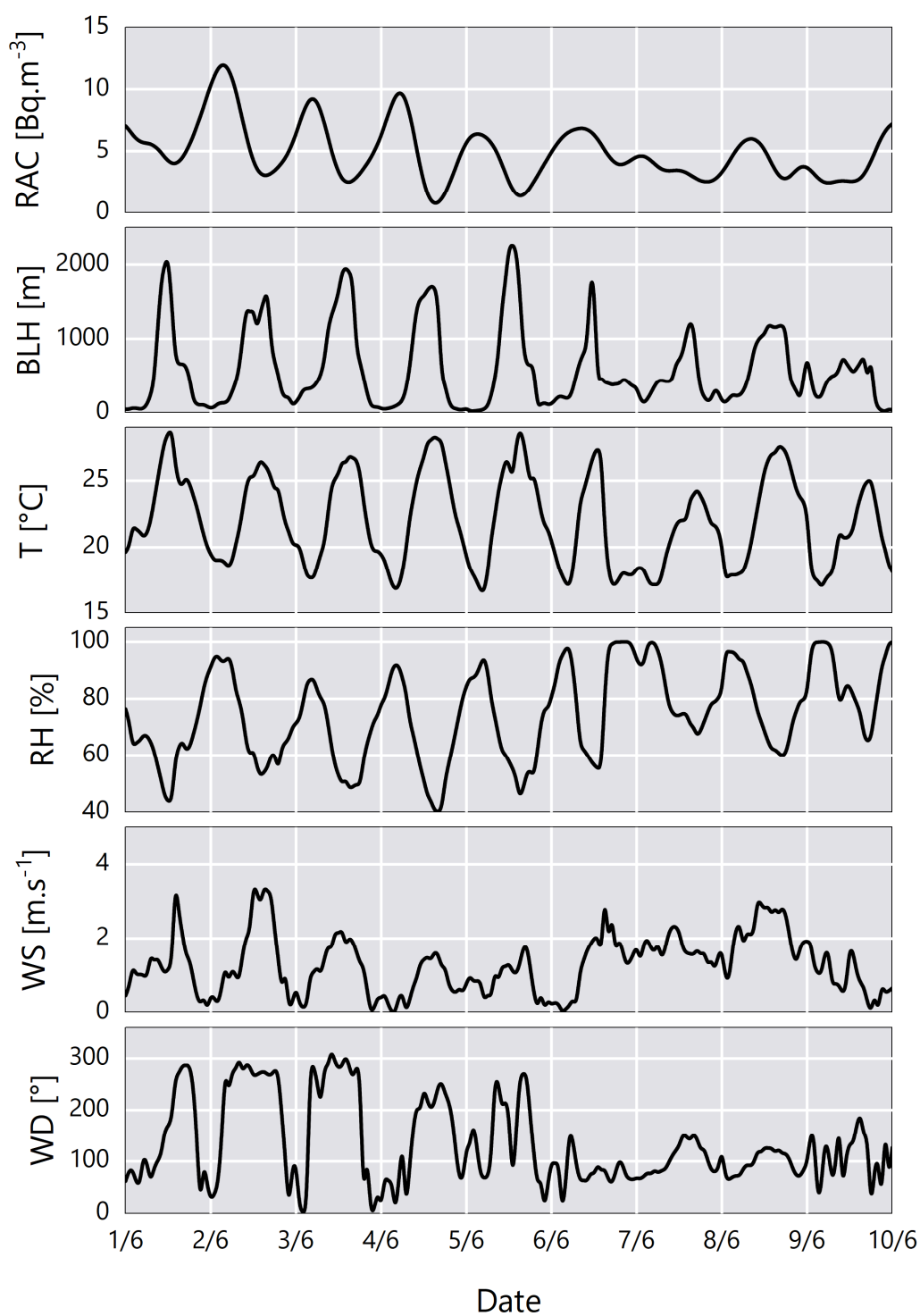


Fig. 4.1 Hourly RAC, BLH and major meteorological parameters for a period of 10 days measured in June 2018.

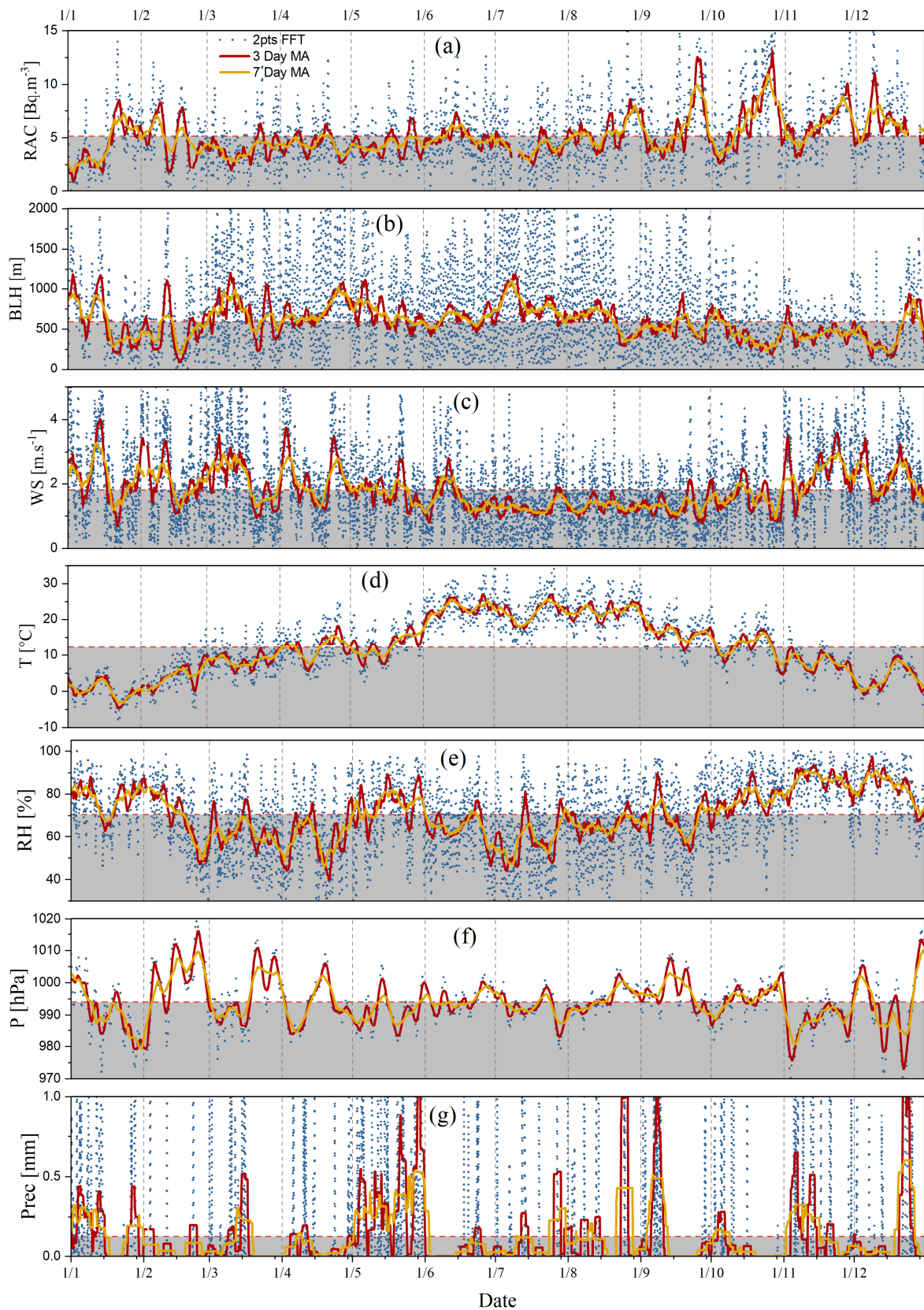


Fig. 4.2 Variations of RAC, BLH and major meteorological parameters over a period of one year (2019). Each graph contains original hourly data that were smoothed using a 2-point FFT, as well as the moving averages of this data smoothed over periods of 3 and 7 days. The grey area below the red dashed line shows the annual mean.

4.1. Factors governing RAC at synoptic scales

In addition to diurnal and seasonal patterns, RAC shows changes on a synoptic scale associated with the weather systems such as cyclones, anticyclones, frontal passages, and persistent temperature inversion events. The synoptic-scale changes are considered to vary from days to about two weeks (Kikaj et al., 2019). To better understand the changes in RAC against BLH and key meteorological factors on the synoptic scale, a one-year data series was plotted using 2-pts FFT smoothing, as well as 3-day and 7-day moving averages (**Fig. 4.2**). A moving average (MA) is a useful smoothing technique that highlights long-term trends by reducing short-term fluctuations. Given that both radon concentrations and environmental parameters exhibit short- and long-term variability, the use of MA smoothing can be useful technique.

The MA values of RAC reveal a clear short- and long-term variability throughout the year 2019, strongly influenced by changes in meteorological conditions. Beyond the diurnal cycle, RAC exhibits synoptic-scale and sub-seasonal variations ranging from a few days to several weeks. The 3-day MA may be useful for revealing changes on a synoptic scale over a period of less than one week, while the 7-day MA should highlight patterns over a period of one week. A closer look at the long-term trend reveals considerably higher RAC from September through February associated with the changes in daylight duration and solar intensity. In this period, the 7-day MA of RAC remains consistently above the annual mean (represented by the red dashed line in the grey-filled area). During this period, the RAC also exhibits several notable synoptic-scale patterns and spikes, each lasting from a few days to approximately three weeks. These spikes maybe attributed to the persistent temperature inversion events occurring during cold months (Kikaj et al., 2019). During this period, lower BLH (**Fig. 4.2b**) and lower temperature (**Fig. 4.2d**) were typically observed. However, high RH is often associated with elevated RAC, reflecting moist, stagnant air masses. Atmospheric pressure plays a more complex role: high-pressure systems can potentially trap radon, while falling pressure may promote its release from the soil. Precipitation generally causes short-term reductions in RAC by blocking the release of radon from the soil, although radon concentrations often rebound once the soil has dried out. Collectively, these factors account for the observed seasonal and synoptic variations in outdoor RAC, with the most pronounced peaks occurring under calm, cold, and stable weather conditions.

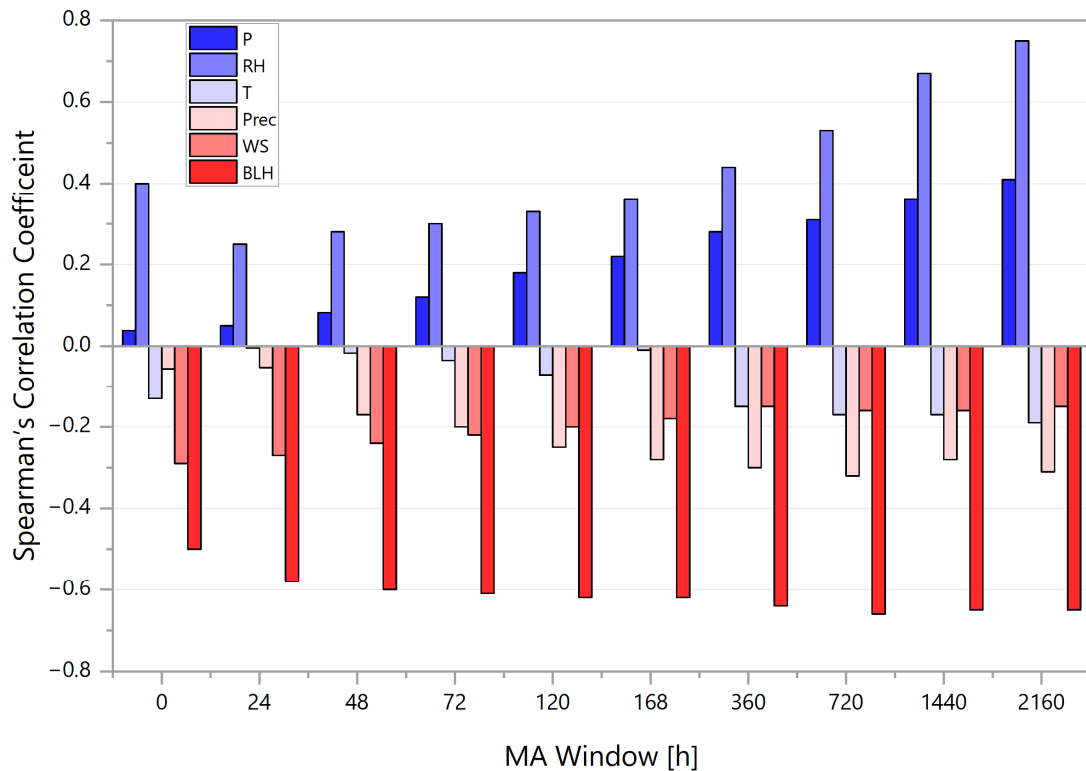


Fig. 4.3 Spearman's correlation coefficient analysis, investigating RAC versus meteorological factors and BLH for various degrees of data smoothing. The first set of bars on the left shows the correlations between the raw (non-smoothed) data. The remaining bars show the correlations after smoothing with moving average windows of 1, 2, 3, 4, 7, 15, 30, 60 and 72 days, respectively.

Spearman's correlation analysis was conducted between radon concentrations and each variable using various MA windows; the results are shown in **Fig. 4.3**. The Spearman's rank correlation coefficient is a non-parametric measure of statistical dependence between two variables (Spearman, 1961). Unlike Pearson's correlation, which assesses linear relationships, Spearman's correlation evaluates the strength and direction of a monotonic relationship between two ranked variables. This makes it particularly useful when the data does not meet the assumptions of linearity. In our case, this approach allows for evaluating the relationship between RAC and any environmental variable.

The RAC exhibits complex relationships with meteorological parameters and BLH. Atmospheric pressure and RH generally show positive correlations with RAC, while BLH, temperature, precipitation, and WS are correlated negatively. The magnitude of correlation coefficients ranges from negligibly small (e.g., $r_s = 0.004$ for RAC vs. temperature smoothed by 24-hour MA) to strong ($r_s = 0.75$ for RAC vs. RH smoothed by 3-month MA). Most of the correlations are statistically significant ($p < 0.05$), except for the weakest one

(the 24-hour MA smoothing of RAC vs. temperature). The strength of correlations increases with longer MA windows for pressure, RH, precipitation and temperature, suggesting that these parameters influence RAC over longer timescales. By contrast, WS correlations decreasing with longer MA windows, implying its effects the radon levels in short-term scales. BLH correlations remain relatively stable across all MA windows, indicating persistent short- and long-term impacts on RAC. RH is an exception: its correlation is high with raw (non-smoothed) data, dips at intermediate MA windows, and strengthens again with longer smoothing windows, possibly reflecting competing short term moisture inhibition (e.g., rainfall sealing the soil pores) and long-term seasonal drivers (e.g., dry vs. wet seasons). Precipitation shows a minimal correlation with raw data, but its influence grows with larger MA windows, likely due to noise reduction or delayed effects such as enhanced radon exhalation caused by soil cracking after rain).

In general, the variability in correlation values as a function of different MA windows is due to two factors:

- 1) *Noise reduction*: Smoothing has a tendency to filter out high-frequency fluctuations (e.g., transient weather events), emphasizing sustained trends.
- 2) *Timescale dependence*: Parameters such as WS act over short timescales (~hours), whereas the seasonal trends of RH correlate strongly with those of RAC.

For example, the raw WS data showed a stronger correlation with the raw RAC data, likely due to radon dispersion caused by short-term wind gusts. However, this correlation weakens after smoothing. By contrast, RAC correlation with precipitation only gains significance for longer MA windows due to the delayed physical impact of precipitation on radon release from the soil. These patterns demonstrate that RAC variability is governed by both immediate meteorological drivers and cumulative environmental conditions.

4.2. Factors governing diurnal variation of RAC

The diurnal variation of RAC, as depicted in **Fig. 4.4**, reveals a distinct pattern driven primarily by boundary layer dynamics and meteorological conditions. The highest RAC values occur during the early morning hours (03:00–05:00), while the lowest values are observed in the midafternoon afternoon (14:00–16:00). This pattern is attributed to the diurnal evolution of the BLH (**Fig. 4.4b**), which is driven by solar radiation, WS, and surface air temperature. Following sunrise, increasing solar radiation warms the ground and the adjacent air layers, enhancing buoyancy and generating convective turbulence. This results

in the progressive erosion of the nocturnal temperature inversion and the development of a convective boundary layer (CBL), which deepens throughout the morning and reaches its maximum during mid-afternoon. The resulting enhanced vertical mixing, supported by rising WS (**Fig. 4.4c**) and elevated temperatures (**Fig. 4.4d**), facilitates the dispersion and dilution of radon from the surface atmosphere into higher atmospheric layers, thereby significantly reducing surface-level RAC concentrations. In contrast, after sunset, the absence of solar heating leads to radiative cooling of the land surface, which promotes the formation of a stable boundary layer (SBL). This SBL is characterized by a shallow BLH, low wind speeds, and suppressed turbulence, all of which inhibit vertical mixing. Under these stable nighttime conditions, radon emitted continuously from the ground accumulates near the surface, resulting in elevated RAC levels by early morning. A graphical depiction of the evolution of the boundary layer throughout a 24-hour period is also shown in **Fig. 2.8** of **Chapter 2**.

The diurnal cycle of RH (**Fig. 4.4e**) exhibits an inverse relationship with temperature, peaking during the early morning and reaching a minimum in the afternoon, further reflecting the thermally stable nighttime atmosphere and more turbulent daytime conditions. Atmospheric pressure (**Fig. 4.4f**) follows a typical semi-diurnal cycle with limited direct influence on RAC, though a weak positive correlation suggests some minor modulation effect might be at work.

Notably, a time lag of around 2h exists between the evolution of BLH and RAC concentrations: the RAC minimum does not coincide exactly with the peak BLH, nor does the RAC maximum align perfectly with the lowest level of BLH. Instead, RAC responds gradually to the changes in atmospheric stability, reflecting the cumulative effects of vertical mixing and sustained accumulation or dilution under prolonged meteorological conditions. This time-lagged response highlights the importance of considering the temporal inertia of boundary layer processes when interpreting pollutant behaviour in near-surface atmospheric layers.

Fig. 4.5 further supports this interpretation by illustrating the heatmap of Pearson correlation coefficients among the studied variables. As shown, RAC is strongly negatively correlated with both BLH and WS (Pearson's $r > 0.7$), while exhibiting a strong positive correlation with RH (Pearson's $r = 0.92$) and a moderate positive correlation with pressure (Pearson's $r = 0.55$). In line with the observed dynamics, BLH shows a very strong positive

correlation with temperature and wind speed, consistent with the role of solar heating and mechanical mixing in boundary layer development and, consequently, in regulating radon dispersion.

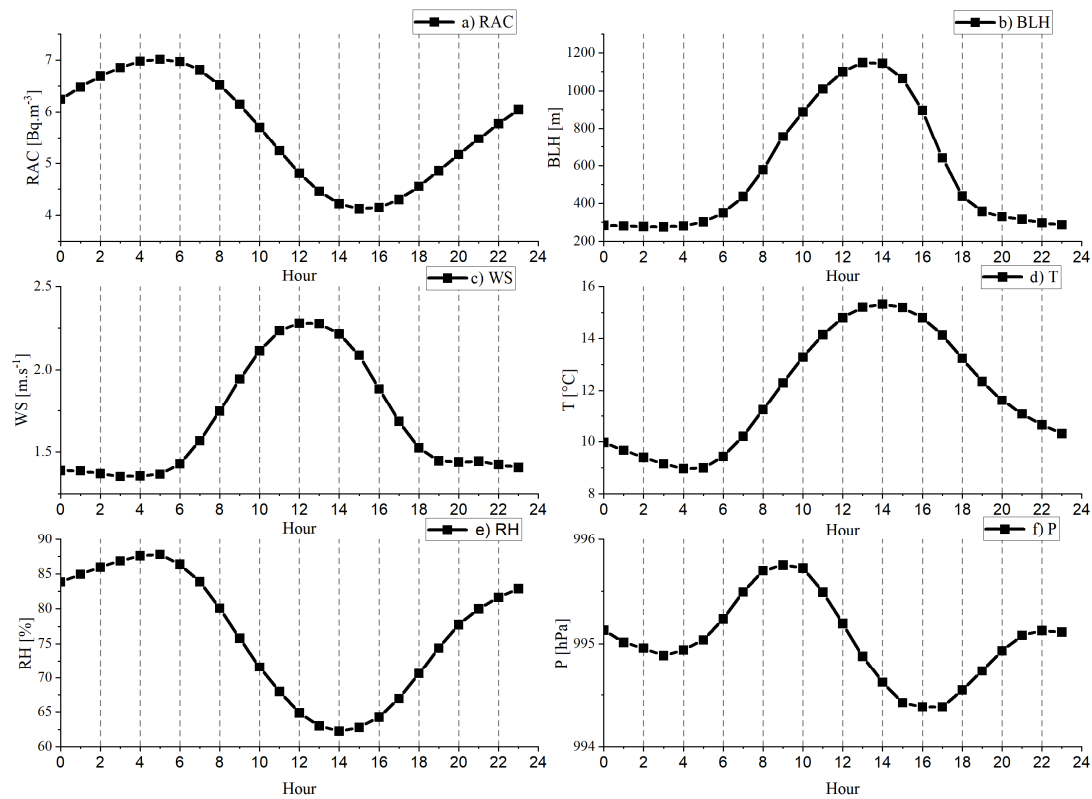


Fig. 4.4 Composite diurnal cycles of RAC, BLH and key meteorological factors averaged over the entire study period.

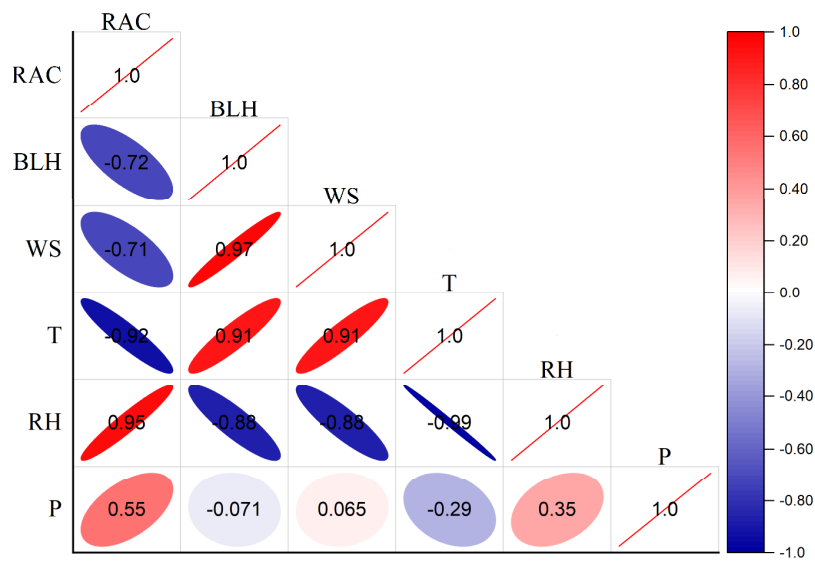


Fig. 4.5 Correlation heatmap depicting the correlations between the diurnal cycles of RAC, BLH and meteorological parameters shown in Fig. 4.4.

4.3. Factors governing seasonal variations of RAC

Seasonal variation of RAC is influenced by several factors, including meteorological and atmospheric mixing conditions, as well as the seasonality of radon exhalation rate. The typical seasonal pattern of RAC shows a minimum in April, followed by an increase throughout the summer, a peak in November, and slightly decline during winter (**Fig. 4.6a**). This pattern is mainly driven by seasonal variations in BLH dynamics (**Fig. 4.6b**) and radon exhalation rates (**Fig. 4.6b, i**). Enhanced solar radiation during spring and summer deepens the boundary layer, promoting vertical mixing and radon dilution. Conversely, in autumn and winter, shallow BLHs, reduced solar input, and persistent surface-based inversions limit vertical dispersion, allowing radon to accumulate near the surface. Measured radon flux shows low variability from April to August, followed by higher variability and a peak in November (**Fig. 4.6i**), aligning well with the seasonal RAC pattern. However, modelled radon flux from the European radon flux map (Karstens et al., 2022) diverges from local measurements, highlighting site-specific variability. Other variables such as temperature, wind speed, pressure, and rainfall may indirectly influence RAC by affecting BLH and radon exhalation rates. As shown in **Fig. 4.7**, RAC exhibits a strong negative correlation with BLH and positive correlations with RH and measured radon flux. In contrast, correlations with temperature, WS, precipitation, pressure, and modelled radon flux are statistically insignificant. These findings are consistent with previous studies such as that of Zimnoch et al. (2014), which attributed seasonal radon variability to changes in atmospheric mixing and stability, air mass history, and radon exhalation rates. While soil parameters may influence radon flux, RAC is predominantly governed by the atmosphere's vertical transport capacity, dictated by BLH and stability.

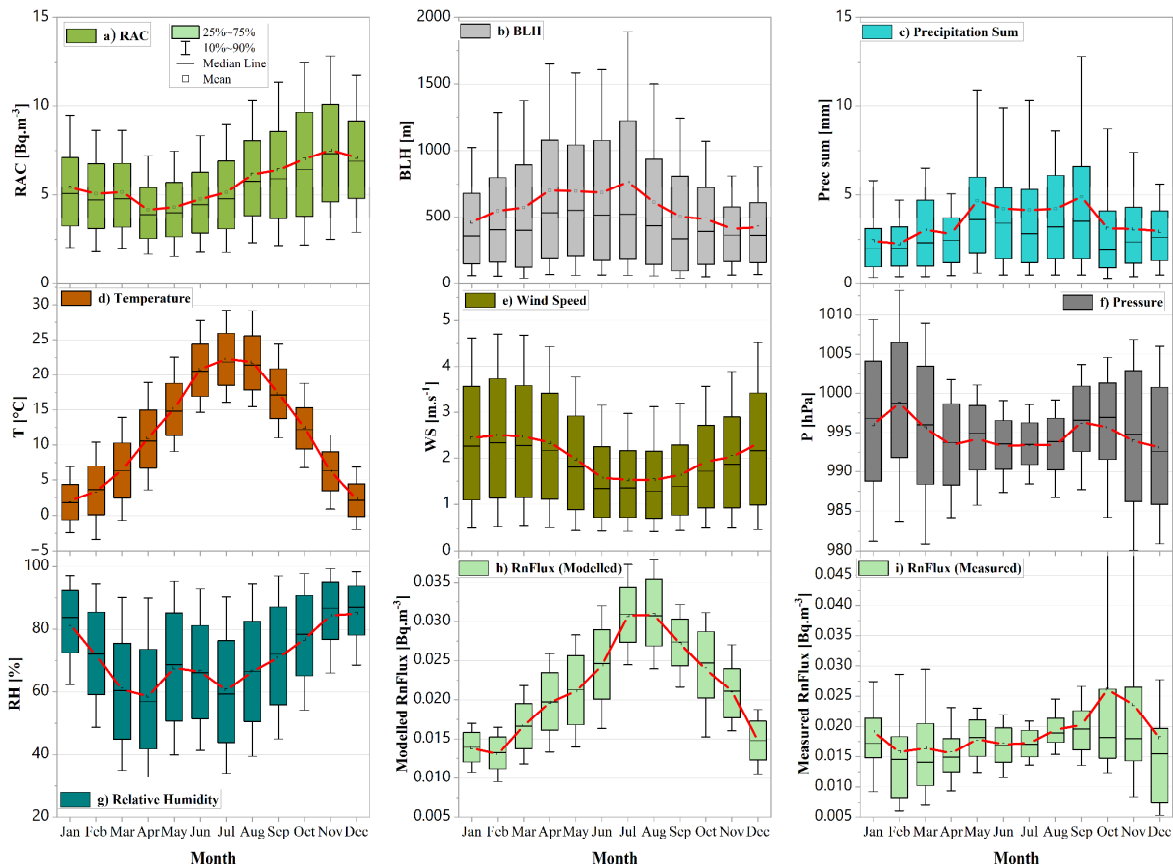


Fig. 4.6 Composite monthly boxplots of RAC against BLH, RnFlux and meteorological factors. The data are aggregated over the entire study period (2018 – 2023).

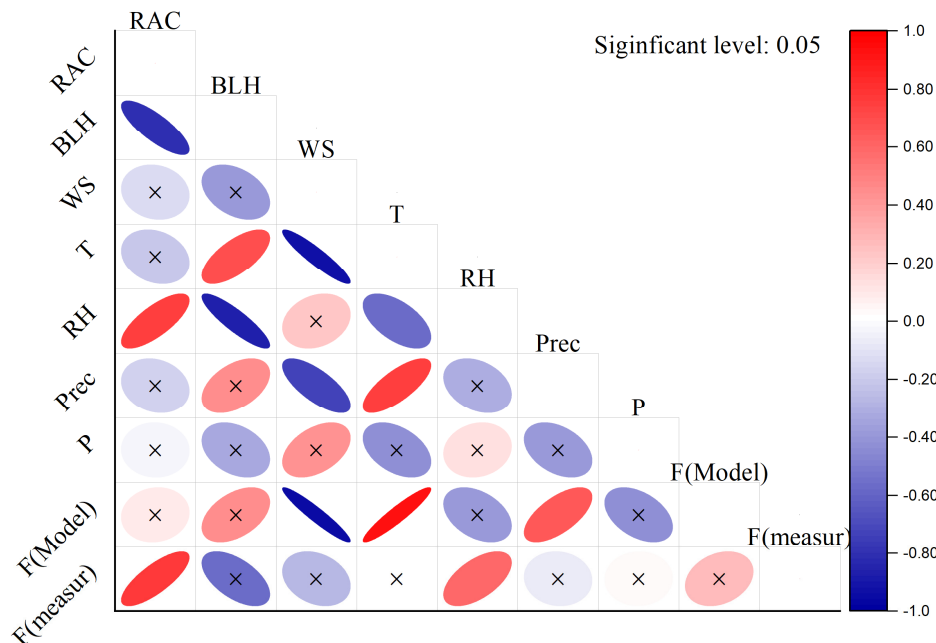


Fig. 4.7 Pearson correlation between composite monthly means of RAC, BLH, meteorological factors and radon flux (F).

Table 4.1 PCA summary including eigenvalue, variance and cumulative percentages.

Number	Eigenvalue	Variance (%)	Cumulative (%)
1	2.2	37.0	37.0
2	1.3	21.6	58.5
3	1.0	17.1	75.7
4	0.7	11.5	87.2
5	0.4	7.5	94.7
6	0.3	5.3	100

Table 4.2 Loadings of each variable for the first three principal components from PCA.

Variable	PC1	PC2	PC3
RAC	-0.44	-0.26	-0.09
RH	-0.49	0.37	0.10
T	0.27	-0.61	0.43
Prec	-0.07	0.36	0.86
WS	0.37	0.53	-0.21
BLH	0.59	0.10	0.06

Statistical and ML analysis of RAC variability

This section employs a comprehensive suite of statistical and machine learning (ML) techniques to analyse the hourly variability of RAC in relation to BLH and key meteorological variables. The methods include PCA, GAM, and MLR, alongside advanced ML algorithms such as XGBoost, GBM, and RF. This diverse modelling approach balances the interpretability of classical statistical techniques with the predictive power of ensemble learning, aiming to identify and quantify the key drivers of RAC variability.

4.4. Principal component analysis

The PCA introduced in **Section 2.8.4**, was applied to the dataset (RAC, BLH and key meteorological parameters) aiming to identify dominant patterns, reduce dimensionality by capturing the most informative combinations of correlated variables, and explore underlying patterns in dataset. The subsequent PCA analysis revealed that three principal components (PC) collectively explain 75.7% of the total variance in the dataset (**Table 4.1**), while a total of six PCs are required to account for the entire variance. The loadings of each original variable for these first three PCs are presented in **Table 4.2** and visually depicted in **Fig. 4.8**. Specifically, **Fig. 4.8** illustrates the loadings for PC1 versus PC2 and PC1 versus PC3, along with their corresponding biplots, which display the distribution of the original data points in

the reduced-dimensional space (**Fig. 4.8c** and **4.8d**, respectively). This approach provides a clear visualization of the relationships between the original variables and the derived PCs, facilitating a more intuitive understanding of the factors influencing outdoor radon variability.

The PCA performed on the dataset revealed distinct patterns in the interplay between atmospheric variables and radon activity concentration (RAC). Since a principal component (PC) is a linear combination of the original variables, direct interpretation can be challenging. However, examining the loadings provides insight into the contribution and influence of each variable within each component.

The first two principal components (PC1 and PC2) explained 58.6% of the total variance, with PC1 alone accounting for 37.0%. PC1 appears to represent atmospheric dispersion processes, as indicated by strong positive loadings from WS, BLH, and temperature, and a clear negative loading from RAC. This inverse relationship suggests that elevated radon concentrations are typically associated with shallow boundary layers and weak winds conditions that inhibit vertical mixing and promote radon accumulation near the surface.

PC2, which accounts for 21.6% of the variance, reflects a secondary axis influenced by RH, temperature, wind speed, and to a lesser extent, precipitation. RAC shows a moderate negative loading on this component as well, further supporting the view that radon levels tend to rise under stagnant, humid, and low-mixing conditions.

The third principal component (PC3), which explains 17.1% of the variance, highlights the more independent role of precipitation, with a strong positive loading, alongside temperature. While its influence is less dominant than PC1, this component points to potential effects of moisture conditions on radon behaviour such as soil saturation altering radon exhalation rates.

Overall, the PCA findings underscore that RAC is most strongly governed by meteorological conditions that control atmospheric mixing and dispersion, particularly BLH. These insights enhance our understanding of the environmental drivers of radon dynamics in the near-surface atmosphere.

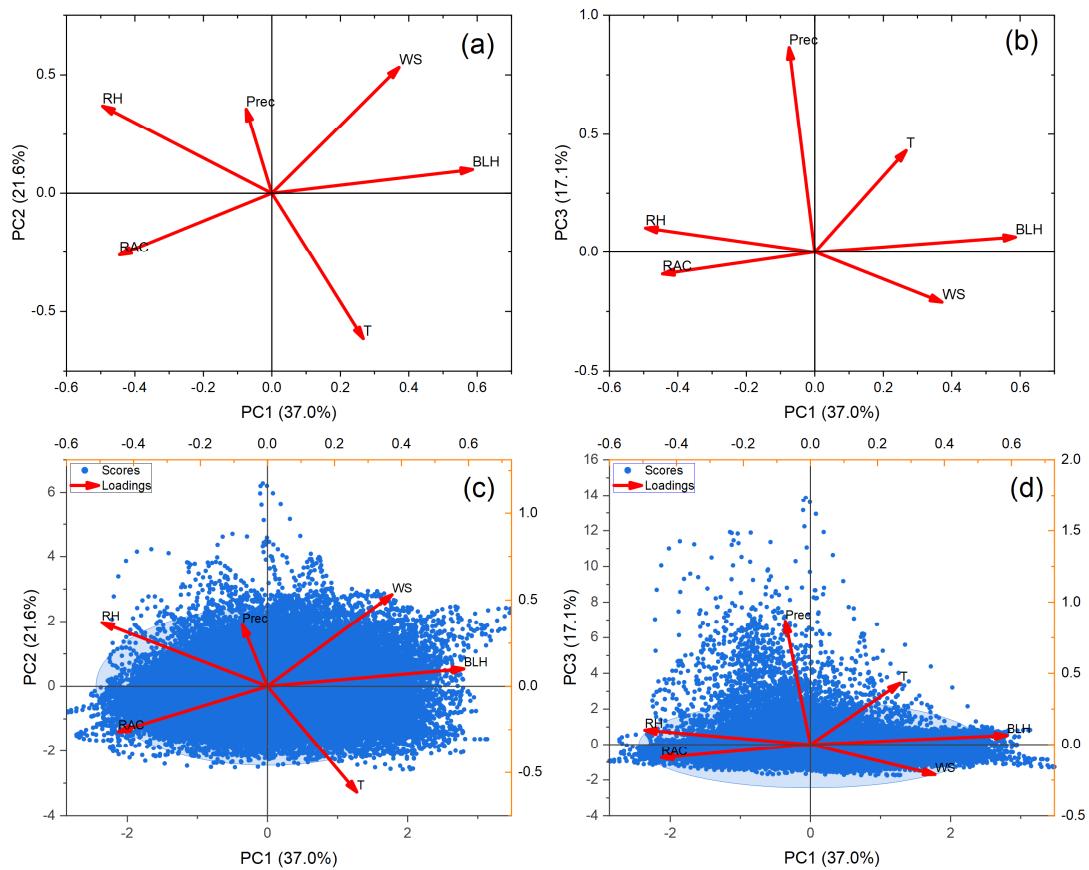


Fig. 4.8 PCA performed on RAC, BLH, and key meteorological variables. Panels (a) and (b) present the loading plots for PC1 vs. PC2 and PC1 vs. PC3, respectively, highlighting the correlations among the variables. Panels (c) and (d) show the corresponding score plots with loadings, illustrating the distribution of observations in the reduced principal component space.

Table 4.3 Results of the multicollinearity check based on the variance inflation factor.

All variables included										
Variable	Prec	SinH	P	WS	CosH	SinDoY	RH	BLH	CosDoY	T
VIF	1.05	1.15	1.16	1.58	1.64	1.77	2.08	2.23	4.42	5.63
Temperature is excluded										
Variable	Prec	SinH	SinDoY	P	CosDoY	CosH	WS	RH	BLH	
VIF	1.04	1.1	1.11	1.11	1.39	1.54	1.57	1.77	2.23	

4.5. Multiple linear regression analysis

MLR as discussed in **Section 2.8.5. Multiple linear regression**, was applied to establish a simpler and more interpretable model between radon and potential predictors. Before applying the MLR analysis, multicollinearity among the independent variables was assessed

using the so-called variance inflation factor (VIF). The VIF quantifies how much the variance of the estimated regression coefficients is inflated due to collinearity with other predictors in the model. Specifically, the VIF is the reciprocal of tolerance, calculated as $1 - R^2$, where R^2 represents the coefficient of determination between a given predictor and all other predictors (O'Brien, 2007). A high VIF indicates that a predictor is highly correlated with other predictors, which can lead to unreliable coefficient estimates and inflated standard errors. Typically, a VIF value greater than 5 or 10 suggests significant multicollinearity (O'Brien, 2007). Therefore, assessing the VIF values prior to model fitting helps identify potential multicollinearity issues and ensures the reliable interpretability of the regression results.

The initial VIF results showed that most variables had relatively low multicollinearity, with values ranging from 1.05 to 5.63 (**Table 4.3**). However, the temperature variable exhibited the highest VIF of 5.63, indicating considerably high degree of collinearity with some predictors. Although the overall VIF analysis indicated that multicollinearity was within acceptable limits; to further enhance model robustness, temperature was excluded from the final MLR analysis. After removing temperature, all VIF values dropped to below 2.3, indicating minimal multicollinearity among the remaining predictors.

Table 4.4 The summary of multiple linear regression analysis of RAC.

Variables	Standardized Coefficients	t-value	p-value
BLH	-0.36	-52.5	0.00
SinDoY	-0.23	-47	0.00
SinH	0.21	43	0.00
WS	-0.11	-19	0.00
CosDoY	0.10	18.7	0.00
CosH	-0.08	-14.6	0.00
P	-0.035	-7.2	0.00
Prec	-0.035	-7.2	0.00
RH	0.021	3.4	0.00

The summary of the MLR results, with RAC as the dependent variable and the selected meteorological parameters and BLH as predictors, is presented in **Table 4.4**. The MLR model demonstrated statistical significance ($p\text{-value} < 0.05$) with a correlation coefficient of $R^2 = 0.31$, indicating that approximately 31% of the variability in RAC was explained by the predictors.

As shown in **Table 4.4**, several variables were statistically significant contributors to the model. BLH had the largest standardized effect ($\beta = -0.36$, $t = -52.5$, $p < 0.001$), followed by SinDoY ($\beta = -0.23$), SinH ($\beta = 0.21$), and WS ($\beta = -0.11$). Other predictors such as CosDoY, CosH, P, Prec, and RH also showed statistically significant but smaller standardized coefficient. Based on these results, it can be concluded that the primary drivers of RAC variability are BLH and WS, both of which are related to the atmospheric mixing processes. Specifically, stronger atmospheric mixing tends to decrease RAC concentrations near the ground, while weaker mixing allows for higher near-surface accumulation. The strong standardized beta coefficients observed for SinDoY and SinH further suggest a significant influence of seasonal and diurnal cycles on RAC variability in the prediction model. By contrast, the contributions of other variables such as RH, P, and precipitation were much smaller. This could be attributed to their weak influence on RAC, or alternatively, to an underlying non-linear relationship between them that is not well captured by the linear MLR approach, leading to a relatively poor predictive performance of the MLR model.

4.6. Generalised additive model analysis

Following the poor performance of the MLR, the generalised additive model discussed in **Section 2.8.6** was applied, using a Gaussian distribution with an identity link function. The response variable (RAC) was modelled as a function of BLH, key meteorological factors, and time-related predictors. Prior to model fitting, concurvity, a measure of nonlinear dependence among smooth terms in GAMs was assessed (Hastie, 1992; Wood, 2017). Similar to multicollinearity in linear models, concurvity indicates that one smooth term can be approximated by others. This can potentially destabilize coefficient estimates, inflate standard errors, and reduce predictive reliability (Fox, 2015).

Table 4.5 presents the concurvity results for all predictors, including transformed time-related variables. The table reports three key metrics for each predictor's smooth term:

- **Worst concurvity:** the highest observed dependence between the smooth term and any other smooth term in the model, indicating the maximum redundancy with other predictors.
- **Observed concurvity:** the actual level of dependence measured between the term and the others based on the data.
- **Estimated concurvity:** a model-based estimate reflecting how well the term can be approximated by the other terms, smoothing out noise in the measurement.

The results showed that the sine and cosine terms for both day of the year (DoY) and hour of day (H) exhibit very high concurvity, with worst-case values reaching 1.00, indicating near-perfect redundancy among these time-based predictors. In contrast, meteorological variables display varying concurvity levels. Notably, precipitation has consistently low concurvity across all metrics, especially in the estimated concurvity (0.06), suggesting a strong, independent contribution to the model. Pressure, WS and BLH also show relatively low concurvity, with worst-case values below 0.65, indicating acceptable levels of dependence and a minimal risk of instability. On the other hand, T and RH showed moderate to high concurvity, particularly temperature, which has an observed value of 0.83 and an estimate of 0.69, highlighting potential issues with shared information among predictors. These findings imply that while some predictors in the model structure are largely independent, others, such as time-related variables, temperature and RH, may require additional evaluation or transformation to reduce multicollinearity and improve the model's interpretability.

Table 4.5 Concurvity analysis results for predictors used in the GAM (part 1).

Concurvity	s(Prec)	s(P)	s(WS)	s(RH)	s(BLH)	s(T)	s(SinDoY)	s(CosDoY)	s(SinH)	s(CosH)
worst	0.12	0.26	0.49	0.63	0.63	0.85	1	1	1	1
observed	0.11	0.1	0.36	0.55	0.62	0.83	0.35	0.8	0.15	0.59
estimate	0.06	0.18	0.35	0.56	0.48	0.69	0.5	0.81	0.2	0.48

For further analysis, the four time-related predictors were transformed using PCA. **Table 4.6** and **Fig. 4.9** present the PCA results for these variables. As shown in **Table 4.6**, the first three principal components explain 75.2% of the total variability among the four time-related predictors. These PCs are orthogonal (i.e., uncorrelated), making them suitable for inclusion in the GAM without introducing multicollinearity. This transformation allows the model to retain the temporal structure of the data while improving numerical stability and interpretability.

Table 4.6 PCA results for time-related predictors.

Principal Component Number	Eigenvalue	Percentage of Variance (%)	Cumulative (%)
1	1.01	25.28	25.28
2	1.00	25.01	50.30
3	0.99	24.93	75.23
4	0.99	24.76	100

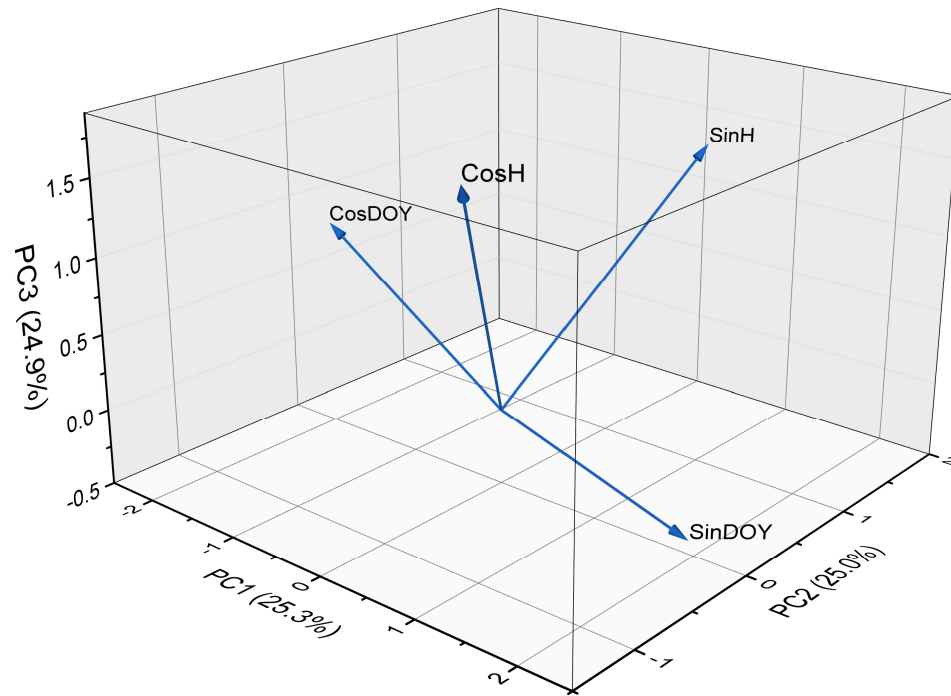


Fig. 4.9 Three-dimensional PCA loading plot of the time-related predictors, illustrating the loadings of principal components and the distribution of observations in the reduced feature space.

The three PCs derived from the time-related predictors, along with the remaining meteorological variables and BLH, were subjected to a concurvity analysis, as presented in Table 4.7. The results show a marked improvement in multicollinearity levels compared to the original time-based predictors. Specifically, all three PCs exhibit moderate concurvity, with worst-case values ranging from 0.55 to 0.76, substantially lower than the perfect concurvity (1.00) observed in the original (untransformed) time variables. Once again, the meteorological variables show low to moderate concurvity, with precipitation showing the lowest values (estimate = 0.06). These results confirm that PCA effectively reduces multicollinearity among time-related variables, making the transformed components suitable for inclusion in the GAM without compromising model reliability.

Table 4.7 Concurvity analysis results of predictors used in the GAM (part 2).

Concurvity	s(Prec)	s(P)	s(WS)	s(T)	s(PC1)	s(RH)	s(BLH)	s(PC2)	s(PC3)
worst	0.11	0.22	0.45	0.49	0.55	0.57	0.64	0.68	0.76
observed	0.11	0.07	0.22	0.48	0.38	0.42	0.62	0.16	0.62
estimate	0.06	0.13	0.32	0.39	0.39	0.51	0.49	0.17	0.58

The summary of the GAM analysis is presented in **Table 4.8**, including model fit statistics, parametric coefficients, and the significance of smooth terms. The GAM demonstrated moderate explanatory power for RAC, with an adjusted R^2 of 0.37 and 37.1% deviance explained. All predictors showed statistically significant non-linear relationships ($p < 0.001$). Among them, BLH emerged as the strongest driver ($F = 448.1$), followed by time related predictors PCs (PC1–PC3) and key meteorological variables. High effective degrees of freedom (edf) for BLH, P, and T indicates complex, non-linear interactions with RAC, while other variables like RH, Prec, and WS also showed notable, though slightly simpler, effects. The model intercept of 5.54 represents the baseline RAC under reference conditions.

Table 4.8 Summary of GAM results for RAC, including model specifications, parametric coefficients, and the significance of smooth terms. edf (effective degrees of freedom) reflects the complexity of smooth terms higher values indicate greater flexibility or nonlinearity. Ref.df (reference degrees of freedom) is used for hypothesis testing of the smooth terms. F represents the F-statistic used to assess the significance of each smooth term.

Family	Link function	Formula	adj. R^2	Deviance explained
Gaussian	Identity	RAC ~ s(PC1)+s(PC2)+s(PC3) +s(P)+s(RH)+s(T)+s(Prec)+s(WS)+s(BLH)	0.37	37.1%
Parametric coefficients				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	5.54	0.014	395.9	<2e-16
Approximate significance of smooth terms				
Predictors	edf	Ref.df	F	p-value
s(BLH)	8.54	8.94	448.1	0.001
s(PC2)	8.4	8.9	265	0.001
s(PC1)	7.88	8.8	207	0.001
s(PC3)	3.84	4.86	120	0.001
S(T)	7.9	8.7	117	0.001
S(RH)	5.4	6.6	72.6	0.001
s(P)	8.32	8.87	21.97	0.001
s(Prec)	5.5	6.5	20.36	0.001
s(WS)	5.8	6.8	14.2	0.001

The estimated smooth functions from the GAM analysis provide a visual representation of the relationship between predictors and RAC (**Fig. 4.10**). Pressure showed a complex relationship, with RAC peaking around 990–1010 hPa and decreasing at both lower and higher pressures (**Fig. 4.10a**). RH exhibits a modest variation across most humidity levels, with a slight increase at high RH (>80%), possibly reflecting atmospheric conditions that limit vertical mixing typical for cold periods of the year (**Fig. 4.10b**). Temperature displays

a strong non-linear pattern, with RAC rising sharply at temperatures above 10 °C, likely driven by enhanced radon exhalation rates during warm periods (**Fig. 4.10c**). However, this contrasts with the fact that higher temperatures intensify thermally driven atmospheric mixing, thus reducing RAC levels near the ground. As shown in **Fig. 4.10d**, precipitation does not have a strong overall effect on RAC variability, although there is a noticeable decrease in RAC around 2.5 mm of rainfall, which is likely due to the short-term reduction of radon exhalation rate. Beyond that point, the scarcity of high intensity precipitation data introduces large uncertainties, making it difficult to clearly interpret its influence. **Fig. 4.10e** demonstrates the suppressing effect of wind speed on near-surface radon levels due to enhanced turbulent mixing. Finally, BLH (**Fig. 4.10f**) shows the strongest inverse and non-linear relationship with RAC. As BLH increases, RAC declines rapidly, particularly at lower heights (<1000 m), which highlights the critical role of vertical atmospheric mixing in controlling radon concentrations near the surface.

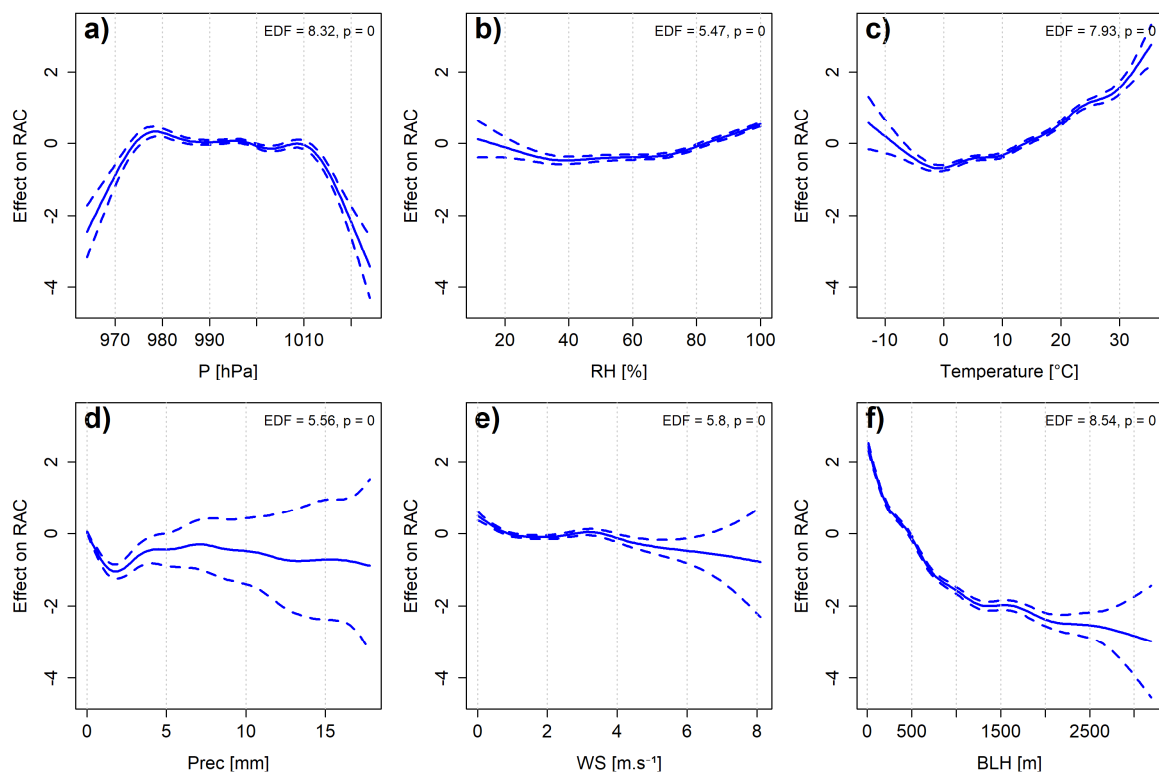


Fig. 4.10 Estimated smooth functions of predictors from the GAM analysis. Solid blue lines represent the estimated smooth terms, and shaded lines indicate 95% confidence intervals. EDF stands for effective degrees of freedom, indicating the model complexity for each variable. EDF stands for effective degrees of freedom showing the smooth functions complexity for each variable.

4.7. Machine learning-based regression models

The GAM and MLR were employed as conventional regression models to model RAC variability in relation to BLH and key meteorological factors. The MLR provided simple and interpretable insights into the most significant predictors and their directional influence on RAC. However, it was unable to capture non-linearity and demonstrated limited predictive performance. While the GAM analysis was able to effectively capture the non-linear relationships between potential predictors and RAC variability, its predictive power was still relatively weak. Both models suffer from conventional challenges inherent in regression analysis, including potential non-stationarity in the time series, multicollinearity among predictors, and deviations from the assumption of normally distributed residuals. To overcome these limitations, machine learning-based ensemble regression models such as GBM, XGBoost and RF were applied (discussed in **Sections 2.9.1 – 2.9.3**). A summary of results for all models, including the MLR, is presented in **Fig. 4.11** and **Fig. 4.12** in the form of scatter plots and feature importance analyses. Furthermore, **Table 4** provides a comparison of model performances based on the evaluation metrics introduced in **Section 2.10**.

As shown in the scatter plots of predicted versus measured RAC (**Fig. 4.11**), the XGBoost model achieved the highest predictive performance with an R^2 value of 0.87, followed by GBM and RF models, both with $R^2 = 0.76$. The MLR model exhibited a much lower predictive capability, with an R^2 of only 0.31, highlighting the greater ability of ML models to capture the underlying complexity and nonlinearity in RAC dynamics compared to linear methods.

The feature importance plots (**Fig. 4.12**) revealed some recurring patterns across all models. In all ML models (XGBoost, GBM, and RF), BLH is identified as the most influential predictor, suggesting that vertical mixing in the atmosphere plays a dominant role in controlling RAC levels. The second most important predictor across all models is SinDoY, emphasizing the strong seasonality embedded in RAC variability. The next four predictors, CosDoY, T, RH and P, show moderate levels of importance and reflect the influence of general meteorological conditions on RAC levels. The relative importance of these features is similar across XGBoost, GBM, and RF models, pointing to a shared understanding of the environmental drivers influencing RAC. Notably, Prec consistently ranks as the least important feature across all models, suggesting that, under the studied conditions, precipitation-driven effects play a relatively minor role compared to atmospheric mixing and

seasonal influences. This may be due to the sporadic nature of rainfall events, which occur only occasionally at a data resolution of one hour. As a result, its impact is minimal compared to that of other datasets that vary continuously. This interpretation is further supported by the application of a larger moving average window, as well as the relatively moderate correlations observed between RAC and Prec (**Fig. 4.3**).

The MLR model showed a roughly similar ranking of feature importance based on the standardized beta coefficients, with BLH emerging as the dominant predictor, followed by SinDoY and other meteorological variables. However, the lower R^2 value for MLR indicates that conventional linear regression is insufficient to fully capture the interactions between RAC and its predictors, further underlining the value of ML approaches for this type of environmental modelling.

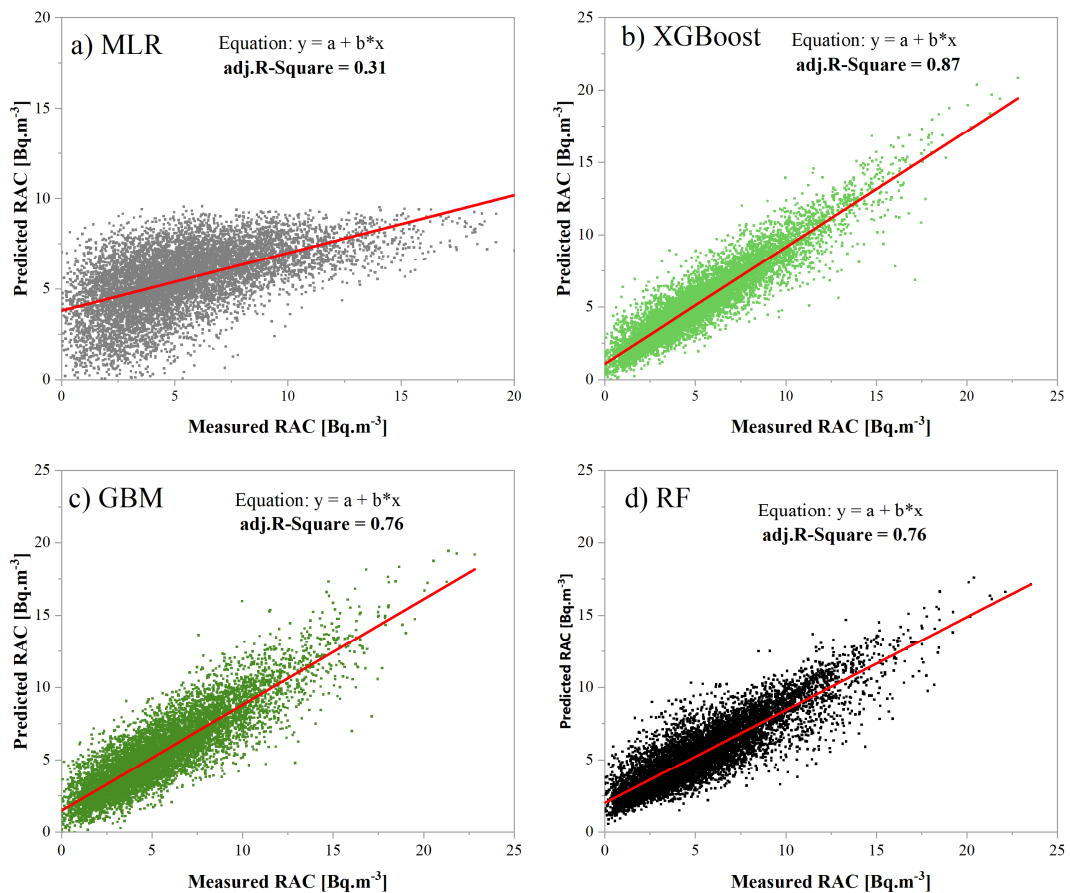


Fig. 4.11 Scatter plots of measured RAC against predicted RAC based on different regression models. The predicted values are based on test dataset for ML models.

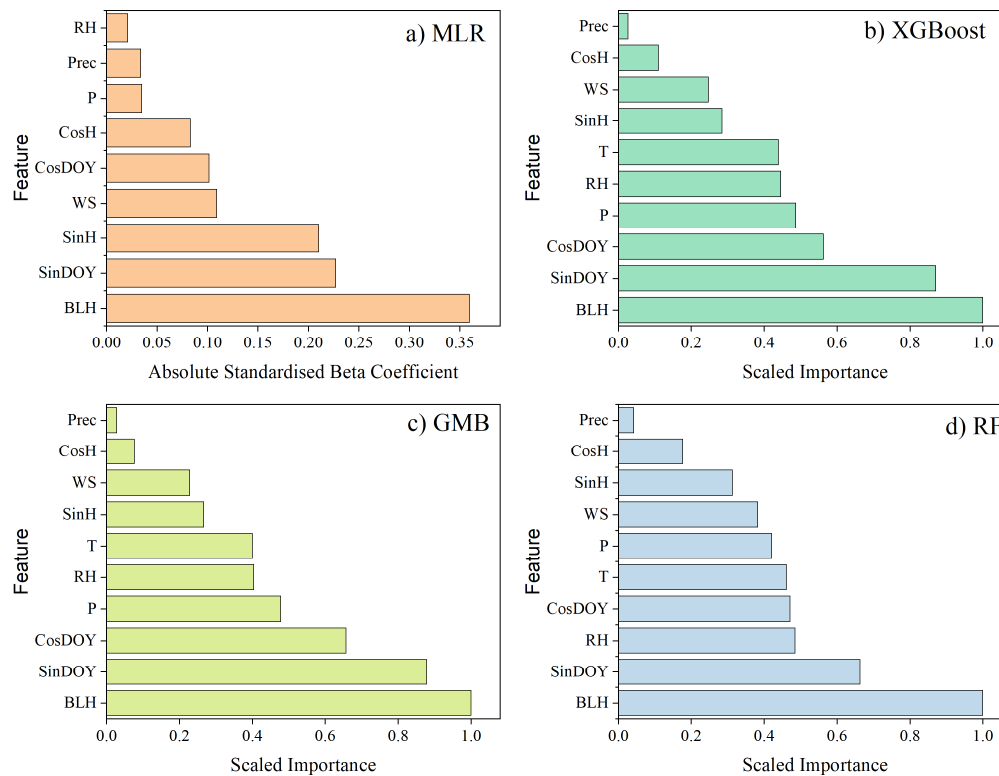


Fig. 4.12 Feature importance plots of different regression models.

4.8. Evaluation of regression models

The performance of regression models used in this study, evaluated based on metrics introduced in **Section 2.10**, are shown in **Table 4.9** and are visually depicted in **Fig. 4.13**. Among the models evaluated, XGBoost demonstrated the best overall performance, achieving the highest cross validation $R^2 = 0.86$ and adj. $R^2 = 0.86$ values, along with the lowest MSE (1.39), RMSE (1.18), and MAE (0.86). GBM and RF models also performed reasonably well, whereas MLR and GAM exhibited comparatively lower predictive capabilities. These results indicate the effectiveness of machine learning ensemble models, particularly XGBoost, in modelling the RAC and its associated environmental factors.

Table 4.9 Evaluation metrics for different regression models.

Model	XGBoost	GBM	RF	GAM	MLR
CV R-square	0.86	0.76	0.73		
adj.R-Square	0.87	0.76	0.76	0.37	0.31
MSE	1.39	2.41	2.6	3.84	6.8
RMSE	1.18	1.55	1.61	1.96	2.6
MAE	0.86	1.19	1.21	2.56	2

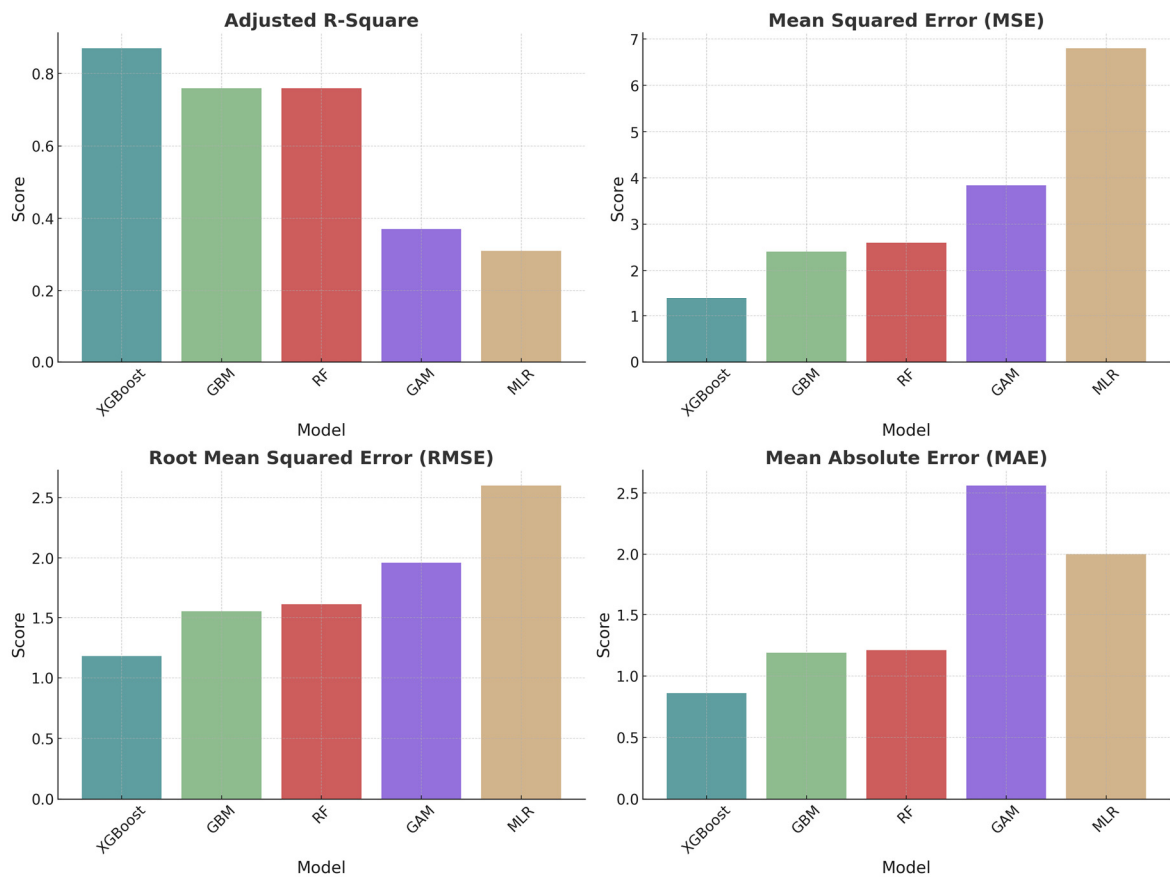


Fig. 4.13 Regression models performance evaluation metrics.

4.9. SHAP analysis of the XGBoost model

The Shapley Additive exPlanations (SHAP) method was used to interpret the performance of the best-performing model used for the prediction of RAC (i.e., XGBoost model). SHAP is a game theory-based approach widely used for interpreting ML models prediction (Lundberg & Lee, 2017). In cooperative game theory, the Shapley value is a method used to fairly distribute gains or costs among players who have collaborated, based on their individual contributions. In machine learning, it is designed to quantify the contribution of individual features to the model's predictions. Typically, ML outputs are treated as "black boxes" due to their complexity and lack of interpretability. By calculating SHAP values, it becomes possible to understand how each feature influences the model's output, providing deeper insights into feature importance and behaviour.

Fig. 4.14 presents the *SHAP summary plot*, showing the distribution and relative importance of each feature in the outcome. Features are ranked according to their mean absolute SHAP values, indicating their overall influence on the model output. As shown, BLH emerged as the most influential predictor, followed by CosDoY, T, and SinH. Lower BLH and higher

temperature values were generally associated with an increase in the predicted RAC, as indicated by their predominantly positive SHAP values. Conversely, features such as RH and precipitation exhibited relatively smaller impacts on the prediction of RAC. The colour gradient in the plot further reveals how high and low feature values influence predictions, providing additional insight into the relationships captured by the model.

For further evaluation, the *SHAP dependence plot* of the XGBoost model is plotted in **Fig. 4.15**. A SHAP dependence plot visualizes the relationship between a feature's value and its SHAP value, showing how changes in the feature affect the model's predictions. In our dataset, features such as temperature, BLH, and WS exhibit the most pronounced effects on radon levels. Specifically, higher temperatures are associated with increased RAC, while greater BLH and WS correspond to lower RAC, likely due to enhanced atmospheric mixing and dispersion. Higher RH is linked to higher radon levels; pressure shows a weak nonlinear relationship. Precipitation has a minimal impact, with most values clustered close to zero. These results highlight the model's ability to capture complex, nonlinear relationships and interactions among meteorological variables influencing RAC levels.

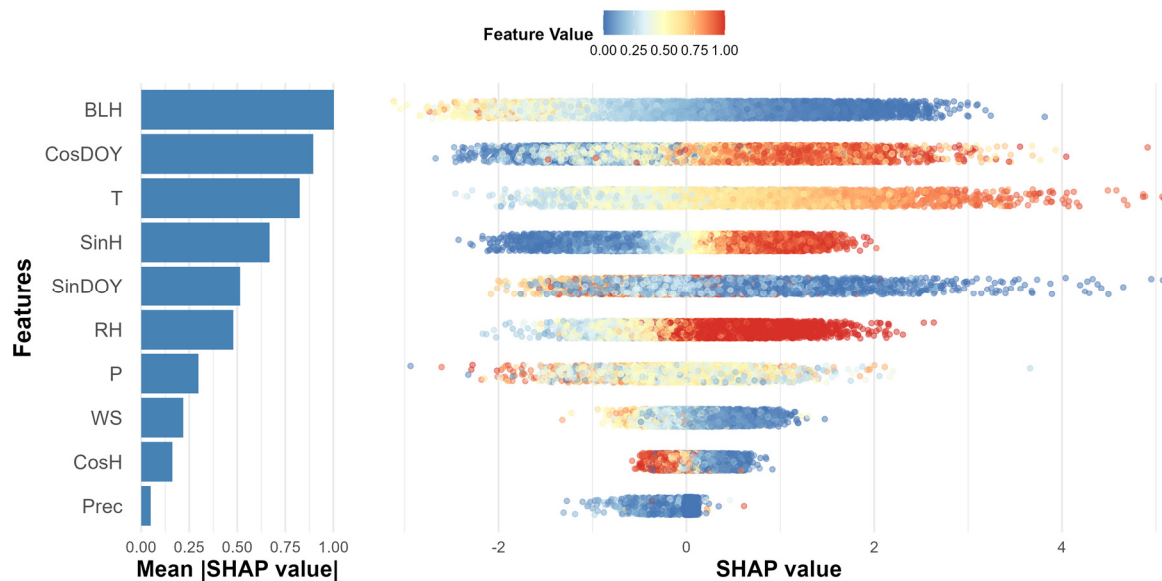


Fig. 4.14 SHAP summary plot, illustrating the contribution of each feature to the predicted RAC for the XGBoost model. The right-hand panel shows a beeswarm plot displaying the distribution and direction of SHAP values for each feature. The plot is coloured according to the normalised feature values, revealing the impact of different feature magnitudes on predictions. The left-hand panel shows a bar chart of the mean absolute SHAP values, ranking features by their overall importance. Together, these visualizations provide a comprehensive understanding of the influence of each feature on the model output.

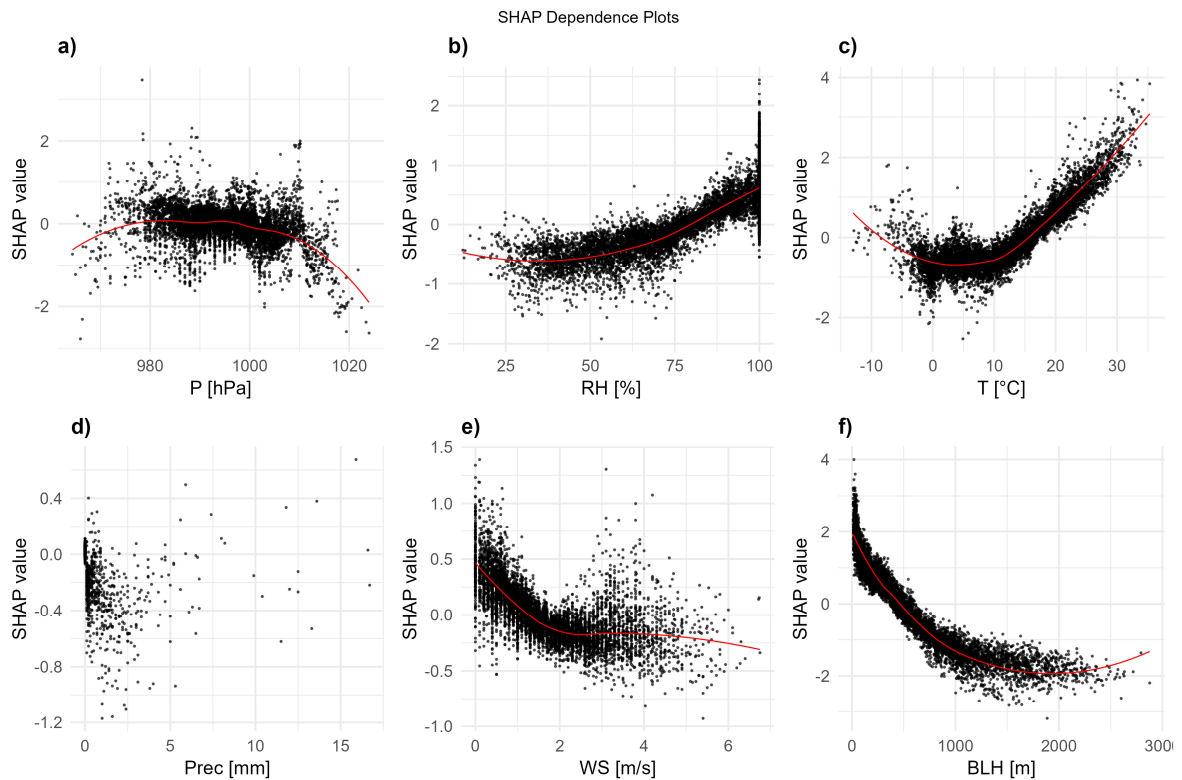


Fig. 4.15 SHAP dependence plots of each feature used in the XGBoost model. Each plot shows the influence of a single feature on the model's output, with SHAP values indicating the magnitude and direction of contribution.

4.10. Partial dependence plots of the XGBoost model

The partial dependence plots (PDPs) of the key environmental features incorporated into the XGBoost model are shown in **Fig. 4.16**. PDPs are widely used model-agnostic interpretation tools that depict the marginal effect of a single feature on a model's predicted outcome, averaging out the influence of all other variables (Friedman, 2001). They show how the model's prediction changes when the value of one feature varies, while the other features remain constant. This enables a global understanding of the direction and strength of each feature's influence on the model output. In this case, the plots reveal the marginal effects of predictors on the predicted radon levels. Similar to the SHAP dependence plots (**Fig. 4.15**), the plots reveal that features such as temperature, BLH, and WS have strong nonlinear effects on RAC, with high temperatures, low BLH and WS associated with elevated radon levels. Relative humidity also exerts a generally positive influence, particularly at high humidity levels. Pressure shows a non-linear relationship, with RAC predictions peaking at intermediate pressures and declining at both low and high extremes. Precipitation shows a

sudden decrease in RAC with an increase in rainfall. Rainfall above 5 mm shows unstable behaviour, which may be due to the limited data available in this range.

Remarkably, all these patterns are consistent with those observed in the SHAP and GAM-based analyses, providing further validation across different interpretive frameworks.

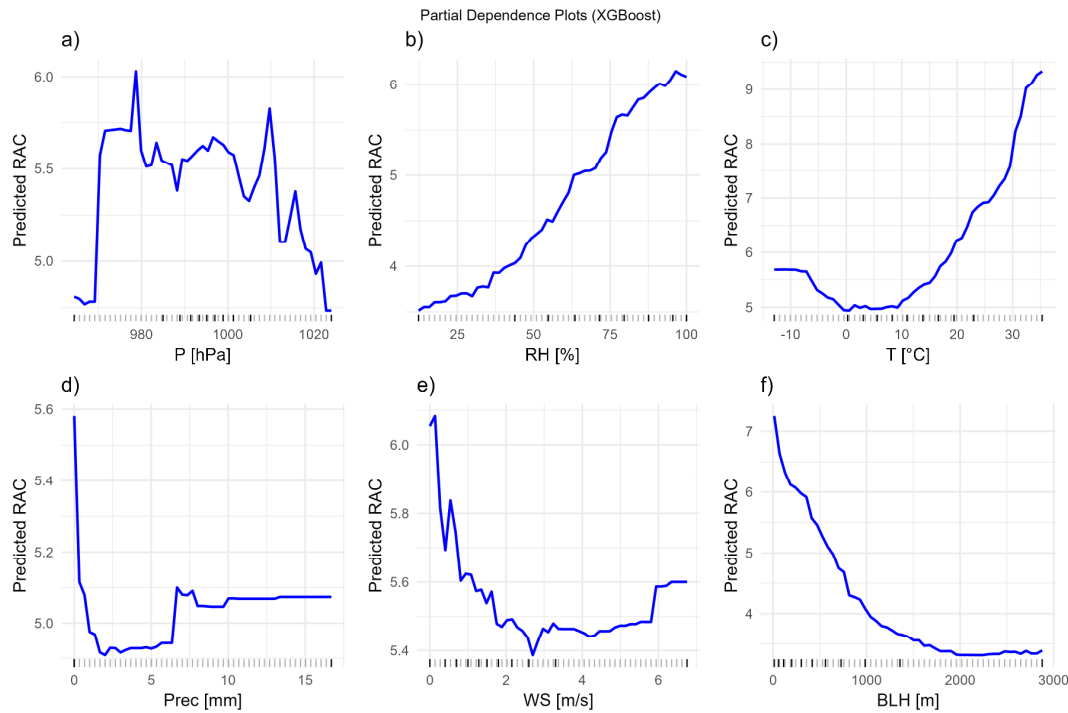


Fig. 4.16 Partial dependence plots for each feature that was used to predict RAC in the XGBoost model.

Chapter 5

5. Radon decay products

In this chapter, the study of radon decay products (i.e., ^{218}Po , ^{214}Pb , ^{214}Bi , and ^{210}Pb), along with the radon equilibrium factor and radon progeny ratios, is presented. Additionally, these radionuclides, together with ^{222}Rn , are used to estimate aerosol residence times. This topic will be discussed in the next chapter. Emphasis is placed on the behaviour and application of these radionuclides as tracers in atmospheric research.

Simultaneous measurements of radon and its decay products provide useful information that can be used in radiation protection and atmospheric research. These data provide information such as the equilibrium equivalent radon concentration (EEC) and the radon equilibrium factor (F_{eq}), which are important parameters for estimating the effective dose from radon. The EEC can be understood as the concentration of radon that is in a state of secular equilibrium with its progeny and poses the same health risk as the actual mixture of radon and progeny in air (Tirmarche et al., 2010). The measurement of radon and its progeny can also be used to determine the residence time of atmospheric aerosols (T_R) (e.g., Baskaran & Shaw, 2001; Długosz–Lisiecka & Bem, 2012; Lambert et al., 1983; Porstendörfer et al., 2000). The aerosols residence time is an important factor which provides insight into the average amount of time aerosols spend in the atmosphere before being deposited on the Earth's surface.

The EEC was calculated using the formulas published in UNSCEAR (2000):

$$EEC = 0.105.C_{Po} + 0.515.C_{Pb} + 0.380.C_{Bi} \quad (5.1)$$

where C_{Po} , C_{Pb} and C_{Bi} are the activity concentrations of ^{218}Po , ^{214}Pb and ^{214}Bi in the atmosphere, respectively. The ratio of EEC and RAC is called the radon equilibrium factor F_{eq} , mathematically defined as

$$F_{eq} = \frac{EEC}{RAC}, \quad (5.2)$$

where RAC is the ^{222}Rn activity concentration measured simultaneously with its short-lived progeny.

5.1. Descriptive statistics of radon progeny

The descriptive statistics of radon short-lived progeny and related calculated variables measured during the study period (December 2019 to October 2022) are presented in **Table 5.1**. Measurements were taken every two hours, resulting in over 9,000 data points for each variable. The mean activity concentrations of ^{218}Po , ^{214}Pb , and ^{214}Bi were 2.78, 3.21, and 2.75, respectively, with ^{214}Pb exhibiting the highest variability (standard deviation = 1.95). The average EEC was 3.01.

Table 5.1 Descriptive statistics of radon short-lived progeny, EEC and F_{eq} measured during the study period (December 2019 to October 2022).

Variable	N	Minimum	Maximum	Mean	Std. Deviation
^{218}Po	9502	0	11.72	2.78	1.65
^{214}Pb	9518	0	13.35	3.21	1.95
^{214}Bi	9541	0	9.36	2.75	1.51
EEC	9423	0	11.38	3.01	1.7
F_{eq}	9346	0	7.50	0.44	0.26

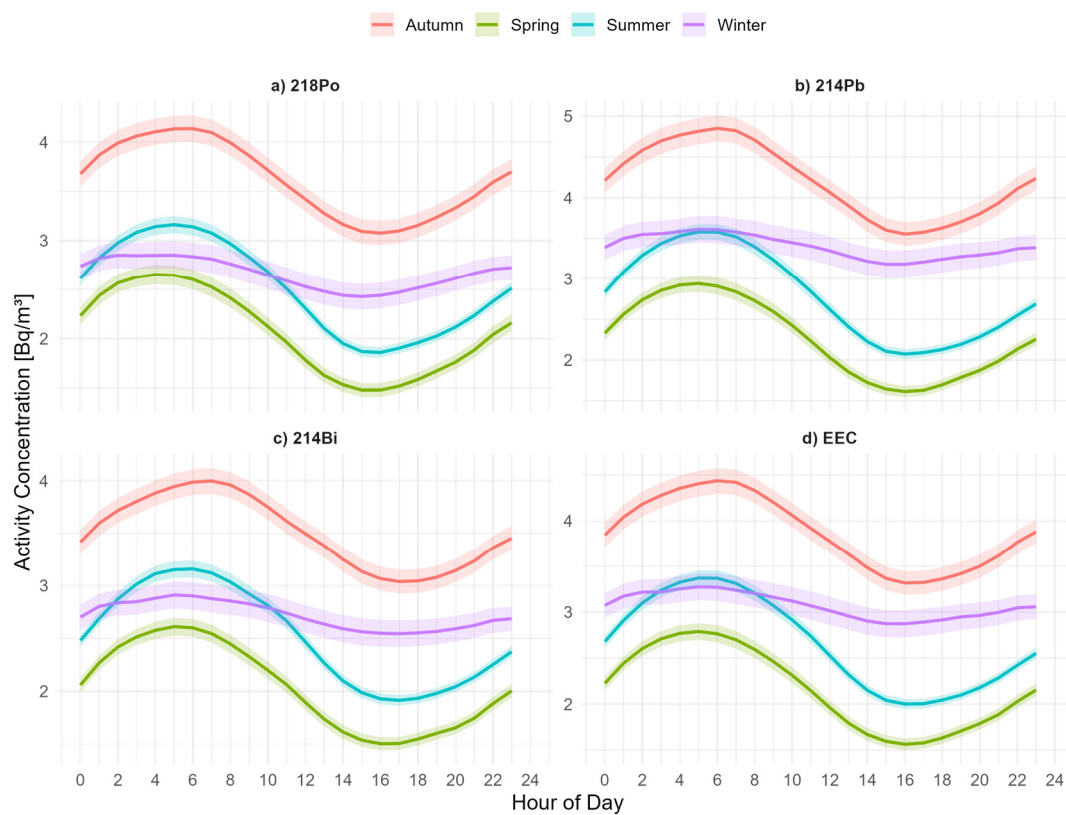


Fig. 5.1 Diurnal cycles of short-lived radon progeny and its equilibrium equivalent concentration.

The diurnal cycle of short-lived radon decay products along with EEC is shown in **Fig. 5.1**. As expected, they follow the same pattern as RAC, discussed in **Section 3.6**, i.e., the highest concentrations were observed in the early morning and lowest in the mid-afternoon. This pattern is mainly attributed to the diurnal changes in BLH influenced by meteorological factors such as temperature and wind speed discussed in **Section 4.2**.

The seasonal variation of short-lived radon progeny concentrations, presented as composite monthly boxplots (**Fig. 5.2**), exhibits a pattern very similar to that of radon, as discussed in **Section 3.7**. The lowest variability in radon progeny concentrations was observed in April, followed by a steady increase, reaching a maximum in November. The observed seasonal variation is mainly governed by changes in BLH and radon flux variability, as previously discussed in the case of radon (**Section 4.3**).

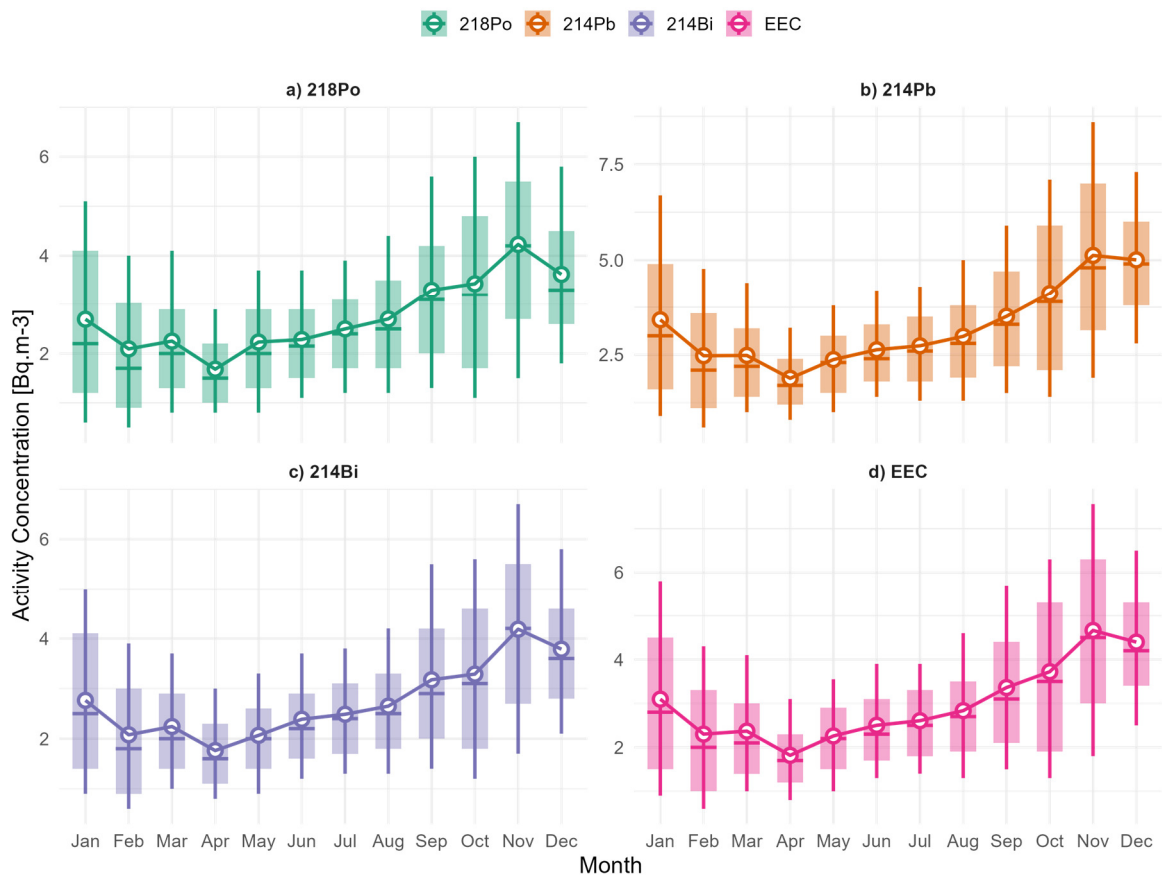


Fig. 5.2 Seasonal boxplots of short-lived radon progeny and EEC. The horizontal lines represent the median, while the circles denote the mean. The boxes represent the 25th to 75th percentile range, while the bars extend this range to the 10th to 90th percentile.

5.2. Radon equilibrium factor

The F_{eq} is an important parameter in radiation protection and atmospheric research. It is widely used to assess the annual effective dose from inhalation of radon and its decay products (UNSCEAR, 2000) and to estimate the residence time of aerosols in the atmosphere based on the ratios of radon and its progeny (Baskaran & Shaw, 2001; Sýkora et al., 2017). Spatial, diurnal and seasonal variations of F_{eq} in the outdoor atmosphere depend mainly on the radon exhalation rate, meteorological and atmospheric mixing conditions (Chen & Harley, 2018).

The radon equilibrium factor calculated according to Equation (5.2) ranges from 0.04 to 7.5, with a mean \pm std. deviation of 0.54 ± 0.28 . The interquartile range for F_{eq} is (0.4–0.6), and the data showed that 5.6% of observations are greater than $F_{eq} > 1$. The range of F_{eq} observed in this study is consistent with the values reported by Chen and Harley (2018), which range from 0.18 to 0.67. However, the mean value of F_{eq} is slightly lower than the typical value of 0.6 given in the UNSCEAR 2000 report. The standard UNSCEAR value of F_{eq} is used when the specific environmental F_{eq} is not known; UNSCEAR acknowledges that this value can vary by more than 50%.

Fig. 5.3a shows the composite diurnal cycle of F_{eq} averaged over the entire study period. F_{eq} exhibits a clear 24-hour pattern: the lowest value observed at 00:00 (~0.45); afterwards, F_{eq} gradually increases and reaches a maximum between 13:00 and 15:00 (~0.58), then it declines steeply. The higher F_{eq} could be related to an increased attachment rate of radon progeny during the day, due to higher aerosol concentrations. As airborne aerosol levels rise primarily because of human activities such as traffic, this leads to an increase in the attached fraction of radon progeny. This is supported by the findings of Abdelfatah Mostafa et al. (2020), who observed a positive correlation between F_{eq} and aerosol concentration in a standard radon box. This is consistent with the observations published in Winkler et al. (2001) for Munich–Neuherberg, Germany. According to these findings, the diurnal cycle of F_{eq} , when averaged over all seasons, starts to increase after sunrise, peaks around 13:00, and gradually decreases towards nighttime.

Fig. 5.4a presents the seasonal variation of the boxplots of F_{eq} . The interquartile range indicates the lowest variability in March, with values increasing through spring and peaking in May and June. A slight decline is observed in August, followed by a gradual increase through autumn, reaching a maximum in November. F_{eq} then declines slightly during winter,

reaching its lowest levels in March. Similar seasonal patterns were reported in Munich–Neuherberg, Germany, where higher F_{eq} values were observed during winter compared to summer (Winkler et al., 2001).

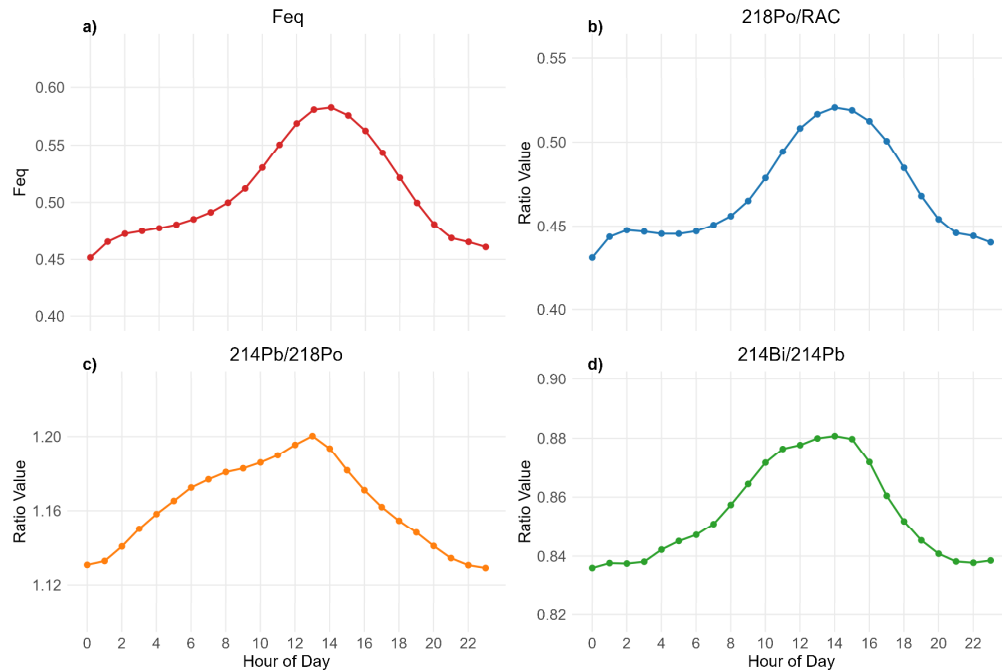


Fig. 5.3 Composite diurnal cycles of F_{eq} and radon progeny ratios averaged over the entire study period.

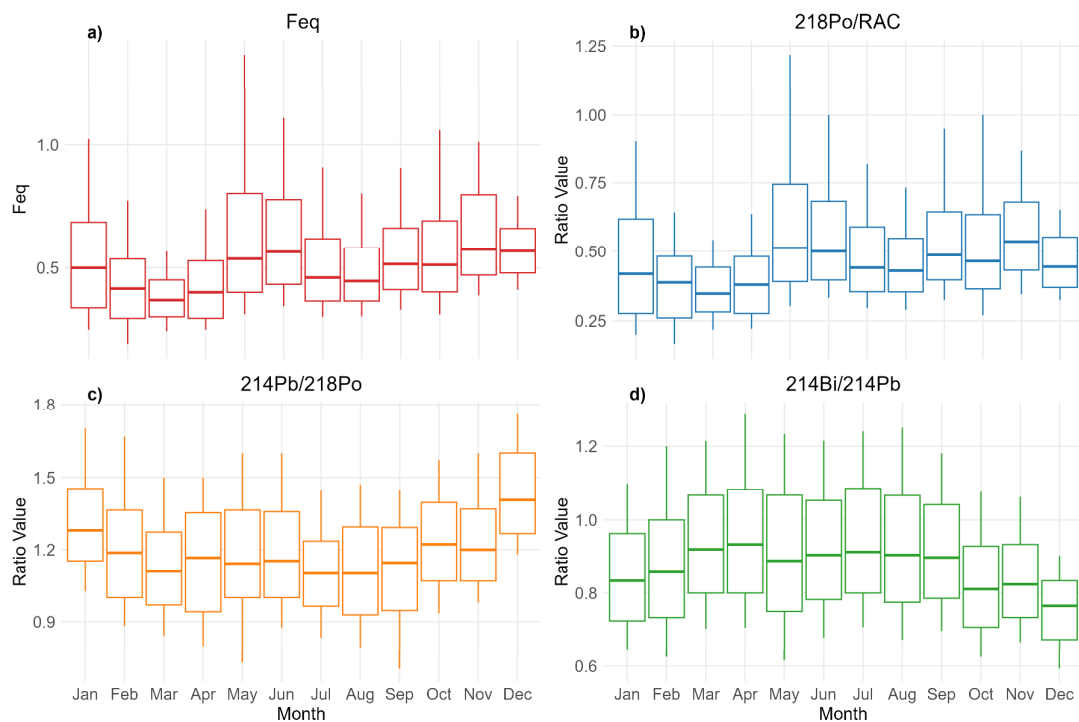


Fig. 5.4 Composite monthly boxplot of F_{eq} and radon progeny ratios over the entire study period. The horizontal lines represent the median, the boxes indicate the 25th to 75th percentile range, and the bars illustrate the 10th to 90th percentile range.

5.3. Radon and its progeny ratios

In this section, the temporal variation of radon progeny ratios is discussed. Studying these ratios is important for understanding radon progeny behaviour, their disequilibrium, and their applications in atmospheric research. Such ratios have been applied in various contexts, including the estimation of aerosol residence times (Barba-Lobo et al., 2024), atmospheric diffusion modelling (Jacobi & André, 1963), and the investigation of wet deposition processes (Liu et al., 2014). Under ideal open-air conditions, where removal occurs only by radioactive decay, the short-lived decay products of radon will reach equilibrium with their parent isotopes within a few half-lives of the decay products. In the real atmosphere, however, this equilibrium is significantly affected by environmental factors, resulting in a disequilibrium between the parent and daughter nuclides. In radioactive decay chains, the daughter/parent activity ratio can vary depending on the equilibrium state. In secular equilibrium, the daughter's activity equals the parent's (ratio ≈ 1). In transient equilibrium, the ratio can temporarily exceed 1 before stabilizing. Under disequilibrium conditions, the activity ratio may deviate significantly from unity. In this study, the average activity ratio of $^{214}\text{Pb}/^{218}\text{Po}/^{214}\text{Bi}$ over the study period was 1:0.86:0.85. Surprisingly, the $^{214}\text{Pb}/^{218}\text{Po}$ ratio tended to be higher than one during the whole measurement period (**Fig. 5.3c**). A similar ratio has been reported in other studies e.g., (Schery & Wasiolek, 1993; Sheets & Lawrence, 1999). The observed higher-than-one ratio can be attributed to several factors, including the different attachment processes of these radionuclides and specific meteorological conditions. The half-life of ^{218}Po of 3.1 minutes is considerably shorter than the half-lives of ^{214}Pb (26.8 minutes) and ^{214}Bi (19.9 minutes). Since the activity concentration measured here represents the attached fraction of these radionuclides, ^{218}Po , with its much shorter half-life, has significantly less time to attach to aerosols than ^{214}Pb and ^{214}Bi . This idea is supported by studies that have reported a much higher fraction of unattached ^{218}Po compared to its attached fractions (Abdelfatah Mostafa et al., 2020; Abdo et al., 2021).

The composite diurnal cycle of three sequential daughter/parent ratios in the radon decay chain ($^{218}\text{Po}/^{222}\text{Rn}$, $^{214}\text{Pb}/^{218}\text{Po}$, and $^{214}\text{Bi}/^{214}\text{Pb}$) is shown in **Fig. 5.3**. The ratios of successive radionuclide pairs exhibit distinct diurnal variations, reflecting changes in atmospheric conditions and radioactive decay dynamics. The $^{218}\text{Po}/^{222}\text{Rn}$ ratio (**Fig. 5.3b**) ranges from 0.43 to 0.53, remains low during the night and high during the day, similar to that of F_{eq} . It begins to increase sharply after 08:00, reaching a peak between 13:00 and 14:00, before declining again and stabilizing at lower values during the night. The $^{214}\text{Pb}/^{218}\text{Po}$ ratio shows

a different pattern, ranging from 1.13 to 1.20. It consistently increases from 0.00, reaches its peak at 13:00 and then gradually decreases throughout the afternoon and evening. The $^{214}\text{Bi}/^{214}\text{Pb}$ ratio diurnal pattern ranges from 0.84 to 0.88 and begins to increase in early morning around 4:00, peaks around 14:00, and then declines steadily, becoming relatively low and stable overnight. These diurnal cycles reflect the interplay between radioactive decay, aerosol attachment rate, and atmospheric mixing, which influence the build-up and dispersion of radon progeny throughout the day.

To evaluate the seasonal variation of these ratios, composite monthly boxplots are presented in **Fig. 5.4**. As expected, the $^{218}\text{Po}/^{222}\text{Rn}$ ratio displays a seasonal pattern similar to that of the F_{eq} (**Fig. 5.4b**). In contrast, the $^{214}\text{Pb}/^{214}\text{Po}$ ratio shows a different trend, remaining relatively stable from March to September, with a slight increase observed during the winter months (**Fig. 5.4c**). Meanwhile, the $^{214}\text{Bi}/^{214}\text{Pb}$ ratio exhibits a more pronounced seasonal variation, with lower values recorded during the colder months while higher values observed during the warmer months (**Fig. 5.4d**). These seasonal trends likely reflect changes in atmospheric mixing conditions and aerosol concentrations, which influence the behaviour and equilibrium of radon progeny in ambient air.

5.4. Radon progeny interaction with particulate matter

As radon progeny attach to fine aerosols, their atmospheric transport, deposition and inhalation dynamics are bound to the host aerosols. This association is important because both radon progeny and particulate matter (PM) are major carcinogens that contribute to air pollution-related health risks (Zoran et al., 2013). The ability of radon progeny to attach to fine PM facilitates its inhalation and deposition in the respiratory system. Despite extensive knowledge of the behaviour of these carcinogens in the atmosphere and their associated health risks, their interactions and relationships remain poorly understood. For example, several studies (Tokonami, 2000; 2000; Yu et al., 2023; 2013; Zoran et al., 2013) have investigated the variability of radon/radon progeny in relation to PM. In this thesis, simultaneous observations of radon progeny with PM_{2.5} and PM₁₀ are considered, and their relationship is investigated.

For visualization purposes, the hourly time series of EEC, along with PM_{2.5} and PM₁₀, is plotted in **Fig. 5.5** for the period of one year (May 2020 – April 2021). Certain periods exhibit high EEC levels, accompanied by similarly elevated PM concentrations. This is further investigated through the Pearson correlation heatmap shown in **Fig. 5.6**. A moderate

positive correlation, with Pearson’s r ranging from 0.48 to 0.50, was observed between short-lived radon progeny and PM. This suggests that high radon progeny concentrations are associated with high pollution events.

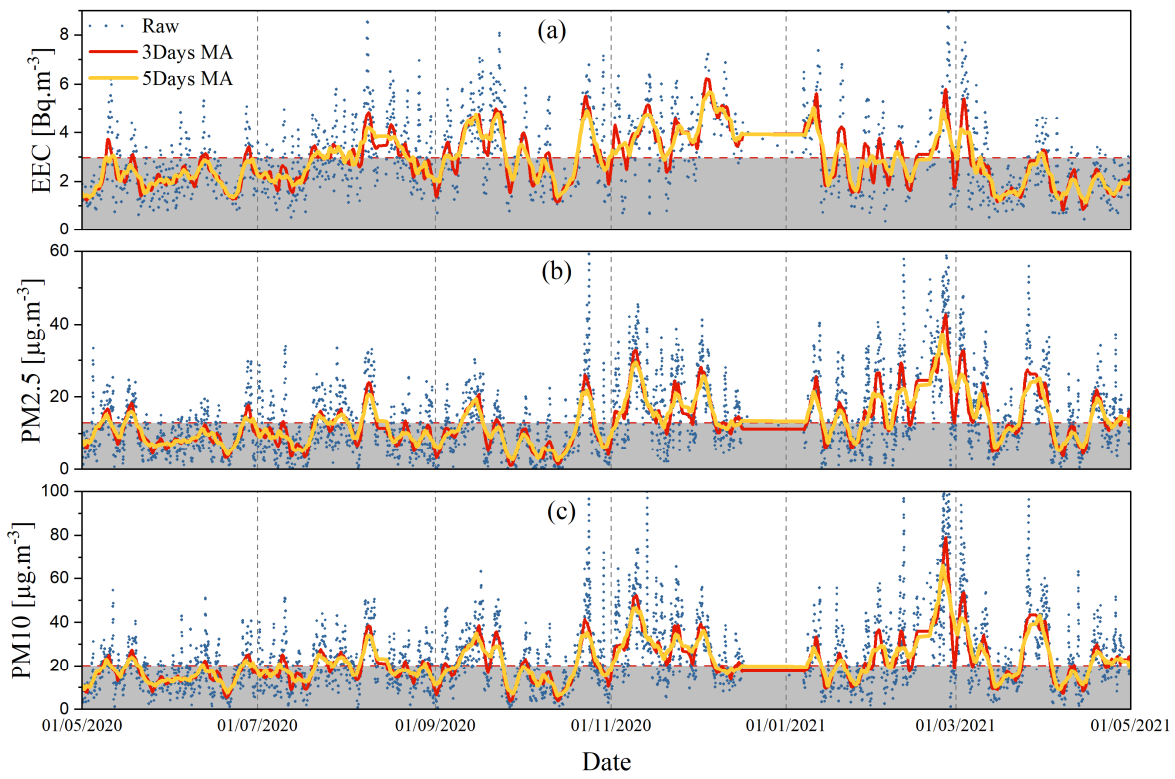


Fig. 5.5 (a) Hourly EEC (a weighted average of three short-lived radon progenies) (b) hourly PM2.5 and (c) hourly PM10 displayed for the period from May 2020 to May 2021.

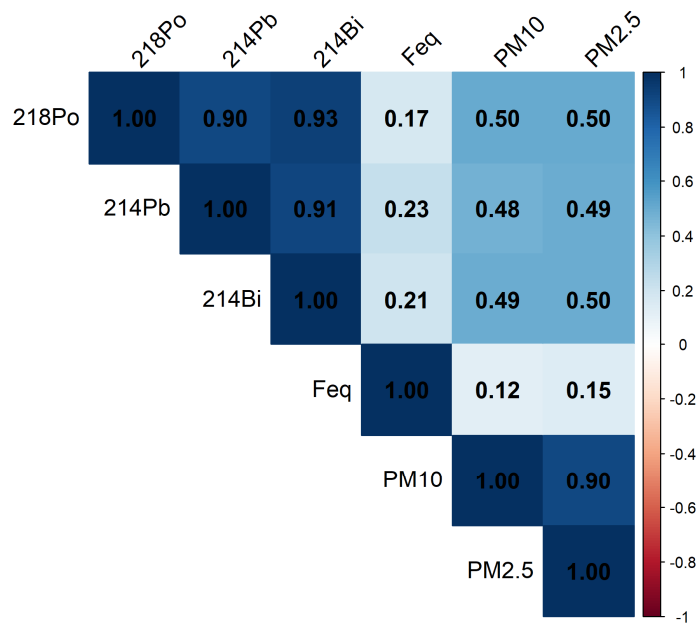


Fig. 5.6 Pearson correlation heatmap among radon progeny, Feq and PM based on the hourly data for the entire measurement period.

To assess the seasonal variation of radon progeny in relation to PM, their monthly variations are shown in **Fig. 5.7**. Both radon progeny and PM concentrations exhibit clear seasonal patterns, with the highest concentrations typically occurring during the colder months (e.g., autumn and winter) and the lowest during the warmer months. The monthly means of PM₁₀ range from $12.9 \mu\text{g.m}^{-3}$ in August 2021 to $35.4 \mu\text{g.m}^{-3}$ in February 2021, with an arithmetic mean of $20.0 \mu\text{g.m}^{-3}$. As expected, the PM_{2.5} concentration is lower than PM₁₀, ranging from $7.26 \mu\text{g.m}^{-3}$ in September 2022 to $22.8 \mu\text{g.m}^{-3}$ in February 2021, with an arithmetic mean of $13.05 \mu\text{g.m}^{-3}$. Periods with higher PM₁₀ and PM_{2.5} concentrations generally correspond to elevated radon progeny levels. This implies a likely positive correlation, where higher PM levels may facilitate attachment of radon progeny to particles, thereby increasing their measurable activity concentrations in the air.

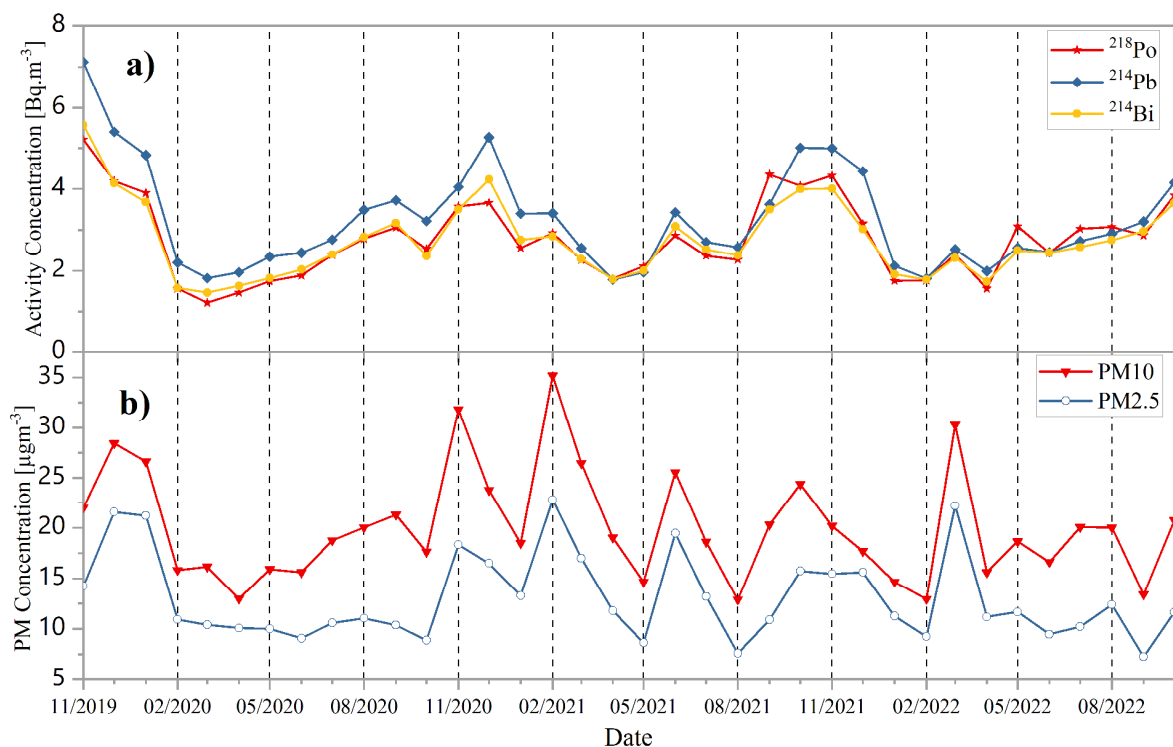


Fig. 5.7 Monthly variation of a) short-lived progeny, and b) PM₁₀ and PM_{2.5}.

To analyse the relationship between the concentrations of PM and radon progeny, in the first step, a Pearson's correlation analysis was performed (**Fig. 5.8**). Additionally, the BLH and key meteorological parameters were also included in further analysis. The short-lived radon progeny shows moderate positive correlation with PM₁₀, PM_{2.5} and RH and a strong negative correlation with BLH. The correlations between radon progeny and other meteorological variables are weak and statistically insignificant. A previous study also

reported a positive correlation between short-lived radon progeny and PM concentration (Yu et al., 2023). This positive correlation could be attributed to the fact that PM acts as a carrier of radon progeny (i.e., the radon progeny readily attaches to fine PM). Several studies have investigated the size distribution of aerosols to which the radon progeny preferentially attach. For example, Porstendörfer et al. (2000) found that the radon progeny is primarily attached to aerosols with aerodynamic diameters between 0.1 and 1 μm , while another study reported that a significant fraction of radon progeny is also attached onto aerosols with aerodynamic diameters greater than 1 μm (Kagerer et al., 2005). Taking this into account, pollution events may lead to higher concentrations of aerosol-attached radon progeny. In addition, the atmospheric behaviour of radon is similar to that of PM, further increasing the positive correlation between them.

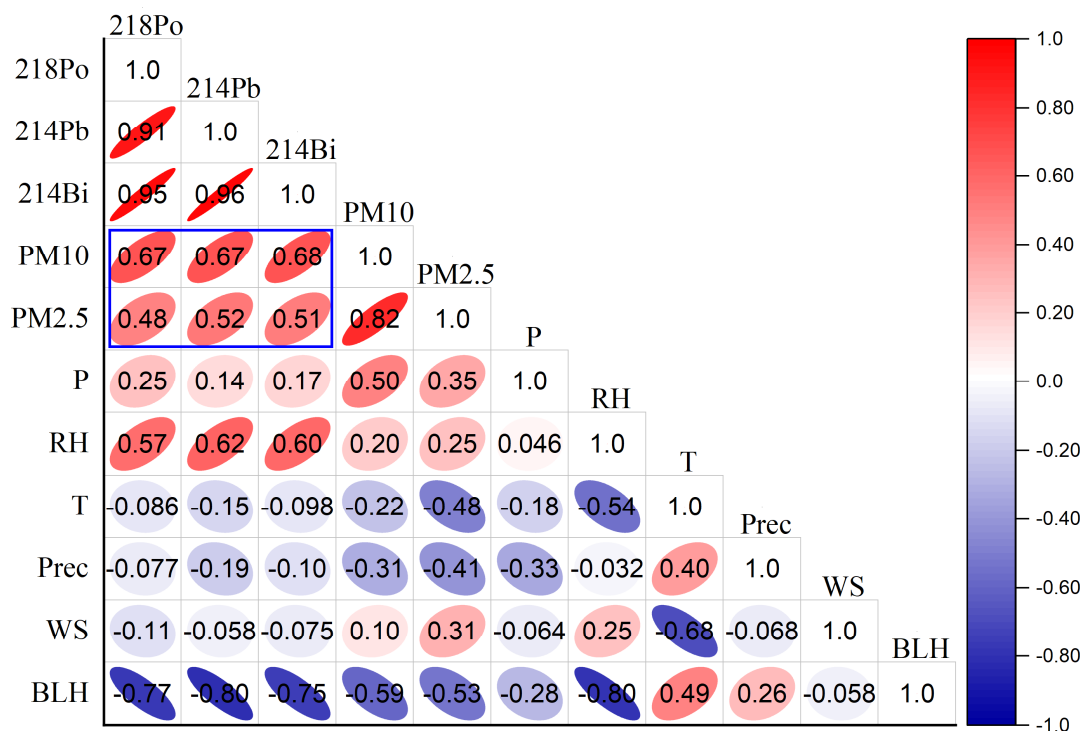


Fig. 5.8 Pearson's correlation heatmap between monthly means of radon, its progeny and particulate matter.

Furthermore, to evaluate the relationship between radon decay products and PM concentrations, canonical correlation analysis (CCA) was applied to the monthly values. CCA is a multivariate statistical method used to examine the relationship between two comprehensive sets of variables (Uurtio et al., 2017). It derives two comprehensive variables from the sets, capturing their overall correlation. The results include canonical correlation, canonical loadings, and standardized/unstandardized coefficients, revealing the relationship

between the two groups and the contribution of each variable to the respective group. In this study, radon decay products (^{218}Po , ^{214}Pb , ^{214}Bi) were considered as one group (Y group) and PM_{2.5}, PM₁₀, BLH and key meteorological factors as another group (X group).

Table 5.2 Results of the canonical correlation analysis (part 1).

Correlation	Eigenvalue	Wilks Statistic	F	Sig.
0.87	3.27	0.09	3.9	<0.001

The CCA reveals a statistically significant and strong positive association between the two groups of variables with a correlation coefficient of 0.87 ($p < 0.001$). This indicates a robust shared variance structure between the two sets of variables. As shown in **Table 5.3** and illustrated in **Fig. 5.9**, the canonical loadings and cross-loadings help to identify which variables contribute significantly to this relationship. Among the variables in group X, the strongest contributor is BLH followed by RH and PM. These variables show the highest canonical loadings in group X, highlighting their importance in shaping the shared variance with group Y. Within the Y group, the strongest contributors to the canonical variate are ^{214}Pb , followed by ^{214}Bi , ^{218}Po . This CCA suggests that meteorological and air quality conditions, especially low BLH, high RH, and high PM concentrations are strongly associated with elevated concentrations of radon progeny (^{214}Bi , ^{214}Pb , and ^{218}Po). Given the positive correlation coefficient, the negative canonical loadings of radon progeny and PM suggest a positive correlation between radon progeny and PM concentration. This relationship is consistent with findings from other studies, such as (Yu et al., 2023), which also reported positive associations between PM levels and radon progeny. One plausible explanation is that atmospheric conditions that favour higher particulate concentrations, such as lower mixing height and higher relative humidity, also promote the accumulation of radon decay products near the ground. The slightly weaker relationship observed for ^{218}Po can be attributed to its short half-life (3.1 minutes), which restricts how much it can attach to aerosol particles before it decays. Nevertheless, its contribution to the canonical structure remains significant.

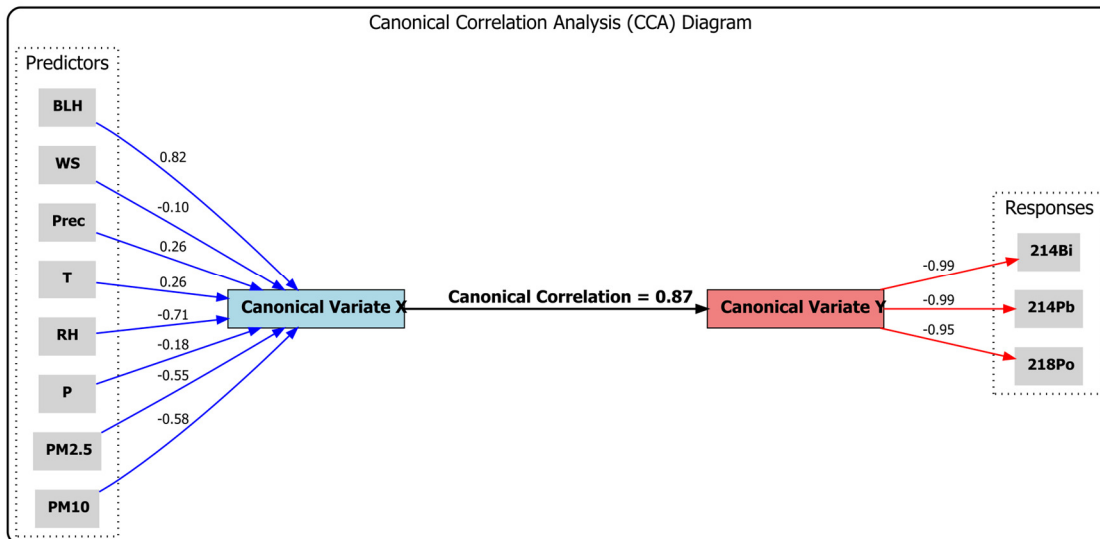


Fig. 5.9 Graphical representation of canonical correlation analysis, including canonical correlation and canonical loadings for each variable.

Table 5.3 Results of the canonical correlation analysis (part 2).

Canonical Weights	Canonical loadings	Canonical variate (X)	Canonical correlation	Canonical variate (Y)	Canonical loadings	Canonical Weights
0.27	-0.58	PM10	0.87	²¹⁸ Po	-0.95	-0.15
-0.55	-0.55	PM2.5		²¹⁴ Pb	-0.99	-0.57
0.078	-0.18	P		²¹⁴ Bi	-0.99	-0.29
-0.45	-0.71	RH				
-0.95	0.26	T				
0.28	0.26	Prec				
-0.53	-0.10	WS				
0.8	0.82	BLH				

5.5. Regression analysis of radon progeny

To further analyse the behaviour of radon progeny in relation to PM concentration and other influencing factors, regression modelling was employed. The regression analysis includes both conventional and ML techniques such as MLR, GAM, RF, and XGBoost, each discussed in detail in **Chapter 2**.

Table 5.4 Multicollinearity between radon progeny predictors.

Predictor	PM10	P	RH	T	Prec	WS	BLH
VIF	1.10	1.19	1.75	1.54	1.02	1.49	1.85

5.6. Multiple linear regression applied to radon progeny

MLR, as discussed in **Section 2.8.5**, was applied to model the relationship between radon progeny, PM, and other environmental factors. For the sake of brevity, only EEC was considered, rather than discussing all three short-lived radon progenies.

Before applying the MLR analysis, multicollinearity among the independent variables was assessed using the variance inflation factor (VIF) discussed in **Section 4.5**. The VIF results showed that most of the variables exhibited a relatively low degree of multicollinearity. Values ranged from 1.1 to 1.85, indicating an acceptable level of multicollinearity between the predictors (**Table 5.4**).

Table 5.5 Summary of MLR for radon progeny.

Variables	Standardized Coefficients	t-value	p-value
PM10	0.515	89.8	0.00
P	-0.082	-13.7	0.00
RH	0.369	51.13	0.00
T	0.145	21.3	0.00
Prec	-0.052	-9.4	0.00
WS	0.077	11.6	0.00
BLH	-0.186	-25	0.00

The summary of MLR for radon progeny is shown in **Table 5.5**. The MLR model for EEC demonstrated statistical significance ($p\text{-value} < 0.05$) with a correlation coefficient of $R^2 = 0.43$, indicating that 43% of the variability in EEC was explained by the predictors. The MLR analysis revealed significant relationships between various environmental variables and EEC, reflecting the physical processes that govern radon progeny behaviour in the atmosphere. Among the predictors, PM10 exhibited the strongest positive association (standardized beta coefficient = 0.51, $p < 0.001$), indicating that higher levels of PM are linked to increased radon concentrations. It is likely due to reduced atmospheric dispersion and the role of particles in carrying radon progeny. RH also showed a substantial positive effect (0.37, $p < 0.001$), while BLH had a notable negative association (-0.186, $p < 0.001$), suggesting that enhanced atmospheric mixing reduces radon progeny accumulation near the surface. The positive effect of RH on radon progeny concentrations may be attributed to the hygroscopic growth of aerosol particles, which enhances the attachment of radon progeny to these particles (Khan et al., 1988). High RH causes hygroscopic aerosol particles to grow, which increases their surface area and enhances the attachment of radon progeny leading to elevated concentrations of attached progeny in the atmosphere. Temperature and wind speed

had moderate and weak positive effects (0.145 and 0.077, respectively), potentially enhancing radon exhalation from soil, whereas atmospheric pressure and precipitation showed weak negative relationships (-0.082 and -0.052), likely due to suppressed soil gas flow and washout mechanisms. All variables were statistically significant ($p < 0.001$), indicating that both meteorological and pollution-related factors play a critical role in influencing radon levels. However, the MLR prediction is not very strong which could be related to the existence of non-linearity between predictors and radon progeny and other complex interactions not captured by the model. These limitations suggest that the MLR approach may not fully account for the intricate relationships between environmental variables and radon EEC, potentially leading to suboptimal predictive accuracy. To better capture the non-linear relationships between radon EEC and the potential predictors, a GAM analysis was conducted discussed below.

Table 5.6 Multicollinearity among the EEC predictors.

Concurvity	s(PM10)	s(P)	s(RH)	s(T)	s(Prec)	s(WS)	s(BLH)
worst	0.23	0.28	0.52	0.50	0.09	0.47	0.55
observed	0.23	0.20	0.49	0.42	0.05	0.38	0.48
estimate	0.16	0.15	0.45	0.40	0.03	0.37	0.41

5.7. Generalised additive model applied to radon progeny

Following the MLR analysis, a penalized GAM was applied to further investigate the modelling of EEC and its potential predictors. A penalized GAM incorporates a penalty term to control the smoothness of the fitted functions, helping to prevent overfitting. This is achieved by adding a "wiggliness" penalty to the model, which allows for more flexible fitting while maintaining generalization to new data. Prior to the GAM analysis, the concurvity diagnostics were performed to evaluate potential dependencies among the smooth terms in the GAM (**Table. 5.6**). The analysis revealed that certain predictors, notably s(RH), s(T), s(WS), s(BLH), exhibited moderate levels of concurvity, with worst-case values ranging from 0.47 to 0.55. This indicates that these smooth terms share some nonlinear predictive structure, which may affect the precision and interpretability of their individual effects. In contrast, variables such as s(PM10), s(P), and s(Prec) demonstrated low concurvity values, suggesting minimal redundancy and stable estimation. As there was no strong multicollinearity, all these predictors were included in the GAM analysis.

The summary of the GAM analysis is shown in **Table 5.7**. The model explained 49% of the deviance with an adjusted R^2 of 0.49, indicating a moderate but meaningful fit. The expected

value of the response variable, also known as the intercept, was statistically significant (estimate = 2.9, $p < 0.001$), and all smooth terms were also highly significant ($p < 0.001$), suggesting each variable substantially contributes to explaining the variations in EEC. Among the predictors, PM10 (edf = 7.5, $F = 1052$), RH (edf = 7.0, $F = 415$), and BLH (edf = 7.5, $F = 85$) showed particularly strong nonlinear effects. These findings underscore the complex interplay between atmospheric conditions and radon EEC, emphasizing the importance of considering PM and other environmental factors in radon exposure assessments.

Table 5.7 Summary of the GAM analysis results for radon progeny.

Family	Link function	Formula		adj. R ²	Deviance explained
Gaussian	Identity	EEC ~ s(PM10)+s(P)+s(RH)+s(T)+ s(Prec)+s(WS)+s(BLH)		0.49	49%
Parametric coefficients					
		Estimate	Std. Error	t value	Pr(> t)
(Intercept)		2.9	0.009	337.5	<0.001
Approximate significance of smooth terms					
Predictors		edf	Ref.df	F	p-value
s(PM10)		7.5	8.3	1051.8	0.001
s(P)		8.6	8.9	34.1	0.001
s(RH)		7.0	8.0	414.6	0.001
s(T)		7.8	8.6	49.8	0.001
s(Prec)		2.8	3.4	27.3	0.001
s(WS)		6.3	7.2	42.7	0.001
s(BLH)		7.5	8.4	85.2	0.001

The estimated smooth terms from the GAM provide insights into the nonlinear effects of meteorological and PM concentration on radon EEC (**Fig. 5.10**). The effect of PM10 shows a strong positive association with EEC, especially at concentrations of up to $\sim 60 \mu\text{g.m}^{-3}$, after which the relationship plateaus, indicating saturation. Relative humidity and temperature also exhibit clear positive trends, suggesting higher EEC values under more humid and warmer conditions. However, the effect of temperature is not as strong as that of RH. In contrast, precipitation shows a consistently negative association with EEC, likely reflecting the washout effect of rain on radon progeny and reduced radon exhalation rate. Atmospheric pressure has a slightly decreasing trend, while WS shows a moderate positive effect up to around 4 m.s^{-1} , after which there is a reversal of the trend. Lower values of BLH

appear to be negatively associated with EEC; however, this trend disappears beyond approximately 700 m. These plots highlight the complex, nonlinear interactions between environmental factors and radon progeny dynamics, reinforcing the necessity of flexible modelling approaches such as GAMs for accurate assessments of radon progeny exposure.

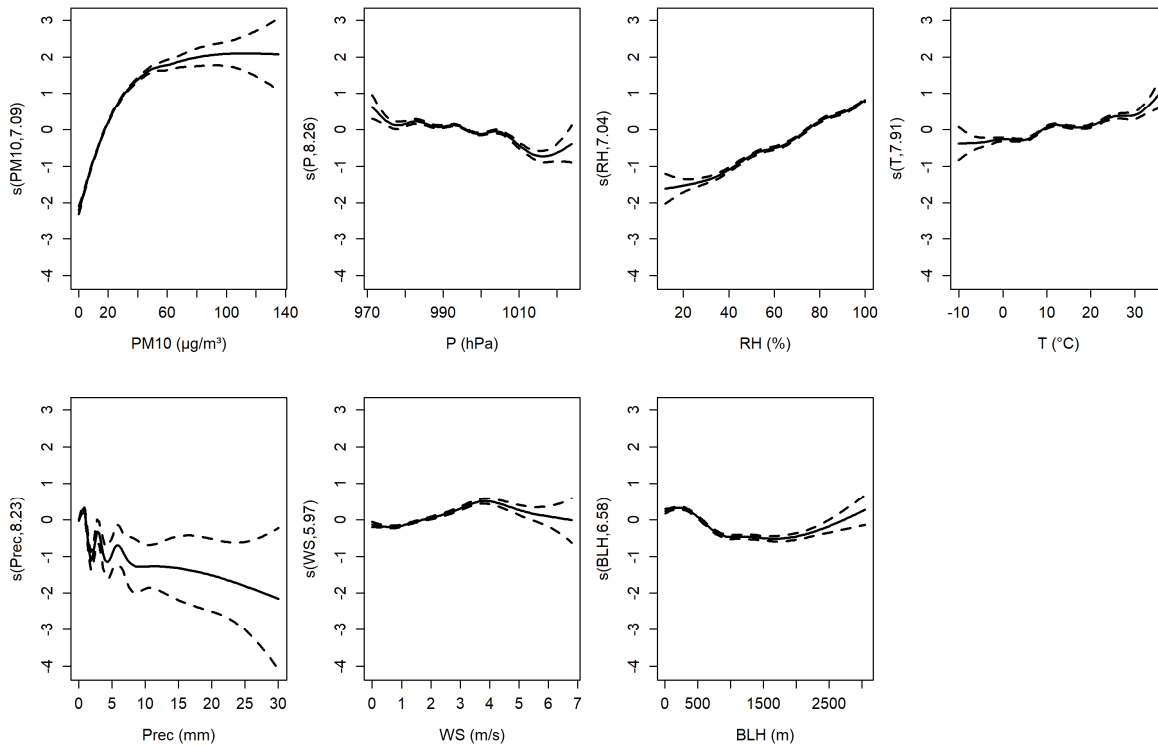


Fig. 5.10 Estimated smooth functions of the predictors in the GAM for radon EEC. The solid lines represent the estimated smooth effects while the dashed lines indicate the 95% confidence intervals.

5.8. Machine learning methods in radon progeny analysis

Conventional regression models such as MLR and GAM revealed important relationships and nonlinear patterns between radon progeny and environmental factors. However, their predictive accuracy remained limited, likely due to inherent model constraints in capturing complex interactions and higher-order dependencies. To gain deeper insights into the relationship between radon progeny, PM, and other environmental parameters, two widely used machine learning regression models (XGBoost and RF) were employed.

The XGBoost model was employed after cleaning the dataset and splitting it into training and testing subsets, at a ratio of 80:20. The XGBoost model was trained with 1000 boosting rounds. Hyperparameters, including a maximum tree depth of 6 and a learning rate of 0.1, were selected to optimize performance. Model evaluation on the testing subset showed

strong predictive accuracy, with performance assessed using metrics such as MSE, RMSE, MAE and R^2 . Feature importance was evaluated using its respective internal importance metrics.

In parallel, the RF regression model was employed, using the same input variables and normalized data. It was trained with 500 trees and a mtry value of 3. In this case, too, the model's performance was evaluated using MSE, RMSE, MAE and R^2 . Feature importance was determined based on the percentage increase in mean squared error (%IncMSE). To enhance interpretability, SHAP (SHapley Additive exPlanations) values were calculated to assess the impact of each feature on response variable. Additionally, Partial Dependence Plots (PDPs) were generated to illustrate how variations in each input variable influence the predicted EEC values.

Both the XGBoost and RF models performed well in predicting the radon EEC based on PM10, BLH, and key meteorological factors, with adjusted R^2 values of 0.70 and 0.74, respectively (prediction based on test dataset). The scatter plots for these models and conventional regression models (MLR and GAM) along with their corresponding feature importance, are shown in **Figs. 5.11** and **5.12**. Among the regression models, the RF exhibited the strongest predictive power (adj. $R^2 = 0.74$), followed by XGBoost, GAM; the weakest predictive power was observed in MLR. The feature importance rankings from all four models showed that PM10 and RH are the two strongest predictors, with precipitation being the weakest. Other features, such as BLH, T, P, and WS, ranked in the middle, with some differences depending on the model.

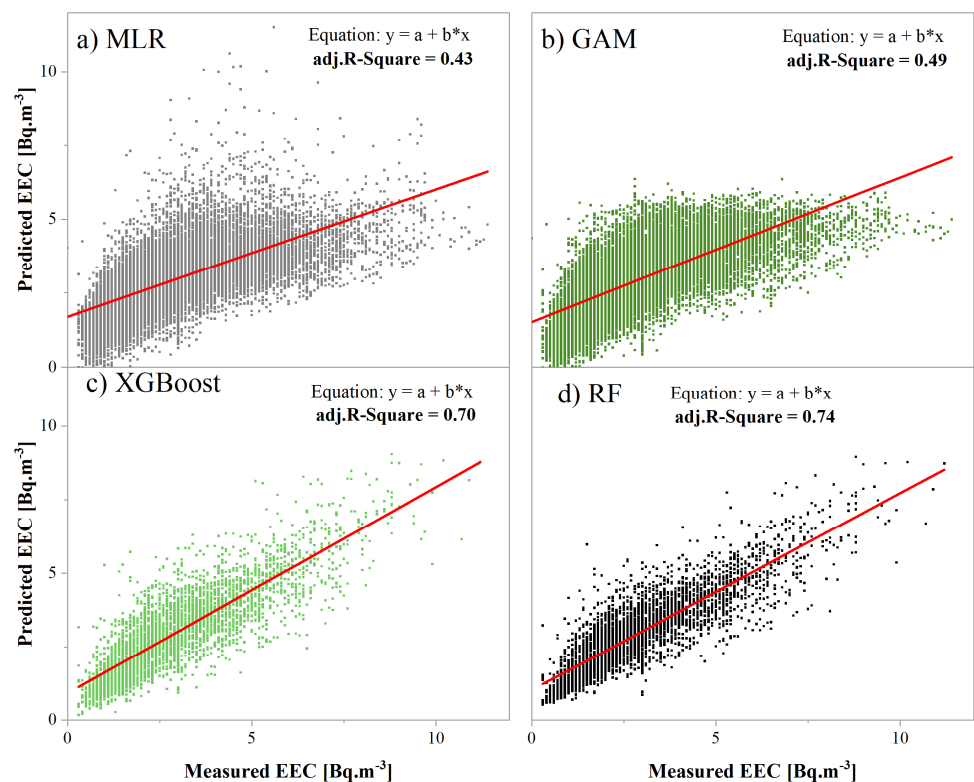


Fig. 5.11 Scatter plots of the regression models showing the predicted vs. measured EEC.

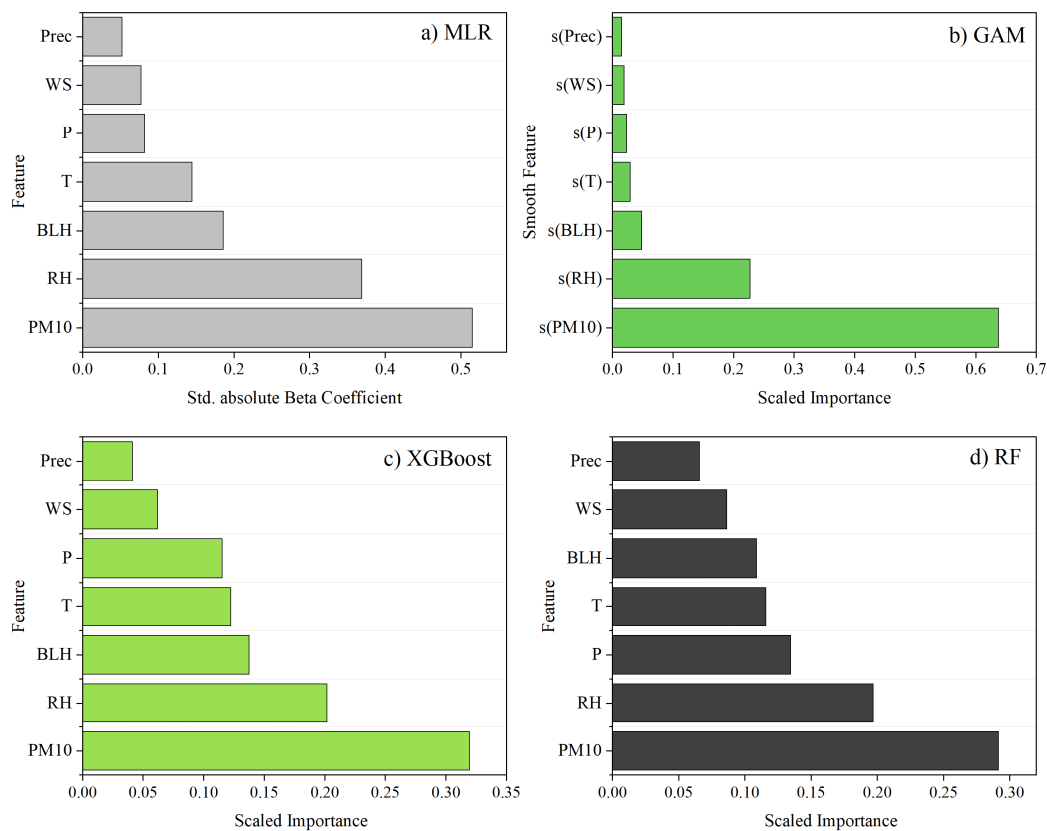


Fig. 5.12 Feature importance rankings from four regression models used in radon progeny analysis: a) MLR, b) GAM, c) XGBoost, and d) RF. PM10 and RH consistently appear as the most influential features across models.

5.9. SHAP analysis of the random forest model

The SHAP analysis, described in **Section 4.9**, was employed to interpret the RF model developed for predicting EEC. **Fig. 5.13** presents the SHAP summary visualization composed of two parts: a bar chart showing the mean absolute SHAP values, and a so-called ‘beeswarm’ plot showing the distribution of SHAP values for each feature. These combined plots provide a comprehensive view of both the direction and magnitude of each feature’s influence on model predictions. The features are ordered by their average contribution to the model output, with PM10 emerging as the most impactful variable, having the highest mean SHAP value and the widest spread of SHAP effects. This implies that PM10 has a consistently strong influence on EEC predictions, with both positive and negative effects depending on its value. RH and BLH follow in importance, while Prec, T, P and WS have moderate influences. The beeswarm plot uses a colour gradient to represent the normalized feature values, providing additional insight into how different magnitudes of a feature impact the prediction. Overall, the SHAP summary enhances model interpretability by quantifying the contribution of each feature and illustrating how they interact with the prediction process.

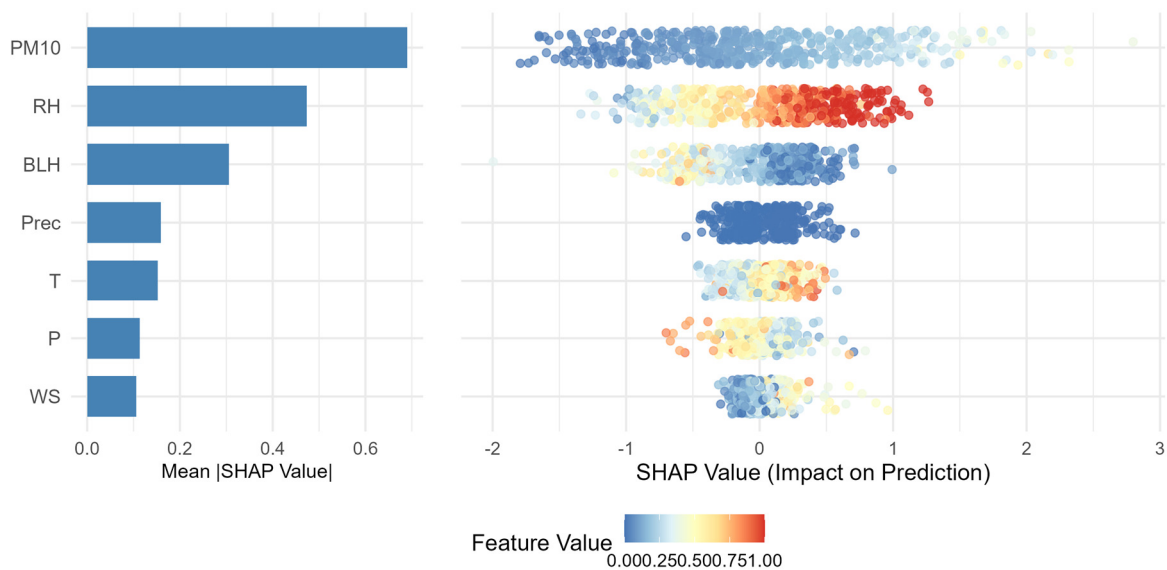


Fig. 5.13 SHAP summary plot, illustrating the contribution of each feature to the predicted EEC for the Random Forest model. The right-hand panel shows a beeswarm plot displaying the distribution and direction of SHAP values for each feature. The left-hand panel shows a bar chart of the mean absolute SHAP values, ranking features by their overall importance.

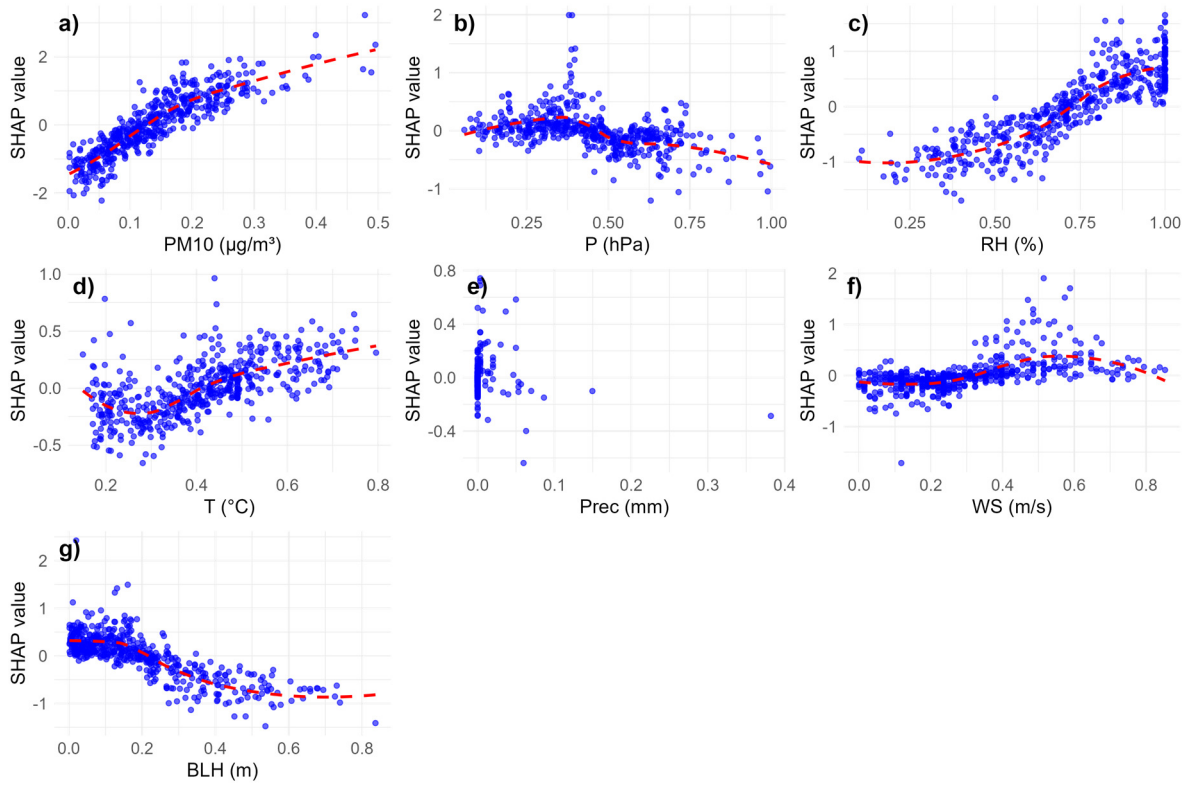


Fig. 5.14 SHAP dependence plots for the RF model used for the prediction of EEC.

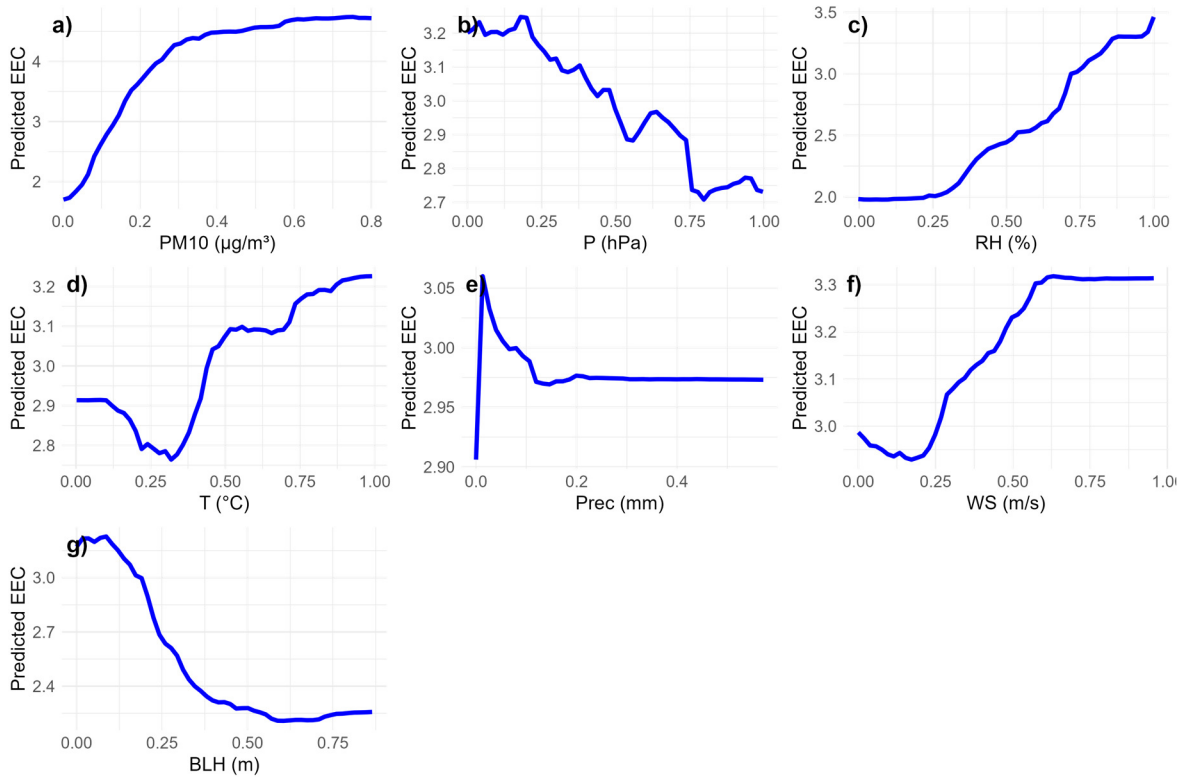


Fig. 5.15 Partial dependence plots for the RF model used in the prediction of EEC. The predictor values on the x-axis have been normalised.

5.10. SHAP dependence plots for the random forest

Fig. 5.14 presents SHAP dependence plots for the RF used to predict EEC, illustrating the impact of various input features on the model's output. Each subplot visualizes how the value of a specific feature influences the predicted EEC, with blue dots representing individual data points and red dashed lines indicating average trends. The plots reveal nonlinear relationships between EEC and almost all features in the RF model. For example, certain features (PM, RH, T) showed a clear positive correlation with EEC, where higher feature values lead to increased predicted radon levels, while others demonstrate a negative or threshold-like effect, suggesting diminishing influence beyond a specific range.

5.11. Partial dependence plot for the random forest

Fig. 5.15 presents the partial dependence plots for the RF model used in predicting EEC, illustrating the marginal effect of each individual feature on the model's output. A similar result was observed in the GAM component smooth function plots and SHAP dependence plot (**Figs. 5.10** and **5.14**), where nearly all predictors exhibited a non-linear relationship with radon EEC. PM10 shows a strong positive effect, with predicted EEC levels increasing sharply at lower concentrations and plateauing at higher levels, indicating a saturation effect. Atmospheric pressure has an approximately linear and negative influence on EEC, where higher pressure values correspond to lower EEC predictions, possibly due to reduced radon exhalation under high-pressure conditions. RH demonstrates a consistent positive effect on EEC, suggesting that more humid conditions favour radon progeny accumulation. WS shows a non-linear trend, with a slight decrease in predicted EEC at lower speeds, followed by an increase at moderate levels, implying complex dispersion dynamics. BLH exhibits a sharp negative effect at low values, indicating that shallow boundary layers trap radon progeny near the surface, whereas higher BLH values promote dispersion. The effect of temperature on the predicted EEC is more complex. At lower temperatures, the relationship is negative, i.e. EEC decreases as temperature increases. However, once a certain threshold is reached, this effect reverses: increased temperatures are associated with higher predicted EEC, potentially due to enhanced radon exhalation from the soil. Finally, precipitation has a negative effect on EEC: increased rainfall is linked to lower EEC levels, likely due to washout or suppression of the radon exhalation rate from the soil. These patterns collectively highlight the non-linear and variable influence of each predictor on radon behaviour.

Chapter 6

6. Radon as a tracer in atmospheric research

In this chapter, the use of radon as tracer will be discussed. Specifically, the measurement outdoor radon will be used to determine the MLH based on the so-called box model. The result will be compared to a different dataset, and the strength and limitation of this model will be discussed. In the second part of this chapter, residence time of aerosols in the atmosphere will be estimated based on the concentration of outdoor radon and its progeny.

6.1. MLH based on radon - results

The MLH was determined using the box model introduced in **Section 2.7.2**. This model takes RAC and radon flux as inputs to calculate the MLH. The accuracy of the model heavily depends on both radon concentration and radon flux measurements. A major challenge in applying this model is the limited availability of continuous and precise radon flux data. Several studies have assumed a constant radon flux throughout the year (e.g., Gregorič et al., 2020; Salzano et al., 2016; Sesana et al., 2003; Vecchi et al., 2019). However, radon flux exhibits distinct seasonal variations that depend mainly on the soil moisture, as reviewed in (Čeliković et al., 2022). Neglecting these variations and using a fixed value throughout the year can result in significant overestimation or underestimation of MLH. To account for the seasonal variability of radon flux, this study utilized the daily radon flux values from the European radon flux map (Karstens et al., 2023). The European radon flux map was created using data on soil uranium content, soil properties and two different soil moisture reanalyses (ERA5 and GLDAS-Noah v2.1). The data for our locality was downloaded from the Integrated Carbon Observation System (ICOS) portal³.

³ traceRadon daily radon flux map for Europe 2017 (based on ERA5-Land soil moisture) | ICOS

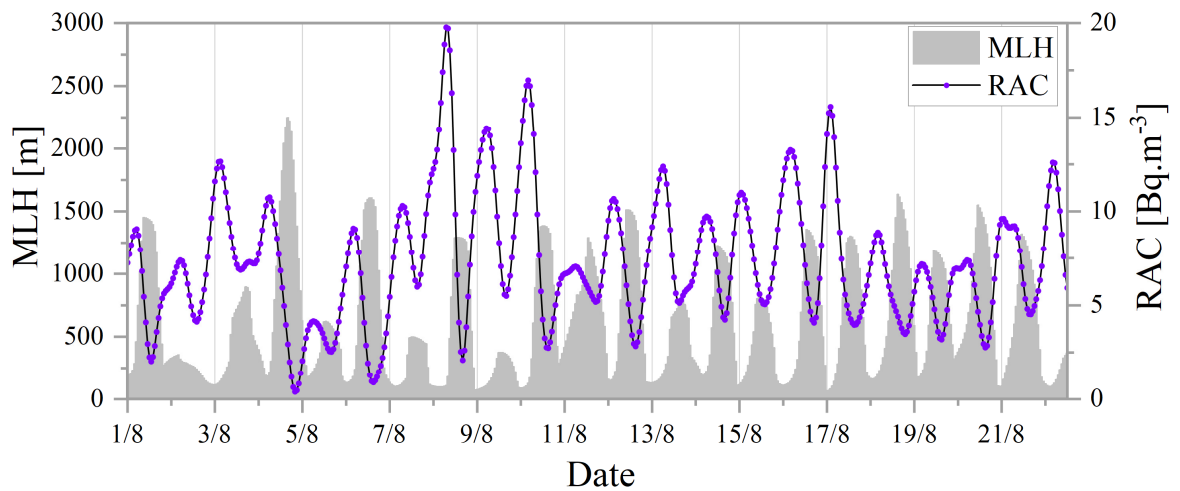


Fig. 6.1 Hourly evolution of MLH and RAC in August 2020.

The temporal variation of MLH in relation to RAC for August 2020 is illustrated in **Fig. 6.1**. On a day-to-day basis, a low RAC corresponds to a high MLH, and vice versa. When the MLH is shallow, radon tends to accumulate near the surface, resulting in higher concentrations. In contrast, when the MLH is high, radon disperses rapidly, causing its concentration near the ground to decrease. **Fig. 6.2** shows the composite diurnal cycles for each month of the year averaged over the entire study period (2018-2023). Generally, the MLH is shallow and stable during the nighttime, starts growing after the sunrise, reaches its peak in the mid-afternoon, and decreases as the solar radiation intensity diminishes. This is due to the fact that after the sunrise the solar radiation warms the Earth's surface, which in turn heats the air parcels near the surface and generates turbulent mixing. This turbulence expands the boundary layer vertically, forming a convective boundary layer (CBL) whose depth peaks in the mid-afternoon. As solar intensity diminishes towards evening, the CBL begins to collapse. After sunset, the surface cools rapidly via thermal radiation, creating a temperature inversion (colder air near the surface, warmer air above). This inversion stabilises the atmosphere, suppresses turbulence and reduces the BLH to a shallow stable boundary layer (SBL), typically 200–300 m deep.

A distinct diurnal cycle of MLH can be observed for each month of the year. While the nighttime MLH remains relatively constant throughout the year, ranging from approximately 200 to 400 m, the daytime MLH varies in both range and duration, depending on solar exposure. For example, the mean daytime MLH peaks at around 500 m during the cold months of December and January, due to lower temperatures and limited mixing. From February onwards, however, the peaks begin to increase, reaching approximately 1500 m in the summer months, before gradually decreasing towards autumn.

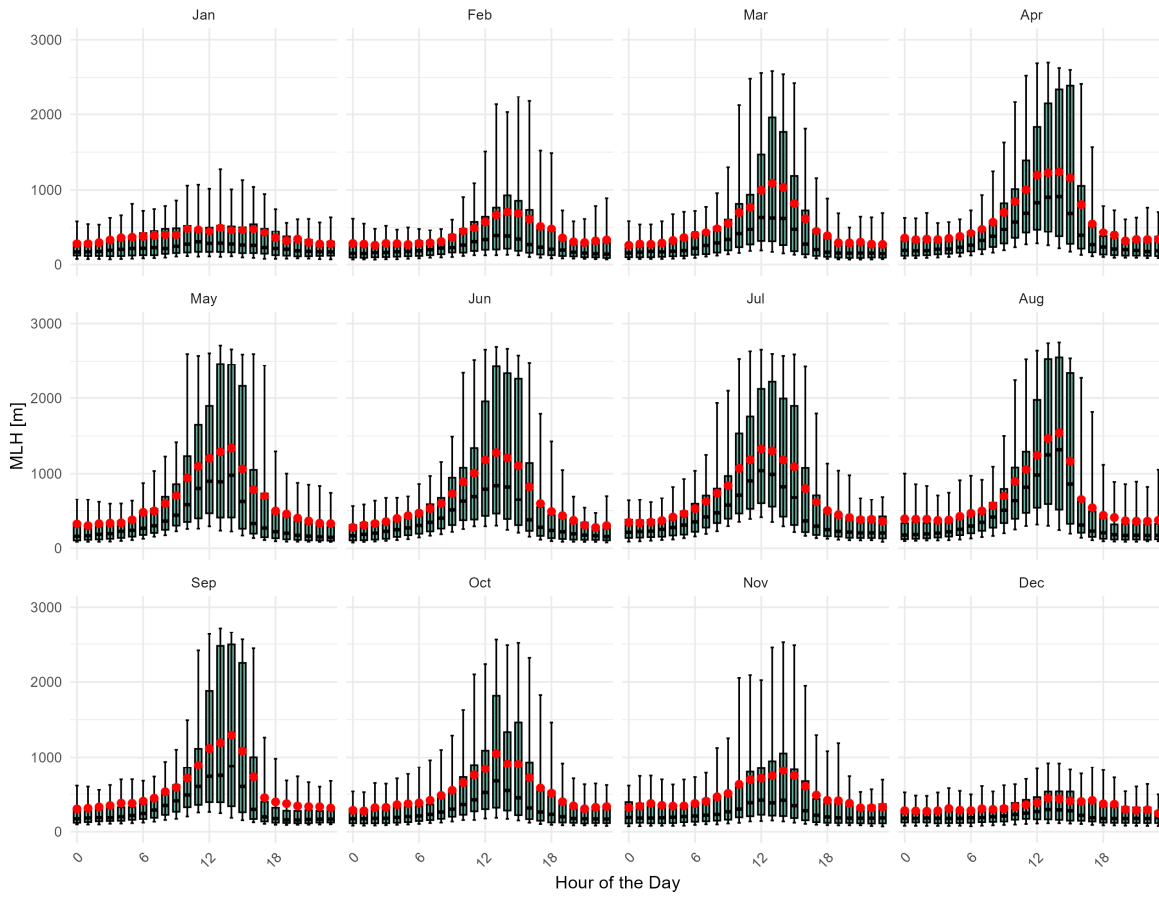


Fig. 6.2 Composite diurnal cycles of MLH for each month of the year averaged over the entire study period (2018-2023). The box shows 25th to 75th percentiles, the whisker line extended from 10th to 90th percentiles. The black dots show the median while the red dots show the mean values.

The seasonal variation of MLH is shown in **Fig. 6.3** in the form of monthly boxplots. A clear seasonal variation can be observed i.e., MLH is usually higher in spring and summer and lower in autumn and winter. The mean MLH is lowest during December-February (around 400 m). It then begins to rise steadily in early spring, reaching its peak of approximately 700 m in July. Over the following months, the mean MLH declines steadily until it reaches its lowest point again in December. The seasonal variation of MLH is mainly governed by the solar radiation, surface heat flux and temperature lapse rate (Lee 1986; Vecchi et al. 2018; Kim 2022).

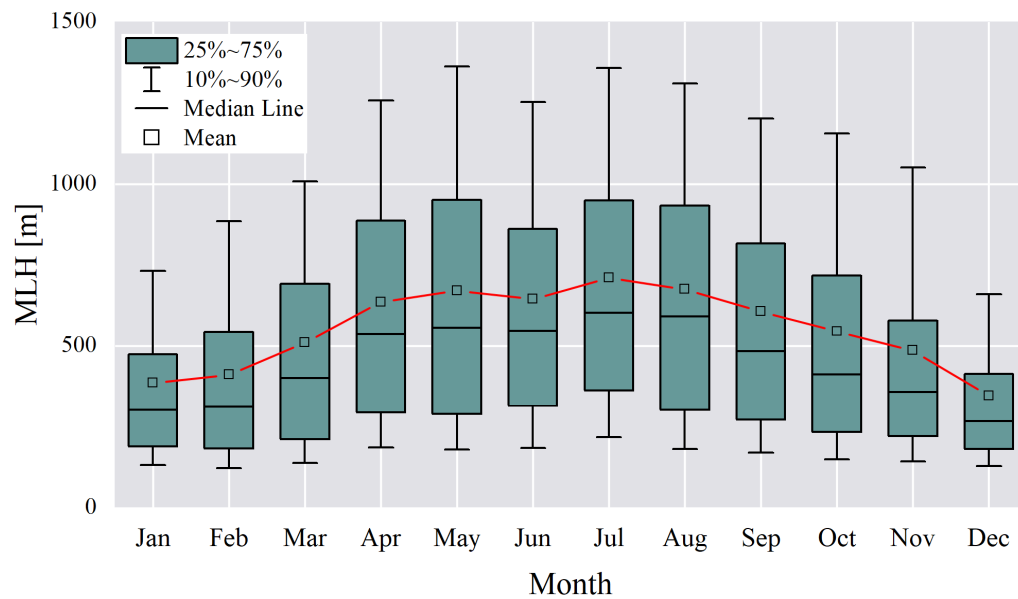


Fig. 6.3 Composite monthly boxplot of MLH created based on data covering the entire study period.

6.2. Comparison of MLH with BLH-ERA5

In this section, the MLH based on radon was compared with the MLH/BLH based on the ERA5 reanalysis dataset (BLH-ERA5). It should be noted that the terms mixing layer height and boundary layer height are used interchangeably but refer to the same atmospheric layer. The ERA5 reanalysis dataset, produced by the European Centre for Medium-Range Weather Forecasts (ECMWF), provides a high-resolution global atmospheric reanalysis from 1940 to the present (Hersbach et al., 2020). One of its valuable products is BLH calculated based on the Bulk Richardson number, with a spatial resolution of $0.25^\circ \times 0.25^\circ$ and a temporal resolution of one hour. The BLH-ERA5 dataset has been used extensively worldwide in atmospheric research in recent years. Several studies conducted the intercomparison between BLH-ERA5 and experimental BLH data obtained using various instruments, such as ground-level LIDARs, sodars, ceilometers and radiosondes. Despite some differences, good agreement is generally reported between the BLH-ERA5 and experimental results e.g., (Allabakash et al., 2020; Madonna et al., 2021; Guo et al., 2021; Sinclair et al., 2022; Li et al., 2023).

The hourly BLH-ERA5 data were extracted for our locality (latitude: $48^\circ 9'N$, longitude: $17^\circ 4'E$), covering the years 2018-2023, which coincides with the period for which radon-based MLH was calculated. The composite diurnal cycles of both approaches are shown in

Fig. 6.4. During the nighttime (20:00–6:00), when the MLH is stable, the mixing height obtained by both methods showed strong agreement, ranging between 200 and 400 m. However, during the daytime, some differences can be observed. For example, the summer trends align closely and reach highest values of ~1500 m around 14:00. However, compared to MLH, the daily peaks of BLH-ERA5 show higher values: up to 100 m in spring and 200 m in winter. Conversely, in autumn, the MLH-Rn peak is up to ~150 m higher than that of BLH-ERA5. Another difference is that the BLH-ERA5 model exhibits more rapid evolution after sunrise compared to the MLH-Rn. This discrepancy aligns with the findings of Vecchi et al. (2019), who observed a similar pattern when comparing the diurnal cycle of MLH based on radon with a turbulence model. This difference may be attributed to the slower measurement rate of radon due to larger sampling interval, compared with the faster measurement rate of variables used in the BLH-ERA5 calculation. Such differences in the evolution of mixing layer height are expected, as the two methods rely on different measured variables with different dynamics, i.e., radon in the box model in one approach versus potential temperature and two wind components in the other approach. A linear regression analysis was conducted between the two datasets to further understand their relationship (**Fig. 6.5**). There was a strong correlation between the two variables (Pearson's $r = 0.90$). Further, a non-parametric Mann-Whitney U test (Mann & Whitney, 1947) was performed on the composite diurnal cycles of both approaches, considering each season independently. At a 95% confidence level, no significant differences were found between the two datasets, except for the winter season.

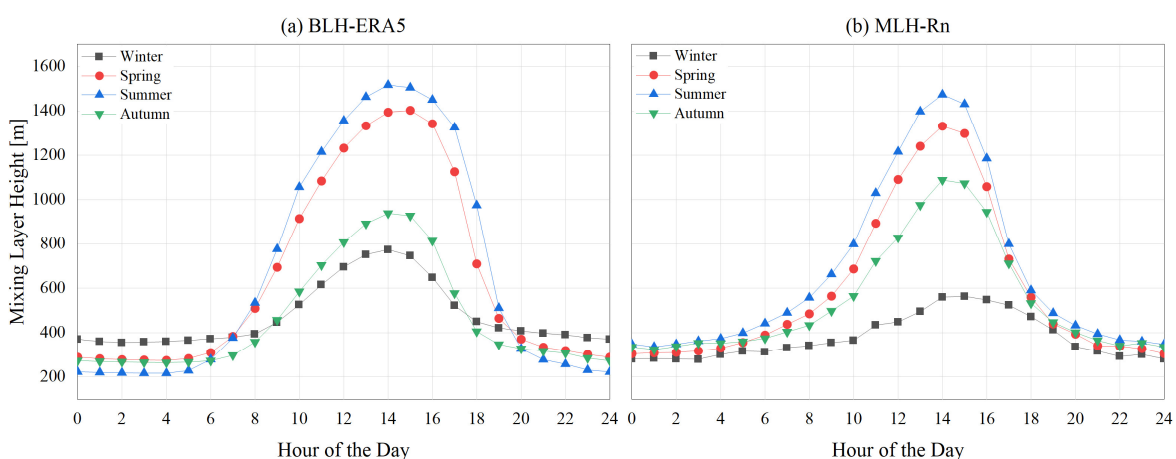


Fig. 6.4 Composite diurnal cycles of a) BLH-ERA5 and b) MLH-Rn for each season of the year. The data were aggregated across the entire study period.

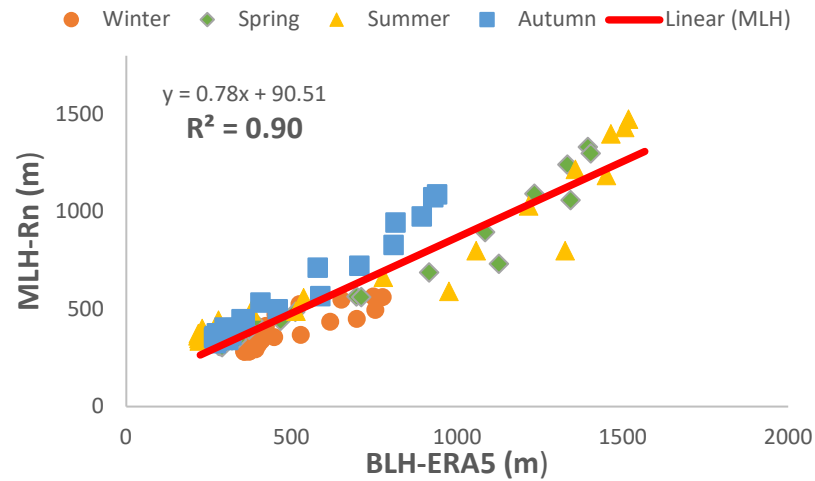


Fig. 6.5 Correlation between the composite diurnal cycles of MLH-Rn and BLH-ERA5 for each season of the year.

The mean values for individual months of 2018-2023, as well as composite monthly means of MLH and BLH-ERA5, are shown in **Fig. 6.6** and **Fig. 6.7**, respectively, along with their respective scatter plots. Despite some differences, the overall seasonal trends are in good agreement. From December to July, the BLH-ERA5 values are higher than the MLH values, but lower during the rest of the year. The higher values of BLH-ERA5 in cold or very stable conditions has also been reported by (Sinclair et al., 2022). However, the non-parametric Mann-Whitney U test at a 95% confidence level revealed no significant difference when applied to the monthly values. Linear regression analysis of the monthly mean and composite monthly mean values for these two approaches shows a positive correlation (Pearson's $r > 0.75$), as illustrated in the right sides of **Fig. 6.6** and **Fig. 6.7**, respectively.

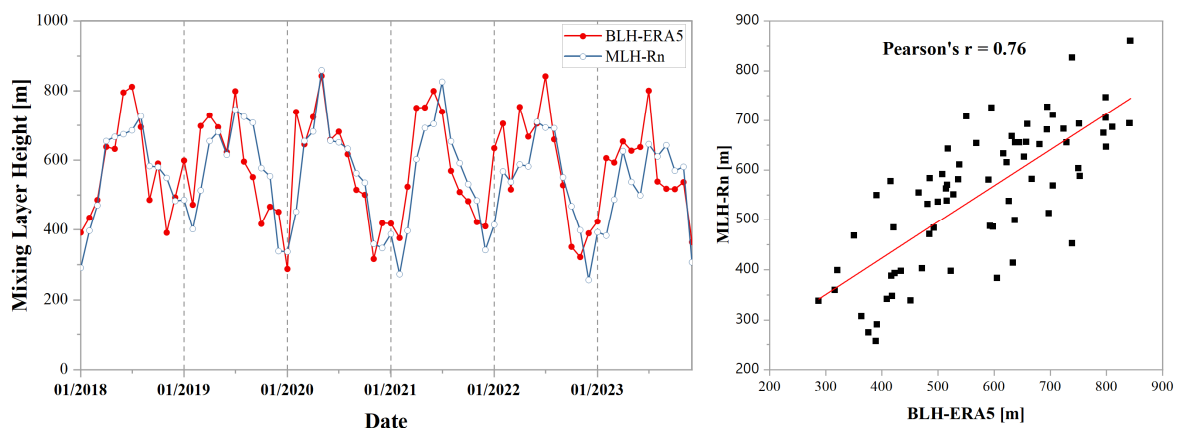


Fig. 6.6 Monthly variations of MLH-Rn compared to BLH-ERA5 for the entire study period (2018-2023).

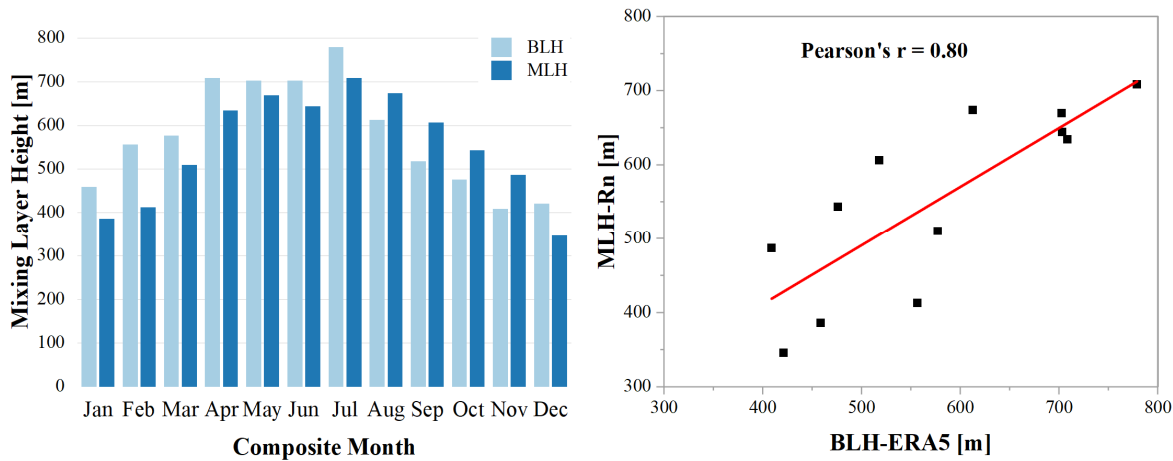


Fig. 6.7 Composite monthly values of MLH-Rn and BLH-ERA5 for the entire study period.

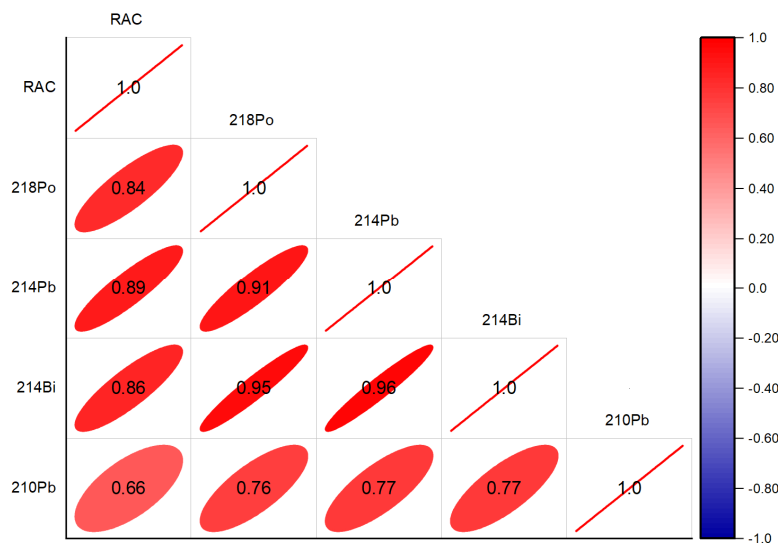


Fig. 6.8 Pearson correlation heatmap between monthly concentrations of radon and its progeny.

6.3. Aerosols residence time - results

The mean residence time of aerosols (T_R) was estimated using the weekly data on radon and radon progeny ratios, in accordance with the methodology outlined in **Section 2.7.3**. A strong correlation (Pearson's $r > 0.7$; **Fig. 6.8**) was observed between the monthly means of ^{210}Pb and short-lived radon progeny, indicating the presence of steady-state conditions necessary for applying the aerosols residence time model (Kim et al. 2000). Descriptive statistics for T_R are summarized in **Table 6.1**. Weekly T_R values range from 0.35 to 6.73 days. **Fig. 6.9** displays the frequency distributions of the residence times calculated using different radon/radon progeny ratios, along with corresponding boxplots. All these

distributions can be approximated by a normal distribution. The boxplots showed consistent T_R estimates across all three ratios, with a slightly higher interquartile range and median for the $^{210}\text{Pb}/^{214}\text{Bi}$ ratio. This variation is likely due to the relatively low variability of ^{214}Bi compared to other short-lived radon decay products.

Table 6.1 Descriptive statistics of T_R [days] based on three ratios of radon and its progeny.

Ratio	No.	Mean	Median	Std. Deviation	Minimum	Maximum
$^{210}\text{Pb}/^{214}\text{Pb}$	129	3	2.86	1.16	0.36	6.73
$^{210}\text{Pb}/^{214}\text{Bi}$	129	3.41	3.17	1.23	0.37	6.7
$^{210}\text{Pb}/^{222}\text{Rn}$	126	3	2.87	1.15	0.35	6.68

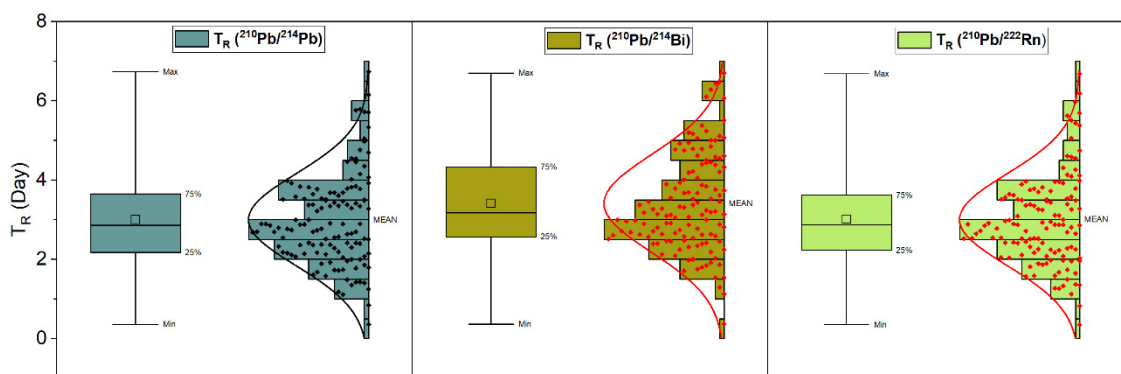


Fig. 6.9 Frequency distribution of aerosol residence times and corresponding boxplots calculated from weekly data using three different radon/radon progeny ratios.

Table 6.2 presents a comparison of our results with those found in the literature. In general, the studies have demonstrated that the T_R ranges from a few hours to several weeks. In most cases, our results are in good agreement with those reported in the literature i.e., methods based on the ratios of $^{210}\text{Pb}/^{214}\text{Pb}$, $^{210}\text{Pb}/^{214}\text{Bi}$ and $^{210}\text{Pb}/^{222}\text{Rn}$ yield a T_R of less than 10 days. As suggested by several authors, these ratios are more reliable for the determination of T_R (Ahmed et al., 2004; Crova et al., 2021; Sýkora et al., 2017). Using the ratios of long-lived radon progeny, such as $^{210}\text{Bi}/^{210}\text{Pb}$ and $^{210}\text{Po}/^{210}\text{Pb}$, longer residence times up to several weeks were observed. The large discrepancies in this case were observed to be due to extraneous sources of radon decay products (e.g., resuspension of soil dust, volcanic eruptions, biomass burning and fossil fuel combustion) (Lambert et al., 1983; Papastefanou, 2009b; Poet et al., 1972; Długosz–Lisiecka, 2016; Baskaran & Shaw, 2001; Moore et al., 1973). Methods employing ratios of radon and its short-lived progeny, such as $^{214}\text{Pb}/^{222}\text{Rn}$, often underestimate aerosols residence time, yielding the T_R values of the order of a few hours. This underestimation could be due to the following reasons: First, the short-lived

radon progeny (^{218}Po , ^{214}Pb , ^{214}Bi) decay rapidly to ^{210}Pb before significant atmospheric removal occurs. Second, the method used in this study assumes a secular equilibrium between radon and its short-lived progeny. To reach this equilibrium and ensure the steady-state conditions, more than 3 hours must elapse (Baskaran, 2011). This implies that the residence time we are tracing is much longer than 3 hours. The $^{210}\text{Pb}/^{214}\text{Pb}$, $^{210}\text{Pb}/^{214}\text{Bi}$ ratios used in this study may be more reliable due to their combination of the short- and long-lived decay products, which contrasts with the limitations of the $^{214}\text{Pb}/^{222}\text{Rn}$ ratio, as well as due to far fewer extraneous sources of ^{210}Pb compared to ^{210}Po . It should also be noted that the aerosols residence time cannot be a fixed value as it depends on several factors, such as sampling altitude, size distribution, air mass history and regional meteorological conditions (Anand & Mayya, 2015; Baskaran, 2011; Moore et al., 1973; Rastogi & Sarin, 2013; Wu et al., 2023).

Table 6.2 Comparison of aerosol residence times obtained in this study with those published in the literature.

Reference	Method	Range (Day)	Mean (Day)	Study location
Present	$^{210}\text{Pb}/^{214}\text{Pb}$	0.36 – 6.73	3	Bratislava, Slovakia
	$^{210}\text{Pb}/^{214}\text{Bi}$	0.37 – 6.7	3.41	Bratislava, Slovakia
	$^{210}\text{Pb}/^{222}\text{Rn}$	0.35 – 6.68	3	Bratislava, Slovakia
Aba et al., (2020)		3 – 5		Kuwait
Sýkora et al. (2017)		3.7 – 5.6	4.5	Bratislava, Slovakia
Kim et al. (2000)		1.3 – 4.5		Delaware, USA
Rastogi & Sarin, (2013)		2 – 8		Ahmadabad, India
Poet et al., (1972)	$^{222}\text{Rn}/^{210}\text{Pb}$	2.2 – 3.4		Belin–Halensee, Germani
Crova et al. (2021)	$^{210}\text{Pb}/^{214}\text{Bi}$	0.5 – 2.3	1.22	Milan, Italy
Vecchi et al. (2005)	$^{210}\text{Pb}/^{214}\text{Pb}$	1 – 2		Milan, Italy
Aba et al., (2020)		3 – 5		Kuwait
Ahmed et al., (2004)		1.5 – 13	10.5	El–Minia, Egypt
Mohery et al., (2016)		1.04 – 15.6	9.9	Jeddah, Saudi Arabia
Gäggeler et al., (1995)		1 – 15	6	Jungfrauoch, Switzerland
Papastefanou, (2006)	$^{210}\text{Bi}/^{210}\text{Pb}$	4.8 – 15.3	8.2	Oak Ridge, Tennessee
Poet et al., (1972)	$^{210}\text{Po}/^{210}\text{Pb}$	11–77		Belin–Halensee, Germani
Baskaran & Shaw, (2001)		0 – 39		Arctic
Długosz–Lisiecka, (2016)	Corrected T_R (^{210}Bi , ^{210}Pb and ^{210}Po)	1 – 25		Lodz, Poland
Barba–Lobo et al., (2024)	$^{214}\text{Pb}/^{222}\text{Rn}$	0.04 – 0.13		Huelva, Spain
Crova et al. (2021)	Deposition velocity and MLH	0.3 – 4.4	2	Milan, Italy
Winkler et al., (2001)	AMD of ^7Be and ^{210}Pb .	4– 5		Neuherberg, Germany
Aba et al., (2020)	AMD of ^7Be		4.2	Kuwait

Conclusion

This study investigated outdoor ^{222}Rn and its short-lived progeny (^{218}Po , ^{214}Pb , and ^{214}Bi) in Slovakia, with the aim of understanding their behaviour and assessing their potential as tracer of environmental processes. Long-term measurements were carried out, with radon activity concentration (RAC) continuously monitored over a six-year period (2018–2023) using a 4-liter scintillation detector. Additionally, the activity of short-lived radon progeny was measured using alpha spectrometry over a three-year span (2020–2022). These datasets enabled a comprehensive analysis of temporal variations and the factors influencing radon and its progeny levels.

The descriptive statistics revealed an annual mean RAC of $5.6 \pm 3.9 \text{ Bq.m}^{-3}$ (range: 0–31 Bq.m^{-3}), which is well below the often-cited global average ($\sim 10 \text{ Bq.m}^{-3}$). This indicates that the study area is a low-background region in terms of radon. The mean activity concentrations of ^{218}Po , ^{214}Pb , and ^{214}Bi were 2.78, 3.21, and 2.75 Bq.m^{-3} , respectively, with ^{214}Pb showing the largest variability ($\sigma = 1.95 \text{ Bq.m}^{-3}$). The mean equilibrium equivalent radon concentration was $3.01 \pm 1.70 \text{ Bq.m}^{-3}$, and the radon equilibrium factor (F_{eq}) averaged 0.54 ± 0.28 (range: 0.04–7.5) over the study period. The observed mean value of F_{eq} is slightly below the 0.6 suggested by UNSCEAR (2000).

A pronounced diurnal cycle was observed in radon and its progeny concentrations: peak values generally occurred in the early morning (03:00–05:00) and minima in the mid-afternoon (14:00–16:00). This diurnal variation is mainly governed by atmospheric mixing processes. During the day, solar heating increases convection and turbulence, increasing the boundary layer height (BLH) and diluting near-surface radon. This interpretation is supported by strong negative correlations between RAC and BLH (and likewise for RAC-wind speed and RAC-temperature). At night, the absence of solar heating leads to surface cooling and stable stratification with temperature inversions; this causes radon and its progeny to accumulate near the ground, thereby producing the observed early-morning concentration peaks.

Seasonal patterns in the concentration of radon and its progeny were also evident. Radon and its progeny concentrations were lowest in spring (April) and increased through autumn, reaching a maximum in November and remaining high during the winter months. This suggests that cooler, more stable conditions favour higher radon levels. Statistically, these seasonal variations were mainly attributed to seasonal changes in radon flux, atmospheric

mixing and stability i.e., RAC correlated strongly with measured radon flux and relative humidity, and negatively with BLH. Neither the air temperature nor the wind speed were significant predictors of the seasonal cycle. A moving-average smoothing revealed synoptic-scale fluctuations in radon levels on timescales ranging from days to weeks, indicating that large-scale weather patterns modulate radon levels beyond the diurnal cycle.

Regression analysis was conducted using both conventional and machine learning approaches to identify the most influential drivers of hourly radon variability. Multiple linear regression quantified the strength and direction of the predictors but exhibited limited predictive power. The generalized additive model (GAM) was effective in revealing the non-linear relationships between RAC and its predictors, resulting in slightly improved performance. However, the best performance was achieved using machine learning methods such as the Gradient Boosting Machine, Extreme Gradient Boosting (XGBoost), and the Random Forest. All these models outperformed the conventional approaches, achieving strong predictive power ($R^2 > 0.75$ for hourly data). Across all models, boundary layer height (BLH) emerged as the strongest predictor of outdoor radon levels, while precipitation was the least influential variable. Both the GAM and the Shapley Additive Explanations (SHAP) analysis of the XGBoost model confirmed that each meteorological predictor affects RAC in a non-linear manner, underscoring why simpler linear models are insufficient to fully describe the observed behaviour.

The use of radon and its progeny as tracers of atmospheric mixing layer height (MLH) and aerosols residence time were investigated. Using the so-called box model with hourly RAC and daily modelled radon flux as inputs, the MLH was retrieved. The radon-derived MLH effectively captured the diurnal and seasonal evolution of atmospheric mixing processes. However, a comparison with independent data revealed both strengths and weaknesses of the radon-based MLH method. Generally, the radon-derived MLH agreed well with BLH from the ERA5 reanalysis dataset on diurnal and seasonal scales. The hourly MLH based on radon was not reliable due to large uncertainties arising from the model's sensitivity to RAC measurement errors. Moreover, the model's seasonal variation depends strongly on the assumed radon flux. Since continuously measured radon flux data are not available, this further distorts the results.

The mean residence time of aerosols obtained from the ratios $^{210}\text{Pb}/^{222}\text{Rn}$, $^{210}\text{Pb}/^{214}\text{Pb}$, and $^{210}\text{Pb}/^{214}\text{Bi}$ in Bratislava, Slovakia was 3.15 days (range: 0.35–6.73 days). This result is consistent with previously reported aerosol residence times based on radon progeny in other

locations. However, the absence of an independent model for estimating aerosol residence time makes it difficult to validate this finding.

In summary, these findings provide deep insights into the dynamics of outdoor radon and its progeny in a low-background environment, shedding light on their interactions with environmental factors. The results support the use of radon and its progeny as tracers of boundary-layer processes and aerosol dynamics. However, precise measurements are essential for their effective use in this way.

Future research directions

The use of radon as an environmental tracer requires the development of highly accurate and traceable instrumentation, alongside the implementation of long-term measurement campaigns. Current study suggests that reliable radon concentration and radon flux measurements are critical, as they significantly influence estimation of atmospheric mixing layer height. Future research should prioritize improving detector sensitivity and automated radon flux monitoring systems that can collect continuous, traceable data to support atmospheric science.

A deeper understanding of radon progeny interactions with various air pollutants is essential, as both are known carcinogens associated with significant health risks. However, these interactions remain poorly characterized. Comprehensive laboratory and field studies under diverse pollution scenarios are needed to better understand these dynamics. Such efforts are crucial for improving radiation dose modelling and for advancing our understanding of the potential synergistic effects between radon progeny and air pollutants.

In parallel, long-term epidemiological investigations should examine the combined health impacts of radon exposure and air pollution. While both radon and fine particulate matter are individually linked to respiratory and cardiovascular diseases, their synergistic effects are not well understood. Emerging evidence suggests that elevated radon levels significantly exacerbate PM-associated mortality (Blomberg et al., 2019). Future research should integrate high-resolution, continuous radon and PM data with health outcome metrics (e.g., hospital admissions, disease incidence, mortality rate) to quantify these interactions across different spatial and temporal scales.

Machine learning offers promising avenues for advancing radon research. Beyond the regression methods used in this thesis, future work could apply machine learning-based clustering to long-term datasets of radon, its progeny, meteorological variables, and air quality data. This would clarify how meteorological and air quality regimes influence radon and progeny variability, providing insights for predictive models and risk mitigation.

List of Publications

1. **Sultani, M. A.**, Bulko, M., Helej, M., Sýkora, I., Müllerová, M., Masarik, J. (2025). *Influence of Atmospheric Mixing Conditions and Particulate Matter on Outdoor Radon and Its Progeny: Implications for Aerosol Residence Time*. Atmospheric Environment, 121307. <https://doi.org/10.1016/j.atmosenv.2025.121307>
2. **Sultani, M. A.**, Bulko, M., Sýkora, I., Müllerová, M., Masarik, J., Tonhauzer, P. (2024). *Impacts of meteorology and mixing height on radioactive and stable aerosols in Bratislava, Slovakia*. Atmospheric Research, 311, 107710. <https://doi.org/10.1016/j.atmosres.2024.107710>
3. **Sultani, M.A.**, Bulko, M., Holý, K., Müllerová, M., Masarik, J., Tonhauzer, P., Helej, M. (2024). *The use of radon as a tracer for air quality assessment: a case study in Bratislava, Slovakia*. Journal of Radioanalytical and Nuclear Chemistry 333, 2515–2527. <https://doi.org/10.1007/s10967-023-08969-3>

Conference contributions

1. **Sultani, M. A.**, et al. (2024). *Factors influencing outdoor radon variability: A time series study in Bratislava, Slovakia*. Presented at the 10th International Conference on Protection against Radon at Home and at Work, Prague, Czech Republic, 9–12 September 2024. Published in Book of Abstracts. Prague: Faculty of Nuclear Sciences and Physical Engineering, National Radiation Protection Institute.
2. **Sultani, M. A.**, et al. (2024). *Outdoor radon and its progeny assessment for use in atmospheric research*. Published in Book of Abstracts, IX. Terrestrial Radioisotopes in Environment: International Conference on Environmental Protection, Hungary, 19–22 November 2024. Hungary: Conference Organizing Committee.
3. **Sultani, M. A.**, et al. (2023). *Investigation of boundary layer height evolution and its implication for air pollution monitoring: A long-term study based on radon in Bratislava, Slovakia*. 7th International Conference on Environmental Radioactivity – New Challenges in the Determination of Environmental Radioactivity, Seville, Spain, 17–22 September 2023. Abstract published in Book of Abstracts, ENVIRA 2023.
4. **Sultani, M. A.**, et al. (2023). *A radon-based approach to characterize the parameters governing the air quality in urban areas*. Published in Book of Abstracts, Student Scientific Conference, Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, p. 264. ISBN 978-80-8147-136-0.
5. **Sultani, M. A.**, et al. (2022). *The use of radon as a tracer for air pollution assessment: A case study in Bratislava, Slovakia*. Published in Book of Abstracts, 8th Terrestrial Radioisotopes in Environment: International Conference on

Environmental Protection, Vonyarcvashegy, Hungary, 4–10 October 2022, p. 80. Veszprém: Social Organization for Radioecological Cleanliness. ISBN 978-615-81632-1-7.

6. **Sultani, M. A.**, et al. (2022). *Radon-based atmospheric mixing layer height and its influence on major air pollutant concentrations in Bratislava, Slovakia*. Published in Book of Abstracts, 43rd Days of Radiation Protection, Stará Lesná, Slovakia, 19–23 September 2022, p. 63. Bratislava: Slovak Medical University. ISBN 978-80-89702-98-5.
7. **Sultani, M. A.**, et al. (2022). *A new approach for determination of mixing layer height using radon*. Published in Book of Abstracts, Student Scientific Conference, Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, 27 April 2022, p. 273. Bratislava: Knižničné a edičné centrum. ISBN 978-80-8147-127-8.

References

- Aba, A., Ismaeel, A., Al-Boloushi, O., Al-Shammari, H., Al-Boloushi, A., & Malak, M. (2020). Atmospheric residence times and excess of unsupported ^{210}Po in aerosol samples from the Kuwait bay-northern gulf. *Chemosphere*, 261, 127690. <https://doi.org/10.1016/j.chemosphere.2020.127690>
- Abdelfatah Mostafa, M. Y., Bader Khalaf, H. N., & Zhukovsky, M. (2020). Radon decay products equilibrium at different aerosol concentrations. *Applied Radiation and Isotopes*, 156, 108981. <https://doi.org/10.1016/j.apradiso.2019.108981>
- Abdo, M. A. S., Boukhair, A., Fahad, M., Ouakkas, S., Arhouni, F. E., Hakkar, M., Belahbib, L., & Al-Suhbani, M. N. (2021). Estimation of unattached and aerosol-attached activities of airborne short-lived radon progeny in indoor environments. *Journal of Environmental Radioactivity*, 237, 106665. <https://doi.org/10.1016/j.jenvrad.2021.106665>
- Adyasari, D., Dimova, N. T., Dulai, H., Gilfedder, B. S., Cartwright, I., McKenzie, T., & Fuleky, P. (2023). Radon-222 as a groundwater discharge tracer to surface waters. *Earth-Science Reviews*, 238, 104321. <https://doi.org/10.1016/j.earscirev.2023.104321>
- Ahmed, A. A., Mohamed, A., Ali, A. E., Barakat, A., Abd El-Hady, M., & El-Hussein, A. (2004). Seasonal variations of aerosol residence time in the lower atmospheric boundary layer. *Journal of Environmental Radioactivity*, 77(3), 275–283. <https://doi.org/10.1016/j.jenvrad.2004.03.011>
- Air Pollution in the Slovak Republic (2018) Air Quality Department Slovak Hydrometeorological Institute Bratislava, Slovakia
- Allabakash, S., & Lim, S. (2020). Climatology of Planetary Boundary Layer Height-Controlling Meteorological Parameters Over the Korean Peninsula. *Remote Sensing*, 12(16), Article 16. <https://doi.org/10.3390/rs12162571>
- Allegrini, I., Febo, A., Pasini, A., & Schiarini, S. (1994). Monitoring of the nocturnal mixed layer by means of particulate radon progeny measurement. *Journal of Geophysical Research: Atmospheres*, 99(D9), 18765-18777.
- Al-Najjar, S. A. R., & Durrani, S. A. (1984). Track profile technique (TPT) and its applications using CR-39. I: Range and energy measurements of alpha-particles and fission fragments. *Nuclear Tracks and Radiation Measurements* (1982), 8(1-4), 45-49.
- Al-Shboul, K. F. (2023). Unraveling the complex interplay between soil characteristics and radon surface exhalation rates through machine learning models and multivariate analysis. *Environmental Pollution*, 336, 122440. <https://doi.org/10.1016/j.envpol.2023.122440>
- Anand, S., & Mayya, Y. S. (2015). Coagulation effect on the activity size distributions of long lived radon progeny aerosols and its application to atmospheric residence time

- estimation techniques. *Journal of Environmental Radioactivity*, 141, 153–163. <https://doi.org/10.1016/j.jenvrad.2014.12.012>
- Appleton, J. D. (2007). Radon: Sources, Health Risks, and Hazard Mapping. *AMBIO: A Journal of the Human Environment*, 36(1), 85–89. [https://doi.org/10.1579/0044-7447\(2007\)36\[85:RSHRAH\]2.0.CO;2](https://doi.org/10.1579/0044-7447(2007)36[85:RSHRAH]2.0.CO;2)
- Ball, T. K., Cameron, D. G., Colman, T. B., & Roberts, P. D. (1991). Behaviour of radon in the geological environment: a review. *Quarterly Journal of Engineering Geology and Hydrogeology*, 24(2), 169–182.
- Banrion, M., Cobelli, M., & Crowley, Q. G. (2023). Applying machine learning to model radon using topsoil geochemistry. *Applied Geochemistry*, 158, 105790. <https://doi.org/10.1016/j.apgeochem.2023.105790>
- Barba-Lobo, A., Gutiérrez-Álvarez, I., Adame, J. A., San Miguel, E. G., & Bolívar, J. P. (2024). Behavior of ^{222}Rn , ^{220}Rn and their progenies along a daily cycle for different meteorological situations: Implications on atmospheric aerosol residence times and Rn daughters' equilibrium factors. *Journal of Hazardous Materials*, 464, 132998. <https://doi.org/10.1016/j.jhazmat.2023.132998>
- Barba-Lobo, A., Gutiérrez-Álvarez, I., San Miguel, E. G., & Bolívar, J. P. (2023). A methodology to determine ^{212}Pb , ^{212}Bi , ^{214}Pb and ^{214}Bi in atmospheric aerosols; Application to precisely obtain aerosol residence times and Rn-daughters' equilibrium factors. *Journal of Hazardous Materials*, 445, 130521. <https://doi.org/10.1016/j.jhazmat.2022.130521>
- BARRETTO, P. M. D. C. (1973). Emanation Characteristics of Terrestrial and Lunar Materials and the Radon-222 Loss Effect on the Uranium—Lead System Discordance. [Ph.D., Rice University]. <https://www.proquest.com/docview/302665821/citation/1E5DF0019A3A41F0PQ/1>
- Baskaran, M. (2011). Po-210 and Pb-210 as atmospheric tracers and global atmospheric Pb-210 fallout: A Review. *Journal of Environmental Radioactivity*, 102(5), 500–513. <https://doi.org/10.1016/j.jenvrad.2010.10.007>
- Baskaran, M. (2016). Radon: A tracer for geological, geophysical and geochemical studies (Vol. 367). Basel: Springer.
- Baskaran, M., & Shaw, G. E. (2001). Residence time of arctic haze aerosols using the concentrations and activity ratios of ^{210}Po , ^{210}Pb and ^7Be . *Journal of Aerosol Science*, 32(4), 443–452.
- Belan, T., Chudy, M., Durana, L., Holy, K., Levaiova, D., Povinec, P., ... & Sivo, A. (1992). Investigation of radionuclide variations in the Bratislava air. In *Rare nuclear processes*.
- Bertin Technologies. (n.d.). AlphaGUARD – Radon monitor. Retrieved May 18, 2025, from AlphaGUARD – Radon monitor - Bertin Technologies
- Blomberg, A. J., Coull ,Brent A., Jhun ,Iny, Vieira ,Carolina L.Z., Zanobetti ,Antonella, Garshick ,Eric, Schwartz ,Joel, & and Koutrakis, P. (2019). Effect modification of

- ambient particle mortality by radon: A time series analysis in 108 U.S. cities. *Journal of the Air & Waste Management Association*, 69(3), 266–276. <https://doi.org/10.1080/10962247.2018.1523071>
- Bossew, P., Benà, E., Chambers, S., & Janik, M. (2024). Analysis of Outdoor and Indoor Radon Concentration Time Series Recorded with RadonEye Monitors. *Atmosphere*, 15(12), Article 12. <https://doi.org/10.3390/atmos15121468>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Bulko Martin. (2010). Radón v atmosfére a jeho aplikácie (Dizertačná práca) Martin bulko. Dizertačná práca. Fakulta matematiky, fyziky a informatiky-Univerzita komenského, Bratislava.
- Bulko, M., Holý, K., & Müllerová, M. (2018). On the relation between outdoor ²²²Rn and atmospheric stability determined by a modified Turner method. *Journal of Environmental Radioactivity*, 189, 79-92.
- Bulko, M., Holý, K., & Müllerová, M. (2020). Analysis of Time Series of Radon Activity Concentration in Outdoor Air and Their Reconstruction. *Radiation Protection Dosimetry*, 191(2), 244-249.
- Butterweck, G., Schuler, Ch., Vezzù, G., Marsh, J. W., & Birchall, A. (2005). In-vivo measurement of deposition and absorption of unattached radon progeny. In *Radioactivity in the Environment* (Vol. 7, pp. 314–325). Elsevier. [https://doi.org/10.1016/S1569-4860\(04\)07035-4](https://doi.org/10.1016/S1569-4860(04)07035-4)
- Calamosca, M., & Penzo, S. (2008). Are ageing and fading really a problem for the quality of the passive radon measurements? *Radiation measurements*, 43, S422-S426.
- Čeliković, I., Pantelić, G., Vukanac, I., Krneta Nikolić, J., Živanović, M., Cinelli, G., Gruber, V., Baumann, S., Quindos Poncela, L. S., & Rabago, D. (2022). Outdoor Radon as a Tool to Estimate Radon Priority Areas—A Literature Overview. *International Journal of Environmental Research and Public Health*, 19(2), Article 2. <https://doi.org/10.3390/ijerph19020662>
- Čeliković, I., Pantelić, G., Vukanac, I., Nikolić, J. K., Živanović, M., Cinelli, G., Gruber, V., Baumann, S., Ciotoli, G., Poncela, L. S. Q., & Rábago, D. (2022). Overview of Radon Flux Characteristics, Measurements, Models and Its Potential Use for the Estimation of Radon Priority Areas. *Atmosphere*, 13(12), Article 12. <https://doi.org/10.3390/atmos13122005>
- Chambers, S. D., Williams, A. G., Crawford, J., & Griffiths, A. D. (2015). On the use of radon for quantifying the effects of atmospheric stability on urban emissions. *Atmospheric Chemistry and Physics*, 15(3), 1175-1190.
- Chambers, S., Williams, A. G., Zahorowski, W., Griffiths, A., & Crawford, J. (2011). Separating remote fetch and local mixing influences on vertical radon measurements in the lower atmosphere. *Tellus B: Chemical and Physical Meteorology*, 63(5), 843-859.

- Chen, C., Thomas, D. M., & Green, R. E. (1995). Modeling of radon transport in unsaturated soil. *Journal of Geophysical Research: Solid Earth*, 100(B8), 15517–15525. <https://doi.org/10.1029/95JB01290>
- Chen, J., & Harley, N. H. (2018). A Review of Indoor and Outdoor Radon Equilibrium Factors—part I: ^{222}Rn . *Health Physics*, 115(4), 490. <https://doi.org/10.1097/HP.0000000000000909>
- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Cheng, W. C., Liu, C. H., & Leung, D. Y. (2009). On the correlation of air and pollutant exchange for street canyons in combined wind-buoyancy-driven flow. *Atmospheric Environment*, 43(24), 3682–3690.
- Cothern, C. R., & Smith, J. E. (1987). *Environmental Radon*. Springer Science & Business Media.
- Crova, F., Valli, G., Bernardoni, V., Forello, A. C., Valentini, S., & Vecchi, R. (2021). Effectiveness of airborne radon progeny assessment for atmospheric studies. *Atmospheric Research*, 250, 105390. <https://doi.org/10.1016/j.atmosres.2020.105390>
- Ćujić, M., Janković Mandić, L., Petrović, J., Dragović, R., Đorđević, M., Đokić, M., & Dragović, S. (2021). Radon-222: environmental behavior and impact to (human and non-human) biota. *International Journal of Biometeorology*, 65, 69–83.
- Cullen, A. C., & Frey, H. C. (1999). *Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs*. Springer Science & Business Media.
- Currie, L. A. (1968). CAS: 528: DyaF1cXnsVyitQ% 3D% 3D: Limits for qualitative detection and quantitative determination. vol. 40. *Appl Radiochem Anal Chem*, 586–593.
- Delignette-Muller, M. L., & Dutang, C. (2015). fitdistrplus: An R Package for Fitting Distributions. *Journal of Statistical Software*, 64, 1–34. <https://doi.org/10.18637/jss.v064.i04>
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, 74(366a), 427–431. <https://doi.org/10.1080/01621459.1979.10482531>
- Dicu, T., Cucos, A., Botoș, M., Burghel, B., Florică, Ș., Baci, C., Ștefan, B., & Bălc, R. (2023). Exploring statistical and machine learning techniques to identify factors influencing indoor radon concentration. *Science of The Total Environment*, 905, 167024. <https://doi.org/10.1016/j.scitotenv.2023.167024>

- Długosz, M., Grabowski, P., & Bem, H. (2009). ^{210}Pb and ^{210}Po radionuclides in the urban air of Lodz, Poland. *Journal of Radioanalytical and Nuclear Chemistry*, 283(3), 719–725. <https://doi.org/10.1007/s10967-009-0407-x>
- Długosz-Lisiecka, M., & Bem, H. (2012). Determination of the mean aerosol residence times in the atmosphere and additional ^{210}Po input on the base of simultaneous determination of ^7Be , ^{22}Na , ^{210}Pb , ^{210}Bi and ^{210}Po in urban air. *Journal of Radioanalytical and Nuclear Chemistry*, 293(1), 135–140. <https://doi.org/10.1007/s10967-012-1690-5>
- Dorn FE (1900) Die von radioactiven Substanzen ausgesandte Emanation. Emanation”. *Abhandlungen der Naturforschenden Gesellschaft zu Halle* 23:1–15
- Duenas, C., & Fernandez, M. C. (1987). Dependence of radon ^{222}Rn flux on concentrations of soil gas and air gas and an analysis of the effects produced by several atmospheric variables. In *Annales geophysicae. Series B. Terrestrial and planetary physics* (Vol. 5, No. 6, pp. 533-539).
- Durrani, S. A., & Ilic, R. (Eds.). (1997). *Radon measurements by etched track detectors-applications in radiation protection, earth sciences*. World Scientific.
- Durridge. (n.d.). RAD7 Radon Detector. Retrieved May 18, 2025, from RAD7 Radon Detector - DURRIDGE Scientific Research radon detectors
- EC (European Council). Council Directive 2013/59/Euratom Laying down Basic Safety Standards for Protection against the Dangers Arising from Exposure to Ionising Radiation. *Off. J. Eur. Union* 2014, 57, 1–73.
- Edsfeldt, C. (2001). The radium distribution in some Swedish soils and its effects on radon emanation (Doctoral dissertation, Institutionen för anläggning och miljö).
- Elío, J., Petermann, E., Bossew, P., & Janik, M. (2023). Machine learning in environmental radon science. *Applied Radiation and Isotopes*, 194, 110684. <https://doi.org/10.1016/j.apradiso.2023.110684>
- Fontan, J., Guedalia, D., Druilhet, A., & Lopez, A. (1979). Une methode de mesure de la stabilite verticale de l’atmosphere pres du sol. *Boundary-Layer Meteorology*, 17, 3-14.
- Fox, J. (2015). *Applied Regression Analysis and Generalized Linear Models*. SAGE Publications.
- Franci, D., Aureli, T., & Cardellini, F. (2015). Study of ageing and fading in CR-39 detectors for different storage conditions. *Radiation protection dosimetry*, 167(4), 425-428.
- Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, 29(5), 1189–1232.
- Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, 29(5), 1189–1232.

- Gäggeler, H. W., Jost, D. T., Baltensperger, U., Schwikowski, M., & Seibert, P. (1995). Radon and thoron decay product and ^{210}Pb measurements at Jungfraujoch, Switzerland. *Atmospheric Environment*, 29(5), 607–616. [https://doi.org/10.1016/1352-2310\(94\)00195-Q](https://doi.org/10.1016/1352-2310(94)00195-Q)
- Galmarini, S. (2006). One year of ^{222}Rn concentration in the atmospheric surface layer. *Atmospheric Chemistry and Physics*, 6(10), 2865–2886.
- Gao, N. P., & Niu, J. L. (2007). Modeling particle dispersion and deposition in indoor environments. *Atmospheric environment*, 41(18), 3862–3876.
- Garzon, L., Juanco, J. M., Perez, J. M., Fernandez, J. M., & Arganza, B. (1986). The universal Rn wave. An approach. *Health Physics*, 51(2), 185–195.
- Gray, R. W., and Ramsay, W. (1910). C. R. 151, 126.
- Gregorič, A., Drinovec, L., Ježek, I., Vaupotič, J., Lenarčič, M., Grauf, D., ... & Močnik, G. (2020). The determination of highly time-resolved and source-separated black carbon emission rates using radon as a tracer of atmospheric dynamics. *Atmospheric Chemistry and Physics*, 20(22), 14139–14162.
- Griffiths, A. D., Parkes, S. D., Chambers, S. D., McCabe, M. F., & Williams, A. G. (2013). Improved mixing height monitoring through a combination of lidar and radon measurements. *Atmospheric Measurement Techniques*, 6(2), 207–218. <https://doi.org/10.5194/amt-6-207-2013>
- Grossi, C., Chambers, S. D., Llido, O., Vogel, F. R., Kazan, V., Capuana, A., ... & Ramonet, M. (2020). Intercomparison study of atmospheric ^{222}Rn and ^{222}Rn progeny monitors. *Atmospheric Measurement Techniques*, 13(5), 2241–2255.
- Grossi, C., Vogel, F. R., Curcoll, R., Àgueda, A., Vargas, A., Rodó, X., & Morguá, J. A. (2018). Study of the daily and seasonal atmospheric CH_4 mixing ratio variability in a rural Spanish region using ^{222}Rn tracer. *Atmospheric chemistry and physics*, 18(8), 5847–5860.
- Grossi, C., Vogel, F. R., Morgui, J. A., Curcoll, R., Àgueda, A., Batet, O., ... & Rodó, X. (2014). First estimation of CH_4 fluxes using the ^{222}Rn tracer method over the central Iberian Peninsula. *WIT Trans. Ecol. Environ*, 183, 233–244.
- Gunning, G. A., Pollard, D., & Finch, E. C. (2014). An outdoor radon survey and minimizing the uncertainties in low level measurements using CR-39 detectors. *Journal of Radiological Protection*, 34(2), 457.
- Guo, J., Zhang, J., Yang, K., Liao, H., Zhang, S., Huang, K., Lv, Y., Shao, J., Yu, T., Tong, B., Li, J., Su, T., Yim, S. H. L., Stoffelen, A., Zhai, P., & Xu, X. (2021). Investigation of near-global daytime boundary layer height using high-resolution radiosondes: First results and comparison with ERA5, MERRA-2, JRA-55, and NCEP-2 reanalyses. *Atmospheric Chemistry and Physics*, 21(22), 17079–17097. <https://doi.org/10.5194/acp-21-17079-2021>

- Gupta, M. L., Douglass, A. R., Kawa, S. R., & Pawson, S. (2004). Use of radon for evaluation of atmospheric transport models: Sensitivity to emissions. *Tellus B: Chemical and Physical Meteorology*, 56(5), 404–412. <https://doi.org/10.3402/tellusb.v56i5.16467>
- H2O.ai. (2020). H2O: Scalable Machine Learning and Deep Learning Platform. <https://www.h2o.ai>
- Hafez, A. F., & Somogyi, G. (1986). Determination of radon and thoron permeability through some plastics by track technique. *International Journal of Radiation Applications and Instrumentation. Part D. Nuclear Tracks and Radiation Measurements*, 12(1-6), 697-700.
- Hardcastle, G. D., & Miles, J. C. H. (1996). Ageing and fading of alpha particle tracks in CR-39 exposed to air. *Radiation protection dosimetry*, 67(4), 295-298.
- Hassan, N. M. (2014). Radon emanation coefficient and its exhalation rate of wasted petroleum samples associated with petroleum industry in Egypt. *Journal of Radioanalytical and Nuclear Chemistry*, 299, 111-117.
- Hassan, N. M., Hosoda, M., Ishikawa, T., SORIMACHI, A., SAHOO, S. K., TOKONAMI, S., & FUKUSHI, M. (2009). Radon migration process and its influence factors; review. *Japanese Journal of Health Physics*, 44(2), 218-231.
- Hastie, T. J. (1992). *Generalized Additive Models*. In *Statistical Models in S*. Routledge.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., ... Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
- Holý, K., Müllerová, M., Bulko, M., Holá, O., & Melicherová, T. (2016). Outdoor ²²²Rn behaviour in different areas of Slovakia. *Nukleonika*, 61.
- Holý, K., Müllerová, M., Palušová, V., & Bulko, M. (2017). Comparison of various approaches to determining CO₂ fluxes from the soil by ²²²Rn calibrated method. *Radiation Protection Dosimetry*, 177(1-2), 144-148.
- Hosoda, M., Shimo, M., Sugino, M., Furukawa, M., & Fukushima, M. (2007). Effect of soil moisture content on radon and thoron exhalation. *Journal of nuclear science and technology*, 44(4), 664-672.
- Hwa Oh, Y., & Kim, G. (2015). A radon-thoron isotope pair as a reliable earthquake precursor. *Scientific Reports*, 5(1), 13084.
- Hwa Oh, Y., & Kim, G. (2015). A radon-thoron isotope pair as a reliable earthquake precursor. *Scientific Reports*, 5(1), 13084. <https://doi.org/10.1038/srep13084>
- IAEA (International Atomic Energy Agency). (2014) *Radiation Protection and Safety of Radiation Sources: International Basic Safety Standards General Safety Requirements Part 3*; IAEA: Vienna, Austria.

- IBM Corp. (2021). IBM SPSS Statistics for Windows, Version 27.0. Armonk, NY: IBM Corp.
- ICRP (International Commission on Radiological Protection). (1993) Protection against radon-222 at home and at work. ICRP Publication 65. Ann. ICRP 23(2). Oxford, UK: Pergamon Press.
- Iskandar, D., Yamazawa, H., & Iida, T. (2004). Quantification of the dependency of radon emanation power on soil temperature. *Applied Radiation and Isotopes*, 60(6), 971-973.
- Jacobi, W., & Andre, K. (1963). The vertical distribution of radon 222, radon 220 and their decay products in the atmosphere. *Journal of Geophysical Research*, 68(13), 3799-3814.
- Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651–666. <https://doi.org/10.1016/j.patrec.2009.09.011>
- Janik, M., Bossew, P., & Kurihara, O. (2018). Machine learning methods as a tool to analyse incomplete or irregularly sampled radon time series data. *Science of The Total Environment*, 630, 1155–1167. <https://doi.org/10.1016/j.scitotenv.2018.02.233>
- Jolliffe, I. T. (Ed.). (2002). Principal Component Analysis for Special Types of Data. In *Principal Component Analysis* (pp. 338–372). Springer. https://doi.org/10.1007/0-387-22440-8_13
- Jonassen, N. (1983). The determination of radon exhalation rates. *Health Physics*, 45(2), 369-376.
- Kagerer, S., Rettenmoser, T., Hofmann, W., Falkensteiner, A., & Steger, F. (2005). Bioaerosols as carriers of radon progeny. In *Radioactivity in the Environment* (Vol. 7, pp. 649–656). Elsevier. [https://doi.org/10.1016/S1569-4860\(04\)07079-2](https://doi.org/10.1016/S1569-4860(04)07079-2)
- Kaimal, J. C., & Finnigan, J. J. (1994). *Atmospheric Boundary Layer Flows: Their Structure and Measurement*. Oxford University Press.
- Karstens, U., Levin, I., 2023. traceRadon daily radon flux maps for Europe for 2017-2023. <https://doi.org/10.18160/CFGD-ZPGG>
- Karstens, U., Schwingshackl, C., Schmithüsen, D., & Levin, I. (2015). A process-based 222radon flux map for Europe and its comparison to long-term observations. *Atmospheric Chemistry and Physics*, 15(22), 12845–12865. <https://doi.org/10.5194/acp-15-12845-2015>
- Khan, A., Phillips, C. R., & Duport, P. (1988). Effect of increase in humidity on the size and activity distributions of radon progeny laden aerosols from hydrocarbon combustion. *Radiat. Prot. Dosim.*; (United Kingdom), 22:1. https://www.osti.gov/etdeweb/biblio/7072180?utm_source=chatgpt.com
- Kikaj, D., Vaupotič, J., & Chambers, S. D. (2019). Identifying persistent temperature inversion events in a subalpine basin using radon-222. *Atmospheric Measurement Techniques*, 12(8), 4455–4477. <https://doi.org/10.5194/amt-12-4455-2019>

- Kim, K. Y. (2022). Diurnal and seasonal variation of planetary boundary layer height over East Asia and its climatic change as seen in the ERA-5 reanalysis data. *SN Applied Sciences*, 4(2), 39.
- Kimball, B. A. (1974). Smoothing Data with Fourier Transformations. *Agronomy Journal*, 66(2), 259–262. <https://doi.org/10.2134/agronj1974.00021962006600020023x>
- Kristiansen, N. I., Stohl, A., Olivié, D. J. L., Croft, B., Søvde, O. A., Klein, H., Christoudias, T., Kunkel, D., Leadbetter, S. J., Lee, Y. H., Zhang, K., Tsigaridis, K., Bergman, T., Evangeliou, N., Wang, H., Ma, P.-L., Easter, R. C., Rasch, P. J., Liu, X., ... Zhang, H. (2016). Evaluation of observed and modelled aerosol lifetimes using radioactive tracers of opportunity and an ensemble of 19 global models. *Atmospheric Chemistry and Physics*, 16(5), 3525–3561. <https://doi.org/10.5194/acp-16-3525-2016>
- Krizman, M., & Stegnar, P. (1992). Environmental impact of the 'Zirovski VRH' uranium mine on the enhancement of outdoor radon concentrations. *Radiation Protection Dosimetry*, 45(1-4), 723-728.
- Kubiak, J. A., & Zimnoch, M. (2022). Assessment of the nocturnal boundary layer height based on long-term atmospheric radon measurements. *Frontiers in Earth Science*, 10. <https://doi.org/10.3389/feart.2022.955791>
- Kumar K, C., & Nagaraja, K. (2024). Time-series analysis of radon concentrations for Bengaluru's atmosphere. *Radiation Protection Dosimetry*, 200(11–12), 1003–1006. <https://doi.org/10.1093/rpd/ncad272>
- Kumar, A., Chauhan, R. P., Joshi, M., & Aggarwal, P. (2015). Implications of variability in Indoor radon/thoron levels: a study of dwellings in Haryana, India. *Environmental Earth Sciences*, 73, 4033-4042.
- Kümmel, M., Dushe, C., Müller, S., & Gehrcke, K. (2014). Outdoor ²²²Rn-concentrations in Germany–part 1–natural background. *Journal of environmental radioactivity*, 132, 123-130.
- Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54(1), 159–178. [https://doi.org/10.1016/0304-4076\(92\)90104-Y](https://doi.org/10.1016/0304-4076(92)90104-Y)
- Lambert, G., Sanak, J., & Polian, G. (1983). Mean residence time of the submicrometer aerosols in the global troposphere. In *Precipitation scavenging, dry deposition, and resuspension*. Volume 2. Proceedings. https://inis.iaea.org/search/search.aspx?orig_q=RN:16006009
- Lee, C. B. (1986). Simple model and climatological aspects of the structure of the convective boundary layer. *Atmospheric Environment* (1967), 20(4), 705-714.
- Levin, I., Glatzel-Mattheier, H., Marik, T., Cuntz, M., Schmidt, M., & Worthy, D. E. (1999). Verification of German methane emission inventories and their recent changes based on atmospheric observations. *Journal of Geophysical Research: Atmospheres*, 104(D3), 3447-3456.

- Levin, I., Karstens, U., Hammer, S., DellaColetta, J., Maier, F., & Gachkivskyi, M. (2021). Limitations of the radon tracer method (RTM) to estimate regional greenhouse gas (GHG) emissions – a case study for methane in Heidelberg. *Atmospheric Chemistry and Physics*, 21(23), 17907–17926. <https://doi.org/10.5194/acp-21-17907-2021>
- Li, Ta-Yung. (1974) Diurnal variations of radon and meteorological variables near the ground. *Boundary-Layer Meteorology* 7, no. 2: 185-198.
- Li, X., Dong, Y., Zhang, Y., Shi, Z., & Yao, J. (2023). Climatology of Planetary Boundary Layer Height over Jiangsu, China, Based on ERA5 Reanalysis Data. *Atmosphere*, 14(9), Article 9. <https://doi.org/10.3390/atmos14091330>
- Liu, H., Daisuke, K., Motokiyo, M., Hirao, S., Moriizumi, J., & Yamazawa, H. (2014). On the characteristics of the wet deposition process using radon as a tracer gas. *Radiation Protection Dosimetry*, 160(1–3), 83–86. <https://doi.org/10.1093/rpd/ncu093>
- Lucas, H. F. (1957). Improved low-level alpha-scintillation counter for radon. *Review of Scientific Instruments*, 28(9), 680-683.
- Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems*, 30. <https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html>
- Lv, W., Wu, Y., & Zang, J. (2021). A Review on the Dispersion and Distribution Characteristics of Pollutants in Street Canyons and Improvement Measures. *Energies*, 14(19), 6155.
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1: Statistics, 5, 281–298. <https://projecteuclid.org/ebook/Download?urlId=bsmsp/1200512992&isFullBook=false>
- Madonna, F., Summa, D., Di Girolamo, P., Marra, F., Wang, Y., & Rosoldi, M. (2021). Assessment of Trends and Uncertainties in the Atmospheric Boundary Layer Height Estimated Using Radiosounding Observations over Europe. *Atmosphere*, 12(3), Article 3. <https://doi.org/10.3390/atmos12030301>
- Magalhães, M. H., Amaral, E. C. S., Sachett, I., & Rochedo, E. R. R. (2003). Radon-222 in Brazil: an outline of indoor and outdoor measurements. *Journal of environmental radioactivity*, 67(2), 131-143.
- Manabe, S., & Terpstra, T. B. (1974). The effects of mountains on the general circulation of the atmosphere as identified by numerical experiments. *Journal of Atmospheric Sciences*, 31(1), 3-42.
- Mann, H. B. (1945). Nonparametric Tests Against Trend. *Econometrica*, 13(3), 245–259. <https://doi.org/10.2307/1907187>

- Mansy, M., Sharaf, M. A., Eissa, H. M., El-Kamees, S. U., & Abo-Elmagd, M. (2006). Theoretical calculation of SSNTD response for radon measurements and optimum diffusion chambers dimensions. *Radiation Measurements*, 41(2), 222-228.
- Marley, N. A., Gaffney, J. S., Drayton, P. J., Cunningham, M. M., Orlandini, K. A., & Paode, R. (2000). Measurement of ^{210}Pb , ^{210}Po , and ^{210}Bi in Size-Fractionated Atmospheric Aerosols: An Estimate of Fine-Aerosol Residence Times. *Aerosol Science and Technology*, 32(6), 569–583. <https://doi.org/10.1080/027868200303489>
- McCormick, R. A., & Xintaras, C. (1962). Variation of carbon monoxide concentrations as related to sampling interval, traffic and meteorological factors. *Journal of Applied Meteorology and Climatology*, 1(2), 237-243.
- McKinley, S., & Levine, M. (1998). Cubic spline interpolation. *College of the Redwoods*, 45(1), 1049-1060.
- Minato, S. (1988). Seasonal variations in radon concentrations in the lower atmosphere at Nagoya. 37(9–10), 233–240.
- Mohery, M., Abdallah, A. M., Ali, A., & Baz, S. S. (2016). Daily variation of radon gas and its short-lived progeny concentration near ground level and estimation of aerosol residence time*. *Chinese Physics B*, 25(5), 050701. <https://doi.org/10.1088/1674-1056/25/5/050701>
- MONTGOMERY, D. C., & and FRIEDMAN, D. J. (1993). Prediction Using Regression Models with Multicollinear Predictor Variables. *IIE Transactions*, 25(3), 73–85. <https://doi.org/10.1080/07408179308964293>
- Moore, H. E., Poet, S. E., & Martell, E. A. (1973). ^{222}Rn , ^{210}Pb , ^{210}Bi , and ^{210}Po profiles and aerosol residence times versus altitude. *Journal of Geophysical Research (1896-1977)*, 78(30), 7065–7075. <https://doi.org/10.1029/JC078i030p07065>
- Morrison, N. (1994). *Introduction to Fourier Analysis*. John Wiley & Sons.
- Naskar, A. K., Akhter, J., Gazi, M., Mondal, M., & Deb, A. (2023). Impact of meteorological parameters on soil radon at Kolkata, India: Investigation using machine learning techniques. *Environmental Science and Pollution Research*, 30(48), 105374–105386. <https://doi.org/10.1007/s11356-023-29769-y>
- Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics*, 7. <https://doi.org/10.3389/fnbot.2013.00021>
- Nazaroff, W. W. (1992). Radon transport from soil to air. *Reviews of geophysics*, 30(2), 137-160.
- Nielson, K. K., Rogers, V. C., Rogers, V., & Holt, R. B. (1994). The Raetrad Model of Radon Generation and Transport from Soils Into Slab-on-grade Houses. *Health Physics*, 67(4), 363.
- Nikolaev, V. A., & Ilić, R. (1999). Etched track radiometers in radon measurements: a review. *Radiation measurements*, 30(1), 1-13.

- Nunes, L. J. R., Curado, A., & Lopes, S. I. (2023). The Relationship between Radon and Geology: Sources, Transport and Indoor Accumulation. *Applied Sciences*, 13(13), Article 13. <https://doi.org/10.3390/app13137460>
- O'Brien, R. M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity*, 41(5), 673–690. <https://doi.org/10.1007/s11135-006-9018-6>
- Oikawa, S., Kanno, N., Sanada, T., Ohashi, N., Uesugi, M., Sato, K., ... & Higuchi, H. (2003). A nationwide survey of outdoor radon concentration in Japan. *Journal of environmental radioactivity*, 65(2), 203-213.
- Omori, Y., & Nagahama, H. (2016). Radon as an Indicator of Nocturnal Atmospheric Stability: A Simplified Theoretical Approach. *Boundary-Layer Meteorology*, 158(2), 351–359. <https://doi.org/10.1007/s10546-015-0089-6>
- Origin(Pro), Version Number (e.g. "Version 2021"). OriginLab Corporation, Northampton, MA, USA.
- Özen, S. A., Celik, N., Dursun, E., & Taskın, H. (2018). Indoor and outdoor radon measurements at lung cancer patients' homes in the dwellings of Rize Province in Turkey. *Environmental geochemistry and health*, 40, 1111-1125. 78.
- Pal, S., Lopez, M., Schmidt, M., Ramonet, M., Gibert, F., Xueref-Remy, I., & Ciais, P. (2015). Investigation of the atmospheric boundary layer depth variability and its impact on the ²²²Rn concentration at a rural site in France. *Journal of Geophysical Research: Atmospheres*, 120(2), 623–643. <https://doi.org/10.1002/2014JD022322>
- Pan, L., Xu, J., Tie, X., Mao, X., Gao, W., & Chang, L. (2019). Long-term measurements of planetary boundary layer height and interactions with PM_{2.5} in Shanghai, China. *Atmospheric Pollution Research*, 10(3), 989-996.
- Pant, P., Kandari, T., Prasad, M., & Ramola, R. C. (2016). A COMPARATIVE STUDY OF DIURNAL VARIATION OF RADON AND THORON CONCENTRATIONS IN INDOOR ENVIRONMENT. *Radiation Protection Dosimetry*, 171(2), 212–216. <https://doi.org/10.1093/rpd/ncw061>
- Papastefanou, C. (2009). Radon Decay Product Aerosols in Ambient Air. *Aerosol and Air Quality Research*, 9(4), 385–393. <https://doi.org/10.4209/aaqr.2009.02.0011>
- Part, N. G. (2011). Radiation protection and safety of radiation sources International Basic Safety Standards.
- Perrino, C., Pietrodangelo, A., & Febo, A. (2001). An atmospheric stability index based on radon progeny measurements for the evaluation of primary urban pollution. *Atmospheric Environment*, 35(31), 5235–5244. [https://doi.org/10.1016/S1352-2310\(01\)00349-1](https://doi.org/10.1016/S1352-2310(01)00349-1)
- Petermann, E., Bossew, P., Kemske, J., Gruber, V., Suhr, N., & Hoffmann, B. (2024). Development of a High-Resolution Indoor Radon Map Using a New Machine Learning-Based Probabilistic Model and German Radon Survey Data.

- Petermann, E., Meyer, H., Nussbaum, M., & Bossew, P. (2021). Mapping the geogenic radon potential for Germany by machine learning. *Science of The Total Environment*, 754, 142291. <https://doi.org/10.1016/j.scitotenv.2020.142291>
- Phong Thu, H. N., Van Thang, N., & Hao, L. C. (2020). The effects of some soil characteristics on radon emanation and diffusion. *Journal of Environmental Radioactivity*, 216, 106189. <https://doi.org/10.1016/j.jenvrad.2020.106189>
- Poet, S. E., Moore, H. E., & Martell, E. A. (1972). Lead 210, bismuth 210, and polonium 210 in the atmosphere: Accurate ratio measurement and application to aerosol residence time determination. *Journal of Geophysical Research (1896-1977)*, 77(33), 6515–6527. <https://doi.org/10.1029/JC077i033p06515>
- Porstendörfer, J. (1994). Properties and behaviour of radon and thoron and their decay products in the air. *Journal of Aerosol Science*, 25(2), 219–263. [https://doi.org/10.1016/0021-8502\(94\)90077-9](https://doi.org/10.1016/0021-8502(94)90077-9)
- Porstendörfer, J., Zock, C., & Reineking, A. (2000). Aerosol size distribution of the radon progeny in outdoor air. *Journal of Environmental Radioactivity*, 51(1), 37–48. [https://doi.org/10.1016/S0265-931X\(00\)00043-6](https://doi.org/10.1016/S0265-931X(00)00043-6)
- Povinec, P., Chudý, M., Sýkora, I., Szarka, J., Pikna, M., & Holý, K. (1988). Aerosol radioactivity monitoring in Bratislava following the Chernobyl accident. *Journal of Radioanalytical and Nuclear Chemistry Letters*, 126(6), 467–478. Scopus. <https://doi.org/10.1007/BF02164550>
- Prichard, H. M., & Gesell, T. F. (1984). Radon in the Environment. In J. T. Lett (Ed.), *Advances in Radiation Biology* (Vol. 11, pp. 391–428). Elsevier. <https://doi.org/10.1016/B978-0-12-035411-5.50014-9>
- Putman, J. L. (1962). An expression for source counting rate required in tracer experiments. *Intern. J. Appl. Radiation and Isotopes*, 13.
- R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Radiation, U. N. S. C. on the E. of A. (2000). Sources and Effects of Ionizing Radiation, United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR) 2000 Report, Volume I: Report to the General Assembly, with Scientific Annexes - Sources. United Nations. <https://doi.org/10.18356/49c437f9-en>
- Radulescu, I., Calin, M. R., Luca, A., Röttger, A., Grossi, C., Done, L., & Ioan, M. R. (2022). Inter-comparison of commercial continuous radon monitors responses. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1021, 165927.
- Rastogi, N., & Sarin, M. M. (2013). Temporal variability in residence time of ambient aerosols using environmental 210 Pb. *Current Science*, 105(8), 1165–1168.

- Rezaie, F., Panahi, M., Bateni, S. M., Kim, S., Lee, J., Lee, J., Yoo, J., Kim, H., Won Kim, S., & Lee, S. (2023). Spatial modeling of geogenic indoor radon distribution in Chungcheongnam-do, South Korea using enhanced machine learning algorithms. *Environment International*, 171, 107724. <https://doi.org/10.1016/j.envint.2022.107724>
- Rezaie, F., Panahi, M., Lee, J., Lee, J., Kim, S., Yoo, J., & Lee, S. (2022). Radon potential mapping in Jangsu-gun, South Korea using probabilistic and deep learning algorithms. *Environmental Pollution*, 292, 118385. <https://doi.org/10.1016/j.envpol.2021.118385>
- Sabbarese, C., Feola, M. L., Ambrosino, F., Roca, V., D'Onofrio, A., La Verde, G., D'Avino, V., Pugliese, M., & Festa, V. (2022). A Preliminary Study of the Characteristics of Radon Data from Indoor Environments and Building Materials in the Campania Region Using PCA and K-Means Statistical Analyses. *Environments*, 9(7), Article 7. <https://doi.org/10.3390/environments9070082>
- Sakoda, A., Ishimori, Y., & Yamaoka, K. (2011). A comprehensive review of radon emanation measurements for mineral, rock, soil, mill tailing and fly ash. *Applied Radiation and Isotopes*, 69(10), 1422–1435. <https://doi.org/10.1016/j.apradiso.2011.06.009>
- Salzano, R., Pasini, A., Casasanta, G., Cacciani, M., & Perrino, C. (2016). Quantitative Interpretation of Air Radon Progeny Fluctuations in Terms of Stability Conditions in the Atmospheric Boundary Layer. *Boundary-Layer Meteorology*, 160(3), 529–550. <https://doi.org/10.1007/s10546-016-0149-6>
- Schery, S. D., & Wasiolek, P. T. (1993). A two-particle-size model and measurements of radon progeny near the Earth's surface. *Journal of Geophysical Research: Atmospheres*, 98(D12), 22915–22923. <https://doi.org/10.1029/93JD02294>
- Schmidt, M., Glatzel-Mattheier, H., Sartorius, H., Worthy, D. E., & Levin, I. (2001). Western European N₂O emissions: A top-down approach based on atmospheric observations. *Journal of Geophysical Research: Atmospheres*, 106(D6), 5507–5516.
- Schumann, R. R., & Gundersen, L. C. S. (1996). Geologic and climatic controls on the radon emanation coefficient. *Environment International*, 22, 439–446. [https://doi.org/10.1016/S0160-4120\(96\)00144-4](https://doi.org/10.1016/S0160-4120(96)00144-4)
- Seibert, P., Beyrich, F., Gryning, S.-E., Joffre, S., Rasmussen, A., & Tercier, P. (2000). Review and intercomparison of operational methods for the determination of the mixing height. *Atmospheric Environment*, 34(7), 1001–1027. [https://doi.org/10.1016/S1352-2310\(99\)00349-0](https://doi.org/10.1016/S1352-2310(99)00349-0)
- Seinfeld, J. H., & Pandis, S. N. (2016). *Atmospheric chemistry and physics: from air pollution to climate change*. John Wiley & Sons.
- Sesana, L., Caprioli, E., & Marazzan, G. M. (2003). Long period study of outdoor radon concentration in Milan and correlation between its temporal variations and dispersion

- properties of atmosphere. *Journal of Environmental Radioactivity*, 65(2), 147–160. [https://doi.org/10.1016/S0265-931X\(02\)00093-0](https://doi.org/10.1016/S0265-931X(02)00093-0)
- Sheets, R. W., & Lawrence, A. E. (1999). Comparative temporal behavior of radon- and thoron-progeny in surface air over the midwestern U.S. *Journal of Radioanalytical and Nuclear Chemistry*, 242(3), 761–767. <https://doi.org/10.1007/BF02347391>
- Siino, M., Scudero, S., & D'Alessandro, A. (2020). Stochastic Models for Radon Daily Time Series: Seasonality, Stationarity, and Long-Range Dependence Detection. *Frontiers in Earth Science*, 8. <https://doi.org/10.3389/feart.2020.575001>
- Sinclair, V. A., Ritvanen, J., Urbancic, G., Erner, I., Batrak, Y., Moisseev, D., & Kurppa, M. (2022). Boundary-layer height and surface stability at Hyytiälä, Finland, in ERA5 and observations. *Atmospheric Measurement Techniques*, 15(10), 3075–3103. <https://doi.org/10.5194/amt-15-3075-2022>
- Singh MA, Ramola RC, Singh SU, Virk HS. (1988) The influence of meteorological parameters on soil gas radon. *Journal of Association of Exploration Geophysicists*.;9(2):85-90.
- Singh, B., Singh, S., Bajwa, B. S., Singh, J., & Kumar, A. (2011). Soil gas radon analysis in some areas of Northern Punjab, India. *Environmental monitoring and assessment*, 174, 209-217.
- Smetanová, I., Holý, K., Müllerová, M., & Polášková, A. (2010). The effect of meteorological parameters on radon concentration in borehole air and water. *Journal of radioanalytical and nuclear chemistry*, 283(1), 101-109.
- Spearman, C. (1961). The Proof and Measurement of Association Between Two Things (p. 58). Appleton-Century-Crofts. <https://doi.org/10.1037/11491-005>
- STANYŠ, T. Meranie objemových aktivít produktov premeny radónu a torónu [diplomová práca]. 1997
- Stull, R. B. (2012). *An Introduction to Boundary Layer Meteorology*. Springer Science & Business Media.
- Sukanya, S., Noble, J., & Joseph, S. (2022). Application of radon (^{222}Rn) as an environmental tracer in hydrogeological and geological investigations: An overview. *Chemosphere*, 303, 135141. <https://doi.org/10.1016/j.chemosphere.2022.135141>
- Sultani, M. A., Bulko, M., Holý, K., Müllerová, M., Masarik, J., Tonhauzer, P., & Helejš, M. (2023). The use of radon as a tracer for air quality assessment: A case study in Bratislava, Slovakia. *Journal of Radioanalytical and Nuclear Chemistry*. <https://doi.org/10.1007/s10967-023-08969-3>
- Sultani, M. A., Bulko, M., Holý, K., Müllerová, M., Masarik, J., Tonhauzer, P., & Helejš, M. (2024). The use of radon as a tracer for air quality assessment: A case study in Bratislava, Slovakia. *Journal of Radioanalytical and Nuclear Chemistry*, 333(5), 2515–2527. <https://doi.org/10.1007/s10967-023-08969-3>

- Sýkora, I., Holý, K., Ješkovský, M., Müllerová, M., Bulko, M., & Povinec, P. P. (2017). Long-term variations of radionuclides in the Bratislava air. *Journal of Environmental Radioactivity*, 166, 27–35. <https://doi.org/10.1016/j.jenvrad.2016.03.004>
- Tchorz-Trzeciakiewicz, D. E., & Solecki, A. T. (2018). Variations of radon concentration in the atmosphere. *Gamma dose rate. Atmospheric Environment*, 174, 54-65.
- Tirmarche, M., Harrison, J. D., Laurier, D., Paquet, F., Blanchardon, E., Marsh, J. W., & International Commission on Radiological Protection. (2010). ICRP Publication 115. Lung cancer risk from radon and progeny and statement on radon. *Annals of the ICRP*, 40(1), 1–64. <https://doi.org/10.1016/j.icrp.2011.08.011>
- Tokonami, S. (2000). Experimental Verification of the Attachment Theory of Radon Progeny onto Ambient Aerosols. *Health Physics*, 78(1), 74.
- Turekian, K. Y., Nozaki, Y., & Benninger, L. K. (1977). Geochemistry of atmospheric radon and radon products. *Annual Review of Earth and Planetary Sciences*, Vol. 5, p. 227, 5, 227.
- UNSCEAR 2000 Report Volume I. (n.d.). United Nations : Scientific Committee on the Effects of Atomic Radiation. Retrieved 10 May 2024, from https://www.unscear.org/unscear/en/publications/2000_1.html
- Unsear, S. (2000). effects of Ionizing Radiation. United Nations, New York, 453-487.
- Urban, M., & Piesch, E. (1981). Low level environmental radon dosimetry with a passive track etch detector device. *Radiation protection dosimetry*, 1(2), 97-109.
- Utkin, V. I., & Yurkov, A. K. (2010). Radon as a tracer of tectonic movements. *Russian Geology and Geophysics*, 51(2), 220–227. <https://doi.org/10.1016/j.rgg.2009.12.022>
- Uurtio, V., Monteiro, J. M., Kandola, J., Shawe-Taylor, J., Fernandez-Reyes, D., & Rousu, J. (2017). A Tutorial on Canonical Correlation Methods (arXiv:1711.02391). arXiv. <https://doi.org/10.48550/arXiv.1711.02391>
- Vaupotič, J., Kobal, I., & Križman, M. J. (2010). Background outdoor radon levels in Slovenia. *Nukleonika*, 55, 579-582.
- Vecchi, R., Piziali, F. A., Valli, G., Favaron, M., & Bernardoni, V. (2019). Radon-based estimates of equivalent mixing layer heights: A long-term assessment. *Atmospheric Environment*, 197, 150–158. <https://doi.org/10.1016/j.atmosenv.2018.10.020>
- Vicanová, M. (2003). Utilisation of solid state nuclear track detectors in the solution of radon problems (Vyuzitie detektorov stôp v pevnej fáze pri riešení radónovej problematiky) (Doctoral dissertation, Ph. D. Thesis, Faculty of Mathematics, Physics and Informatics, Comenius University of Bratislava, 82 p.(in Slovak)).
- Virk, J., & Virk, H. S. (2021). Radon Estimation in Water of Surrey Region of British Columbia, Canada using LR-115 Type II Nuclear Track Detector. *Earth*, 12, 50.

- Vogelezang, D. H. P., & Holtslag, A. A. M. (1996). Evaluation and model impacts of alternative boundary-layer height formulations. *Boundary-Layer Meteorology*, 81(3-4), 245-269.
- Vukovich FM, Scarborough J (2005) Aspects of ozone transport, mixing, and chemistry in the greater Maryland area. *Atmos Environ* 39:7008–7019
- Wada, A., Matsueda, H., Murayama, S., Taguchi, S., Hirao, S., Yamazawa, H., ... & Sawa, Y. (2013). Quantification of emission estimates of CO₂, CH₄ and CO for East Asia derived from atmospheric radon-222 measurements over the western North Pacific. *Tellus B: Chemical and Physical Meteorology*, 65(1), 18037.
- Wagner, P., & Schäfer, K. (2017). Influence of mixing layer height on air pollutant concentrations in an urban street canyon. *Urban Climate*, 22, 64-79.
- Ward, D. C., & Borak, T. B. (1991). Determination of time-varying ²²²Rn concentrations using flow-through scintillation flasks. *Health physics*, 61(6), 799-807.
- Wasikiewicz, J. M. (2018). Impact of environmental factors on PADC radon detector sensitivity during long term storage. *Radiation Physics and Chemistry*, 142, 141-145.
- Wilkening, M. (1990). *Radon in the Environment*. Elsevier.
- Williams, A. G., Zahorowski, W., Chambers, S., Griffiths, A., Hacker, J. M., Element, A., & Werczynski, S. (2011). The Vertical Distribution of Radon in Clear and Cloudy Daytime Terrestrial Boundary Layers. *Journal of the Atmospheric Sciences*, 68(1), 155–174. <https://doi.org/10.1175/2010JAS3576.1>
- Winkler, R., Ruckerbauer, F., Trautmannsheimer, M., Tschiersch, J., & Karg, E. (2001). Diurnal and seasonal variation of the equilibrium state between short-lived radon decay products and radon gas in ground-level air. *Radiation and Environmental Biophysics*, 40(2), 115–123. <https://doi.org/10.1007/s004110100096>
- Woith, H. (2015). Radon earthquake precursor: A short review. *The European Physical Journal Special Topics*, 224(4), 611–627. <https://doi.org/10.1140/epjst/e2015-02395-9>
- Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R*, Second Edition (2nd ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781315370279>
- World Health Organization. (2009). WHO handbook on indoor radon: A public health perspective. WHO Press. <https://www.who.int/publications/i/item/9789241547673>
- Wu, Z., Liu, G., Liu, Y., Ding, M., Su, L., & Feng, J. (2023). Seasonal residence time variations for ²¹⁰Po/²¹⁰Pb aerosol particles in Shenzhen, China. *Journal of Radioanalytical and Nuclear Chemistry*, 332(5), 1459–1467. <https://doi.org/10.1007/s10967-022-08576-8>
- Xu, R., & Wunsch, D. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16(3), 645–678. <https://doi.org/10.1109/TNN.2005.845141>

- Yang, Z., Prox, L., Meernik, C., Raveendran, Y., Press, D. J., Gibson, P., Koch, A., Ajumobi, O., Chen, R., Zhang, J. (Jim), & Akinyemiju, T. (2025). Identifying predictors of spatiotemporal variations in residential radon concentrations across North Carolina using machine learning analytics. *Environmental Pollution*, 367, 125592. <https://doi.org/10.1016/j.envpol.2024.125592>
- Yu, C., Sun, Y., & Wang, N. (2023). Outdoor Radon and Its Progeny in Relation to the Particulate Matter during Different Polluted Weather in Beijing. *Atmosphere*, 14(7), Article 7. <https://doi.org/10.3390/atmos14071132>
- Zahorowski, W., Chambers, S. D., & Henderson-Sellers, A. (2004). Ground based radon-222 observations and their application to atmospheric studies. *Journal of environmental radioactivity*, 76(1-2), 3-33.
- Zhang, J., & Rao, S. T. (1999). The role of vertical mixing in the temporal evolution of ground-level ozone concentrations. *Journal of Applied Meteorology*, 38(12), 1674-1691.
- Zhang, Y., Sun, K., Gao, Z., Pan, Z., Shook, M. A., & Li, D. (2020). Diurnal climatology of planetary boundary layer height over the contiguous United States derived from AMDAR and reanalysis data. *Journal of Geophysical Research: Atmospheres*, 125(20), e2020JD032803.
- Zimnoch, M., Wach, P., Chmura, L., Gorczyca, Z., Rozanski, K., Godłowska, J., Mazur, J., Kozak, K., & Jeričević, A. (2014). Factors controlling temporal variability of near-ground atmospheric ²²²Rn concentration over central Europe. *Atmospheric Chemistry and Physics*, 14(18), 9567–9581. <https://doi.org/10.5194/acp-14-9567-2014>
- Zmazek, B., Todorovski, L., Džeroski, S., Vaupotič, J., & Kobal, I. (2003). Application of decision trees to the analysis of soil radon data for earthquake prediction. *Applied Radiation and Isotopes*, 58(6), 697–706. [https://doi.org/10.1016/S0969-8043\(03\)00094-0](https://doi.org/10.1016/S0969-8043(03)00094-0)
- Zoran, M. A., Dida, M. R., Zoran, A., Zoran, L. F., & Dida, A. (2013). Outdoor ²²²Radon concentrations monitoring in relation with particulate matter levels and possible health effects. *Journal of Radioanalytical and Nuclear Chemistry*, 296(3), 1179–1192. <https://doi.org/10.1007/s10967-012-2259-z>
- Zoran, M. A., Savastru, R. S., Savastru, D. M., & Penache, M. C. V. (2019). Temporal trends of carbon monoxide (CO) and radon (²²²Rn) tracers of urban air pollution. *Journal of Radioanalytical and Nuclear Chemistry*, 320(1), 55-70.
- Zoran, M., Savastru, R., & Savastru, D. (2012). Ground based radon (²²²Rn) observations in Bucharest, Romania and their application to geophysics. *Journal of Radioanalytical and Nuclear Chemistry*, 293(3), 877-888.
- Žunic, Z. S., Yarmoshenko, I. V., Birovljev, A., Bochicchio, F., Quarto, M., Obryk, B., ... & Waligorski, M. P. R. (2007). Radon survey in the high natural radiation region of Niška Banja, Serbia. *Journal of environmental radioactivity*, 92(3), 165-174.