Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Risk-oriented interpretable prediction of regional dust exposure in a fully mechanized heading face under long-forcing and short-exhaust ventilation

Anfei Yu
China Railway Beijing Engineering Bureau Group Second Engineering Co., Ltd.
Hu Bai
China Railway Beijing Engineering Bureau Group Second Engineering Co., Ltd.
Hongrui Ren
China Railway Beijing Engineering Bureau Group Second Engineering Co., Ltd.
Fuqiang Liu
China Railway Beijing Engineering Bureau Group Second Engineering Co., Ltd.
Li Chen
China Railway Beijing Engineering Bureau Group Second Engineering Co., Ltd.
Yang Yang
China Railway Beijing Engineering Bureau Group Second Engineering Co., Ltd.
Jinzhang Jia
Liaoning Technical University
Tao Lei
2327067123@qq.com

Liaoning Technical University https://orcid.org/0009-0001-0785-2586

Research Article

Keywords: fully mechanized heading face, long-forcing and short-exhaust ventilation, regional dust exposure, explainable machine learning, SHAP, risk-oriented ventilation control

Posted Date: September 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10762286/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

To improve near-field dust exposure risk identification and ventilation-parameter regulation in fully mechanized heading faces, this study proposes an interpretable prediction method for dust exposure in key zones. Based on 1296 computational fluid dynamics (CFD) cases under long-forcing and short-exhaust ventilation, regional average concentrations were extracted for the driver zone, return-side breathing zone, heading hazardous zone, and floor-sensitive zone. The over-threshold fraction in the heading hazardous zone was further used to characterize the spatial expansion of high-concentration dust near the source. Source-strength, particle-size, ventilation, and interaction features were constructed to establish machine-learning mappings between operating parameters and regional exposure indicators. Shapley additive explanations (SHAP) analysis was then used to identify dominant factors and their physical response mechanisms. The results show that the model achieved satisfactory predictive performance for the average concentrations in the floor-sensitive zone, return-side breathing zone, and heading hazardous zone, with test-set R^2 values of 0.93, 0.88, and 0.84 and mean absolute percentage error (MAPE) values of 7.61%, 8.82%, and 6.72%, respectively. The over-threshold fraction in the heading hazardous zone achieved an R^2 of 0.80, indicating its potential for assisting the identification of high-concentration source-zone expansion. SHAP results indicate that source-strength parameters dominated all five core indicators, particle-size parameters mainly affected particle transport-deposition redistribution, and the pressure-exhaust ratio modified dust migration paths and risk transfer through ventilation parameters and interaction features. Using the average concentration in the heading hazardous zone as the criterion, the low-, medium-, and high-risk classification achieved an accuracy of 0.862 and a macro-F1 score of 0.851. The proposed framework provides an interpretable basis for regional exposure assessment, risk screening, and ventilation regulation.

1. Introduction

With the development of rapid excavation equipment and high-power cutting machines in coal mines, dust generation intensity in fully mechanized heading faces has increased, and near-field dust transport has become more complex. Coal dust particles are continuously released into the confined roadway space by cutting, transfer, and equipment-induced disturbance. At the same time, forcing airflow, exhaust airflow, machine obstruction, and roadway walls jointly control particle transport, dilution, deposition, and resuspension near the heading source zone, driver working area, return-side breathing zone, and floor. Colinet et al.[1] summarized engineering practices for coal-mine dust control and emphasized source suppression, capture, and removal as core measures for reducing underground dust exposure. From an occupational-health perspective, Laney and Weissman[2], Perret et al.[3], and Cohen et al.[4] reported that long-term exposure to coal dust and silica-bearing dust significantly increases the risk of coal workers' pneumoconiosis and related respiratory diseases. Blackley et al.[5] further observed that the prevalence of coal workers' pneumoconiosis has continued to increase in some mining regions. Therefore, identifying high-exposure near-field zones and clarifying the effects of source strength, particle size, and ventilation parameters on regional risk are essential for dust-risk prevention and ventilation optimization in fully mechanized heading faces.

Numerical simulation has become an important tool for investigating coupled airflow-dust transport in confined roadways. Wang et al.[6] used a CFD modelling approach to analyze respirable dust pollution in an underground heading face driven by a continuous miner and revealed the spatial distribution of dust near the

working face and along the return side. Wang et al.[7] numerically investigated dust movement and distribution in a laneway with a hybrid ventilation system, showing that the combination of forcing and exhaust airflow can substantially alter dust transport paths. Cai et al.[8] further evaluated the dust-suppression performance of a hybrid ventilation system during tunnel excavation and confirmed the key role of ventilation organization in dust dilution and removal. Under long-forcing and short-exhaust, or far-forcing and near-exhaust, ventilation conditions, Du et al.[9], Li et al.[10], and Gong et al.[11] investigated dust distribution, pressure-exhaust ratio effects, and exhaust-airflow optimization, respectively, demonstrating that the matching relationship between forcing and exhaust airflow affects dust accumulation in the heading source zone and downstream migration. Cheng et al.[12], Yu et al.[13], and Wang et al.[14] focused on dust-removal air curtains and air-curtain dust control, showing that auxiliary airflow structures can reshape dust diffusion boundaries and pollutant migration paths.

These studies provide an important basis for understanding dust dispersion and optimizing local ventilation and dust-control measures in fully mechanized heading faces. However, most existing studies still use whole-field concentration contours, single monitoring lines, or overall dust-removal efficiency as the main evaluation objects. Such results are useful for visualizing dust distribution, but they do not directly answer key engineering questions: which zones pose higher exposure risk, which parameters are most sensitive to regional risk, and whether ventilation adjustment transfers risk among the heading source zone, driver zone, return-side breathing zone, and floor-sensitive zone. In practical mine management, decision makers generally require a small number of physically meaningful and controllable indicators, such as the average exposure level in the driver zone, the average concentration in the return-side breathing zone, the high-concentration range in the heading hazardous zone, and the floor dust-deposition risk. Therefore, relying only on whole-field contours or single-point concentration evaluation is insufficient to support regional and risk-oriented ventilation-parameter optimization.

In recent years, data-driven methods have increasingly been applied to nonlinear prediction problems under complex operating conditions. XGBoost, proposed by Chen and Guestrin[15], has strong capability for nonlinear fitting and feature-interaction representation, making it suitable for response prediction under multiparameter coupling. Random forests proposed by Breiman[16], support vector regression described by Smola and Schölkopf[17], and ridge regression proposed by Hoerl and Kennard[18] also provide a basis for multi-model comparison and robustness analysis. Compared with case-by-case CFD computation, machine-learning models can rapidly establish mappings between operating parameters and target responses based on parameterized samples, thereby improving multi-scenario prediction and risk-screening efficiency.

Nevertheless, machine-learning models that focus only on prediction accuracy often behave as black boxes, making it difficult to interpret dust-risk formation mechanisms or derive actionable ventilation-control suggestions. The SHAP method proposed by Lundberg and Lee[19] assigns feature contributions to model predictions based on Shapley values and explains model outputs from both global-importance and local-contribution perspectives. Lundberg et al.[20] further extended this concept to tree-model interpretation, enabling efficient understanding of local explanations and global structure. Reviews by Guidotti et al.[21] and Arrieta et al.[22] highlighted that explainable artificial intelligence should not only improve prediction transparency but also support trustworthy engineering decision-making. Longo et al.[23] further argued that XAI for real systems should move from model explanation toward interdisciplinary decision support and risk

control. For dust control in fully mechanized heading faces, SHAP analysis can identify dominant parameters and reveal how source strength, particle size, ventilation parameters, and their interactions contribute differently to regional exposure risk.

Despite progress in CFD-based dust simulation and engineering application[24], ventilation dust-control optimization based on numerical simulation[25], and machine-learning prediction, regional exposure-indicator prediction and risk-oriented regulation for fully mechanized heading faces under long-forcing and short-exhaust ventilation remain insufficiently explored. First, existing CFD studies mainly focus on the overall dust-field distribution and dust-removal performance, whereas the indicator-based expression of key zones, including the driver zone, return-side breathing zone, heading hazardous zone, and floor-sensitive zone, remains limited. Second, some data-driven studies emphasize full-field concentration reconstruction or overall prediction accuracy but provide insufficient physical interpretation of controllable parameters such as dust source strength, particle size, and pressure-exhaust ratio. Third, few studies have integrated regional-indicator prediction, interpretable machine learning, and risk-level identification into a unified framework that directly supports ventilation-parameter selection and high-risk case screening.

To address these gaps, this study proposes an interpretable prediction and regulation framework for key-zone dust exposure risk in a fully mechanized heading face under long-forcing and short-exhaust ventilation. Rather than aiming at full-field concentration reconstruction, the proposed framework converts 1296 CFD cases into regional average concentrations for the driver zone, return-side breathing zone, heading hazardous zone, and floor-sensitive zone, and uses the over-threshold fraction in the heading hazardous zone to characterize the expansion of high-concentration dust near the source. Source-strength, particle-size, ventilation, and interaction features are then constructed to establish a machine-learning mapping from operating parameters to regional exposure indicators. SHAP analysis is used to explain the contribution mechanisms of different parameters to regional risk. Finally, the average concentration in the heading hazardous zone is used as the primary criterion for low-, medium-, and high-risk identification, and ventilation-control suggestions are proposed for source-zone suppression, operator-exposure reduction, return-side protection, floor dust management, and pressure-exhaust ratio adjustment. The objectives of this study are to:

(1) establish a key-zone dust exposure indicator system for a fully mechanized heading face under long-forcing and short-exhaust ventilation;

(2) develop a rapid prediction model linking operating parameters to regional exposure indicators;

(3) reveal the differential effects of source strength, particle size, and pressure-exhaust ratio on regional exposure risk using SHAP analysis; and

(4) develop a risk-level identification method based on the average concentration in the heading hazardous zone and provide risk-oriented ventilation-control implications.

2. Numerical Model, Sample Database, and Regional Indicator System

2.1. Geometric model and long-forcing/short-exhaust ventilation layout

The numerical object is a fully mechanized heading face equipped with a roadheader, a forcing duct, and an exhaust duct. Noncritical structural details were simplified while preserving the main aerodynamic obstruction effects and near-field dust-transport characteristics. The computational domain was 40 m × 6 m × 4 m. A long-forcing and short-exhaust ventilation arrangement was adopted: the outlet of the forcing duct was 10 m from the working face, and the inlet of the exhaust duct was 5 m from the working face. After establishing a unified coordinate system, the 0–15 m region behind the working face was defined as the near-field study domain for regional exposure assessment.The numerical model and the near-field ventilation layout are shown in Fig. 1.

2.2. Gas-solid two-phase numerical model and boundary conditions

Dust migration was treated as a gas-solid two-phase flow, with air as the continuous phase and dust particles as the discrete phase. Based on the assumption of a low particle volume fraction and negligible interparticle interaction, the Euler-Lagrange approach was adopted for coupled solution. The airflow field was solved using the standard k-ε turbulence model, whereas dust particles were tracked using a Lagrangian discrete-phase approach with a discrete random walk model to account for turbulent dispersion. Pressure-velocity coupling was solved using the SIMPLE algorithm. The particle boundary condition was set as trap for the floor and reflect for the other walls. The main numerical settings are summarized in Table 1. These numerical treatments are consistent with commonly used RANS-based CFD, finite-volume solution, and Lagrangian particle-tracking approaches for gas–solid two-phase flow [26–28]. In addition, because the present simulations assume a low particle volume fraction, particle–particle collisions were not explicitly resolved, which is consistent with dilute particle-laden turbulent-flow modeling[29].

Table 1
Numerical model and boundary-condition settings.

Category	Setting	Value/description
Continuous phase	Turbulence model	Standard k-ε model
Continuous phase	Solver	Pressure-based steady solver; SIMPLE pressure-velocity coupling
Discrete phase	Particle material and injection mode	coal-mv; surface injection
Discrete phase	Turbulent dispersion model	Discrete Random Walk model
Discrete phase	Phase coupling and coupling frequency	Enabled; 20
Discrete phase	Integral scale and particle-size discretization	0.15; 10 particle-size classes
Boundary conditions	Forcing duct / exhaust duct	velocity-inlet
Boundary conditions	Roadway outlet / walls / floor	outflow; reflect; trap
Physical condition	Gravity acceleration	-9.81 m·s − 2

2.3. Parameterized CFD sample database and data-quality control

To cover the typical parameter space of dust source strength, particle-size combination, and local ventilation structure, four mass-flow-rate levels, three minimum particle-size levels, three maximum particle-size levels, six forcing-duct velocity levels, and six exhaust-duct velocity levels were specified, generating 1296 numerical cases. This study did not perform POD reduction or concentration-field reconstruction for whole-field regular grid snapshots. Instead, regional statistical indicators were extracted from the processed CFD data for subsequent regional exposure prediction, risk identification, and parameter-regulation analysis.

Table 2
Parameter levels and input-feature construction.

Parameter/feature	Symbol	Value or construction	Physical meaning
Mass flow rate	Q□	0.005, 0.008, 0.012, 0.015 kg·s − 1	Dust source strength controlling the overall dust load
Minimum particle size	d_min	0.5, 2, 5 μm	Airflow-following tendency of fine particles
Maximum particle size	d_max	50, 100, 150 μm	Inertia and deposition potential of coarse particles
Forcing-duct velocity	v_press	8.5, 8.7, 8.9, 9.2, 9.6, 10.0 m·s − 1	Transport capacity of the forcing jet
Exhaust-duct velocity	v_suction	8.4, 8.6, 8.8, 9.0, 9.2, 9.4 m·s − 1	Dust-extraction capacity near the heading face
Pressure-exhaust ratio	λ	v_press / |v_suction|	Control parameter for airflow structure and risk migration
Particle-size derived features	d_mean, d_span, d_ratio	Calculated from d_min and d_max	Particle-size distribution width and deposition tendency
Interaction features	Q□×λ, Q□×d_mean, λ × d_mean	Products of base features	Source-ventilation and particle-size-ventilation coupling

The continuous-phase calculation was judged by residual convergence and the stability of monitoring-section velocity. After the continuous flow field reached a stable state, discrete particle tracking was performed. Because this study focuses on the relative variation of regional indicators and risk grading across multiple operating conditions rather than transient vortex-resolving analysis, the standard k-ε model provides a balance between computational cost and engineering applicability. Grid independence was verified using the dust-concentration distribution along the return-side breathing zone. Three grids with approximately 570,000, 650,000, and 790,000 cells were compared. The concentration trends were generally consistent for the three grids, with higher concentrations near the working face followed by rapid attenuation and stabilization. Considering both computational accuracy and cost, the 570,000-cell grid was used in subsequent simulations.

2.4. Key-zone division and exposure indicators

Based on dust generation, transport, accumulation, and deposition processes, and considering personnel working locations and ventilation-structure characteristics, the near-field region was divided into the driver zone, return-side breathing zone, heading hazardous zone, and floor-sensitive zone. Figure 3 shows the dust distribution and key-zone locations under a representative case.

The regional average concentration was used to characterize the overall dust exposure level in each key zone. The over-threshold fraction in the heading hazardous zone was used to describe the spatial expansion of the high-concentration region near the source. To ensure consistency and avoid test-set information leakage, the threshold for the heading over-threshold fraction was defined as the 90th percentile of the heading-zone concentration distribution in the training set.

For any region *R*, the regional average concentration is defined as:

$$C_R = \frac{1}{N_R}\sum_{i=1}^{N_R} C_i$$

where C_R is the average dust concentration in region R, N_R is the number of sampling points in region R, and C_i is the dust concentration at the i-th sampling point.

The over-threshold fraction in the heading hazardous zone is defined as:

$$P_{th} = \frac{1}{N_H}\sum_{i=1}^{N_H} I(C_i > T_{90})$$

where P_{th} is the over-threshold fraction in the heading hazardous zone, N_H is the number of sampling points in the heading hazardous zone, T90 is the 90th percentile of the heading-zone concentration distribution in the training set, and I(·) is an indicator function that equals 1 when $C_i > T_{90}$ and 0 otherwise.

Table 3
Key-zone division and exposure indicators.

Zone	Spatial range/m	Zone type	Dominant physical process	Extracted indicator	Engineering implication
Driver zone	x ∈ [7.0, 8.0] y ∈ [2.5, 3.5] z ∈ [1.5, 2.5]	Personnel exposure zone	Jet transport + local recirculation	Average concentration	Represents the exposure level of the roadheader operator
Return-side breathing zone	x ∈ [0, 15] y ∈ [4.7, 5.3] z ∈ [1.4, 1.7]	Transport exposure zone	Mainstream transport + turbulent diffusion	Average concentration	Represents along-route breathing-zone exposure and migration risk
Heading hazardous zone	x ∈ [0, 5] y ∈ [0, 6] z ∈ [0, 2]	High-risk source zone	Dust release + recirculation accumulation	Average concentration; over-threshold fraction	Represents source-zone hazard intensity and high-concentration range
Floor-sensitive zone	x ∈ [0, 15] y ∈ [0, 6] z ∈ [0, 0.5]	Deposition-sensitive zone	Gravitational deposition + resuspension potential	Average concentration	Represents floor dust deposition and secondary-dust potential

3. Interpretable Regional Prediction and Risk Assessment Method

3.1. Methodological framework

The overall technical route is shown in Fig. 4. The framework is designed for regional exposure assessment. It converts the CFD dust field into key regional indicators with explicit engineering meaning, which can be used for ventilation-parameter selection and high-risk case screening. The workflow consists of five steps. First, a dust-dispersion database is constructed based on 1296 CFD cases. Second, regional average concentrations and over-threshold indicators are extracted from the driver zone, return-side breathing zone, heading hazardous zone, and floor-sensitive zone. Third, source-strength, particle-size, and ventilation-related features are constructed to develop a rapid regional-indicator prediction model. Fourth, SHAP analysis is used to identify key influencing factors and their response patterns. Finally, risk levels are classified using the average concentration in the heading hazardous zone as the core criterion, and ventilation-control suggestions are proposed for source-zone suppression, operator-exposure reduction, return-side protection, floor dust management, and pressure-exhaust ratio adjustment.

3.2. Model inputs, outputs, and evaluation metrics

The model inputs include Q□, d_min, d_max, d_mean, d_span, d_ratio, v_press, |v_suction|, λ, and several interaction features. The model outputs are not the complete dust field but the key exposure indicators of the driver zone, return-side breathing zone, heading hazardous zone, and floor-sensitive zone. This setting reduces the influence of local grid-level fluctuations on the prediction results and provides better correspondence with field dust measurement and ventilation-regulation targets.

Table 4
Input-output settings for machine-learning prediction and SHAP interpretation.

Category	Variable	Description
Input feature	Q□, d_min, d_max, v_press, |v_suction|, λ	Base operating parameters
Input feature	d_mean, d_span, d_ratio	Particle-size derived features used to interpret transport-deposition effects
Input feature	Q□×λ, Q□×d_mean, λ × d_mean	Interaction features used to characterize multiparameter coupling
Core output	driver_avg, return_avg, heading_avg, bottom_avg	Average exposure levels in the four key zones
Auxiliary output	heading_over_th_frac	Expansion of high-concentration region in the heading hazardous zone
Model evaluation	R^2, RMSE, MAE, MAPE	Regression prediction accuracy
Risk evaluation	Accuracy, F1-score, confusion matrix	Low-, medium-, and high-risk identification

To avoid interference among regional indicators with different response mechanisms, an independent regression model was established for each output indicator. Candidate models included XGBoost, random forest, support vector regression, ridge regression, and kernel ridge regression. The samples were divided into

training and test sets at an 8:2 ratio. The training set was used for model training, hyperparameter tuning, and risk-threshold determination, whereas the test set was used only for final performance evaluation. Hyperparameter selection was conducted within the training set to prevent test-set information from participating in model selection. For scale-sensitive models, including support vector regression, ridge regression, and kernel ridge regression, input features were standardized only using the training set, and the same transformation was applied to the test set. Model performance was evaluated using R^2, RMSE, MAE, and MAPE. To ensure the independence of risk assessment, all percentile thresholds were calculated from the true CFD indicators in the training set, and the test set was not involved in threshold determination. Because XGBoost showed more stable performance in nonlinear fitting and feature-interaction representation, it was used as the base model for SHAP interpretation.Considering its stable nonlinear fitting performance and compatibility with tree-based SHAP interpretation, XGBoost was selected as the base model for SHAP analysis.

3.3. SHAP-based interpretation of dominant factors

SHAP analysis was used to explain the contributions of operating parameters to the predicted regional dust indicators. The global mean absolute SHAP value was used to identify dominant factors. Feature-group contributions were used to compare the relative effects of source-strength parameters, particle-size parameters, ventilation parameters, and coupled features. SHAP dependence analysis was further used to examine the influence directions of key parameters on the average concentrations in the heading hazardous zone, return-side breathing zone, and floor-sensitive zone. According to the SHAP interpretation results, parameter effects can be summarized as source-strength amplification, transport-deposition redistribution, and flow-field reconstruction.

3.4. Risk-level classification method

To rapidly identify dust exposure risk in key zones, the average concentration in the heading hazardous zone was used as the main criterion for risk-level classification. The heading hazardous zone is close to the dust-release source and is jointly affected by cutting disturbance, the forcing jet, and exhaust-induced recirculation; thus, it directly reflects the near-field high-concentration dust accumulation level in the heading face. Therefore, the average concentration in the heading hazardous zone was selected as the basis for low-, medium-, and high-risk classification. To avoid test-set information leakage, the risk thresholds were determined only from the true CFD indicators in the training set. Let T50 and T80 denote the 50th and 80th percentiles of the average concentration in the heading hazardous zone in the training set, respectively. The classification rule is expressed as follows:

$$\mathrm{Risk}\left(C_{\mathrm{head,avg}}\right)=\begin{cases} Low, C_{\mathrm{head,avg}}<T_{50} \\ Medium, T_{50}\le\ C_{\mathrm{head,avg}}<T_{80} \\ High, C_{\mathrm{head,avg}}\ge\ T_{80} \end{cases}$$

Here, $\mathrm{C}_{\mathrm{head,avg}}$ is the average concentration in the heading hazardous zone in mg·m^{-3}. T50 and T80 are the 50th and 80th percentiles of $\mathrm{C}_{\mathrm{head,avg}}$ in the training set, respectively. In this study, T50 = 540.23 mg·m^{-3} and T80 = 855.09 mg·m^{-3}. The heading over-threshold fraction P_{th} was used as an auxiliary indicator to determine whether the high-concentration dust region expanded significantly.

4. Results and Discussion

4.1. Prediction performance of key regional indicators

To verify the effectiveness of the regional-indicator prediction model, the test-set predicted values were compared with the CFD-extracted values, as shown in Fig. 5. The dashed line in each subplot represents the ideal 1:1 relationship between predicted and true values. Points closer to this line indicate smaller prediction errors.

Figure 5 shows that, except for the driver zone where local disturbance causes a relatively scattered distribution, the predicted values of the other regional indicators are generally close to the CFD-extracted values along the 1:1 line. This indicates that the model can stably capture the main variation trends of regional exposure indicators.

Table 5
Predictive performance of core regional indicators on the test set.

Indicator	R^2	RMSE	MAE	MAPE(%)	Description
Average concentration in driver zone	0.64	65.56	41.91	32.62	Higher uncertainty due to equipment obstruction and local recirculation
Average concentration in return-side breathing zone	0.88	35.17	21.50	8.82	Represents along-route exposure on the return side
Average concentration in heading hazardous zone	0.84	104.93	46.32	6.72	Can be used as the main criterion for risk grading
Average concentration in floor-sensitive zone	0.93	56.63	36.81	7.61	Represents floor deposition and dust-accumulation risk
Heading over-threshold fraction	0.80	0.03	0.01	12.41	Assists in identifying the expansion of high-concentration source-zone dust

Note: Except for the heading over-threshold fraction, the units of RMSE and MAE are $mg \cdot m^{-3}$. The heading over-threshold fraction is dimensionless.

Table 5 summarizes the prediction results for the regional indicators on the test set. The average concentrations in the floor-sensitive zone, return-side breathing zone, and heading hazardous zone achieved R^2 values of 0.93, 0.88, and 0.84, respectively, with MAPE values of 7.61%, 8.82%, and 6.72%. These results indicate that the model captured the major concentration variations in the main exposure zones with acceptable accuracy. The heading over-threshold fraction achieved an R^2 of 0.80 and a MAPE of 12.41%, suggesting that it can assist in determining the expansion of high-concentration dust near the source. By contrast, the R^2 for the average concentration in the driver zone was only 0.64, mainly because this zone is spatially small and strongly affected by equipment obstruction, local recirculation, and jet deflection. This also indicates that driver-zone exposure prediction may require additional information on local geometric obstruction, operator location, or local airflow velocity.

4.2. SHAP interpretation of dominant factors

Figures 6 and 7 present SHAP contributions at the individual-feature and feature-group levels, respectively. The results show that the mass flow rate Q has the highest contribution for all five core indicators and is the dominant factor controlling the near-field dust load. Particle-size-related features, especially d_ratio, contribute more strongly to the floor-sensitive zone and return-side breathing zone, indicating that particle-size variation affects the partitioning between airflow-following transport and gravitational deposition. The overall contributions of ventilation parameters and coupled features are relatively lower, but their effects are more evident for the heading over-threshold fraction, implying that the pressure-exhaust ratio influences risk distribution mainly by altering dust migration paths and high-concentration range expansion.

Table 6
Dominant SHAP factors and physical mechanisms.

Region/indicator	Dominant feature group	Main features	Secondary feature group	Interpretation focus
Average concentration in driver zone	Source-strength parameters, 55.70%	Q, Q×λ, d_ratio	Particle-size parameters, 30.12%	Source-strength amplification dominates personnel-zone exposure; particle size affects dust migration capability
Average concentration on return side	Source-strength parameters, 61.05%	Q, d_ratio, Q×λ	Particle-size parameters, 25.67%	Source strength controls along-route exposure; fine-particle transport enhances breathing-zone risk
Average concentration in heading hazardous zone	Source-strength parameters, 62.18%	Q, d_ratio, Q×λ, λ	Particle-size parameters, 20.76%	Source-zone concentration is mainly controlled by dust-release intensity; ventilation structure regulates local accumulation
Average concentration in floor-sensitive zone	Source-strength parameters, 55.20%	Q, d_ratio, Q×λ	Particle-size parameters, 31.06%	Source strength raises overall load; particle size controls deposition and floor dust accumulation
Heading over-threshold fraction	Source-strength parameters, 51.56%	Q, d_ratio, λ × d_mean	Particle-size parameters, 26.48%; coupled features, 12.74%	High-concentration range is jointly affected by source strength, particle size, and pressure-exhaust coupling

The SHAP results in Table 6 indicate that source-strength parameters dominate all five core indicators, with relative contributions ranging from 51.56% to 62.18%. The source-strength contributions to the average concentrations in the heading hazardous zone and return-side breathing zone reached 62.18% and 61.05%, respectively, confirming that mass flow rate is the primary factor controlling near-field dust load and along-route exposure. Particle-size parameters contributed 20.76%-31.06% across the indicators, with the highest contribution to the floor-sensitive zone (31.06%), suggesting that particle-size variation affects floor dust accumulation and regional exposure through transport-deposition redistribution. Although the individual contributions of ventilation parameters and coupled features were generally lower, coupled features contributed 12.74% to the heading over-threshold fraction, indicating that the pressure-exhaust ratio affects the expansion of the high-concentration range and risk migration mainly through its interaction with source strength and particle size.

4.3. Regional response mechanisms under multiparameter effects

To verify the physical consistency of the SHAP interpretation, the variation rates of key regional indicators under different parameter perturbations were further calculated, as shown in Fig. 8. Increasing the mass flow rate significantly increased all regional indicators, which is consistent with the SHAP result that source-strength parameters dominate the model response. Particle-size variation led to different responses among zones, indicating that particle scale changes the balance between airflow-following transport and gravitational deposition. When the pressure-exhaust ratio deviated from the near-optimal state, the regional indicators exhibited both increases and decreases, suggesting that ventilation regulation mainly changes dust migration paths rather than simply reducing concentrations in all zones.

Combining the prediction results and SHAP interpretation, the effects of mass flow rate, particle size, and pressure-exhaust ratio vary among different regional indicators. The mass flow rate mainly acts as a source-strength amplification factor, raising the average concentrations in the return side, heading zone, and floor zone. Particle-size-related features mainly affect the partitioning between particle transport with airflow and gravitational deposition, making the floor-sensitive zone more sensitive to particle-scale variation. The pressure-exhaust ratio changes dust migration paths through interactions with mass flow rate and particle size, leading to risk redistribution among the heading source zone, driver zone, and return-side breathing zone. Therefore, ventilation regulation should not pursue the minimum concentration in a single zone alone, but should consider both source-zone control and personnel exposure reduction.

4.4. Risk identification and ventilation-control suggestions

According to the risk-level classification rule defined in Section 3.4, the predicted average concentration in the heading hazardous zone on the test set was used to identify low-, medium-, and high-risk cases. Figure 9 shows the distribution of test samples under the risk thresholds, and Fig. 10 presents the corresponding confusion matrix.

Figures 9 and 10 show that most samples fall within the corresponding risk-level regions. Misclassifications mainly occurred between adjacent risk levels, and only one high-risk sample was directly classified as low risk. This result suggests that the proposed method can effectively screen high-risk cases, although uncertainty remains near the risk boundaries. However, uncertainty remains near the risk thresholds, and field dust-measurement data should be incorporated for calibration in future applications.

Table 7
Risk-identification performance for the heading hazardous zone.

Risk criterion	Low/medium threshold / (mg·m − 3)	Medium/high threshold / (mg·m − 3)	Accuracy	Macro-F1	Low-risk samples	Medium-risk samples	High-risk samples
Average concentration in heading hazardous zone	540.23	855.09	0.862	0.851	133	83	44

The risk-classification results show that the overall accuracy and macro-F1 score for the low-, medium-, and high-risk levels were 0.862 and 0.851, respectively. This demonstrates that the average concentration in the heading hazardous zone can serve as an effective main criterion for screening high-risk cases. The misclassifications were concentrated between adjacent risk levels, indicating that some uncertainty still exists near the risk boundaries.

Based on the prediction performance, SHAP interpretation, and parameter-response patterns, typical ventilation-control strategies for different operating scenarios are summarized in Table 8. These strategies are intended mainly for risk screening and preliminary ventilation-parameter selection. In practical applications, they should be calibrated using field dust measurements, airflow constraints, and equipment layout. It should also be emphasized that pressure-exhaust ratio adjustment should aim to reduce the overall regional risk rather than simply minimize the concentration in the heading zone.

Table 8
Typical risk scenarios and ventilation-control suggestions.

Risk scenario	Key zone of concern	Main response	Dominant mechanism	Control suggestion
High mass-flow-rate condition	Heading hazardous zone; floor-sensitive zone	Average concentrations increase in all zones; high-concentration range expands in the heading zone	Source-strength amplification	Prioritize source suppression; strengthen cutting spray, local extraction, and transfer-point dust control
Fine-particle-dominated condition	Return-side breathing zone	Breathing-zone exposure becomes more prominent	Enhanced airflow-following transport	Strengthen breathing-zone monitoring and prevent fine-particle diffusion along the return side
Coarse-particle-dominated condition	Floor-sensitive zone	Floor deposition and dust-accumulation risk increase	Enhanced inertial deposition	Improve floor cleaning, low-level airflow organization, and anti-resuspension measures
Low pressure-exhaust ratio condition	Heading zone; return side	Heading-zone concentration may decrease, but downstream migration risk requires attention	Path change caused by exhaust dominance	Avoid excessive exhaust that induces recirculation or transports dust into personnel zones
High pressure-exhaust ratio condition	Driver zone; return-side breathing zone	Forcing jet may push dust toward the driver zone or return side	Enhanced jet penetration	Appropriately enhance exhaust matching or reduce forcing dominance
Balanced pressure-exhaust ratio condition	Multiple zones	Source control and downstream exposure are relatively balanced	Coordination between forcing and exhaust airflow	Use as an initial regulation window and adjust further based on regional indicators

5. Engineering Implications and Limitations

The engineering significance of the proposed framework lies in converting a complex three-dimensional dust field into a small number of interpretable key-zone indicators. The driver-zone indicator corresponds to operator exposure, the return-side breathing-zone indicator corresponds to along-route transport risk, the heading hazardous-zone indicator corresponds to source-zone control effectiveness, and the floor-sensitive-zone indicator corresponds to dust deposition and secondary-dust potential. Compared with whole-field concentration contours, regional indicators are more suitable for ventilation-parameter selection, monitoring-point arrangement, and high-risk case screening. It should be noted that the risk levels in this study are based on percentile thresholds from the training set and are mainly suitable for relative risk ranking among different operating conditions. If the method is used for occupational-health compliance assessment, it must be further combined with respirable-dust exposure limits, time-weighted average concentration, and site-specific management standards. Because the samples in this study were generated from CFD simulations under a fixed geometry and fixed duct layout, field application still requires calibration using measured data. Future work should incorporate dynamic dust sources, spray efficiency, and online monitoring data to improve the applicability of the method in intelligent ventilation control.

6. Conclusions

(1) A key-zone dust exposure indicator system was established for a fully mechanized heading face under long-forcing and short-exhaust ventilation. The near-field region was divided into the driver zone, return-side breathing zone, heading hazardous zone, and floor-sensitive zone. Regional average concentration was used to characterize overall exposure, and the over-threshold fraction in the heading hazardous zone was used to represent the expansion of high-concentration dust near the source, providing an indicator basis for exposure-risk assessment and ventilation regulation.

(2) A regional-indicator prediction method different from full-field concentration reconstruction was developed. The test-set results show that the model achieved high prediction accuracy for the average concentrations in the floor-sensitive zone, return-side breathing zone, and heading hazardous zone, with R^2 values of 0.93, 0.88, and 0.84, respectively. The over-threshold fraction in the heading hazardous zone achieved an R^2 of 0.80. The relatively lower prediction accuracy for the average concentration in the driver zone reflects the stronger influence of local obstruction and recirculation on small-scale personnel exposure zones.

(3) SHAP results indicate that source-strength parameters dominated all core indicators, with contributions ranging from 51.56% to 62.18%, confirming that mass flow rate is the primary factor controlling near-field dust exposure. Particle-size parameters contributed 20.76%-31.06% and mainly reflected transport-deposition redistribution. The pressure-exhaust ratio affected dust migration paths and high-concentration range expansion through ventilation parameters and coupled features such as $Q_{\square} \times \lambda$ and $\lambda \times d_mean$.

(4) Ventilation regulation should be risk oriented. Under high-source-strength conditions, source suppression should be prioritized. Under fine-particle-dominated conditions, return-side breathing-zone exposure should be controlled. Under coarse-particle-dominated conditions, floor dust accumulation and resuspension should receive more attention. The pressure-exhaust ratio should avoid both excessively low and excessively high values that may cause risk migration. A near-optimal pressure-exhaust ratio can be used as an initial regulation window and further adjusted according to the heading hazardous-zone and personnel-exposure indicators.

Declarations

Ethics Statement

This study was based on computational fluid dynamics simulations and machine-learning analysis and did not involve human participants, human data, or animal experiments. Therefore, ethical approval and informed consent were not required.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability Statement

Subject to project data-management requirements, the CFD-extracted regional indicator dataset and the trained interpretable prediction scripts are available from the corresponding author upon reasonable request.

References

1. Colinet JF, Rider JP, Listak JM, Organiscak JA, Wolfe AL (2010) Best Practices for Dust Control in Coal Mining. National Institute for Occupational Safety and Health, DHHS (NIOSH) Publication No. 2010–110
2. Laney AS, Weissman DN (2014) Respiratory diseases caused by coal mine dust. J Occup Environ Med 56(10 Suppl):S18–S22
3. Perret JL, Plush B, Lachapelle P, Hinks TSC, Walter C, Clarke P, Irving L, Brady P, Dharmage SC, Stewart A (2017) Coal mine dust lung disease in the modern era. Respirology 22(4):662–670
4. Cohen RA, Petsonk EL, Rose C, Young B, Regier M, Najmuddin A, Abraham JL, Churg A, Green FHY (2016) Lung pathology in U.S. coal workers with rapidly progressive pneumoconiosis implicates silica and silicates. Am J Respir Crit Care Med 193(6):673–680
5. Blackley DJ, Halldin CN, Laney AS (2018) Continued increase in prevalence of coal workers' pneumoconiosis in the United States, 1970–2017. Am J Public Health 108(9):1220–1222
6. Wang Z, Li S, Ren T, Wu J, Lin H, Shuang H (2019) Respirable dust pollution characteristics within an underground heading face driven with continuous miner: A CFD modelling approach. J Clean Prod 217:267–283
7. Wang Y, Luo G, Geng F, Li Y, Li Y (2015) Numerical study on dust movement and dust distribution for hybrid ventilation system in a laneway of coal mine. J Loss Prev Process Ind 36:146–157
8. Cai X, Nie W, Yin S, Liu Q, Hua Y, Guo L, Cheng L, Ma Q (2021) An assessment of the dust suppression performance of a hybrid ventilation system during the tunnel excavation process: Numerical simulation, vol 152. Process Safety and Environmental Protection, pp 304–317
9. Du C, Wang H, Jiang Z, He Z, Hu G (2010) Numerical simulations of dust distribution in a fully mechanized excavation face with far-pressing-near-absorption ventilation. J Univ Sci Technol Beijing 32(8):957–962

10. Li X, Wang W, An N, Deng L (2018) Study on dust distribution law under different pressure-exhaust ratios in fully mechanized excavation face of coal mine. Coal Technol 37(10):195–196
11. Gong X, Fan J, Liu Z, Tong D, Chen Y, Wei Y (2021) Optimization analysis of dust field under comprehensive control of air outlet and exhaust air flow in fully mechanized excavation face. J China Coal Soc 46(S2):800–809
12. Cheng W, Nie W, Yao Y, Zhou G, Yu Y, Yang J (2011) Numerical simulation on the flow field of swirling flow air curtain aspiration control dust in fully mechanized workface. J China Coal Soc 36(8):1342–1348
13. Yu H, Cheng W, Wang H, Peng H, Xie Y (2017) Formation mechanisms of a dust-removal air curtain in a fully-mechanized excavation face and an analysis of its dust-removal performances based on CFD and DEM. Adv Powder Technol 28(11):2830–2847
14. Wang H, Nie W, Cheng W, Liu Q, Jin H (2018) Effects of air volume ratio parameters on air curtain dust suppression in a rock tunnel's fully-mechanized working face. Adv Powder Technol 29(2):230–244
15. 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
16. Breiman L (2001) Random forests. Mach Learn 45:5–32
17. Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. Stat Comput 14:199–222
18. Hoerl AE, Kennard RW (1970) Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 12(1):55–67
19. Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. Adv Neural Inf Process Syst, 30
20. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B et al (2020) From local explanations to global understanding with explainable AI for trees. Nat Mach Intell 2:56–67
21. Guidotti R, Monreale A, Ruggieri S, Turini F, Giannotti F, Pedreschi D (2018) A survey of methods for explaining black box models. ACM-CSUR 51(5):1–42
22. Arrieta AB, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A et al (2020) Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Inform Fusion 58:82–115
23. Longo L, Brcic M, Cabitza F, Choi J, Confalonieri R, Del Ser J et al (2024) Explainable Artificial Intelligence (XAI) 2.0: A manifesto of open challenges and interdisciplinary research directions. Inform Fusion 106:102301
24. Zhou G, Zhang Q, Bai R, Fan T, Wang G (2017) The diffusion behavior law of respirable dust at fully mechanized caving face in coal mine: CFD numerical simulation and engineering application. Process Saf Environ Prot 106:117–128
25. Liu Q, Nie W, Hua Y, Peng H, Liu C, Wei C (2019) Research on tunnel ventilation dust suppression technology based on CFD numerical simulation. Tunn Undergr Space Technol 90:44–54
26. Launder BE, Spalding DB (1974) The numerical computation of turbulent flows. Comput Methods Appl Mech Eng 3(2):269–289
27. Versteeg HK, Malalasekera W (2007) An Introduction to Computational Fluid Dynamics: The Finite Volume Method, 2nd edn. Pearson Education

28. Tsuji Y, Tanaka T, Ishida T (1992) Lagrangian numerical simulation of plug flow of cohesionless particles in a horizontal pipe. Powder Technol 71(3):239–250
29. Elghobashi S (1994) On predicting particle-laden turbulent flows. Appl Sci Res 52:309–329

Figures

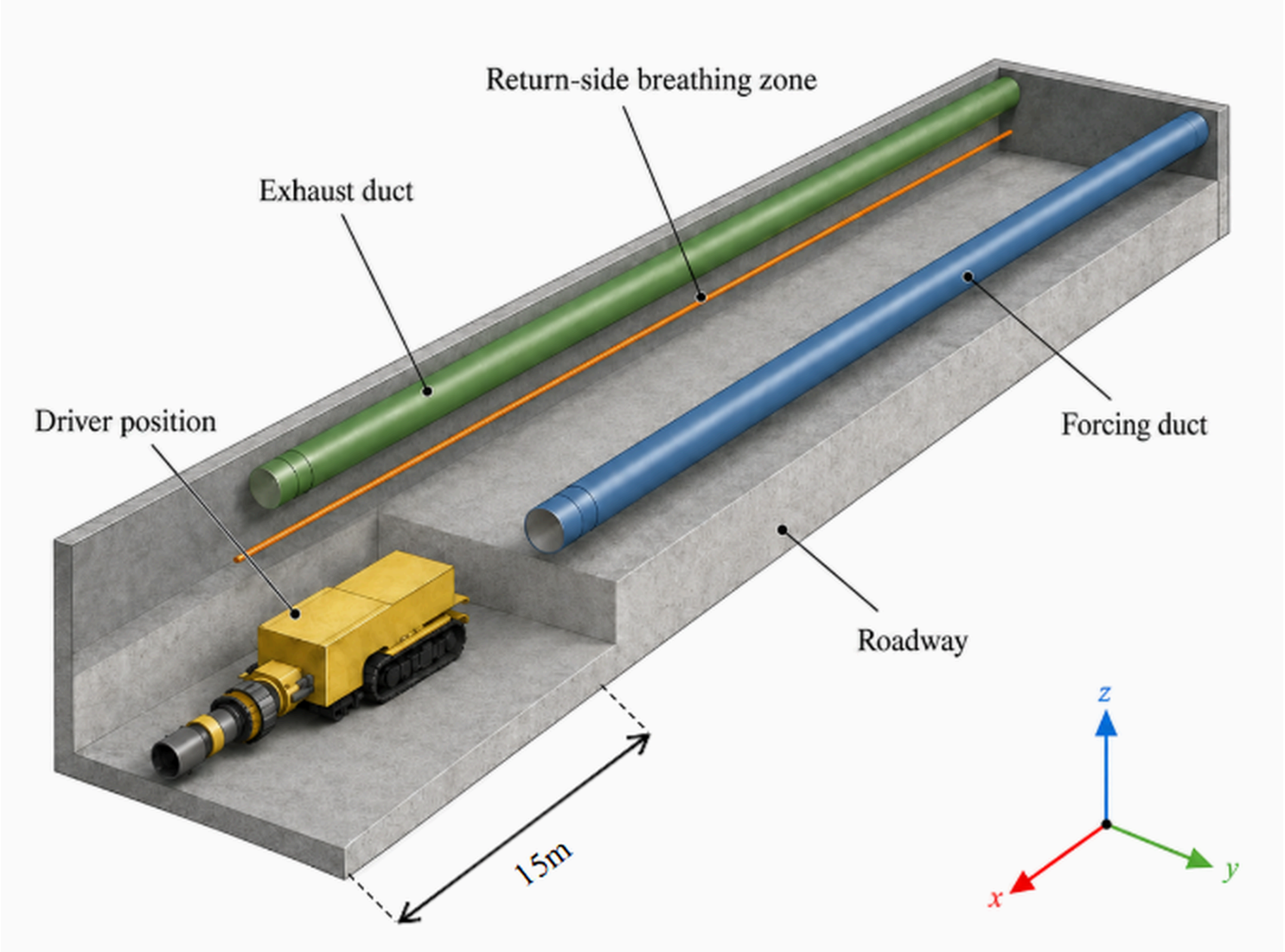


Figure 1

Numerical model and long-forcing/short-exhaust ventilation layout.

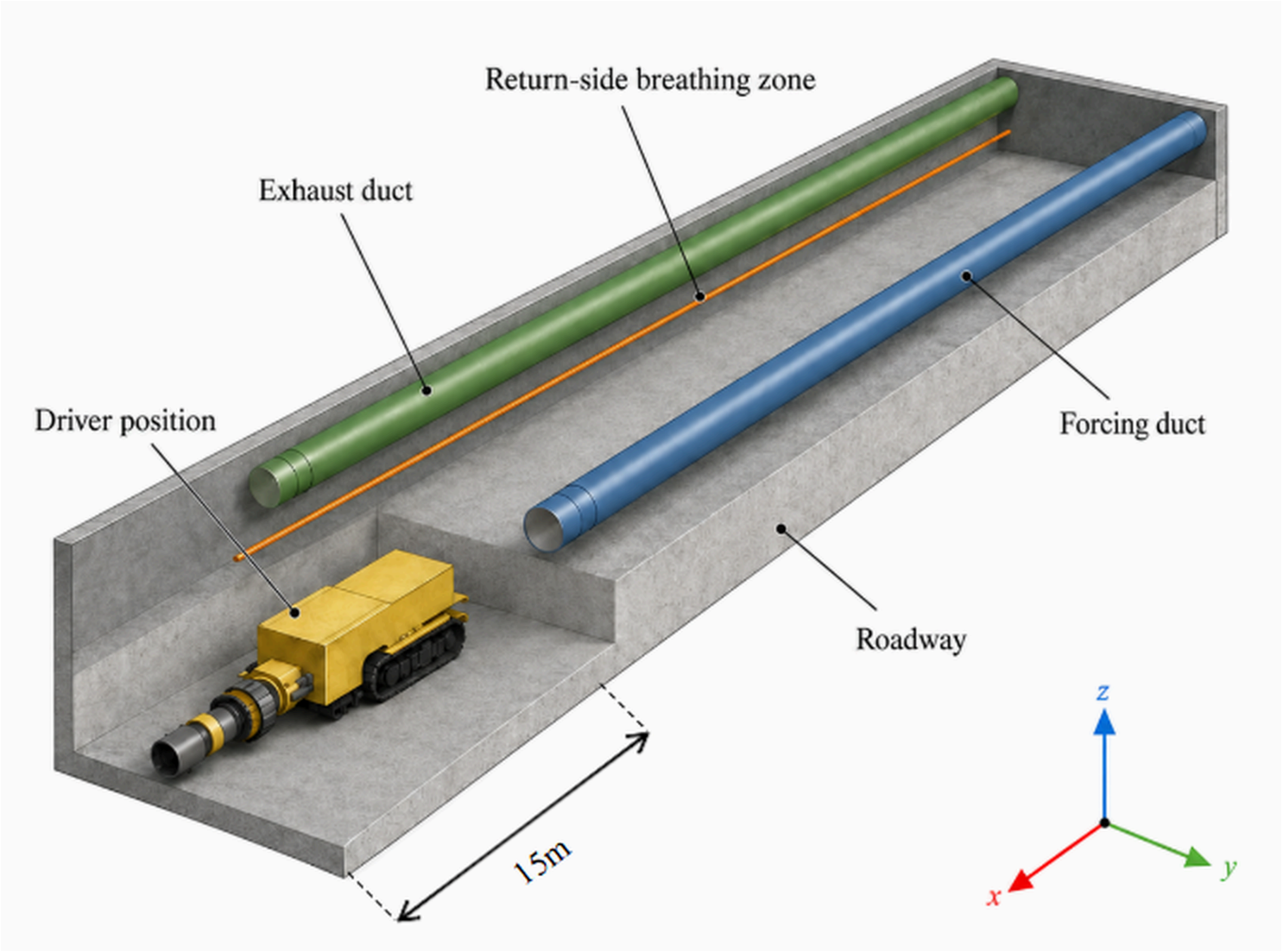


Figure 1

Numerical model and long-forcing/short-exhaust ventilation layout.

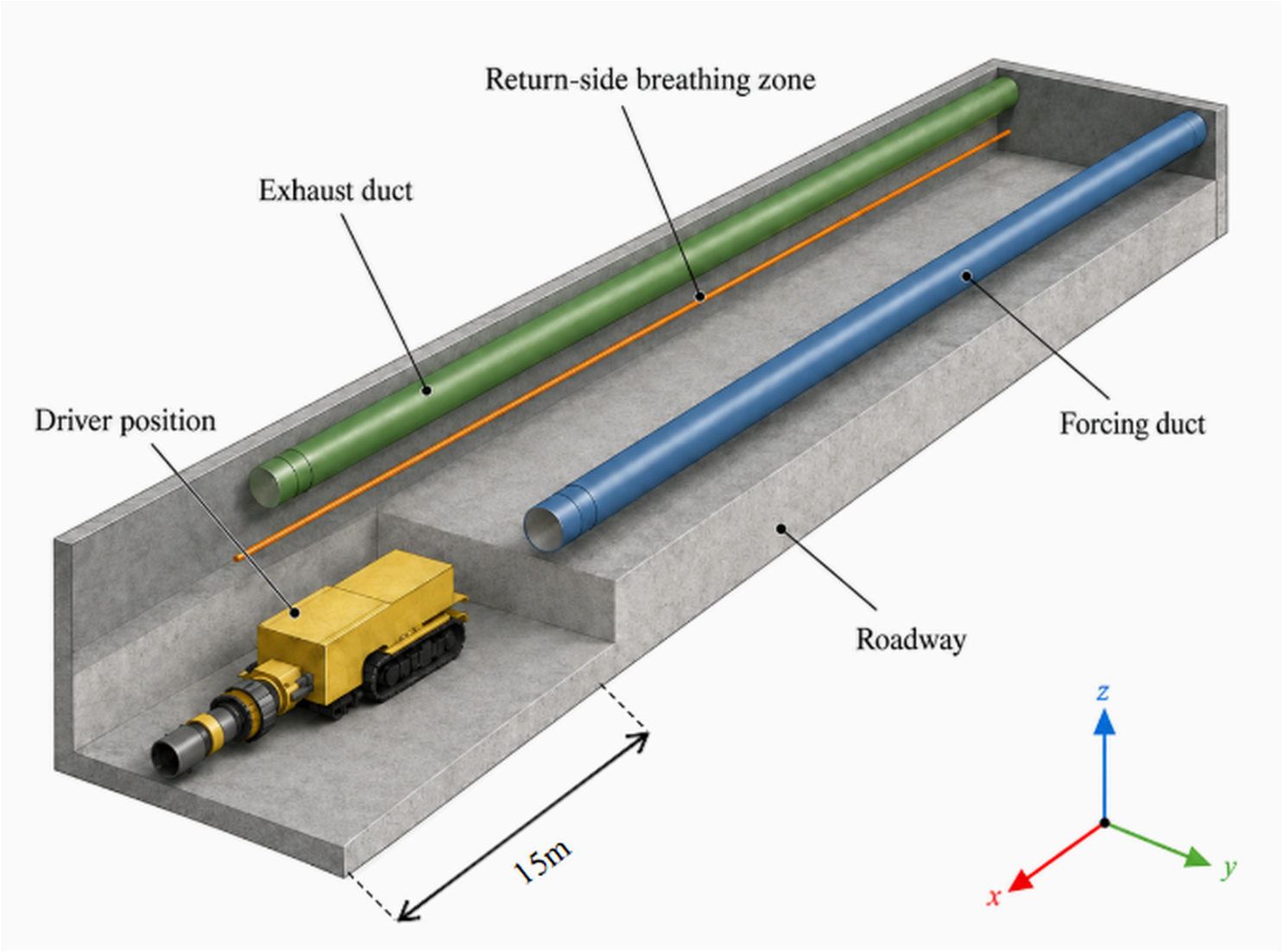


Figure 1

Numerical model and long-forcing/short-exhaust ventilation layout.

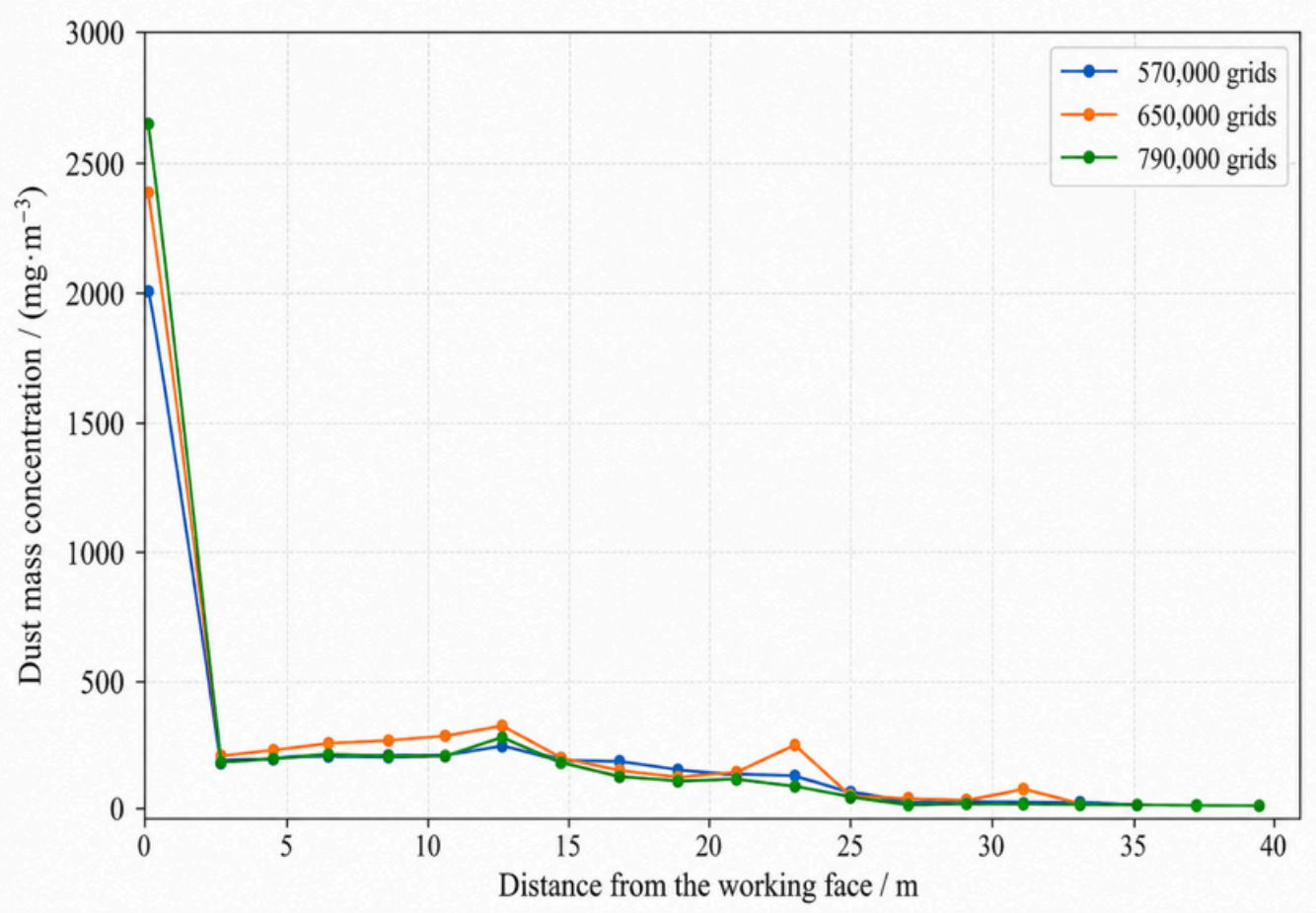


Figure 2

Grid independence verification.

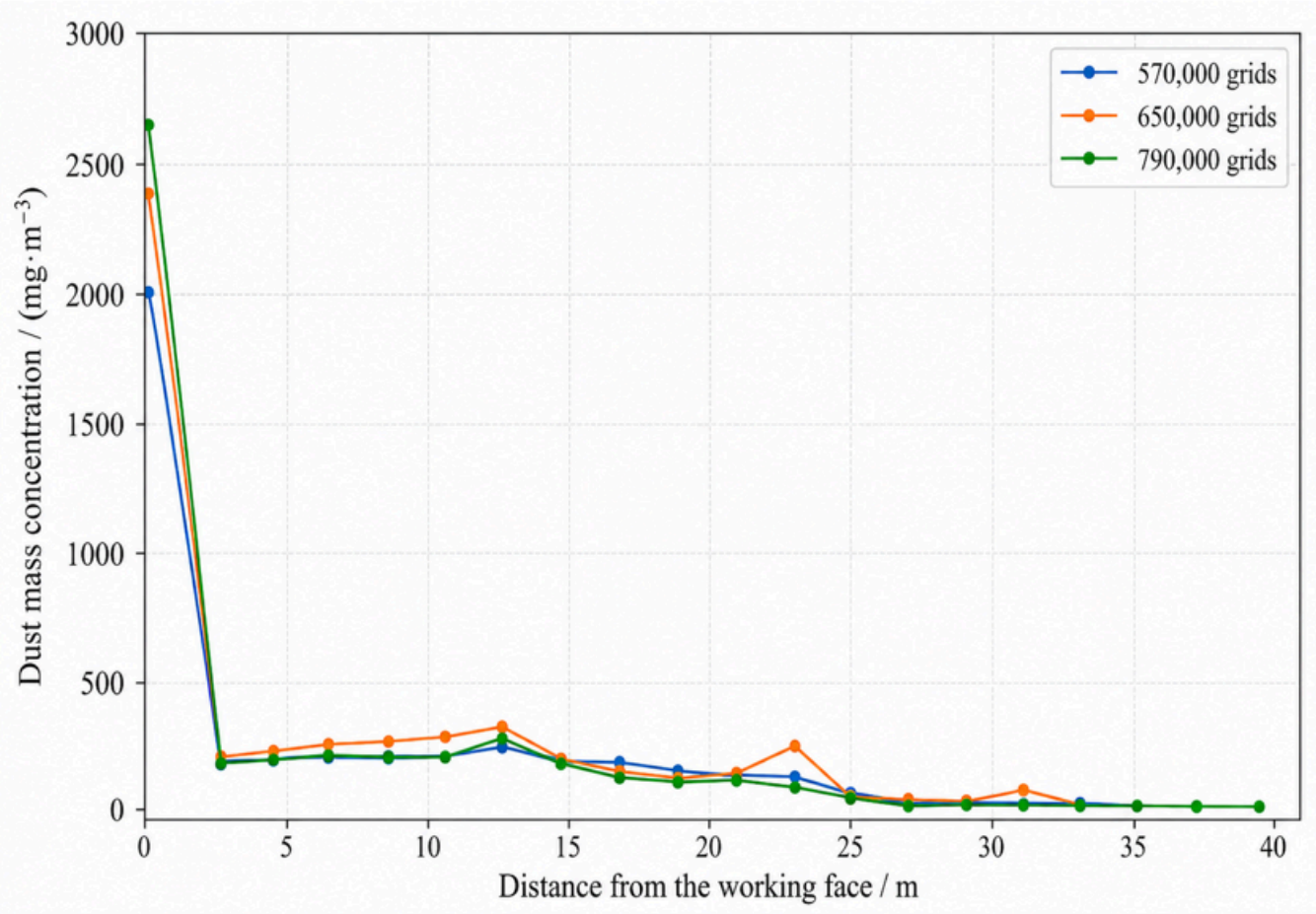


Figure 2

Grid independence verification.

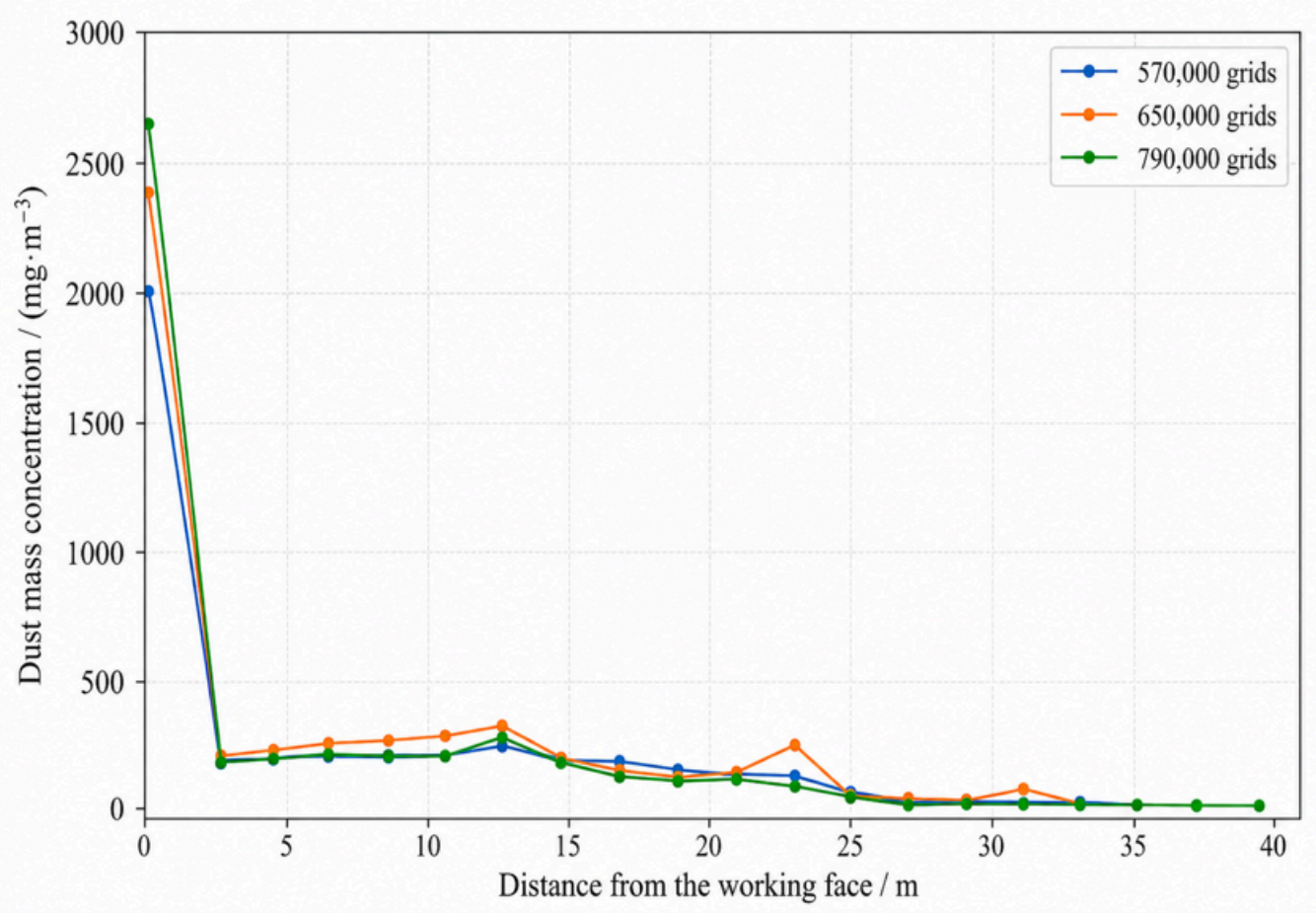


Figure 2

Grid independence verification.

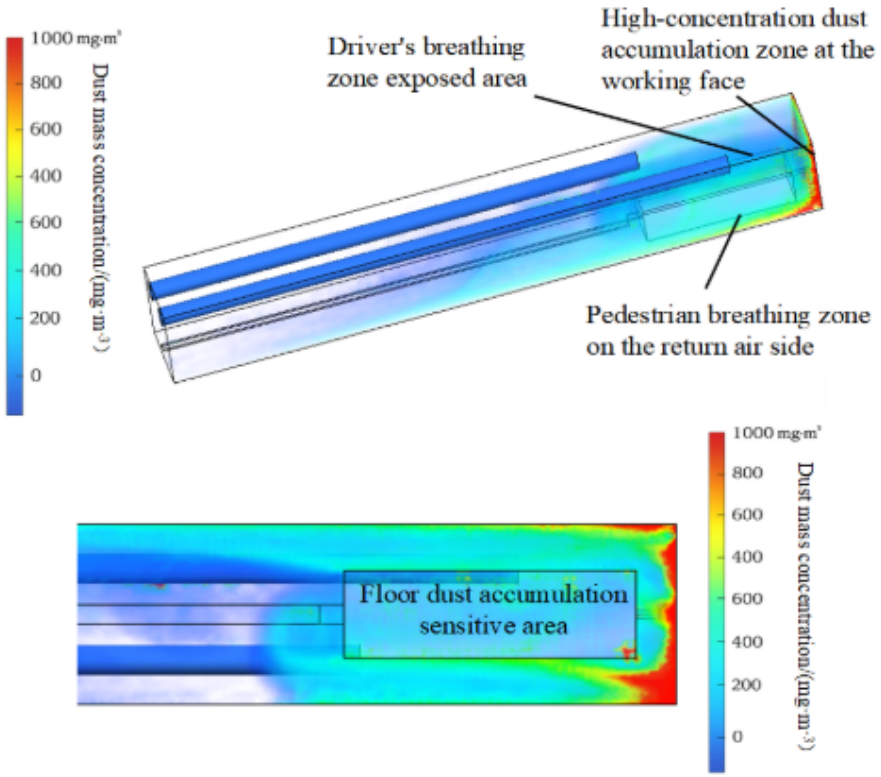


Figure 3

Dust distribution and key-zone definition under a representative case.

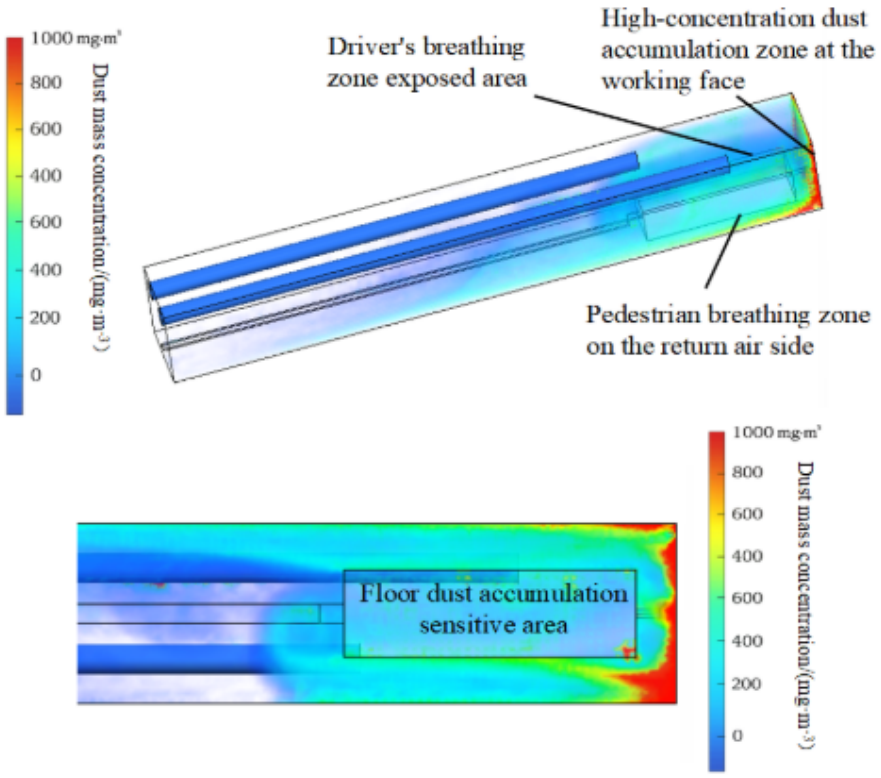


Figure 3

Dust distribution and key-zone definition under a representative case.

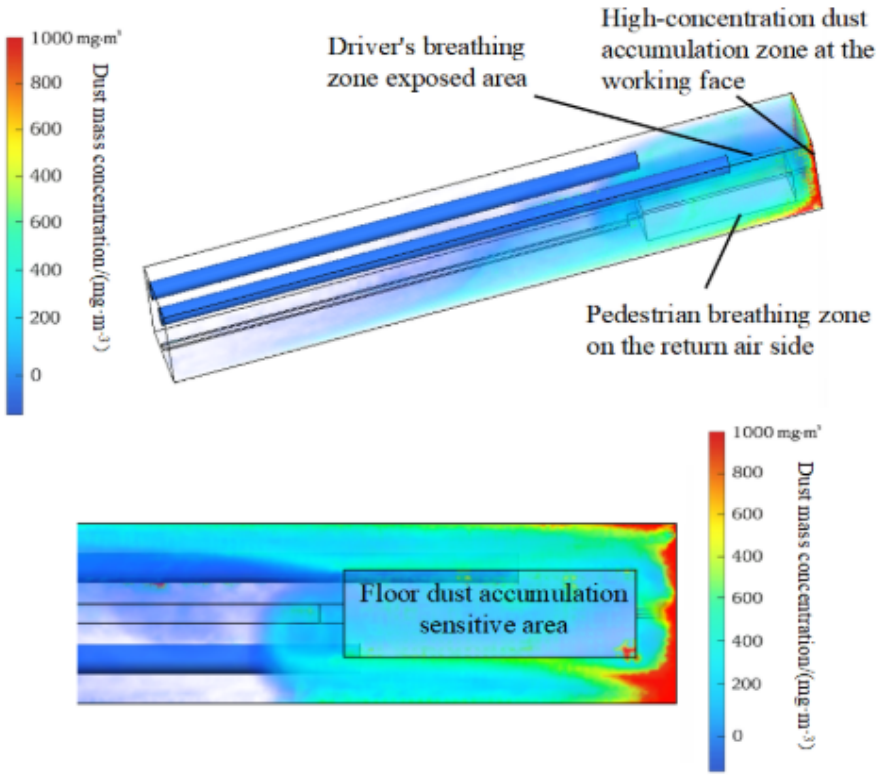


Figure 3

Dust distribution and key-zone definition under a representative case.

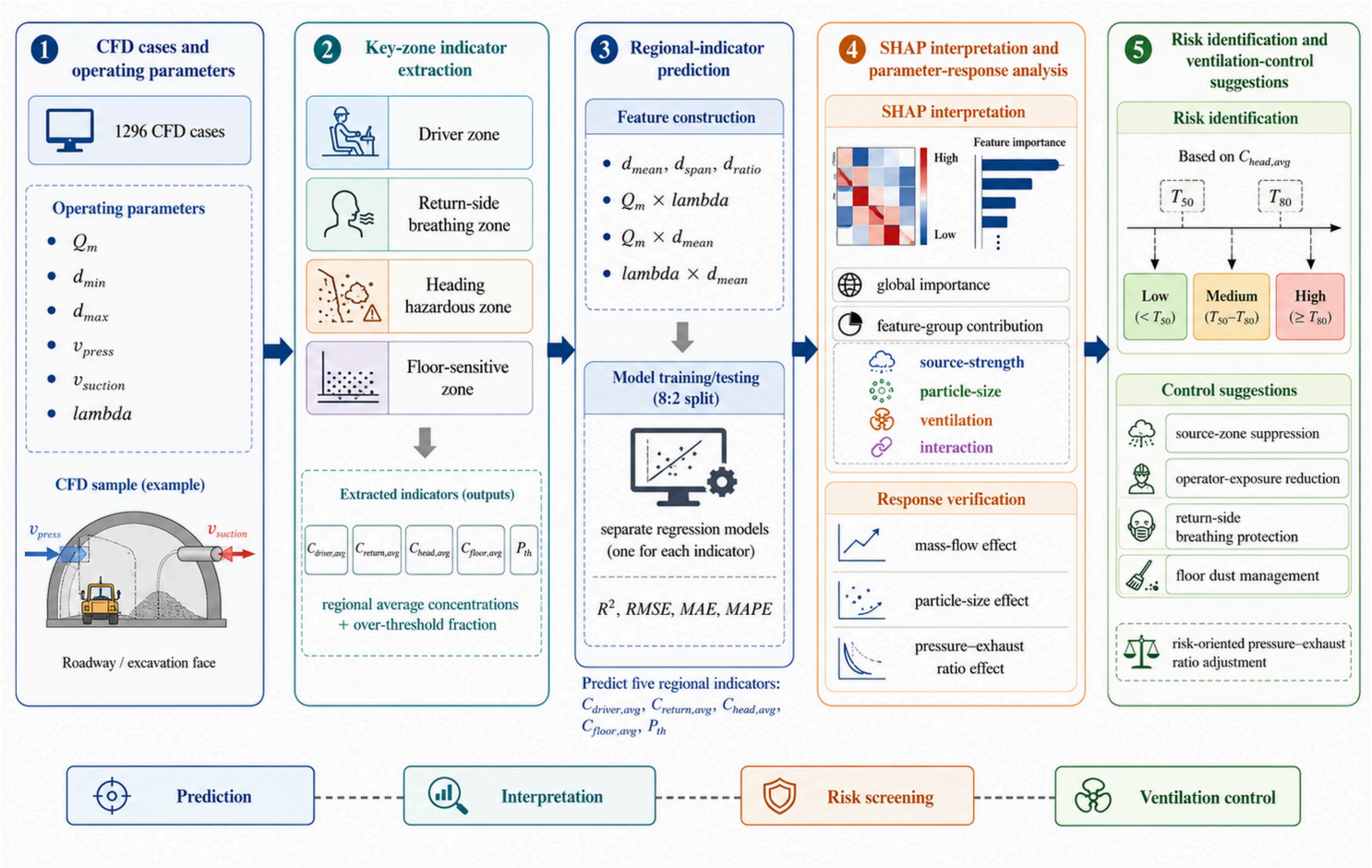


Figure 4

Framework of interpretable regional prediction and risk-oriented ventilation control.

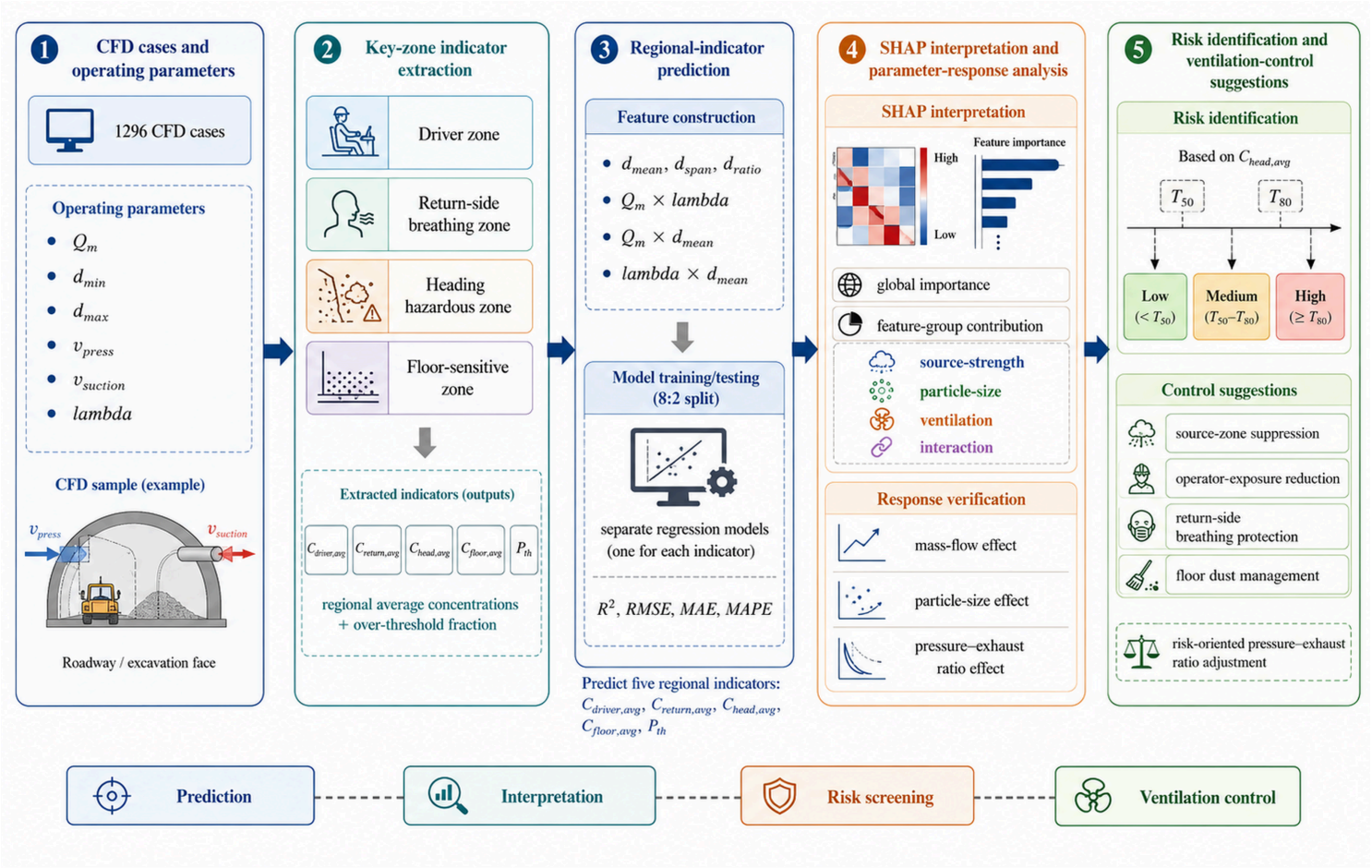


Figure 4

Framework of interpretable regional prediction and risk-oriented ventilation control.

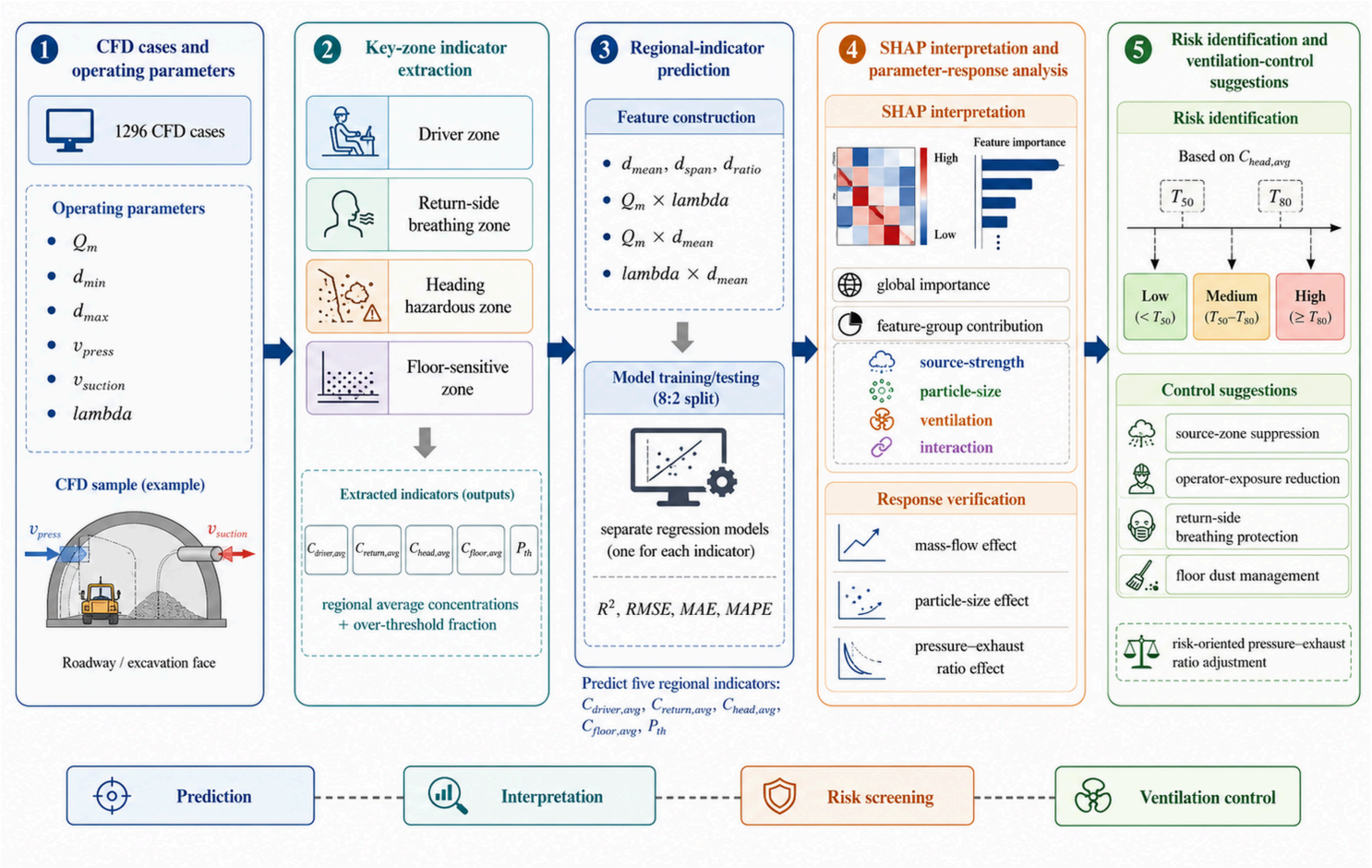


Figure 4

Framework of interpretable regional prediction and risk-oriented ventilation control.

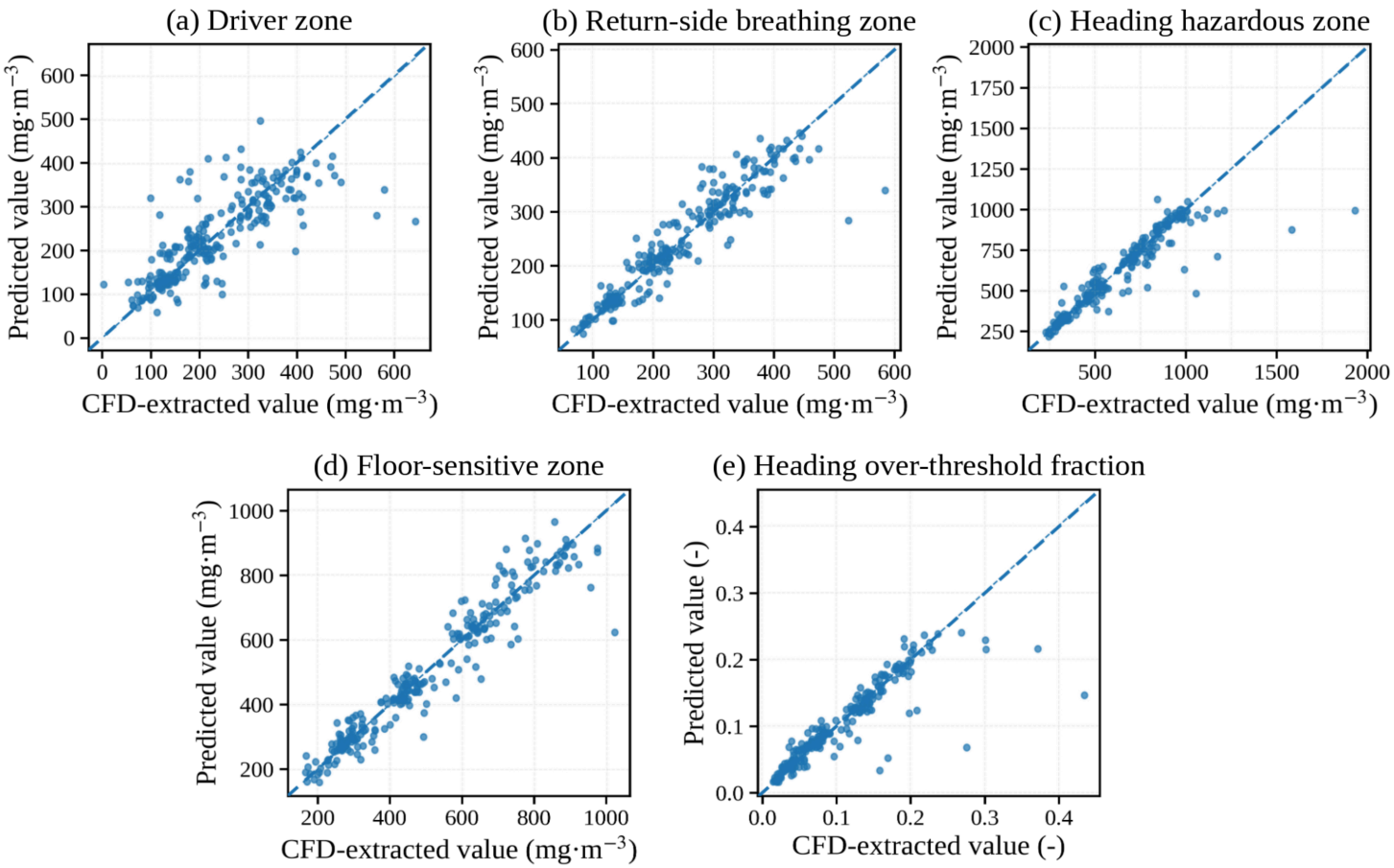


Figure 5

Comparison between predicted and CFD-extracted values of core regional indicators.

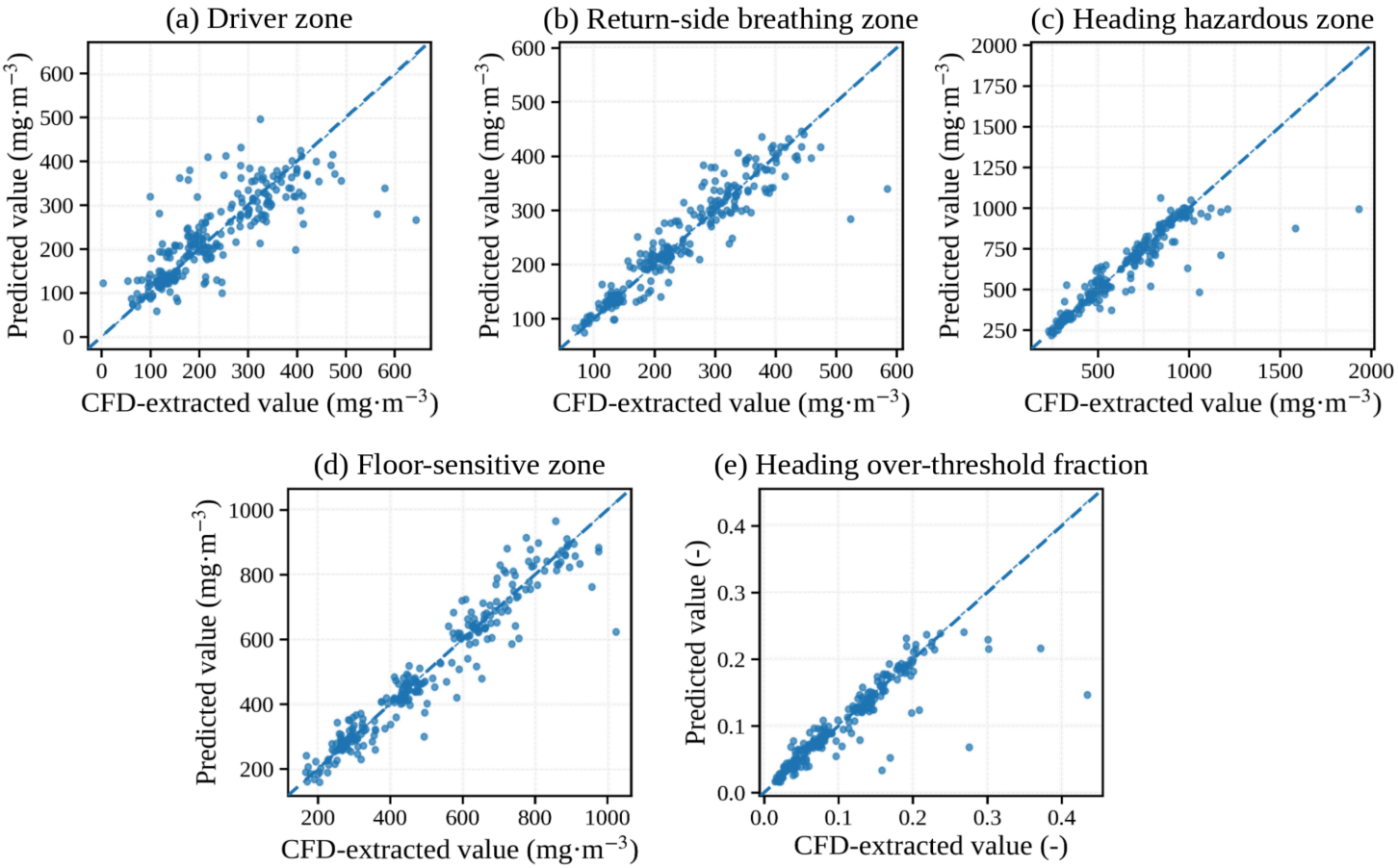


Figure 5

Comparison between predicted and CFD-extracted values of core regional indicators.

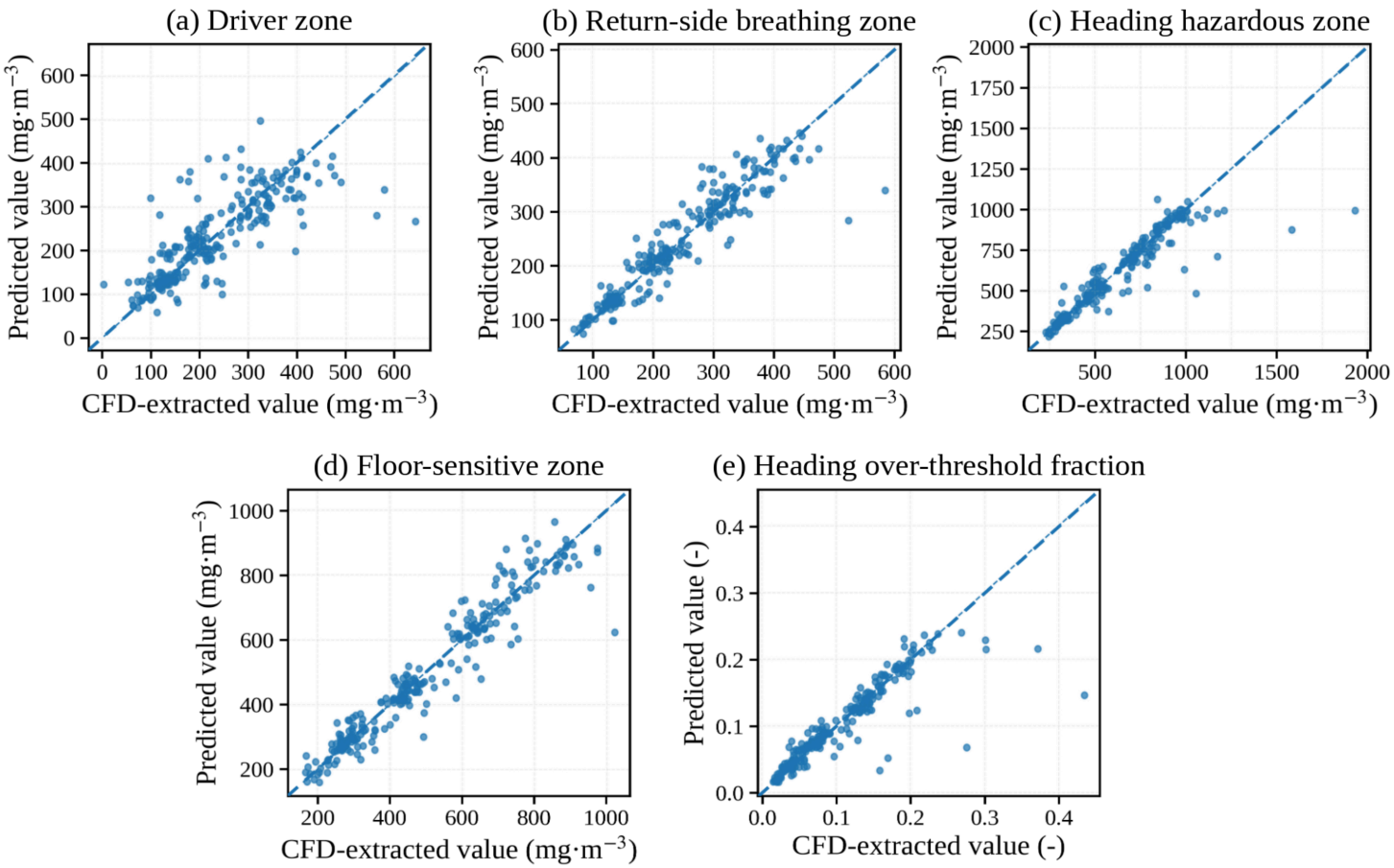


Figure 5

Comparison between predicted and CFD-extracted values of core regional indicators.

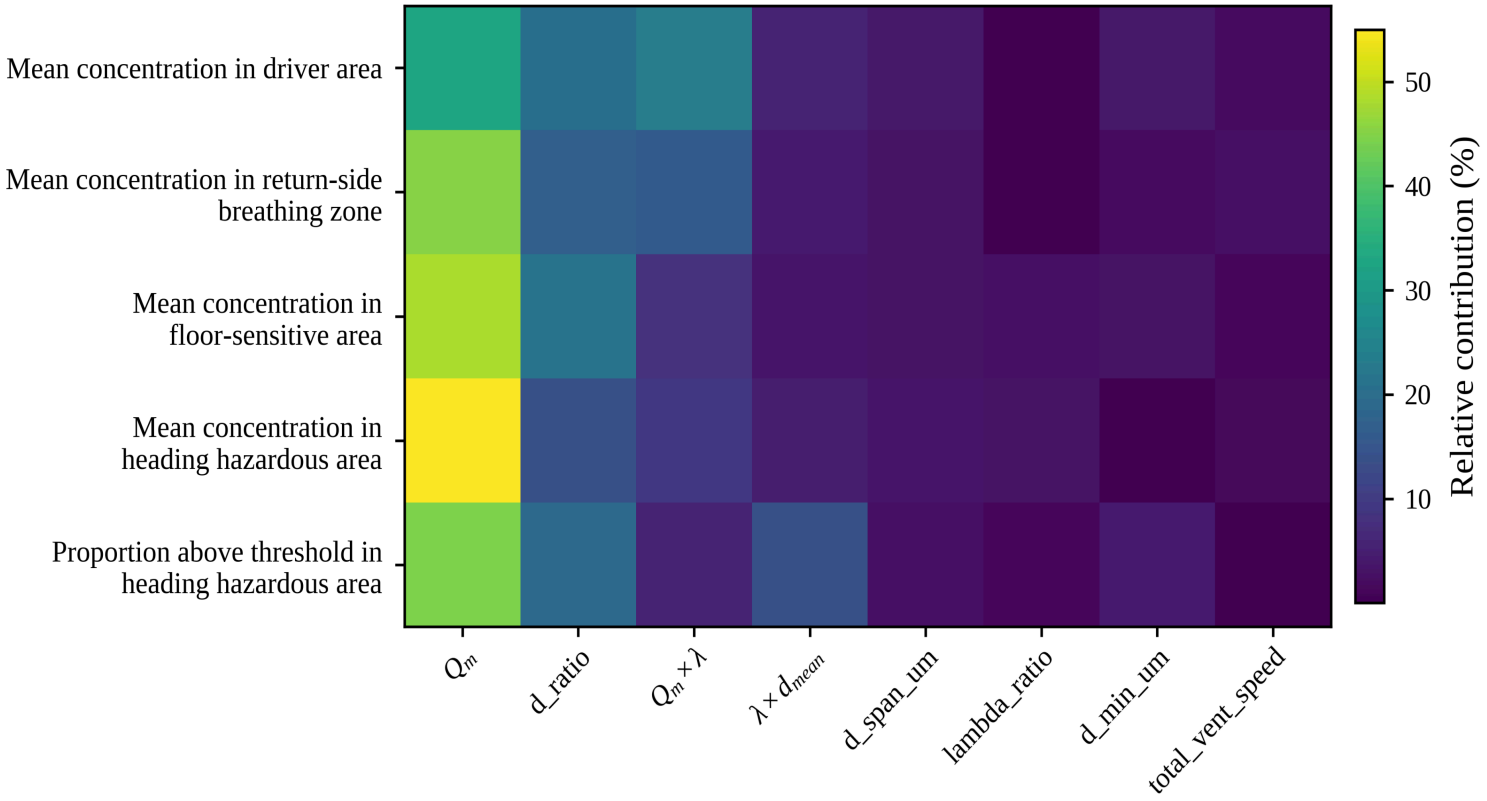


Figure 6

SHAP global importance heatmap for core indicators.

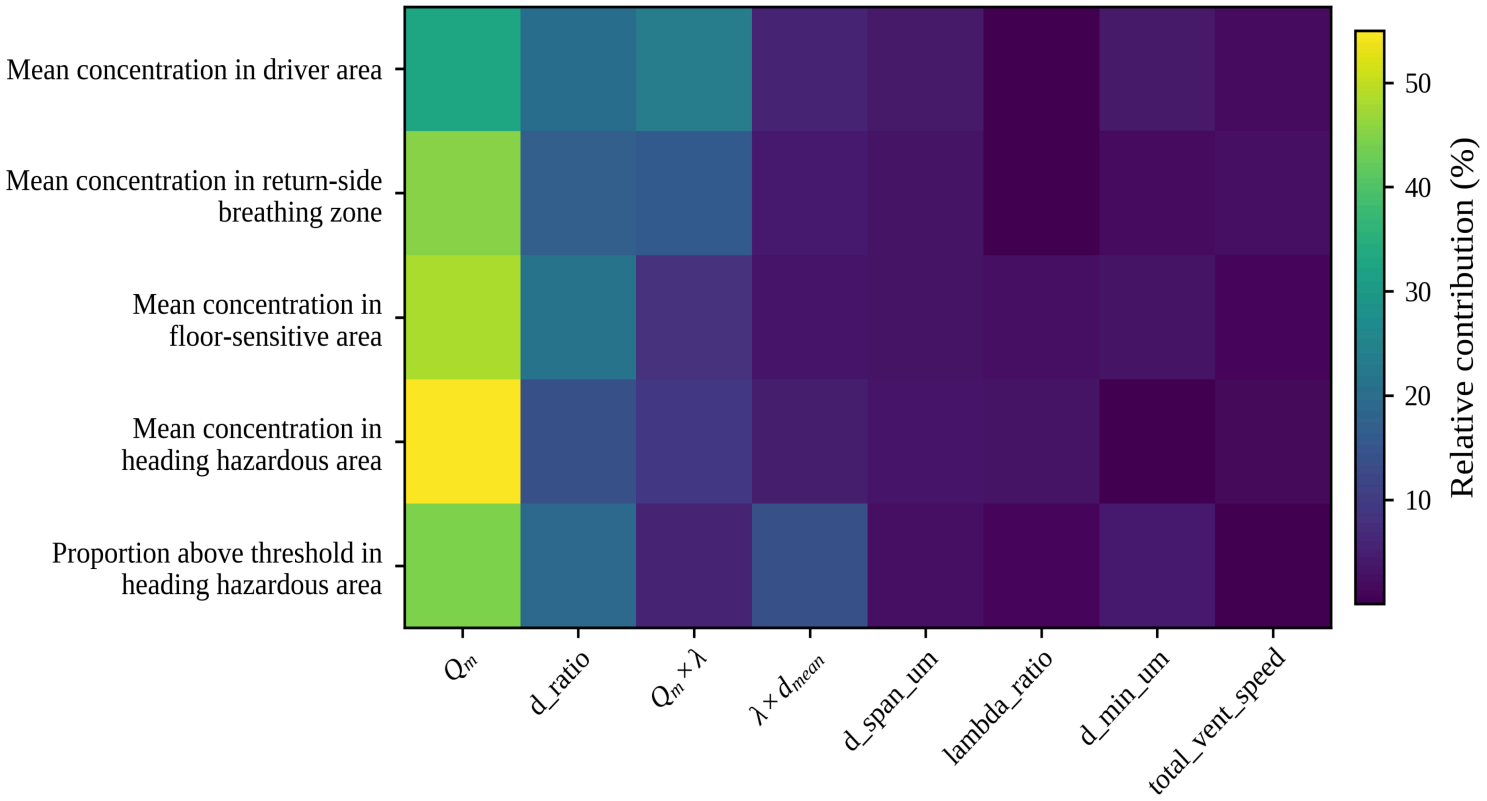


Figure 6

SHAP global importance heatmap for core indicators.

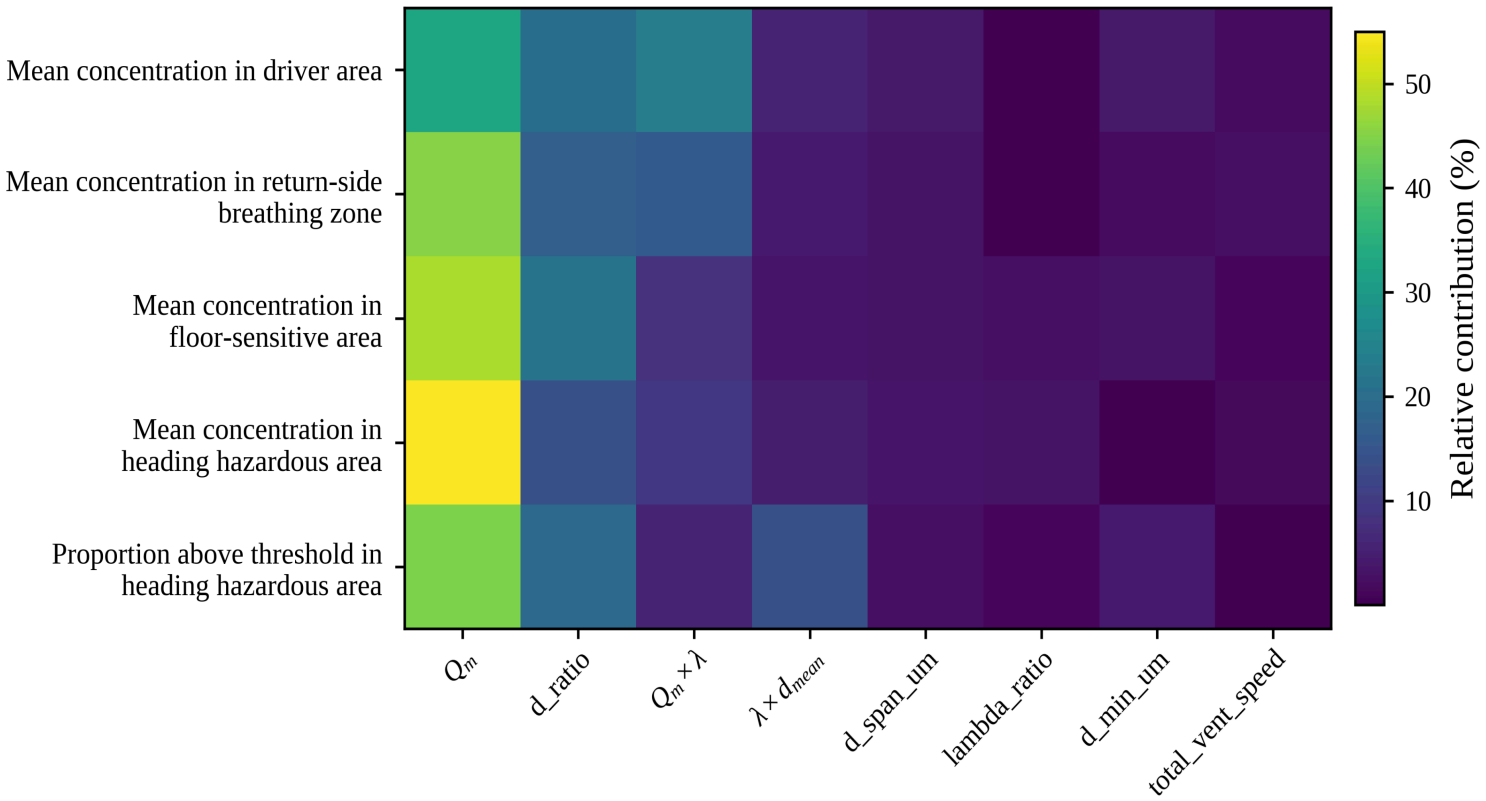


Figure 6

SHAP global importance heatmap for core indicators.

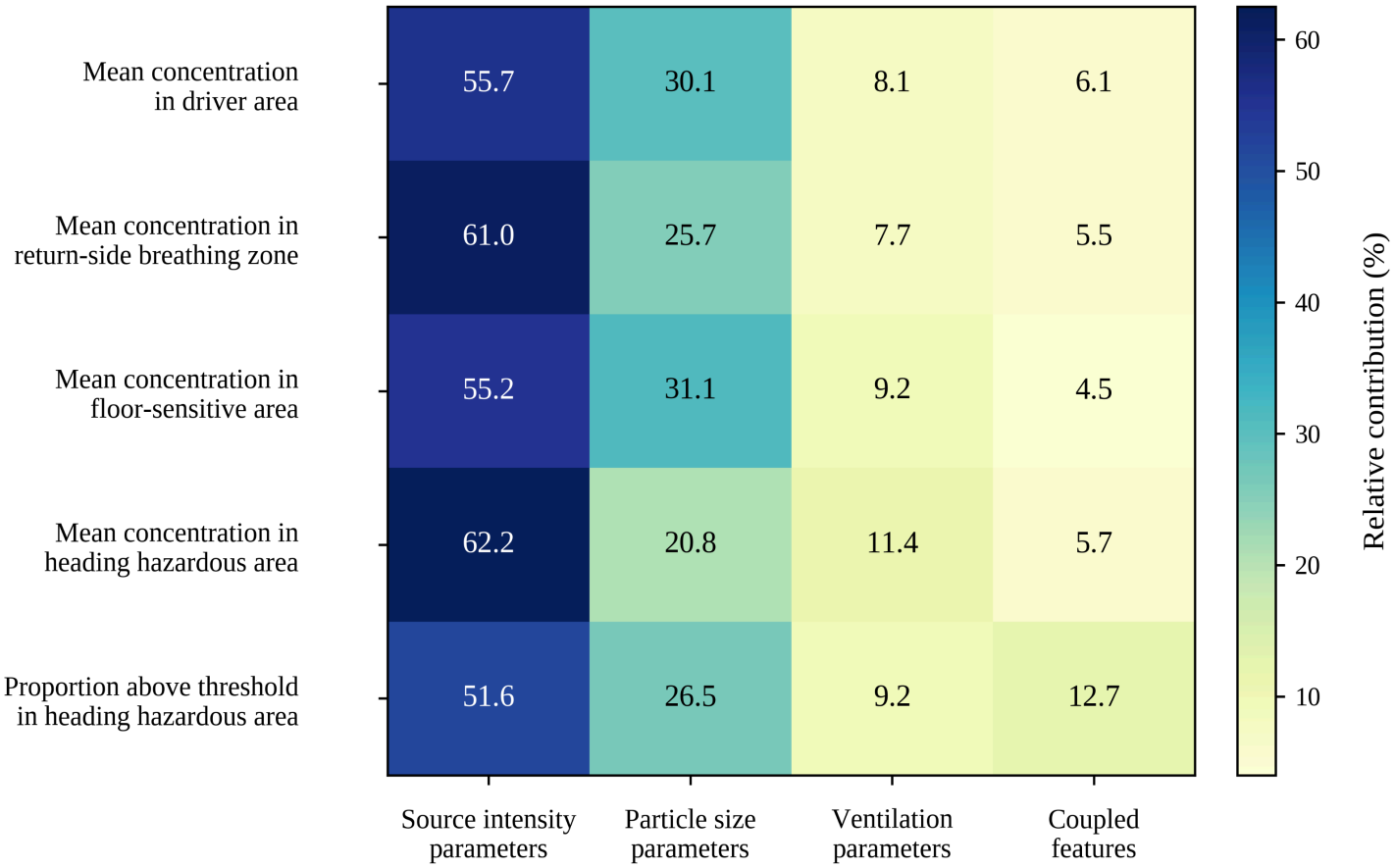


Figure 7

Relative SHAP contributions of feature groups to regional indicators.

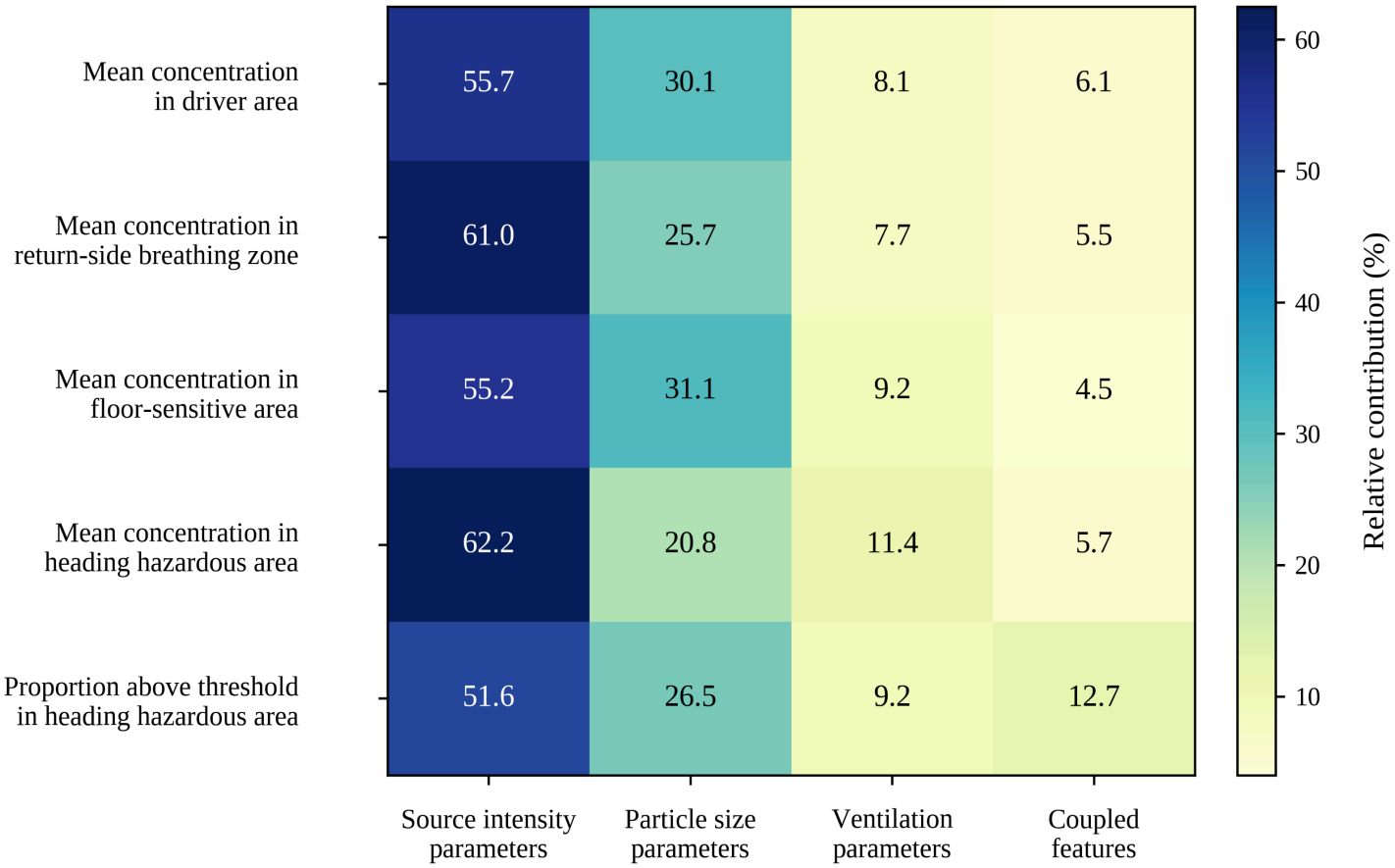


Figure 7

Relative SHAP contributions of feature groups to regional indicators.

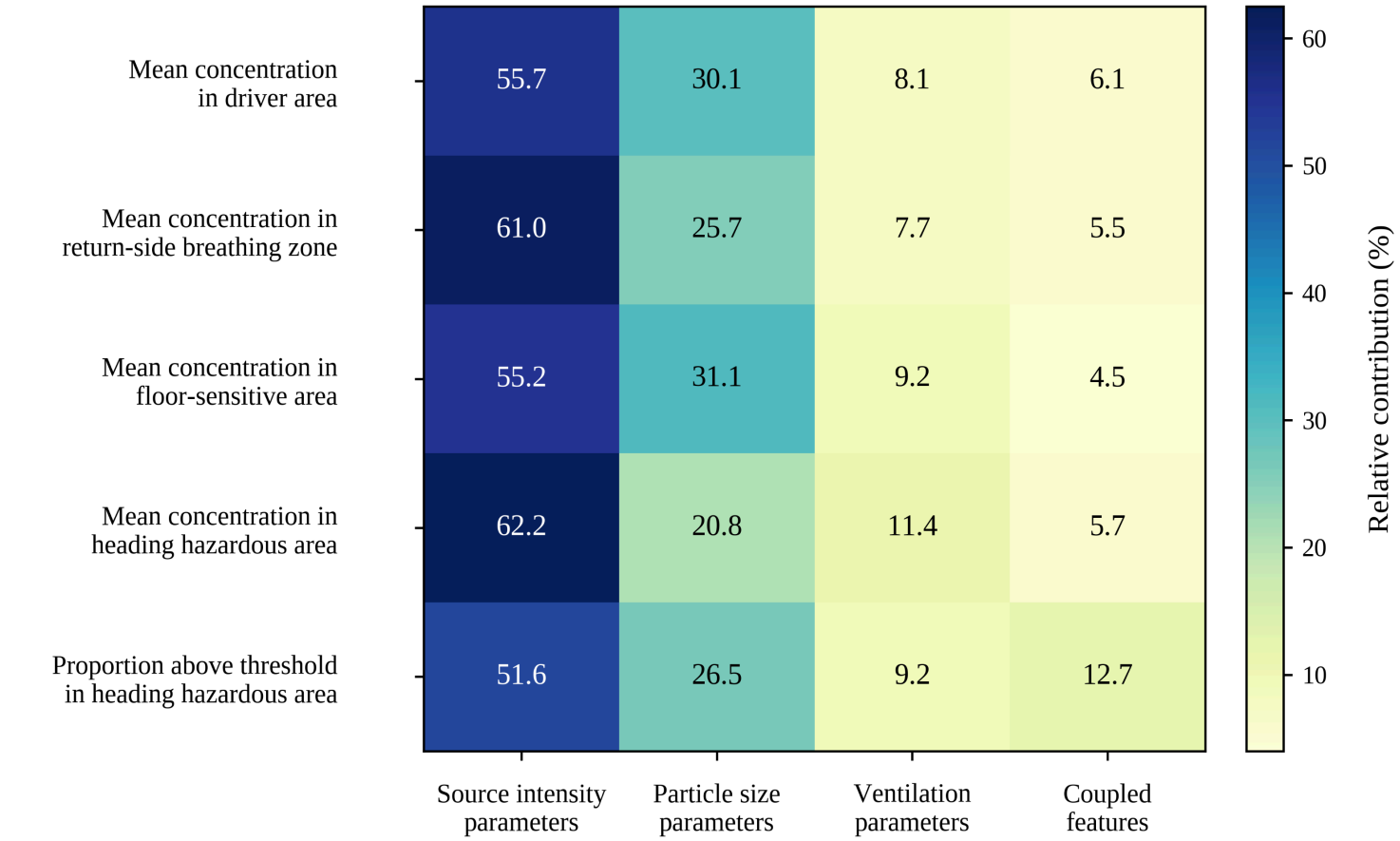


Figure 7

Relative SHAP contributions of feature groups to regional indicators.

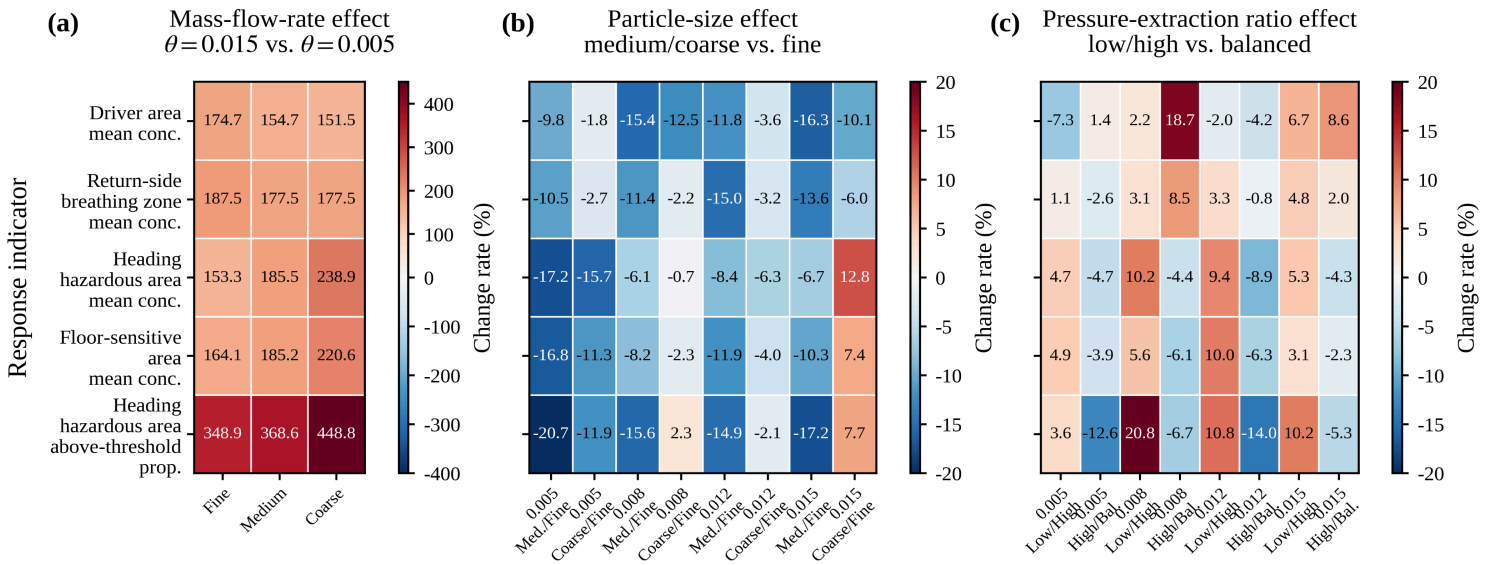


Figure 8

Response heatmap of core regional indicators under multiparameter variations.

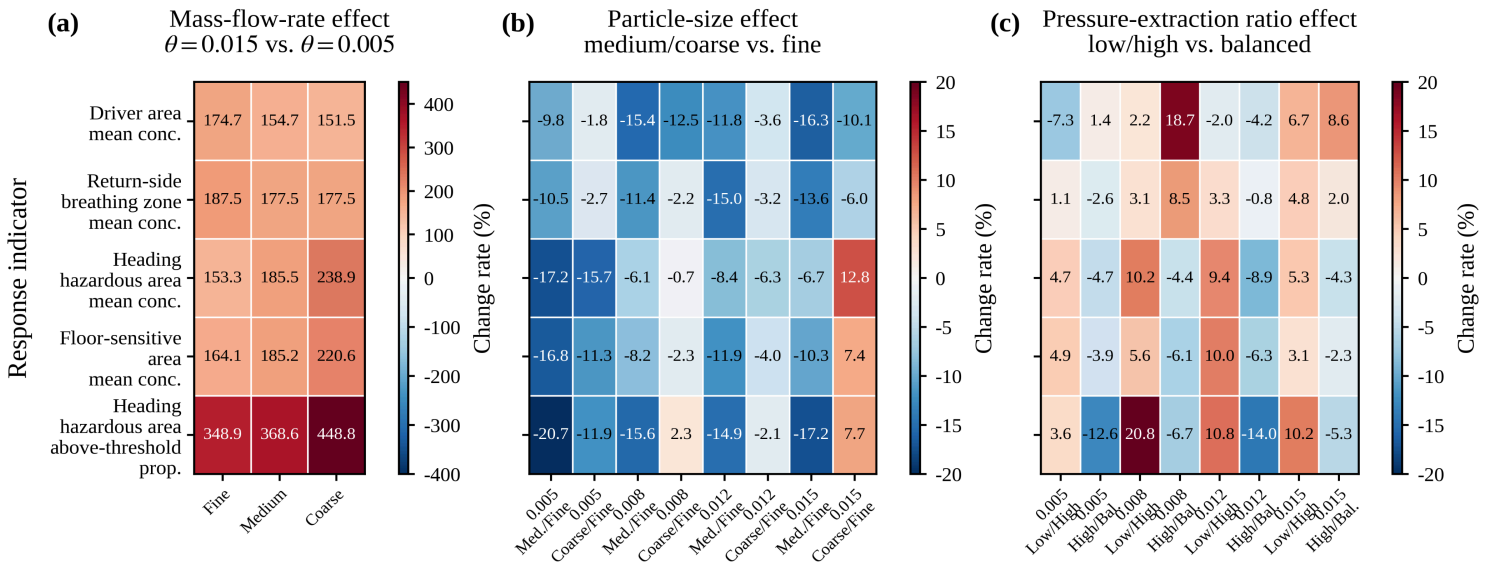


Figure 8

Response heatmap of core regional indicators under multiparameter variations.

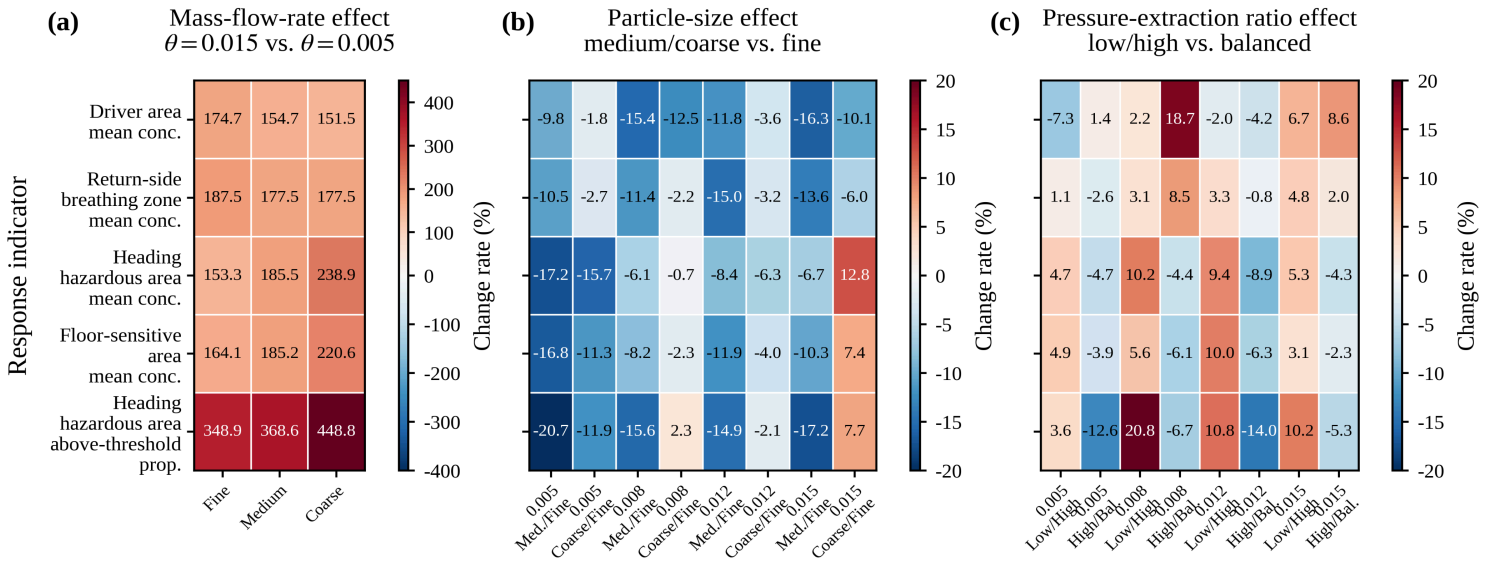


Figure 8

Response heatmap of core regional indicators under multiparameter variations.

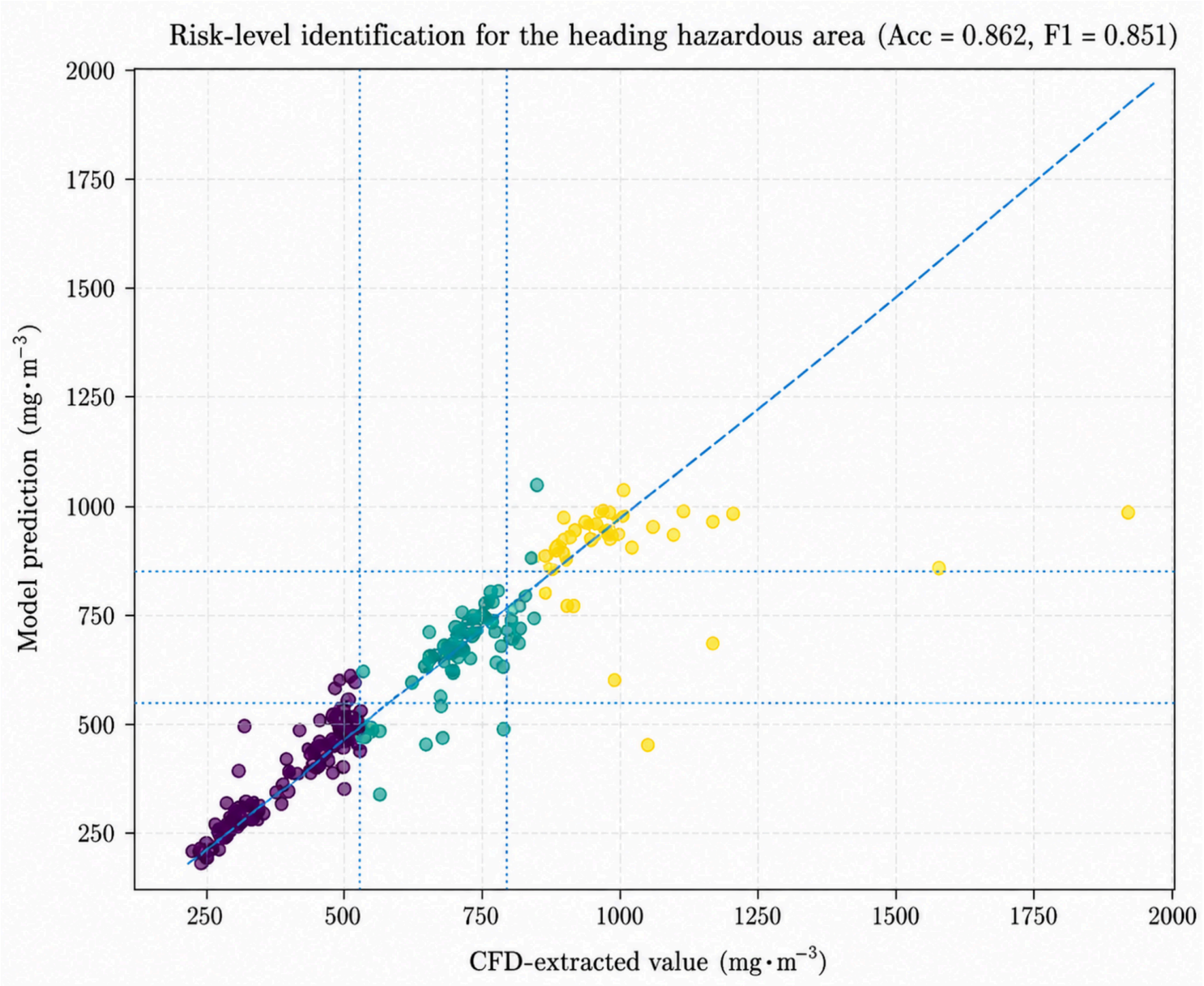


Figure 9

Risk-level identification results for the heading hazardous zone.

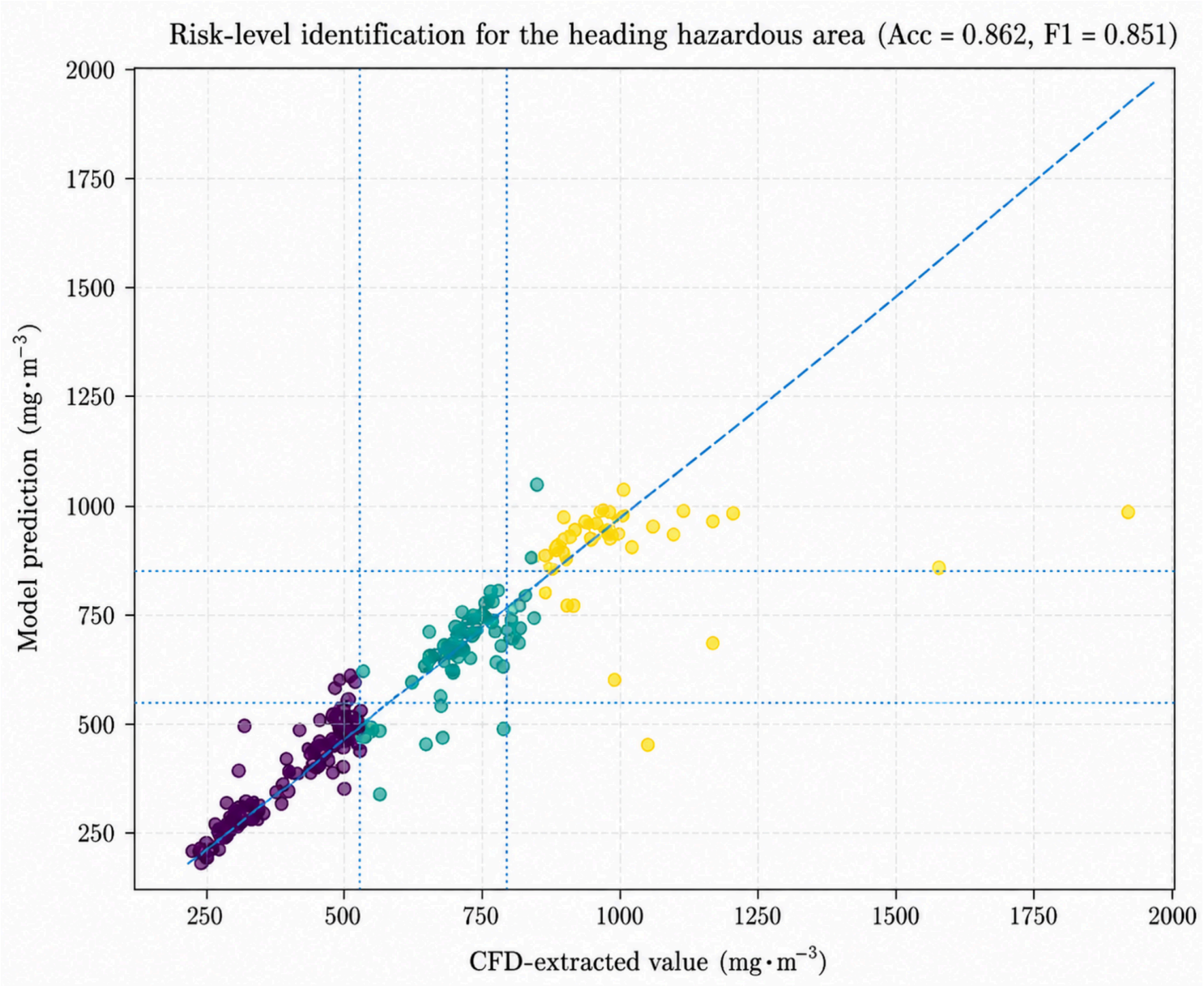


Figure 9

Risk-level identification results for the heading hazardous zone.

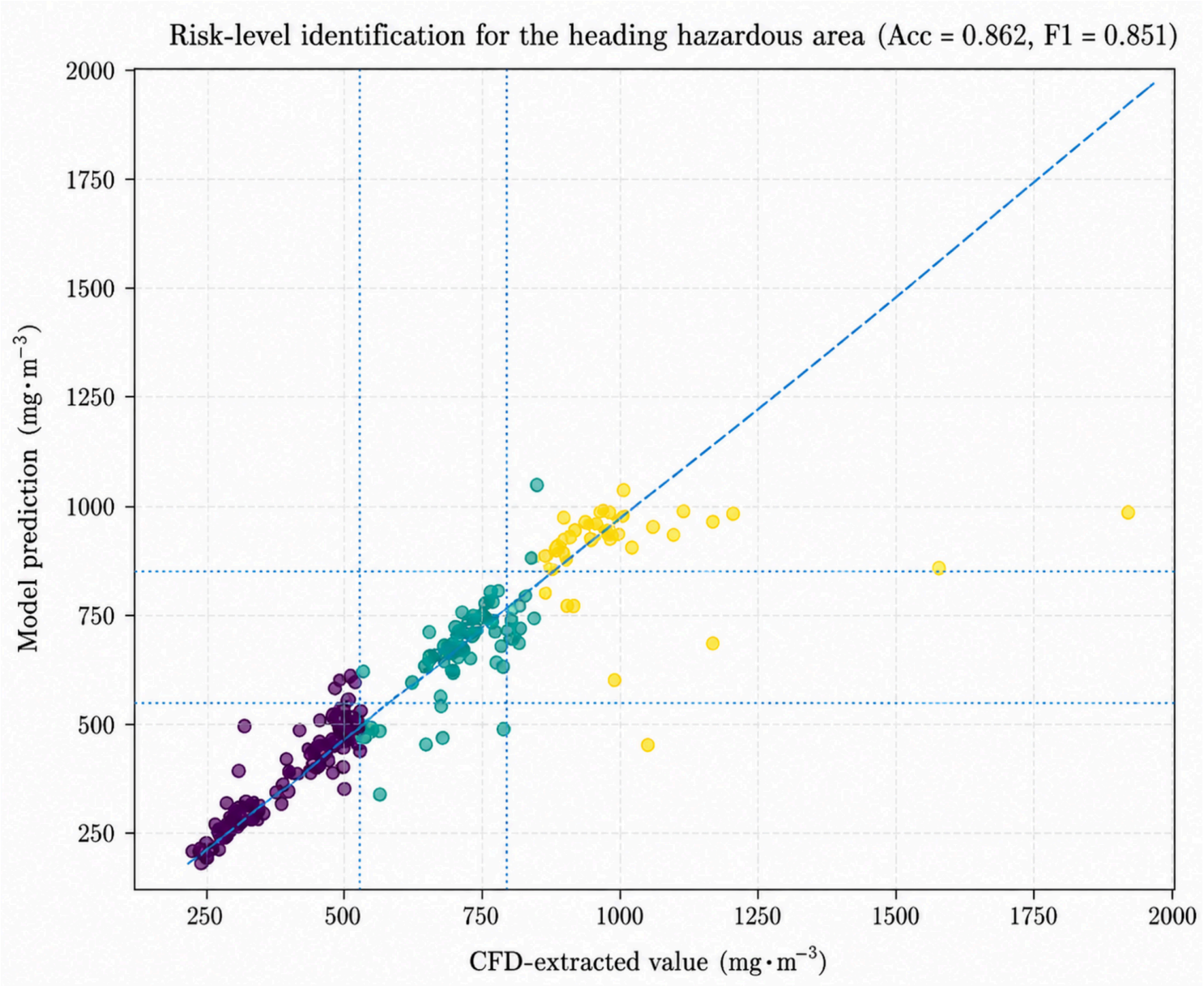


Figure 9

Risk-level identification results for the heading hazardous zone.

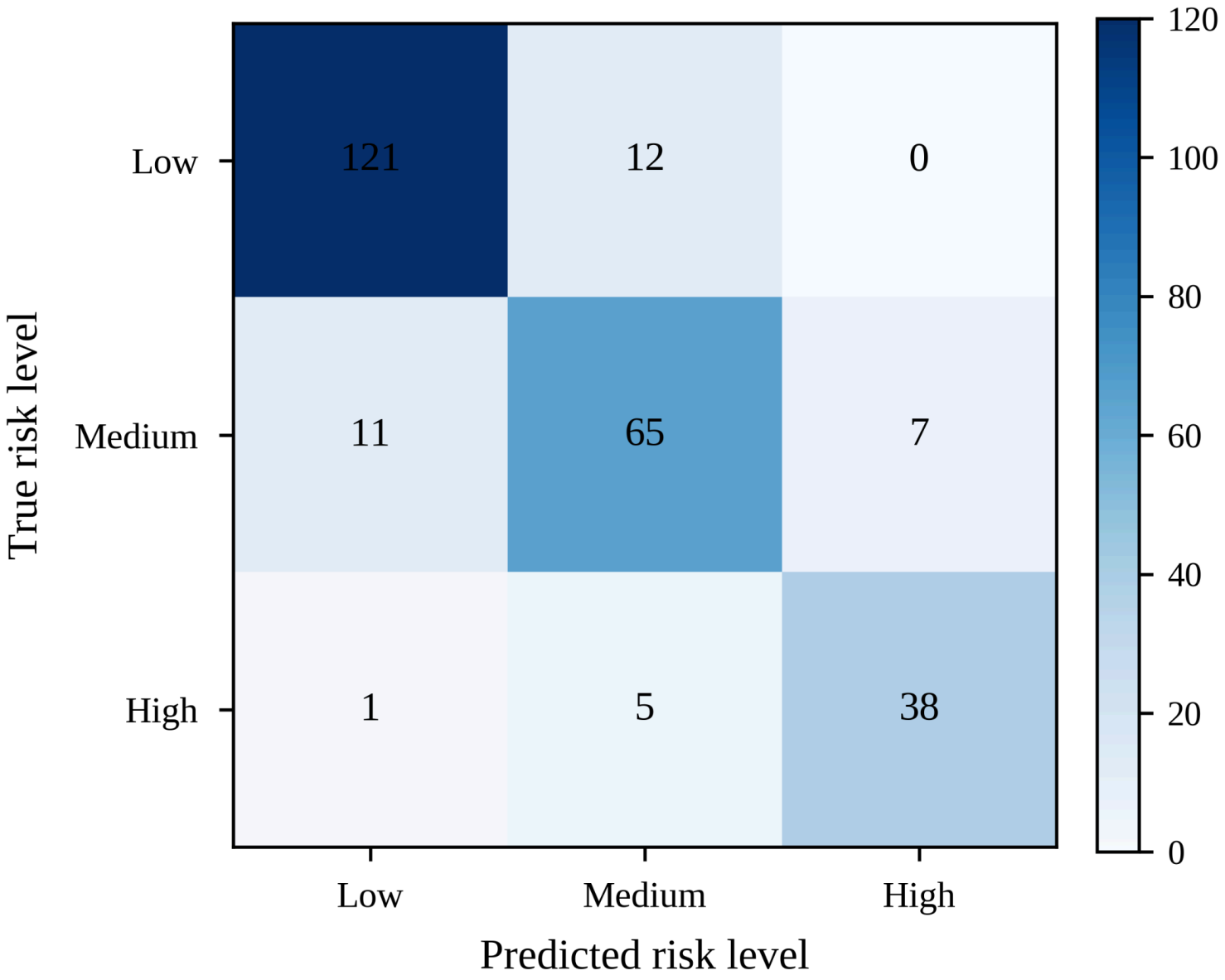


Figure 10

Confusion matrix of risk-level identification.

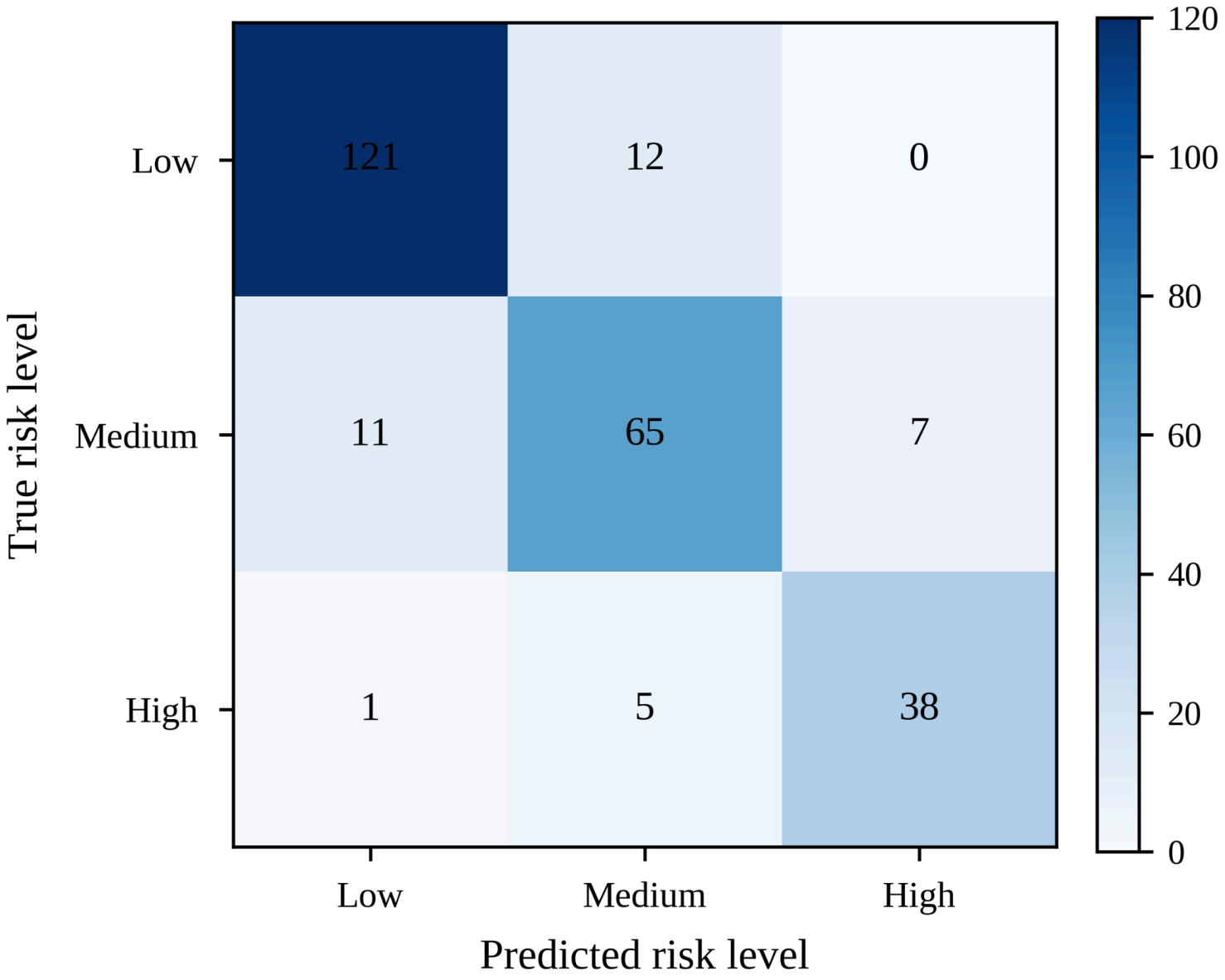


Figure 10

Confusion matrix of risk-level identification.

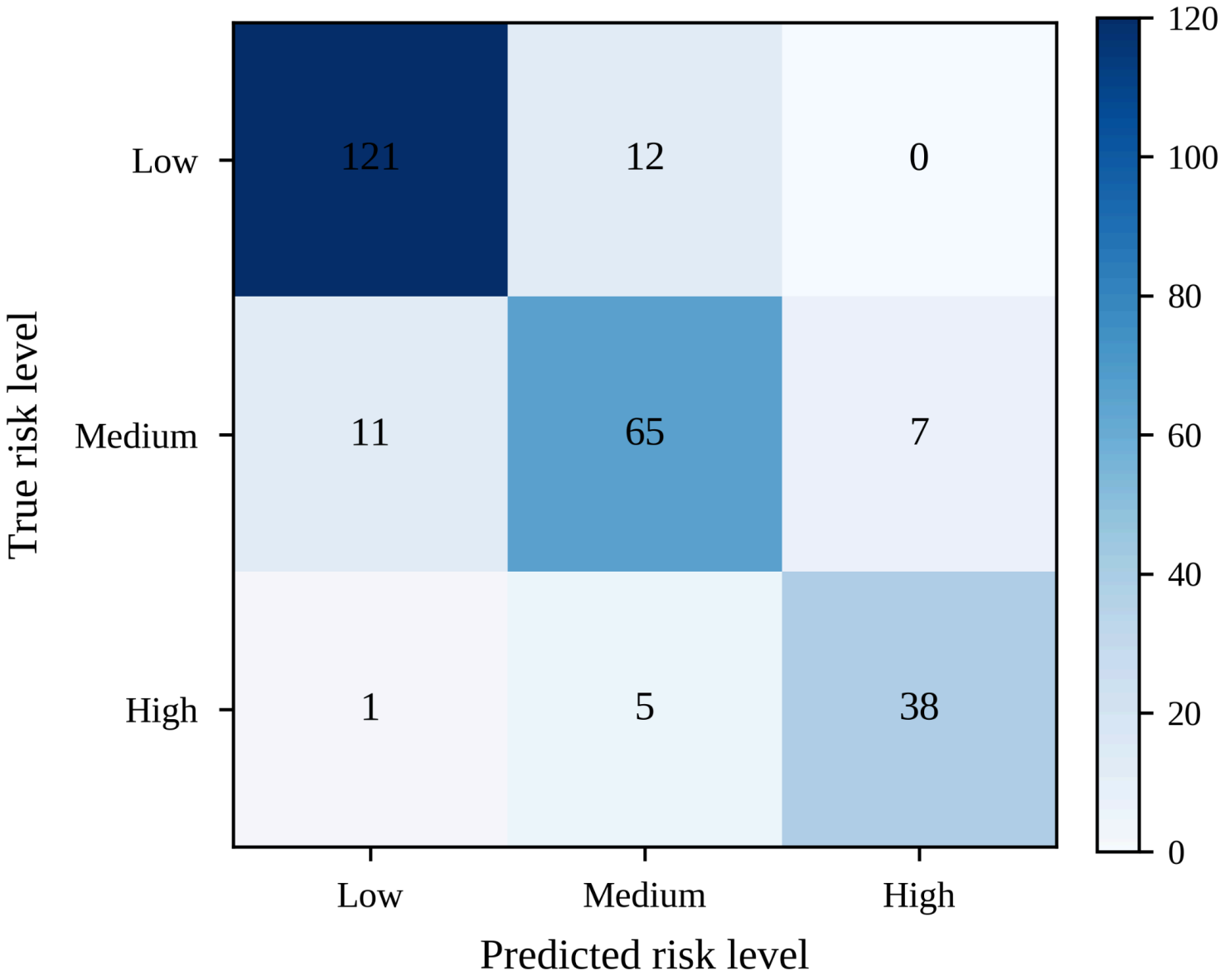


Figure 10

Confusion matrix of risk-level identification.