Frontiers in Environmental Research
Print ISSN: 2960-0251
Online ISSN: 2960-026X
DOI: https://doi.org/10.61784/fer3050

INTELLIGENT EARLY-WARNING AND DECISION-MAKING FOR COALFIELD FIRE RISK BASED ON RANDOM FOREST: A CASE STUDY OF ABANDONED GOAFS IN NORTHERN CHINA

Rui Sun, ZiXue Wang*
Ningxia Coal Exploration Engineering Co., Ltd., Yinchuan 750000, Ningxia, China.
**Corresponding Author: ZiXue Wang*

Abstract: Coalfield fire risk identification is challenged by the strong concealment of goafs, the development of surface fissures, the unclear distribution of residual coal in abandoned workings, and uncertainties associated with multi-source detection data. Conventional approaches based on a single temperature threshold or expert empirical judgement are therefore insufficient for accurate early warning. To address field problems observed in an abandoned goaf of a coal mine in northern China, including temperature anomalies, abnormal CO concentrations at small-mine openings, smoke emissions from surface fissures, and the potential hydraulic or fracture connection between upper and lower goafs, this study develops an intelligent coalfield fire risk early-warning model based on the random forest algorithm. Thermal infrared imaging, isotopic radon measurements, and geological-goaf spatial information were used as the primary data sources. Surface temperature anomalies were extracted from UAV-based thermal infrared images, gas-migration signals along underground fracture pathways were obtained using alpha-cup radon measurements, and spatial-distance features to Jurassic abandoned-goaf boundaries and small-mine openings were quantified. Using expert interpretation results and field-observed anomaly points as labels, a random forest classification model was established to quantitatively classify coalfield fire risk levels in the study area. The results show that thermal anomaly features directly represent surface temperature responses, radon anomaly features effectively reveal concealed fractures and gas-migration pathways, and geological and goaf-related spatial features characterize the controlling conditions for fire development. Fusion of the three types of features substantially improves the stability and interpretability of coalfield fire risk identification. The proposed model provides a scientific basis for delineating temperature anomalies in abandoned goafs, optimizing verification borehole layout, zoning fire prevention and control measures, and ensuring safe longwall panel retreat.
Keywords: Coalfield fire; Radon measurement; Thermal infrared imaging; Random forest; Risk early warning

1 INTRODUCTION

Coalfield fires and spontaneous combustion of residual coal in goafs are typical concealed hazards in coal mine safety management. Under continuous oxygen supply, residual coal in abandoned goafs undergoes low-temperature oxidation and gradually accumulates heat. When heat generation exceeds heat dissipation, high-temperature anomalous zones or even active fire zones may form. Such hazards are characterized by strong concealment, unclear spatial boundaries, slow development, and significant control by geological structures and mining-induced disturbances. Once a high-temperature abandoned goaf becomes connected with production goafs, fracture zones above working faces, or surface fissures, large amounts of greenhouse gases, including CO_2 and CH_4, and toxic or harmful gases, such as CO, SO_2, and H_2S, may be released. This can further induce land degradation, vegetation damage, surface subsidence, and groundwater contamination. Therefore, accurate and efficient delineation of coalfield fire zones and high-temperature anomalous areas is of both theoretical significance and practical engineering value for fire-control design, extinguishing engineering layout, and coal-resource conservation.

Detection of coalfield fires and spontaneous combustion in goafs is a highly complex problem involving coupled multiphysical processes. The main difficulties arise from hazard concealment, low signal-to-noise ratios, and spatial uncertainty. Existing methods include infrared thermometry, magnetic exploration, geological investigation, borehole verification, and radon measurement, each of which has its own physical basis and applicability. For infrared thermometry, underground combustion points often generate weak and diffuse surface thermal signals, which are susceptible to atmospheric interference, vegetation cover, soil-moisture variations, and thermal diffusion attenuation. As a result, the spatial resolution and effective detection depth are limited [1-3]. Although UAV-based thermal infrared remote sensing can markedly improve near-surface spatial resolution, its detection capability is mainly restricted to shallow or middle-to-late-stage fire zones [4]. Magnetic exploration can rapidly delineate approximate fire zones by exploiting high-temperature-induced changes in magnetic minerals [5]. However, this method is nearly ineffective for low-temperature oxidation zones below the Curie point and is strongly affected by surface ferromagnetic interference and primary variations in magnetic minerals, leading to non-unique interpretations. Geological investigation provides regional constraints, whereas borehole verification remains the only gold-standard method for directly obtaining in situ temperature, gas composition, and combustion-state information within a fire zone [6]. Nevertheless, boreholes provide point-scale information, are expensive and time-consuming, and have limited spatial coverage, making them unsuitable

for large-scale dynamic monitoring. They are generally used only for final verification in key areas or for precise deployment of fire-extinguishing projects [7].

Radon measurement provides an alternative approach by utilizing the enhanced migration of radon (Rn), a decay product in the uranium series, during spontaneous combustion. When low-temperature oxidation occurs in coal, changes in pore structure and permeability facilitate upward migration of radon from deep strata to the surface through fractures, thereby producing radon-concentration anomalies. This method is sensitive to deep oxidation zones and has been successfully applied to fire-source localization in abandoned small-mine goafs [8-9]. Owing to its advantages of wide-area coverage, low cost, and sensitivity to deep anomalies, radon measurement is one of the preferred approaches for detecting coalfield fire zones. However, its engineering application is still constrained by lagging interpretation methods and insufficient localization accuracy. Traditional interpretation commonly relies on empirical thresholds based on fixed multiples of background values, ignoring the spatial autocorrelation and local heterogeneity of radon data, which can increase misclassification [10]. In recent years, machine learning has made substantial progress in geochemical anomaly recognition, mineral-resource prediction, and environmental monitoring. Common anomaly-detection models include DBSCAN clustering, local outlier factor, isolation forest, and kriging-residual methods [11-14]. These methods can incorporate spatial features and overcome the limitations of traditional thresholding, providing a technical foundation for intelligent interpretation of radon measurement data. However, systematic research and engineering validation of machine learning methods for radon-based coalfield fire detection remain limited.

To address these issues, this study takes a typical coalfield fire area in northern China as the engineering background and conducts machine-learning-based anomaly identification and risk assessment using large-scale in situ radon measurements. A two-stage sampling strategy, including coarse measurements on a 50 m grid and refined measurements on a 10 m grid in key areas, was used to obtain a complete radon dataset over an approximately 2.4 km² fire-prone area. On this basis, UAV thermal infrared thermometry and goaf spatial-relationship analysis were integrated to construct an anomaly-recognition model based on the random forest algorithm. The resulting intelligent early-warning model was used to grade anomalous zones, forming a three-source fusion method that integrates thermal anomalies, radon anomalies, and spatial features. The proposed method enables quantitative classification of fire risk levels in the study area and provides decision support for abnormal-zone identification, fire-range delineation, verification borehole placement, prevention and control zoning, and safe working-face retreat.

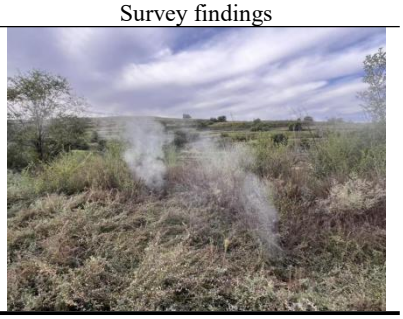
2 DATA COLLECTION AND MODEL BUILDING

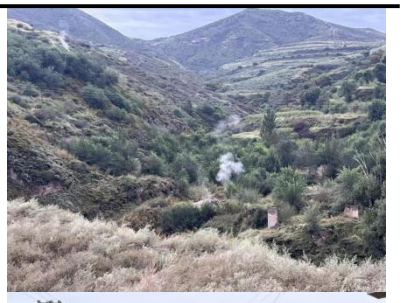
2.1 Engineering Background

The study area is located in a large coal-production base in northern China and belongs to a typical Carboniferous-Permian coalfield characterized by thick coal seams, shallow burial depth, and complex overlying strata. The terrain is dominated by loess hills and gullies, with pronounced relief and alternating hills and valleys, representing a typical hilly landform of the Loess Plateau. Historically, years of disorderly small-mine extraction produced numerous abandoned goafs with complex geometries and poorly constrained boundaries. Under the combined effects of mining disturbance and long-term strata pressure, fracture systems developed extensively in the study area, providing oxygen-supply pathways and heat-storage spaces for residual-coal spontaneous combustion while also facilitating upward radon migration.

In recent years, as fully mechanized caving mining progressed in the lower coal seams, residual coal in the overlying abandoned goafs underwent low-temperature oxidation under continuous oxygen supply, gradually forming high-temperature anomalous zones. Field investigations identified several typical indicators of coalfield fire activity (Table 1), suggesting that the area has entered an early-to-middle development stage of coalfield fire and urgently requires systematic detection and risk assessment

Table 1 Ground Investigation Observations in the Study Area

Problem	Phenomenon	Survey findings
Temperature anomaly	Distinct ground-temperature anomalies occurred inside and around abandoned goafs; local borehole temperatures ranged from 45 to 156 °C.	

CO anomaly	Abnormal CO concentrations were detected at several small-mine openings, with maximum values reaching several hundred ppm.	
Smoke emission from surface fissures	White smoke and a sulfur odor were observed at some surface fissures.	
Potential connection risk	Fracture-zone connection may exist between overlying abandoned goafs and underlying production goafs.	

2.2 Data Measurement

To accurately arrange measurement points and collect detection information, an initial field survey was conducted over the entire target area. The survey focused on obvious smoke-emission points and fissure zones and compiled the local topographic distribution (Figure 1). The northwestern and southwestern parts of the area are mainly cultivated land, whereas the northeastern part is dominated by hillslopes. The exploration area covers approximately 2.4 km², with coordinate ranges of X = 415800-418000 and Y = 4423000-4427000. The terrain is typical loess hill-gully topography, with a relative elevation difference of approximately 80-120 m. Owing to terrain constraints, some steep slopes and subsidence areas could not be instrumented; the effective survey area was therefore approximately 2.2 km².

A multi-source data fusion strategy was adopted. Thermal infrared imaging, radon measurement, and spatial information acquisition were combined to characterize surface thermal responses, deep oxidation signals, and geological constraints. The main field instruments included a thermal imager, an alpha-cup radon detector, and a global positioning system.

Figure 1 Distribution of Survey Points in the Exploration Area
Source: http://bzdt.ch.mnr.gov.cn

2.2.1 Infrared thermometry

UAV-mounted WIRIS Pro Sc thermal imaging equipment was used to acquire thermal infrared data (Figure 2). The flight altitude was set to 300 m, corresponding to a ground sampling resolution of 0.65 m for the WIRIS Pro Sc sensor. The forward and side overlaps were set to 80% and 75%, respectively, and the thermal image acquisition interval was 2 s. The UAV takeoff and landing point was located on a nearby road with flat ground and no surrounding obstructions, providing suitable flight-operation conditions.

Figure 2 Equipment Used for Exploration and Spatial Information Acquisition

2.2.2 Radon measurement

Under comparable geological and stratigraphic conditions, differences in radon concentration can be used to infer the relationship between underground radon distribution and spontaneous combustion, thereby constraining the approximate extent of coal-mine fire zones. When underground coal seams undergo oxidative heating or spontaneous combustion, radon exhalation from surrounding and overlying rocks increases. Through ion-exchange processes associated with radon decay, this signal can be transferred to the surface and form a radioactive anomaly, which can be detected as an indirect indicator of temperature-related processes (Figure 3).

Alpha cups were buried at pre-designed measurement points. At each point, a 30-40 cm pit was excavated, the cup was placed in the pit, and the top was covered with plastic film and soil. After 4 h of burial, the cup was retrieved and measured using an HS04 alpha-cup detector for 3 min. The readings and corresponding lithology, time, and other parameters were recorded. The radon survey covered approximately 2.4 km². Measurement points were distributed across the survey area, and a 50 m x 50 m grid was used during the coarse-measurement stage. Because the area is mountainous and includes steep peaks, deep gullies, legacy mine facilities, former residential buildings, and subsidence zones, some measurement points were reasonably shifted or omitted to ensure field safety.

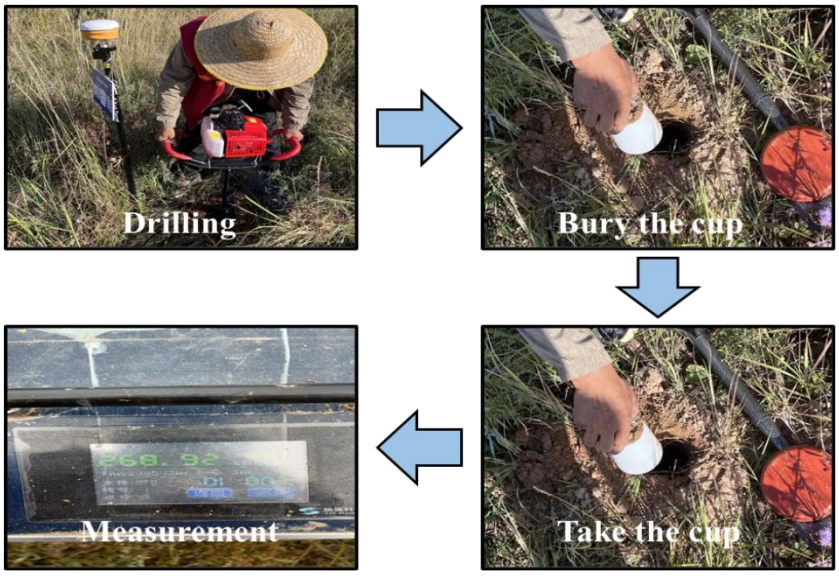


Figure 3 Field Radon-Measurement Procedure.

2.3 Intelligent Early-Warning Model

To identify potential coalfield fire zones and classify risk levels, a random forest-based anomaly-recognition model was developed using thermal infrared measurements, coarse radon measurements, and goaf spatial-relationship analysis. Considering that fire risk in the study area is jointly controlled by underground oxidative heating of coal, gas escape through fracture pathways, and spatial proximity to overlying Jurassic abandoned goafs and small-mine openings, four input features were selected: surface temperature, radon concentration, distance to the Jurassic abandoned-goaf boundary, and distance to the nearest small-mine opening. Surface temperature represents near-surface thermal anomaly responses; radon concentration reflects fracture development and abnormal gas escape from goafs; distance to the Jurassic abandoned-goaf boundary represents the spatial influence of the overlying goaf; and distance to small-mine openings characterizes the influence of legacy mining pathways and local gas emission on anomaly development.

Random forest is an ensemble learning algorithm based on decision trees. It generates multiple sub-training sets through bootstrap sampling, trains multiple decision trees, and produces final outputs by majority voting for classification or averaging for regression. The model has strong nonlinear fitting capacity and resistance to overfitting, making it suitable for coalfield fire anomaly identification involving multi-factor coupling and complex feature relationships. Hyperparameter optimization is important for improving anomaly-recognition accuracy. In this study, GridSearchCV was used to identify the optimal number and depth of decision trees by maximizing the F1 score. Radon concentration, surface temperature, distance to the Jurassic abandoned-goaf boundary, and distance to small-mine openings were used as input variables, whereas anomaly-recognition labels and risk-level labels were used as output variables to establish binary anomaly-recognition and multiclass risk-classification models, respectively. The dataset was divided into training and testing subsets using stratified sampling, with 75% of samples used for training and 25% for testing, ensuring that the proportions of different risk levels remained stable during model development and validation.

3 RESULTS AND ANOMALY ANALYSIS

3.1 Measurement Results

3.1.1 Infrared thermometry results

UAV thermal infrared and multispectral images were processed using Pix4D Mapper, which can rapidly generate two-dimensional orthophotos and three-dimensional models from large batches of UAV images. According to the conversion relationship from WIRIS Pro Sc digital number values to brightness temperature, a difference of 100 DN corresponds to a temperature difference of 2.5 K. Therefore, the UAV thermal infrared images exhibited substantial temperature drift and required correction through thermal image normalization.

As shown in Figure 4, temperature anomalies were mainly concentrated in the northern part of the minefield. Because buildings, concrete surfaces, and human activities are present in the central northern area, related temperature anomalies were treated as interference and excluded. After eliminating interference from roads, crops, buildings, photovoltaic panels, shrubs, legacy structures, and abandoned coal mines, 14 temperature-anomaly sites were identified. The maximum temperature reached 34 °C, whereas the average background temperature was 10 °C. Based on the infrared imagery, six temperature-anomaly zones were delineated and labelled A1-A6; their areas are listed in Table 2.

Table 2 The Areas of Temperature-Anomaly Zones

Zone	A1	A2	A3	A4	A5	A6
Area(m^2)	7354	29765	1218	503	3864	1135

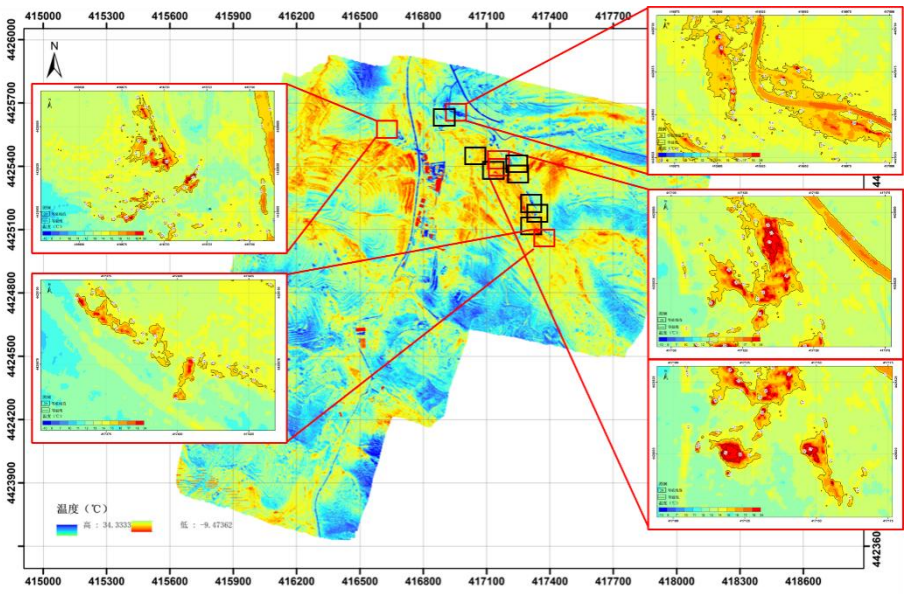


Figure 4 Thermal Map and Delineated Temperature-Anomaly Zones
Source: http://bzdt.ch.mnr.gov.cn

3.1.2 Coarse radon-measurement results

A total of 1001 valid measurement points were obtained during the coarse radon survey. The statistical characteristics of radon concentration are summarized in Table 3. The coarse radon values exhibited a strongly right-skewed distribution, with a skewness of 4.23, and a high kurtosis of 24.60, indicating the presence of pronounced extreme high-value anomalies. The median value (7.40 Bq/m³) was substantially lower than the mean value (16.80 Bq/m³), further confirming the upward influence of high-value anomalies on the mean.

Table 3 Statistical Characteristics of Radon Concentration in Rough Measurement

Statistical indicators	Value
Number of points	1001
Minimum (Bq/m³)	0.00
Maximum (Bq/m³)	217.70
Mean (Bq/m³)	16.80
Median (Bq/m³)	7.40
Standard deviation (Bq/m³)	22.40
Coefficient of variation	1.33
Skewness	4.23
Kurtosis	24.60

After data preprocessing, Surfer 8.0 was used to simulate and analyze the measurement results and to generate plan-view and three-dimensional anomaly maps (Figure 5). The spatial distribution of coarse radon values indicated that high-radon zones occurred as multiple isolated patches rather than as a single continuous zone. Major high-value clusters were distributed in the northeastern, central-western, and southwestern parts of the study area, and approximately 8-10 potential anomalous clusters were preliminarily identified. A clear concentration gradient existed between high-value and background zones, with radon values rapidly decreasing from the center outward, consistent with a point-source-plus-diffusion radon migration pattern. Ten anomalous radon zones, labelled A1-A10, were identified; their areas are listed in Table 4.

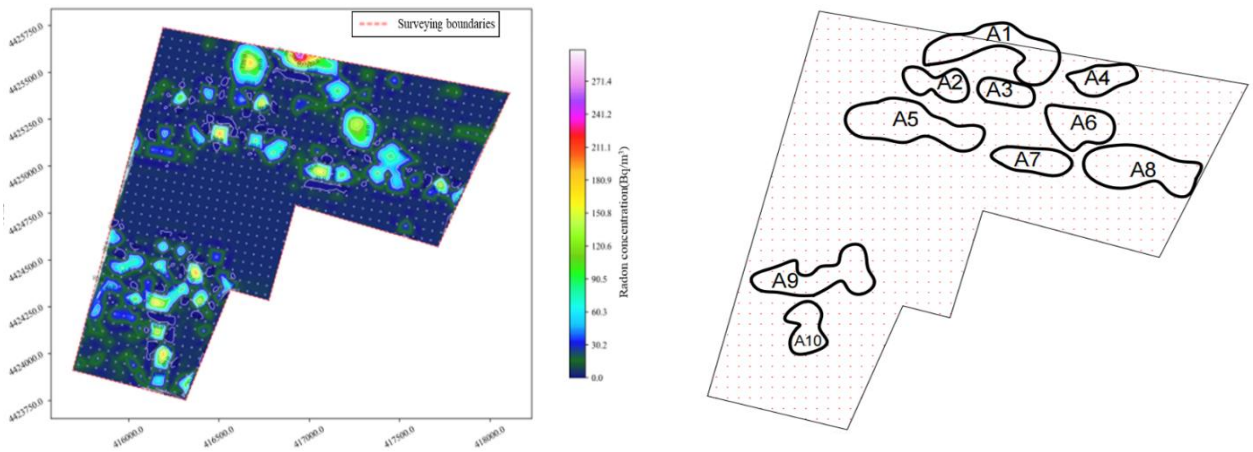


Figure 5 Radon-Measurement Heat Map and Anomalous Zones
Source: http://bzdt.ch.mnr.gov.cn

Table 4 Areas of Radon-Anomaly Zones

Number	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
Area (m^2)	81065	3186	2574	3143	8872	46229	35318	83828	7052	31735

3.2 Risk-Level Early-Warning Results

Based on the random forest anomaly-recognition model, binary anomaly identification and multiclass risk-level prediction were performed, and the model performance was evaluated using the test dataset. The binary anomaly-recognition results are presented in Table 5. The random forest model achieved high overall accuracy on the test set, with an accuracy of 0.992, indicating that it can effectively distinguish normal points from anomalous points. The precision was 0.960, suggesting that most samples predicted as anomalous were truly anomalous and that the false-alarm rate was low. The recall was also 0.960, indicating that the model identified most true anomalous samples and produced few missed detections. The F1 score of 0.960 demonstrates a good balance between precision and recall. The ROC-AUC reached 0.998, close to 1, indicating strong discriminative ability between normal and anomalous samples and high reliability of anomaly-probability outputs.

Table 5 Binary Anomaly-Recognition Performance

Metric	Accuracy	Recall	F1 score
Value	0.992	0.960	0.960

The coalfield fire temperature-anomaly risk values were further output by the model, and a risk heat map was generated using kriging interpolation(Figure 6-7). The fused indicators showed that high-risk areas were mainly distributed in the northeastern part of the study area.

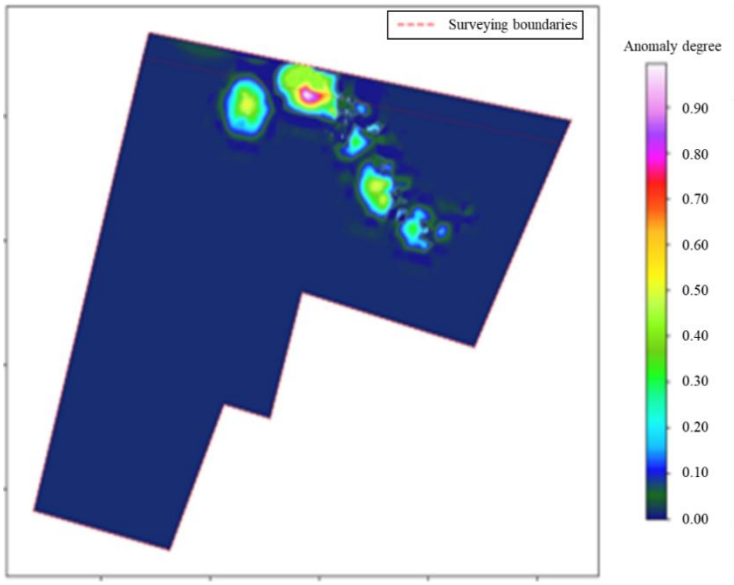


Figure 6 Heat Map of Anomaly-Related Risk
Source: http://bzdt.ch.mnr.gov.cn

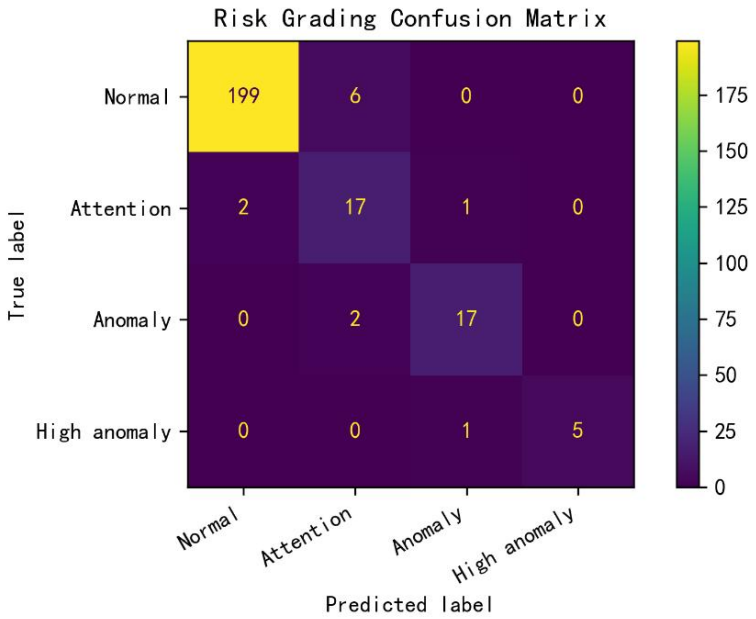


Figure 7 Risk-Level Classification Confusion Matrix

The confusion matrix for risk-level classification indicates that the normal zone was identified with the best performance. Among 205 normal samples, 199 were correctly classified, and only six were misclassified as attention zones; no normal samples were directly misclassified as warning or high-risk zones. This indicates that the model has strong discrimination capability for low-risk background areas. Among 20 attention-zone samples, 17 were correctly identified, two were misclassified as normal zones, and one was misclassified as a warning zone, suggesting a transitional relationship between the attention zone and adjacent classes. Among 19 warning-zone samples, 17 were correctly identified and two were misclassified as attention zones, indicating that the model can effectively identify warning areas with relatively distinct anomaly characteristics, although it may underestimate some marginal warning points with weaker anomaly intensity. Among six high-risk samples, five were correctly identified and one was misclassified as a warning zone; no high-risk samples were misclassified as normal or attention zones. This result indicates that the model provides favorable safety performance for high-risk-zone identification.

To evaluate the consistency and complementarity of thermal infrared thermometry and radon measurement in coalfield fire detection, the anomalous zones identified by radon measurement and the thermal anomalies delineated from thermal infrared imagery were jointly projected onto the orthophoto of the study area for spatial overlay analysis. The results are shown in Figure 8. The overlay results indicate a distinct spatial correspondence between the anomalous zones identified by the two methods. Four overlapping zones were identified and labelled X1, X2, X3, and X4. The basic information for each overlapping zone is summarized in Table 6. Combined with in situ borehole temperature measurements in the anomalous areas, these results further confirm the accuracy of the risk-level early-warning model.

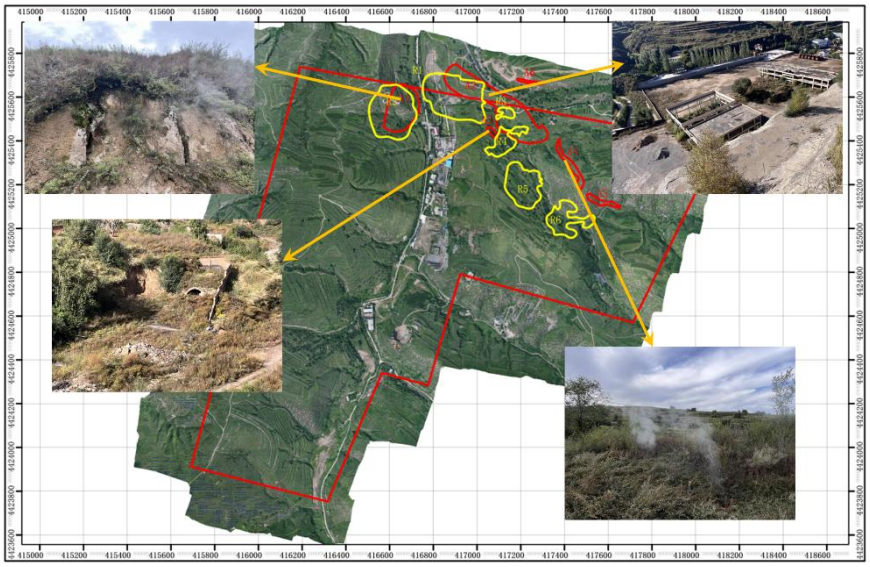


Figure 8 Overlay of Radon Anomalies, Thermal Anomalies, and Field Investigation Results
Source: http://bzdt.ch.mnr.gov.cn

Table 6 Areas of Overlapping Radon and Thermal Anomaly Zones

Zone	X1	X2	X3	X4
Area(m^2)	21869.63	12750.41	2671.81	2093.29

4 CONCLUSIONS

This study developed a random forest early-warning model that integrates thermal infrared, radon, and spatial features. Multi-source feature fusion significantly improved recognition accuracy, with an accuracy of 0.992 and an F1 score of 0.960, demonstrating strong classification capability. Based on thermal infrared thermometry, coarse radon measurement, and goaf spatial-relationship analysis, the random forest model automatically identified high-risk areas requiring refined measurement. Spatial overlay analysis confirmed significant coupling between thermal anomalies and high-radon zones in densely distributed abandoned goafs, validating the heat-gas synergistic mechanism. The model achieved refined classification of risk levels and performed well in identifying high-risk and warning zones, with no severe missed detections, further demonstrating its value for early-stage hazard identification. Although the model accurately distinguished low-risk and anomalous areas, uncertainty remained in marginal zones, particularly in the transitional boundary between attention and warning zones. Future studies should incorporate time-series monitoring data and feedback from field treatment measures to develop dynamically updated coalfield fire risk early-warning models that can adapt to working-face advance and on-site mitigation progress.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Nath S, Agarwal S, Pandey G N. Evaluation of Knowledge Gaps in Mathematical Applications of Thermal Image Processing Techniques for Fire Prevention. ACM Computing Surveys, 2017, 50(1): 1–43. DOI: 10.1145/3009967.

[2] Chen Y, Liao S, Qin D.Study on inversion of coal seam temperature in mining area --Pingshuo mining area of Shanxi Province. E3S Web of Conferences, 2020, 165: 03014. DOI: 10.1051/e3sconf/202016503014.

[3] Yan S, Shi K, Li Y, et al. Integration of satellite remote sensing data in underground coal fire detection: A case study of the Fukang region, Xinjiang, China. Frontiers of Earth Science, 2019, 14(1): 1–12. DOI: 10.1007/s11707-019-0757-9.

[4] He X, Yang X, Luo Z, et al. Application of unmanned aerial vehicle (UAV) thermal infrared remote sensing to identify coal fires in the Huojitu coal mine in Shenmu city, China. Scientific Reports, 2020, 10(1). DOI: 10.1038/s41598-020-70964-5.

[5] Zhang B, Xiao F, Jin W. Burnt coal field detection via magnetic exploration. Environmental Earth Sciences, 2023, 82(7). DOI: 10.1007/s12665-023-10843-0.

[6] Ma Z, Qin B, Shi Q, et al. The location analysis and efficient control of hidden coal spontaneous combustion disaster in coal mine goaf: A case study. Process Safety and Environmental Protection, 2024, 184: 66–78. DOI: 10.1016/j.psep.2024.01.054.

[7] Hu H, Xing Z, Chen M. Application of Surface Drilling Grouting in Fire Extinguishing of Small Mine Goaf. In Safety in Coal Mines, 2019.

[8] Duan S, Fang L, Shi Q, et al. Application of Rapid Identification Technology in Shallow Coal Seam Fire Detection. Combustion Science and Technology, 2023, 197(7): 1535–1549. DOI: 10.1080/00102202.2023.2288218.

[9] Zhou B, Wu J, Wang J, et al. Surface-based radon detection to identify spontaneous combustion areas in small abandoned coal mine gobs: Case study of a small coal mine in China. Process Safety and Environmental Protection, 2018, 119: 223–232. DOI: 10.1016/j.psep.2018.08.011.

[10] Zhou W, Wang J, Zhou C, et al. Technology of Detecting Spontaneous Combustion Source in Goaf of Coal Mine by Isotopic Measurement of Polonium. In Safety in Coal Mines, 2019.

[11] Yu B, She J, Liu G, et al. Coal fire identification and state assessment by integrating multitemporal thermal infrared and InSAR remote sensing data: A case study of Midong District, Urumqi, China. ISPRS Journal of Photogrammetry and Remote Sensing, 2022, 190: 144–164. DOI: 10.1016/j.isprsjprs.2022.06.007.

[12] Karanam V, Motagh M, Garg S, et al. Multi-sensor remote sensing analysis of coal fire induced land subsidence in Jharia Coalfields, Jharkhand, India. International Journal of Applied Earth Observation and Geoinformation, 2021, 102: 102439. DOI: 10.1016/j.jag.2021.102439.

[13] Zou J, Zhang R, Zhou F, et al. Hazardous Area Reconstruction and Law Analysis of Coal Spontaneous Combustion and Gas Coupling Disasters in Goaf Based on DEM-CFD. ACS Omega, 2023, 8(2): 2685–2697. DOI: 10.1021/acsomega.2c07236.

[14] Zheng Y, Li S, Xue S, et al. Study on the evolution characteristics of coal spontaneous combustion and gas coupling disaster region in goaf. Fuel, 2023, 349: 128505. DOI: 10.1016/j.fuel.2023.128505.